МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
Харьковский национальный университет радиоэлектроники

НАУКОЕМКИЕ ТЕХНОЛОГИИ В ИНФОКОММУНИКАЦИЯХ: ОБРАБОТКА ИНФОРМАЦИИ, КИБЕРБЕЗОПАСНОСТЬ, ИНФОРМАЦИОННАЯ БОРЬБА

Монография

Под общей редакцией В. М. Безрука, В. В. Баранника

Харьков
Издательство «ЛИДЕР»
2017

УДК 004.56+004.67
ББК 32.973.20.-018.2
Б 39

Рецензенты:
О.Е. Федорович – докт. техн. наук, проф., зав. кафедрой информационно-управляющих систем Национального аэрокосмического университета «ХАИ»;
С.И. Приходько – докт. техн. наук, проф., проректор по научно-педагогической работе Украинского государственного университета железнодорожного транспорта.

Печатается по решению научно-технического совета Харьковского национального университета радиоэлектроники (протокол № 3 от 19.05.2017 г.)

Б 39 Наукоемкие технологии в инфокоммуникациях: обработка информации, кибербезопасность, информационная борьба : Монография / под общей редакцией В. М. Безрука, В. В. Баранника. – Х. : Издательство «Лидер», 2017. – 600 с.
ISBN 978-966-2732-78-8

Коллективная монография содержит материалы по актуальным направлениям наукоемких инфокоммуникационных технологий. Рассматриваются вопросы планирования и управление в инфокоммуникационных сетях, эффективного хранения, обработки, интеллектуализации инфокоммуникационного простанства, распознавания образов, распределенной обработки информации и облачных вычислений, многопрофильного кодирования, кибербезопасности и информационной борьбы с использованием инфокоммуникаций.

УДК 004.56+004.67
ББК 32.973.20.-018.2

ISBN 978-966-2732-78-8

СОДЕРЖАНИЕ

Часть 2 Обработка информации в инфокоммуникациях, ителлектуализированое кодирование и многомерные синтаксические преобразования

ПРЕДИСЛОВИЕ

Стремительное развитие отраслей информатизации и коммуникации обусловило кардинальные изменения в мировоззрении человека, общества, государственных структур. Массовая компьютеризация, развитие и внедрение новейших информационно-телекоммуникационных технологий, привели к поражающему рывку ведущих стран мира в разных областях жизни, включая производство, управление, образование, медицину, финансы, банки, торговлю.

Интегральным результатом процессов информатизации являются глобальные качественные изменения в жизни человечества, которые заключаются в трансформации современного общества от постиндустриального к информационному и в глубоких изменениях современной картины мира.

Формируется и ускоренно развивается такая интеграционная отрасль как инфокоммуникации, образующаяся на стыке информационных технологий телекоммуникационных и радиотехнических систем, социологии. Понятие «инфокоммуникации» является доминирующим при обсуждении современного уровня научно-технического прогресса общества. Инфокоммуникации – это совокупность методов и средств накопления, хранения, обработки, защиты и передачи информации в пространстве.

Развитие инфокоммуникаций является необходимым условием создания информационного обмена и построения информационной структуры общества. Здесь требуется учитывать постулат относительно того, что Украина, в силу ее геополитического расположения, является объектом интересов многих государств. Это обусловливает сущность втягивания нашего государства в информационное противостояние. Соответственно остро стоит необходимость разработки собственных технологий и методов информационного противоборства. Это является важным фактором достижения информационного баланса и превосходства, что в конечном итоге обеспечивает выполнения национальных интересов и как вытекающая отсюда киберагресивность современного инфокоммуникационного пространства.

Наличие проблем информатизации общества, информационной безопасности, которые составляют важную компоненту в общей системе национальной безопасности государства, определяет необходимость учитывать их на нынешнем этапе развития инфокоммуникаций.

В современном мировом пространстве для Украины чрезвычайно актуальной является проблема построения и развитие собственных инфокоммуникационных технологий и методов. В этом направлении должен учитываться системный подход относительно проблемных вопросов в построении отрасли инфокоммуникаций, формирования нормативно-

правовой базы и подготовки специалистов. Соответствующее отражение такие процессы получают в определении направлений и приоритетов во внешней и внутренней политике государства.

Это дает мощный толчок относительно развития наукоемких технологий в инфокоммуникациях. Наукоемкость технологий обусловлена использованием современного математического аппарата и новых достижений в области высокоинтегрированных вычислительных средств для достижения информационного превосходства в направлении реализации национальных интересов.

Собственно отражению вопросов разработки и внедрения в инфокоммуникациях новых наукоемких технологий и посвящена данная коллективная монография, что подчеркивает ее актуальность.

В коллективную монографию вошли материалы ведущих ученых, представителей научных школ Киева, Харькова, Одессы, Львова, Черкасс, Винницы, развивающих наукоемкие технологии обработки информации, кибербезопасности, защиты информации и информационной борьбы в инфокоммуникационном пространстве.

Материал монографии условно разделен на три части, содержащих самостоятельные разделы отдельных авторов.

Первая часть книги включает разделы по общим проблемным вопросам инфокоммуникаций, предоставления информационных услуг, принятия оптимальных решений в задачах инфокоммуникаций, распределенной обработки информации и облачных вычислений.

Вторая часть книги посвящена проблемным вопросам создания и развития новых подходов по интеллектуальной обработки многоуровневой информации, многопрофильного объектного кодирования и представления видеоинформационных ресурсов, распознавания образов и идентификации объектов на уровне синтаксических структур.

Третья часть книги содержит результаты исследований в области кибербезопасноти, защиты информации и информационной борьбы. Включены новые подходы относительно защиты статических и динамических информационных ресурсов для систем критической инфраструктуры государства, основанные на принципах интегрирования нескольких научно-прикладных направлений, вопросы оценки рисков и угроз в информационных пространствах, методы выявления и распознавания информационно-пихологических атак третьего поколения на социум в инфокоммуникационном пространстве.

Научные редакторы сборника проф. Баранник В.В., проф. Безрук В.М.

ЧАСТЬ 1
ОБЩИЕ ВОПРОСЫ ИНФОКОММУНИКАЦИЙ

МЕТОДОЛОГИЧЕСКАЯ БАЗА ПОСТРОЕНИЯ АЛГОРИТМОВ КОНТРОЛЯ БИТОВОЙ СКОРОСТИ ВИДЕОПОТОКА С ОБЕСПЕЧЕНИЕМ ТРЕБУЕМЫХ ХАРАКТЕРИСТИК КАЧЕСТВА

Баранник В.В., Твердохлеб В.В., Хаханова А.В., Харченко Н.А.

Введение

На сегодняшний день современные инфокоммуникации характеризуются стремительным ростом объема передаваемых видеоданных. Так же происходит постоянный рост числа пользователей систем видеоконференций и сервисов трансляции потокового видео. Одновременно с этим, увеличение пропускной способности каналов запаздывает, что является причиной частых перегрузок сетей.

В таких условиях возможность адаптируемости интенсивности видеопотока к пропускной способности канала является актуальной.

Данная возможность способна обеспечить эффективную передачу видеопотока, предотвратить возникновение потерь и задержек передачи видеоданных.

Целью данного исследование является построение методики управления битовой скоростью видеопотока с целью согласования ее величины с пропускной способностью канала инфокоммуникационной сети.

Основными задачами построения метода управления контроля битовой скоростью являются: определение условий эффективной передачи видеопотока на фоне изменяющейся пропускной способности канала, построение механизма управления битовой скоростью видеопотока и контроля уровня ошибки, а также способов обеспечения быстродействия механизма управления.

1. Условия эффективной передачи видеопотока

Эффективной можно считать такую передачу видеопотока, при которой обеспечивается выполнение следующих условий:

– соответствие требованиям QoS касательно величин задержки и потерь данных;

– поддержание уровня ошибки, не превышаеющего заданного значения;

– обеспечение визуально приемлимого качества видео на приеме.

Таким образом, наряду с управлением битовой скоростью R, необходимо также обеспечить значение ошибки, в качестве которой будем

рассматривать уровень среднеквадратического отклонения, на требуемом уровне [1].

Принимая во внимание тот факт, что величина пропускной способности изменяется во времени произвольным образом, метод управления должен быть способен обеспечивать минимальный уровень битовой скорости при минимальных значениях ошибки.

Тогда условия эффективной передачи видеопотока при изменяющейся пропускной способности канала могут быть представлены следующим образом:

$$\begin{cases} R \to \min; \\ d \le d_{\min}. \end{cases} \tag{1}$$

При полученных условиях эффективной передачи видеопотока, рассмотрим способ организации данных, дающий возможность построения управляющего метода.

2. Суть метода контроля битовой скорости

Исходный видеокадр F, после выполнения ДКП и преобразования цветовой модели RGB в модель YCbCr, рассматривается как множество P трансформант, определяемое следующим образом:

$$P = \sum_{p=1}^{Q} Y_p, \tag{2}$$

где Y_p – $p-$я трансформанта кадра.

В свою очередь, каждая трансформанта Y_p представлена совокупностью $(h;w)$ – х компонент $Y_p = \left\| y(p)_{hw} \right\|$.

Каждая компонента $y(p)_{hw}$ трансформанты Y_p представляется в двоичном виде, на основе последовательности $\alpha(p)_{hw}^{(\mu)}$ бит (рис.1). Это эквивалентно преобразованию:

$$\left\| y(p)_{hw} \right\| \to \left\| \left\langle \alpha(p)_{hw}^{(\mu)}, \alpha(p)_{hw}^{(\mu-1)} \ldots \alpha(p)_{hw}^{(m)} \ldots \alpha(p)_{hw}^{(0)} \right\rangle^{T} \right\|, \tag{3}$$

$$\alpha(p)_{hw}^{(\mu)} \in \{0,1\},\ h = \overline{0,7};\ w = \overline{0,7};\ \mu = \overline{7,0},$$

где $\alpha(p)_{hw}^{(\mu)}$ – $\mu-$й бит двоичного разложения $(h; w)-$й компоненты $p-$й трансформанты.

Множество всех бит $\mu-$го разряда $p-$й трансформанты составляет битовую плоскость $Y(p)_{\mu}$.

В свою очередь, совокупность двоичных представлений всех элементов матрицы Y_p составляет битовый куб $Y_p^{(3d)}$, пример которого представлен на рисунке 1.

При рассматриваемом способе организации данных, верхний слой данного куба образуют старшие биты $\alpha(p)_{hw}^{(\mu)}$ двоичного представления.

Представление трансформанты Y_p в трехмерном пространстве позволяет осуществлять передачу данных отдельными битовыми плоскостями, аналогично подходу, который используется методом последовательного приближения технологии Progressive JPEG [2].

В этом случае появляется возможность контролировать объем передаваемой информации в зависимости от требований пропускной способности B_w канала.

В зависимости от требуемого объема бит для представления кадра, используются либо все n битовых плоскостей $Y(p)_\mu$ трансформанты Y_p, либо только $(n-\mu)$ битовых плоскостей, чтобы обеспечить битовую скорость R_F кадра на уровне, не превышающем некоторое требуемое значение.

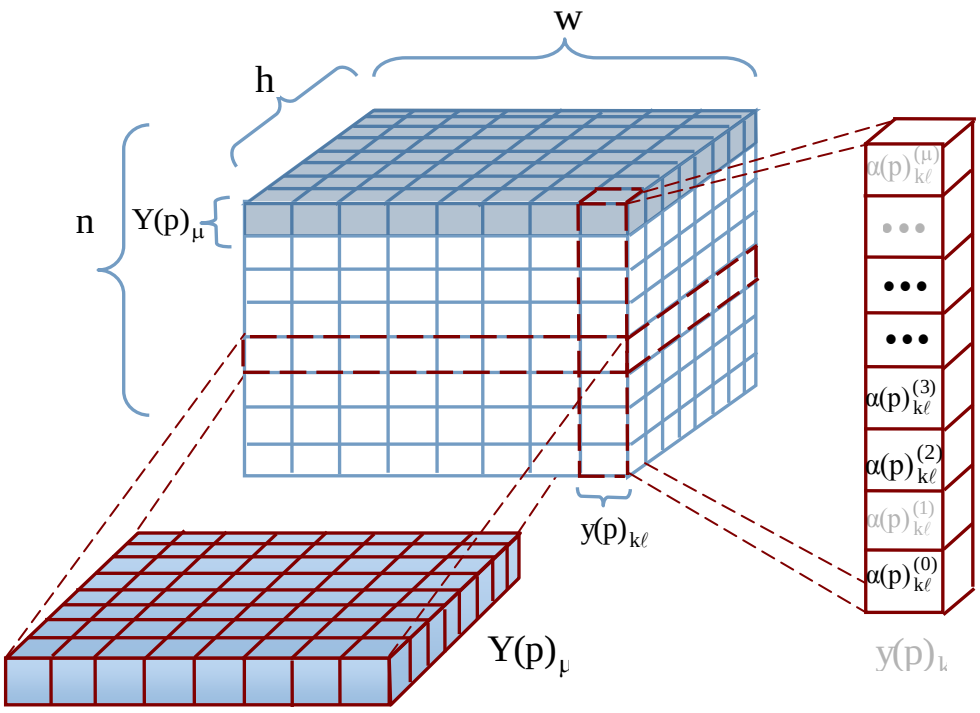


Рис. 1. Представление трансформанты Y_p *в виде битового куба* $Y_p^{(3d)}$

Таким образом, данный способ представления данных обеспечивает возможность управления компрессией и может быть использован в качестве базового.

Принцип работы метода управления на базе технологии Progressive JPEG может быть описан следующим образом:

1. Данные о величине пропускной способности канала поступают кодеру.

2. Исходя из количества кадров, которые необходимо передать в единицу времени, а также их суммарной битовой скорости, кодер оценивает возможность передачи кадров без дополнительной обработки.

3. Вычисляется разница между фактической и требуемой битовыми скоростями серии кадров.

4. Определяется требуемая битовая скорость для каждого кадра.

5. Вычисляется требуемая битовая скорость трансформант кадра. Если битовая скорость трансформанты превышает требуемую, часть битовых плоскостей исключаются из рассмотрения.

6. Формируется последовательность, состоящая из необходимого для передачи количества кадров, каждый из которых является совокупностью трансформант с битовыми скоростями, не превышающими требуемые значения.

В начальный момент времени t_0 буферное устройство отправляет в канал тестовые пакет R_{start} с известной величины.

Используя значение времени двусторонней задержки RTT, определяется величина полосы пропускания B_w в момент t_0:

$$B_w = \frac{R_{start}}{RTT}. \quad (4)$$

Также, используя значение RTT, вычисляется фактическое число кадров, которое необходимо поместить в буфер:

$$\varpi(t_0) = \varpi RTT. \quad (5)$$

При полученном значении $\varpi(t_0)$, для соответствующего количества кадров происходит оценка битовых скоростей $R_{p\mu}$ и величины ошибки $d_{p\mu}$ по всем битовым плоскостям каждой из трансформант кадра.

На этапе, предшествующем нахождению $R_{p\mu}$ и $d_{p\mu}$ битовых плоскостей трансформант, выполняется оценка насыщенности НЧ-областей трансформант [3], как показано следующей формулой:

$$\chi_{m,n} = \log_2\left(\prod_{\gamma=1}^{L}\prod_{\lambda=1}^{\Lambda} y(p)_{\gamma\lambda}\right), \quad (6)$$

где γ – количество диагоналей НЧ-области трансформанты, λ – число элементов диагонали, $y(p)_{\gamma\lambda} - \gamma, \lambda$ – я компонента НЧ-области трансформанты.

Для кадра, состоящего из $m \times n$ трансформант, величины $\chi_{m,n}$ вычисляются по всем строкам.

Если в последовательности $\overline{\chi_{m,1}, \chi_{m,n}}$ выявлены значения насыщенностей, для которых разность $|\Delta\chi| = \chi_{m,k} - \chi_{m,k+1}$ имеет несущественную величину и справедливо соотношение

$\chi_{m,1} \approx \chi_{m,k} \ldots \approx \chi_{m,k+1}$, данные трансформанты составляют вектор стабилизации $S_{j,\delta}$. Индекс j при этом определяет позицию трансформанты в кадре, а δ- количество входящих в него трансформант.

В пределах вектора стабилизации $S_{j,\delta}$, используя подобие между трансформантами, существует возможность сократить количество выполняемых арифметических операций при обработке трансформант.

В частности, часть значений битовых скоростей $R_{p\mu}$ и $d_{p\mu}$ битовых плоскостей одного разряда для трансформант вектора $S_{j,\delta}$ может быть интерполировано.

Интерполированные значения битовых скоростей $R_{p,\mu}^{инт}$ определяются формулой:

$$R_{p,\mu}^{инт} = \varphi(R_{p-1,\mu}, R_{p+1,\mu}, R_{p+m,\mu}), \qquad (7)$$

где $R_{p-1,\mu}, R_{p+1,\mu}$ и $R_{p+m,\mu}$- значения битовых скоростей, соответствующие $Y^{(p-1,\mu)}$, $Y^{(p,\mu)}$ и $Y^{(p+m,\mu)}$ битовым плоскостям,

Интерполяция значений $d_{p,\mu}^{инт}$ при этом аналогична (7) и определяется выражением:

$$d_{p,\mu}^{инт} = \varphi(d_{p-1,\mu}, d_{p+1,\mu}, d_{p+m,\mu}), \qquad (8)$$

где $d_{p-1,\mu}, d_{p+1,\mu}$ и $d_{p+m,\mu}$- значения битовых скоростей для $Y^{(p-1,\mu)}$, $Y^{(p,\mu)}$ и $Y^{(p+m,\mu)}$ битовых плоскостей.

Итоговая битовая скорость и СКО для серии из $\varpi(t_0)$ кадров определяется следующими выражениями:

$$R_{seq} = \sum_{i=1}^{\varpi}\sum_{p=1}^{Q} R_{p,i}, \qquad d_{seq} = \sum_{i=1}^{\varpi}\sum_{p=1}^{Q} d_{p,i}, \qquad (9,10)$$

где $R_{p,i} = \sum_{\mu=1}^{n} R_{p\mu}$ – битовая скорость $p-$й трансформанты $i-$го кадра,

$d_{p,i} = \sum_{\mu=1}^{n} d_{p\mu}$ – уровень СКО $p-$й трансформанты $i-$го кадра

При полученных значениях R_{seq} и d_{seq} для последовательности $\varpi(t_0)$ кадров определяется разность $\Delta R = B_w - R_{seq}$ между суммарной фактической битовой скоростью кадров серии и требуемой битовой скоростью, величина которой равна $R_{seq}^{треб} = B_w$ [4].

Если $\Delta R \le 0$, то вся последовательность $\varpi(t_0)$ передается в буфер передатчика без дополнительной обработки.

В случае, когда $\Delta R < 0$, битовую скорость R_{seq} необходимо снизить на величину $|\Delta R|$ для обеспечения требуемой битовой скорости $R_{seq}^{треб}$ последовательности кадров $\varpi(t_0)$.

Используя значение величины требуемой битовой скорости $R^{треб}$, соотношение (1) можем представить в следующем виде:

$$\begin{cases} R \le R^{треб}; \\ d \le d_{min}. \end{cases} \tag{11}$$

Очевидно, что величина ΔR определяется следующим выражением:

$$\Delta R = \sum_{i=1}^{\varpi(t_0)} \Delta R_i, \tag{12}$$

где ΔR_i – величина, на которую необходимо снизить битовую скорость каждого кадра последовательности $\varpi(t_0)$.

В то же время, распределение битовых скоростей в серии $\varpi(t_0)$ кадров имеет неравномерный характер. Как правило, серия $\varpi(t_0)$ состоит из кадров различных типов - I, B и P, при этом максимум битовых скоростей соответствует I и P кадрам.

Для такого случая, снижение величины $|\Delta R|$ достигается путем уменьшения битовых скоростей B – кадров последовательности. Обусловлено это тем, что потеря части информации в кадрах данного типа внесет минимальную ошибку в суммарное значение СКО последовательности из $\varpi(t_0)$ кадров при восстановлении.

Величина ΔR_i, на которую необходимо снизить битовую скорость каждого из входящих в последовательность B – кадров, в этом случае определяется следующей формулой:

$$\Delta R_i = \frac{\Delta R}{\varpi(t_0)}, \quad i = \overline{1, j} \tag{13}$$

где j – количество B – кадров в рассматриваемой последовательности

В то же время, последовательность кадров в начале видеопотока, а также сцены с высокой динамикой, могут содержать только I – кадры, обладающие высокой интенсивностью, либо совокупность I и P кадров.

В этом случае вычисление величины ΔR_i производится пропорционально величинам битовых скоростей каждого кадра последовательности $\varpi(t_0)$, как показано выражением:

$$\Delta R_i = \frac{\Delta R}{\varpi} \cdot \frac{R_i}{R_{cp}} = \frac{\Delta R \, R_i}{R_{seq}}, \tag{14}$$

где $\varpi(t_0)$ – число кадров в серии, R_i – битовая скорость i–го кадра,

$R_{cp} = \dfrac{R_{seq}}{\varpi(t_0)}$ – средняя битовая скорость кадра в серии.

В свою очередь, требуемая битовая скорость кадра последовательности $\varpi(t_0)$ определяется формулой:

$$R_i^{треб} = R_i - \Delta R_i . \qquad (15)$$

После того, как величина $R_i^{треб}$ для $i-$го кадра найдена, определяется значение требуемых битовой скоростей для трансформант кадра, в сумме дающих величину битовой скорости кадра, равную $R_i = \sum_{p=1}^{Q} R_p^{треб} \le R_i^{треб}$.

При определении требуемой битовой скорости трансформанты кадра используются следующие подходы.

В случае, когда в пределах кадра величины $\chi_{m,n}$ для всех трансформант отличаются незначительно, применяется подход, использующий усреднение битовой скорости по кадру:

$$R_p^{тр} = \frac{R_i^{треб}}{Q}, \qquad (16)$$

где $Q-$ число трансформант в кадре.

Для более точного учета битовой скорости трансформант в пределах кадра, используется подход, учитывающий характер распределения битовой скорости в кадре:

$$R_p^{треб} = \frac{R_i^{треб}}{Q}\gamma_p, \qquad (17)$$

где γ_p – коэффициент, зависящий от степени насыщенности $p-$й трансформанты.

После того, как для каждой трансформанты кадра найдены величины $R_p^{треб}$, определяются битовые плоскости, которые будут исключены, чтобы обеспечить значения битовых скоростей в соответствии с формулами (16, 17).

Для битовых плоскостей трансформант, битовые скорости которых необходимо снизить до величины $R_p^{треб}$, определяется порядок ранжирования, при котором первыми обрабатываются битовые плоскости $Y^{(p,\mu)}$, вносящие максимальные значения $d_{p\mu}$ в общее СКО трансформанты.

В первую очередь, это относится к старшим битовым плоскостям $Y^{(p,\mu)}$. Далее обрабатываются битовые плоскости в порядке снижения величин $d_{p\mu}$, вносимых ими в общее СКО.

После определения ранжирования битовых плоскостей, на каждом $p-$м шаге происходит вычисление суммарных значений СКО и битовой скорости трансформант $Y_p^{(\mu)}$ и $Y_{p+1}^{(\mu)}$, путем сложения значений $R_{p,\mu}$ и $R_{p+1,\mu}$ а также $d_{p,\mu}$ и $d_{p+1,\mu}$.

Сложение значений СКО и битовых скоростей трансформант $Y_p^{(\mu)}$ и $Y_{p+1}^{(\mu)}$ происходит попарно, с учетом порядка обработки. Суммируются при этом только величины, имеющие одинаковые индексы ранжирования.

На каждом $p-$м шаге вычисления требуемой битовой скорости трансформанты определяется условно-оптимальное $R_{F,p}^*$ значение битовой скорости кадра, состоящего из p трансформант, исходя из условий:

$$\begin{cases} R_{i,p}^* \in \left\{R_{i,p}^*\right\} \mid R_{i,p} \le R_p^{\text{треб}}; \\ d_{i,p} \to \min. \end{cases} \tag{18}$$

Величина $R_{F,p+1}$ для кадра F на $p+1-$м шаге в этом случае будет определяться следующим способом:

$$R_{F,p+1} = \sum_{i=p}^{p+1} \sum_{u=1}^{n} R_{i,u}, \tag{19}$$

где $u-$индекс очередности обработки битовой плоскости, $R_{i,u}-$битовая скорость $u-$й битовой плоскости трансформанты в порядке снижения вносимого уровня СКО,

В свою очередь, СКО на $p+1-$м шаге будет определяться выражением:

$$d_{F,p+1} = \sum_{i=p}^{p+1} \sum_{u=1}^{n} R_{i,u}, \tag{20}$$

где $d_{i,u}-$битовая скорость $u-$й битовой плоскости трансформанты.

В результате сложения битовых скоростей и СКО по трансформантам $Y_p^{(\mu)}$ и $Y_{p+1}^{(\mu)}$ на $p+1$ шаге, результирующий порядок обхода полученного множества будет определяться суммарными значениями СКО по уменьшению, как показано на рис. 2.

При таком способе перераспределения данных, крайними в кодограмме трансформанты располагаются битовые плоскости, вносящих минимальные величины СКО для трансформанты.

Это позволяет более эффективно производить коррекцию битовой скорости трансформант, исключая необходимый объем битовых плоскостей для снижения битовой скорости трансформанты, не прибегая к дополнительным операциям.

Действия, описанные выражениями (16-20), выполняются на каждом из Q шагов для всех $\varpi(t_0)$ кадров серии. Результирующая битовая скорость R_{seq} последовательности кадров, подлежащих передаче при этом будет эквивалентна следующему выражению:

$$R_{seq}^{треб} = \sum_{i=1}^{v} \sum_{p=1}^{Q} R_{i,p}^{треб}, \qquad (21)$$

где $R_{p,i}^{треб}$ - полученное в соответствии с (16, 17) значение битовой скорости p-й трансформанты i-го кадра последовательности.

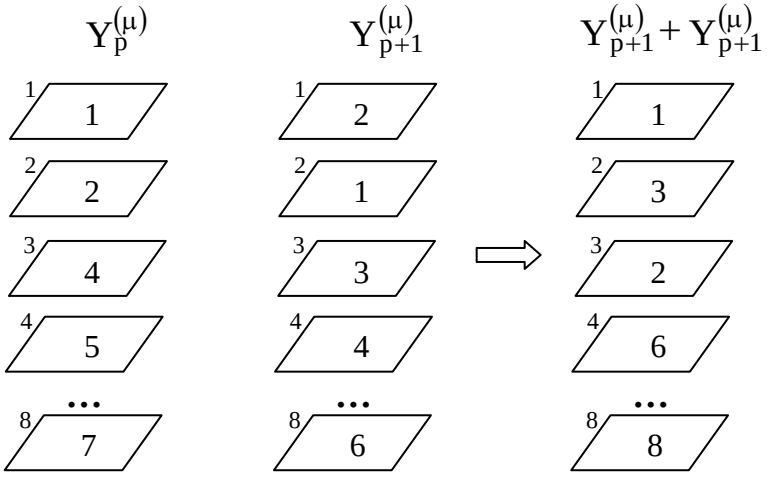


Рис. 2. Порядок обхода отдельных трансформант $Y_p^{(\mu)}$ и $Y_{p+1}^{(\mu)}$ а также их суммарного множества, полученного в результате, в зависимости от величин СКО битовых плоскостей

Выбор на каждом p−м шаге вычисления требуемой битовой скорости трансформанты в соответствии с условиями (18) гарантирует, что уровень СКО, соответствующий полученной последовательности из $\varpi(t_0)$ кадров, будет минимально возможным в момент времени t_0.

Далее, сформированная последовательность $\varpi(t_0)$ кадров, помещается в выходной буфер, размер которого рассчитывается согласно следующей формуле:

$$R_{буф} = \varpi RTT\, R_{ср}. \qquad (22)$$

После помещения $\varpi(t_0)$ кадров в буфер происходит отправка всей серии кадров в канал, а также определение величины RTT, по величине которой, определяется величина полосы пропускания B_w в момент t_1, в соответствии с (4), а также количество кадров $\varpi(t_1)$, из которого будет состоять передача в момент t_1.

Повышению быстродействия рассматриваемого метода управления битовой скоростью способствует учет особенностей распределения уровня

СКО в пределах каждой трансформанты. Величина ошибки в пределах трансформанты распределяется таким образом, что старшим битовым плоскостям соответствуют высокие уровни $d_{p\mu}$ [5]. В то же время, распределение битовой скорости между битовыми плоскостями в трансформанте имеет случайный характер. Данные особенности представлены следующим соотношением:

$$\begin{cases} d_{p,1} < d_{p,2} < d_{p,3} < \dots d_{p,n} \\ R_{p,\mu} = \text{var} \end{cases}. \qquad (22)$$

Используя значения $R_{p\mu}$ и $d_{p\mu}$, полученные для трансформанты $Y_p^{(\mu)}$, а также зависимость (21), для трансформанты $Y_p^{(\mu)}$ могут быть определены битовые плоскости, которые могут быть исключены еще на подготовительном этапе работы метода. Такими являются битовые плоскости $Y^{(p,\mu)}$, соответствующие значениям битовых скоростей, стремящихся к R_p^{max}, при незначительных величинах СКО.

При этом, если для трансформанты $Y_p^{(\mu)}$ выявлены битовые плоскости, которые могут быть исключены, отбрасываются также битовые плоскости соседних трансформант вектора $S_{i,m}$, имеющих одинаковые индексы разрядности, если их интерполированные значения $R_{p\mu}$ и $d_{p\mu}$ отличаются от величин $R_{p\mu}$ и $d_{p\mu}$ незначительно.

Определение множества $Y_p^{(\mu)'}$ отсекаемых БП трансформанты $Y_p^{(\mu)}$ при этом может быть описано выражением:

$$\begin{gathered} Y_p^{(\mu)'} = \left\{ Y^{(p,\mu)} \right\} \mid R_{p,\mu} \to R_p^{max}, d_{p,\mu} \to d_p^{min} \\ R_p^{max} = \max\left\{ R_{p,\mu} \right\}, d_p^{min} = \min\left\{ d_{p,\mu} \right\} \end{gathered}, \qquad (23)$$

где R_p^{max} – максимальное значение битовой скорости битовой плоскости трансформанты $Y_p^{(\mu)}$, d_p^{min} – минимальное значение СКО битовой плоскости трансформанты $Y_p^{(\mu)}$.

Кроме подхода, заключающегося в усечение битовых плоскостей одного разряда, при незначительном отличии их интерполированных величин $R_{p\mu}$ и $d_{p\mu}$, с целью снижения общего объема вычислений может быть использовано подобие между трансформантами.

Суть данного подхода заключается в следующем. Если для интерполированных значений СКО и битовых скоростей по всем битовым плоскостям трансформант $Y_{p-1}^{(\mu)}$, $Y_p^{(\mu)}$ и $Y_{p+m}^{(\mu)}$ вектора $S_{j,m}$ справедливо:

$$\begin{cases} R_{m,\mu} \approx R_{m+1,\mu} \approx ... \approx R_{m+g,\mu} \\ d_{m,\mu} \approx d_{m+1,\mu} \approx ... \approx d_{m+g,\mu} \end{cases}, \quad \mu = \overline{1,n}, \tag{24}$$

где $m-$ индекс трансформанты в векторе $S_{j,m}$, то вычислений сочетаний битовых плоскостей для трансформант $Y_p^{(\mu)}$ и $Y_{p+m}^{(\mu)}$ вектора $S_{j,m}$ не производится. В этом случае, вместо обработки трансформант на $p-$м и $p+g-$м шагах, кодеру сообщается признак подобия, определяющий стратегию обработки, которая будет применяться к трансформантам.

Выводы

Разработана методологическая база построения алгоритмов управления интенсивностью битовой скорости видеопотока, способствующая адаптации интенсивности транслируемого видео к изменяющейся пропускной способности канала.

Управляемыми параметрами видеопотока при этом являются битовые плоскости трансформант кадра, а именно – их количество и позиция в трансформанте, в зависимости от величины требуемой битовой скорости последовательности кадров, транслируемых в единицу времени.

Учет степени информативности трансформант в пределах каждого кадра передаваемой последовательности, а также типа обрабатываемого кадра, способствует снижению вносимой погрешности при обработке, минимизируя ее.

Рассмотрены условия и способы обеспечения эффективной передачи видеопотока с использованием рассмотренной методики. С целью повышения производительности управляющего метода, предложены пути повышения быстродействия, основанные на снижении количества информации, подлежащей обработке, за счет исключения из рассмотрения наименее информативных составляющих трансформант кадров, дающих минимальный уровень СКО на стороне приема.

Литература

1. Баранник В., Двухглавов Д.,Твердохлеб В. Метод динамического управления битовой скоростью видеопотока с использованием трехмерного представления трансформант. // Автоматизированные системы управления и приборы автоматики, 2014. – №176. – С. 37 – 43.
2. Сэломон, Д. Сжатие данных, изображений и звука / Д. Сэломон. – М.: Техносфера, 2004. – 368 с.
3. Баранник В.В. Методологические рекомендации по совершенствованию технологии снижения интенсивности кодового представления базовых кадров / В.В. Баранник, О.Ю. Отман Шади, А.А. Подорожняк // Системи обробки інформації, 2014. – № 8(124). – С. 87-92.
4. Ян Ричардсон. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. – Москва: Техносфера, 2005. – 368 с.
5. Гонсалес Р.С. Цифровая обработка изображений/ Р.С. Гонсалес, Р.Э. Вудс. – М.:Техносфера, 2006. – 1072 с.

МНОГОКРИТЕРИАЛЬНЫЙ ВЫБОР ПРОЕКТНЫХ ВАРИАНТОВ СИСТЕМ МОБИЛЬНОЙ СВЯЗИ 3 и 4 ПОКОЛЕНИЯ

Безрук В.М., Чеботарева Д.В., Скорик Ю.В.

1. Общая характеристика задач многокритериального выбора оптимальных проектных решений

Рассмотрим особенности многокритериального выбора оптимальных проектных решений с учетом совокупности показателей качества на основе применения методов многокритериальной оптимизации [1-8]. Считают, что выбор оптимальных проектных решений выполняет некоторое лицо, принимающее решение (ЛПР), которое преследует вполне определенные цели. В зависимости от конкретной ситуации в роли ЛПР может выступать как отдельный человек (инженер-проектировщик), так и целый коллектив (группа специалистов, занятая решением задачи проектирования системы).

Каждое возможное проектное решение характеризуется определенной степенью достижения поставленной цели. В соответствии с этим у ЛПР имеется свое представление о достоинствах и недостатках проектных решений, на основании которого одно решение предпочитается другому. Оптимальное проектное решение – это решение, которое с точки зрения ЛПР предпочтительнее других возможных решений. Таким образом, понятие оптимального решения связано с предпочтениями ЛПР. Формализация задачи принятия оптимальных проектных решений и состоит в выборе математических средств, которые помогли бы ЛПР наиболее полно и точно выразить свои предпочтения в рамках соответствующей математической модели выбора. Это позволить сформулировать соответствующую оптимизационную задачу и в конечном итоге обосновано выбирать действительно оптимальные проектные решения.

Предпочтения ЛПР на практике выражаются в различной форме и их математическая формализация может составить непростую задачу. Это связано с тем, что ЛПР, как правило, не может ясно и четко сформулировать свои предпочтения в таком виде, чтобы их можно было описать некоторыми математическими категориями и применить некоторою формализованную процедуру выбора оптимальных вариантов системы.

Таким образом, оптимальное проектное решение – это наилучшее решение, которое с точки зрения ЛПР предпочтительнее других проектных решений. Задание критерия оптимальности для выбора наилучших решений из множества допустимых проектных решений связано с формализацией понимания ЛПР про оптимальность проектных решений. При этом существуют два подхода к определению понятия оптимальности решений: ординалистический и кардиналистический.

Ординалистический подход к определению понятия оптимальности решений аппелирует к порядку (лучше-хуже) и основан на введении понятия бинарных отношений, что позволяет формализовать операции попарного сравнения альтернатив и выбора оптимальных решений. В частном случае, когда при выборе «наилучших» (оптимальных) решений ЛПР руководствуется отношением строгого предпочтения $\succ$, из всего из множества возможных решений X выделяются решения, недоминируемые по этому бинарному отношению. В множество оптимальных по отношению строгого предпочтения $\succ$ решений включают такие решения $x^{(o)} \in X$, для которых не существует других предпочтительных решений $x \in X$, чтобы было справедливо бинарное отношение $x \succ x^{(o)}$. Это множество называется множеством Парето-оптимальных решений, которое обозначается через $\mathrm{opt}_{\succ} X$ или $P(X)$. Остальные решения являются безусловно худшими, поскольку для них имеются более предпочтительные альтернативные решения. Они удаляются из множества X, поскольку их заведомо нельзя считать оптимальными. В результате исключения из множества X безусловно худших решений по бинарному отношению $\succ$ останутся только те оптимальные решения, для которых выполняется отношение неразличимости.

Кардиналистический подход к определению понятия оптимальности решений основан на введении некоторой целевой функции $f(x)$, значение которой интерпретируется как полезность (ценность) решения x и оно определяет предпочтение ЛПР. Выбранная целевая функция задает соответствующее отношение порядка на множестве X. Значение целевой функции $f(x)$ является индикатором предпочтения ЛПР. В частности, при задании скалярной целевой функции считается, что решение x' предпочтительнее альтернативного решения x'' тогда и только тогда, когда выполняется условие $f(x') \geq f(x'')$. При таком подходе может быть задана формализованная процедура выбора оптимальных решений $x^{(o)}$ из условия экстремума целевой функции $f(x)$ на множестве X.

Однако из-за недостаточно строгой определенности представления ЛПР про оптимальность проектных решений часто не удается задать скалярную целевую функцию и соответствующий скалярный критерий оптимальности. Поэтому решения характеризуют векторной целевой функцией, включающей совокупностью частных целевых функций $f_1(x), f_2(x), ..., f_m(x)$, которые определяют полезность (ценность) решения x с точки зрения разных требований. В зависимости от содержания задачи выбора эти функции называют критериями оптимальности, критериями эффективности, целевыми функциями, показателями или критериями качества системы. Совокупность целевых функций образуют векторный критерий оптимальности

$$\vec{f}(x) = (f_1(x), f_2(x), ..., f_m(x)), \qquad (1)$$

который принимает значения в пространстве m-мерных векторов R^m.

В случае введения векторной целевой функции (1), наряду с множеством допустимых решений X, рассматривают множество соответствующих им оценок значений этой векторной функции

$$Y = \vec{f}(X) = \{\vec{y} \in Y \mid \vec{y} = \vec{f}(x),\ x \in X\},\ Y \subset R^m, \qquad (2)$$

которое называют также множеством векторных оценок или критериальным пространством. При этом возникают более сложные задачи оптимизации проектных решений по совокупности показателей качества, которые называются задачами многокритериальной оптимизации

$$x^{(0)} = \arg\operatorname*{extrem}_{x \in X} \left[\vec{f}(x) = (f_1(x), f_2(x), \ldots, f_m(x))\right]. \qquad (3)$$

В многокритериальной оптимизационной задаче (3) возможны следующие варианты: частные целевые функции независимы между собой; функции связаны между собой и являются согласованными; функции связаны между собой и являются антагонистическими. В первых двух случаях оптимизационная задача (3) сводится к совокупности независимых скалярных оптимизационных задач для частных целевых функций. В последнем случае, который встречается часто в практических приложениях проектирования, такая оптимизационная задача сводится к нахождению компромисса - согласованного экстремума частных целевых функций. Этот компромисс означает, что дальнейшее улучшения значения каждой целевой функции может быть достигнуто лишь за счет ухудшения значений других целевых функций. В результате этого находится множество решений, оптимальных по совокупности показателей качества решений, отражающих значения целевых функций. Соответствующие решения $x^{(o)} \in X$ называют Парето-оптимальными (оптимальными по Парето) относительно векторной целевой функции $\vec{f}(x)$. Множество всех таких решений обозначают через $P_{\vec{f}}(X)$.

Каждому решению $x \in X$ соответствует одна векторная оценка $\vec{y} = \vec{f}(x) \in Y$. С другой стороны, каждой оценке отвечают альтернативные решения $x \in X$ (их может быть и более одного), для которых $\vec{f}(x) = \vec{y}$. Таким образом, между множествами X и Y имеется тесная связь и поэтому выбор решения из множества X в указанном смысле равносилен выбору соответствующей оценки в критериальном пространстве Y.

Для векторных оценок $\vec{y}'$ и $\vec{y}''$ на множестве Y можно рассматривать разные типы бинарных отношений между оценками. В частности, широко используются такие бинарные отношения:

– отношение нестрогого предпочтения, называемое также отношением Парето $\vec{f}(x') \geq \vec{f}(x'')$. Оно означает, что $f_i(x') \geq f_i(x''), i = \overline{1,m}$ и по крайней мере для одного i выполняется строгое неравенство;

– отношение строгого предпочтения, называемое также отношением Слейтера $\vec{f}(x') > \vec{f}(x'')$, что означает $f_i(x') > f_i(x'')$, для всех $i = \overline{1,m}$.

Следует отметить, что отношение нестрогого предпочтения $\geq$, которое имеет место для векторных оценок, превращается при $m = 1$ в отношение строгого неравенства $>$ для скалярных оценок. При этом Парето-оптимальная оценка совпадает с максимальным элементом множества R^1, которому отвечает екстремум скалярной целевой функции $f(x)$. Таким образом, понятие Парето-оптимальности следует рассматривать как обобщение понятия екстремума на случай нескольких целевых функций. При этом оптимум по Парето – это согласованный оптимум связанных между собой и конкурирующих между собой целевых функций.

Важнейшим инструментом решения многокритериальных задач оптимизации является принцип Эджворта–Парето (принцип Парето) [7]. Принцип Эджворта-Парето формулируется в виде утверждения о том, что множество выбираемых решений содержится лишь в подмножестве Парето-оптимальных решений. Иначе говоря, каждое выбираемое разумно решение является Парето-оптимальным. Применение принципа Эджворта-Парето позволяет из множества всех возможных исключить заведомо неприемлемые решения, т. е. те, которые никогда не могут оказаться выбранными, если выбор осуществляется достаточно «разумно». После такого исключения остается множество, которое называют множеством Парето или областью компромиссов.

Парето-оптимальные проектные решения могут быть найдены как непосредственно на исходном множестве X с применением введенных бинарных отношений предпочтения, так и в критериальном пространстве оценок значений целевых функций Y. Отношению строгого предпочтения $\succ$ на множестве X соответствует отношение нестрого предпочтения $\geq$ на множестве Y. Для любых двух решений $x', x'' \in X$, для которых выполняется векторное неравенство $\vec{f}(x') \geq \vec{f}(x'')$, всегда имеет место отношение $x' \succ x''$.

Выбор Парето-оптимальных решений. Существуют разные методы нахождения подмножества Парето-оптимальных решений, в частности, метод последовательных уступок, метод рабочих характеристик, весовой метод, которые фактически сводят решение задачи многокритериальной оптимизации к решению совокупности задач скалярной оптимизации [6-8]. В настоящей работе для нахождения подмножества Парето-оптимальных решений использован метод дискретного выбора, который применяется,

когда множество X является дискретным и имеет конечную мощность. В этом методе включение решений $x^{(o)}$ в подмножество Парето $x^{(o)} \in P_f(X)$ имеет место тогда и только тогда, когда не существует других альтернативных решений $x \in X$, чтобы выполнялось неравенство $\vec{f}(x) \geq \vec{f}(x^{(o)})$. Формализованная процедура нахождения подмножества Парето-оптимальных оценок, которые соответствуют недоминируемым проектным решениям определяется соотношением

$$P(Y) = \mathrm{opt}_{\geq} Y = \left\{ \vec{f}(x^o) \in Y : \exists \vec{f}(x) \in Y : \vec{f}(x) \geq \vec{f}(x^o) \right\}. \qquad (4)$$

При нахождении подмножества Парето-оптимальных оценок, согласно процедуре (4), исключаются безусловно худшие оценки, а следовательно, и соответствующие им безусловно худшие проектные решения из множества X. Выбранным Парето-оптимальным проектным решениям соответствует потенциально возможное значение каждой из целевых функций, которое может быть достигнуто при фиксированных, но произвольных значениях других целевых функций. Это свойство m–кратного согласованного экстремума векторной целевой функции. Совокупность таких оптимальных значений целевых функций определяет потенциальные значения соответствующих показателей качества системы и называются многомерными потенциальными характеристиками системы.

Сужение множество Парето-оптимальных решений до единственного варианта. Как правило, множество Парето-оптимальных решений является достаточно широким и в процессе многокритериального выбора оптимальных решений неизбежно возникает вопрос о том, какое именно единственное возможное решение выбрать среди Парето-оптимальных [1]. В принципе все Парето-оптимальные решения являются несравнимыми между собой. Поэтому существует проблема сужения множества Парето, связанная с выбором того или иного Парето-оптимального варианта в качестве «наилучшего». Сужение множества Парето до единственного решения может быть осуществлено только при наличии дополнительной информации о предпочтениях ЛПР относительно проектних вариантов системы.

Для выбора единственного предпочтительного проектного варианта системы с учетом совокупности показателей качества могут быть использованы разные методы, которые базируються, в частности, на основных положениях теории полезности, теории нечетких множеств, теории лексографических отношений [6-8]. В настоящей работе для выбора единственного предпочтительного проектного варианта системы использован метод анализа иерархий [5], в котором используется исходная дополнительная информация от экспертов в виде результатов попарного сравнения проектних вариантов и показателей качества систем.

Метод анализа иерархий (МАИ) состоит в декомпозиции проблемы выбора единственного предпочтительного варианта системы на простые

составляющие части и получении суждений экспертов по парным сравнениям различных элементов проблемы выбора. В результате обработки полученных численных данных суждений экспертов формируется матрица парных сравнений. Для этой матрицы вычисляется главный собственный вектор и согласно определенной математической процедуры получают компоненты глобального вектора приоритетов, компоненты которого характеризуют приоритетность выбора вариантов проектируемой системы. Единственному предпочтительному варианту системы из заданного множества вариантов соответствует максимальное значение компонент глобального вектора приоритетов.

Принцип декомпозиции предусматривает структурирование проблемы выбора в виде иерархии уровней с вершины (цель выбора) через промежуточные уровни (показатели качества системы) к самому низкому уровню (альтернативные варианты построения системы).

Принцип сравнительных суждений экспертов в МАИ состоит в том, что объекты проблемы выбора сравниваются экспертами попарно по важности. Попарно сравниваются важности разных вариантов систем (на уровне 3) и разных показателей качества (на уровне 2). Результаты парных сравнений элементов приводятся к матричной форме

$$A = \begin{pmatrix} a_{11}a_{12}...a_{1j} \\ a_{21}a_{22}...a_{2j} \\ ... \\ a_{i1}a_{i2}...a_{ij} \end{pmatrix}, \qquad (5)$$

где $a_{ij} = \frac{w_s}{w_j}$ – оценки парных сравнений элементов выбора.

Оценки парных сравнений элементов a_{ij} находятся с использованием субъективных суждений экспертов, численно определяемых по шкале относительной важности элементов. Диагональ этой матрицы заполняется значениями "1", а элементы матрицы, лежащие ниже диагонали, заполняются соответствующими обратными значениями.

Далее выполняется некоторая обработка сформированных матриц парных сравнений элементов иерархий (5) на уровнях 2 и 3 проблемы выбора. С математической точки зрения эта задача обработки сводятся к вычислению главного собственного вектора, соответствующего максимальному собственному значению матрицы.

В частности, вычисляются компоненты главного собственного вектора для матрицы парных сравнений показателей качества как среднее геометрическое значение в строке матрицы парных сравнений [5]

$$V_j = \sqrt[n]{\prod_{i=1}^{n} a_{ij}}, \quad j = \overline{1,n}, \qquad (6)$$

где n – число показателей качества проектных решений.

Через компоненты главного собственного вектора вычисляются соответствующие компоненты вектора приоритетов показателей качества как нормированные значения

$$P_j = \frac{V_j}{S},\ j = \overline{1,n},\ \text{где}\ S = \sum_{j=1}^{n} V_j . \qquad (7)$$

Аналогично находятся оценки матриц парных сравнений вариантов систем на уровне 3 в отдельности по отношению к каждому показателю качества системы. На основе этих матриц вычисляются компоненты соответствующих главных собственных векторов и векторов приоритетов систем $\vec{Q}_j$ по отношению к отдельным показателям качества систем.

С использованием полученных данных вычисляются значения компонент вектора глобальных приоритетов $\vec{C}$ согласно [5]

$$C_i = \sum_{j=1}^{n} P_j Q_{ij},\ i = \overline{1,N}, \qquad (8)$$

где N - число сравниваемых вариантов систем.

По максимальному значению компонент вектора глобальных приоритетов (8) выбирается соответствующий предпочтительный вариант системы.

2. Многокритериальный выбор вариантов СМС 3-го поколения

Была поставлена и решена следующая задача выбора оптимальных вариантов СМС на этапе номинального планирования [9,10]. Для планируемой СМС третьего поколения стандарта UMTS задано некоторое множество возможных вариантов построения сети. Необходимо было найти подмножество Парето-оптимальных проектных решений с учетом совокупности показателей качества, а затем выполнить сужение подмножества Парето до единственного предпочтительного варианта СМС.

Выбор подмножество Парето-оптимальных вариантов СМС. Для решения поставленной задачи было сформировано множество из 10 вариантов построения СМС стандарта UMTS, которые определялись различными исходными данными, в качестве которых использовались: планируемое количество абонентов в сети, плотность обслуживаемой территории, предполагаемая активность абонентов, допустимая вероятность блокировки вызова.

Для заданных исходных данных были рассчитаны значения показателей качества для каждого из 10 вариантов СМС. При решении оптимизационной задачи среди всех показателей качества были выбраны показателя качества, которые характеризуют соответственно вероятность блокировки ($P_{бл}$), плотность обслуживаемых абонентов (N_a/S_0) и

необходимое количество базовых станций в сети (N_{BTS}). Абсолютные значения этих показателей качества и нормированные к их максимальным значениям (соответственно $k_{1н}$, $k_{2н}$ и $k_{3н}$) представлены в табл. 1.

При нахождении подмножества Парето-оптимальных оценок, согласно формальной процедуре (4), исключаются безусловно худшие оценки, а следовательно, и соответствующие им безусловно худшие проектные решения из множества X. Выбранные Парето-оптимальные проектные решения являются нехудшими по отношению нестрогого предпочтения. В рассмотренном примере подмножество Парето-оптимальных вариантов включало 5 вариантов построения СМС - 1П, 5(П), 8(П), 9(П), 10(П). Остальные 5 вариантов СМС оказались безусловно худшими и были исключены из дальнейшего рассмотрения.

Таблица 1

№ варианта СМС	Рассчитанные значения показателей качества					
	($P_{бл}$)	(N_a/S_0)	(N_{BTS})	$k_{1н}$	$k_{2н}$	$k_{3н}$
1 (П)	0,1	166	11	1	0,135	0,458
2	0,05	133	15	0,5	0,307	0,625
3	0,05	160	24	0,5	0,166	1
4	0,09	172	18	0,9	0,104	0,75
5 (П)	0,07	192	21	0,7	0	0,875
6	0,03	144	20	0,3	0,25	0,833
7	0,04	140	19	0,4	0,271	0,791
8 (П)	0,04	142	15	0,4	0,26	0,625
9 (П)	0,02	183	18	0,2	0,047	0,75
10 (П)	0,02	189	22	0,2	0,015	0,916

Программная реализация многокритериального выбора оптимальных вариантов СМС. Для многокритериального выбора оптимальных проектных решений была разработана программа МСО (Multi Criterial Optimization). Программа МСО реализована на языке Java. Интерфейс программы многокритериального выбора проектных вариантов СМС показан на рис. 1. Программа решает следующие задачи: выбор и задание значений показателей качества для допустимых проектных вариантов, нормирование и приведение показателей качества к сопоставимому виду, выбор подмножества Парето-оптимальных вариантов с использованием метода дискретного выбора по безусловному критерию предпочтения Парето, сужение множества Парето до единственного проектного варианта построения сети с использованием методов на основе функций ценности или метода на основе лексикографических отношений.

Выбор единственного предпочтительного варианта СМС. Для выбора единственного предпочтительного варианта СМС 3-го поколения

из множества Парето-оптимальных был использован метод анализа иерархий [5].

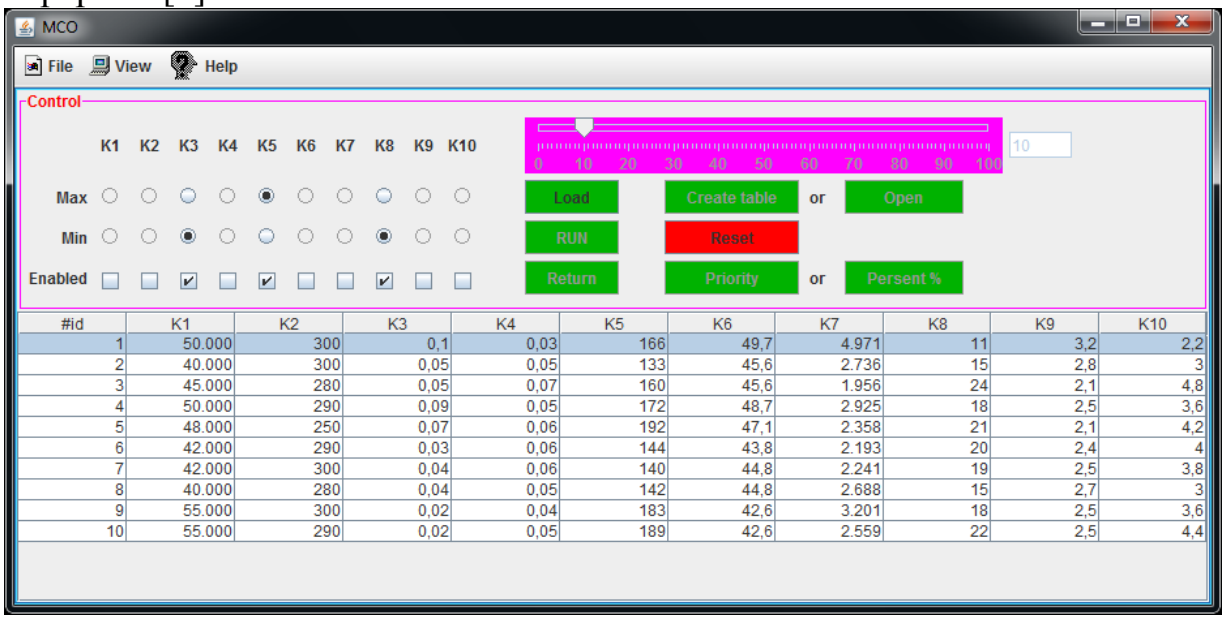


Рис. 1. Интерфейс программы многокритериального выбора проектных вариантов СМС

С использованием суждений экспертов построены матрицы парных сравнений для СМС на втором уровне иерархии (важности показателей качества), а также на третьем уровне (альтернативных вариантов систем по отношению к показателям качества - соответственно к вероятности блокировки, плотность обслуживаемых абонентов, количеству базовых станций). В результате обработки полученных матриц парных сравнений вычислены согласно (6) и (7) главные собственные векторы и векторы приоритетов.

Для примера в табл. 2 приведена матрица парных сравнений показателей качества и вычисленные оценки компонент главного собственного вектора и вектора приоритетов показателей качества. В табл. 3 приведена матрица парных сравнений вариантов СМС по отношению к вероятности блокировки и вычисленные оценки компонент главного собственного вектора и вектора приоритетов

Таблица 2

	$К_1$	$К_2$	$К_3$	V_j	P_j
$К_1$	1	5	3	2,464	0,6173
$К_2$	1/5	1	1/5	0,3424	0,0858
$К_3$	1/3	5	1	1,1854	0,297

В табл. 4 приведены вычисленные векторы приоритетов показателей качества и вариантов построения СМС по отношению к каждому

показателю качества, а также вычисленные согласно (8) компоненты вектора глобальных приоритетов $\vec{C}$ (здесь N_i – номера вариантов СМС).

Таблица 3

	N_1	N_2	N_3	N_4	N_5	V_i	P_i
N_1	1	1/3	1/7	1/9	1/9	0,23	0,03
N_2	3	1	1/7	1/9	1/9	0,35	0,04
N_3	7	7	1	1/7	1/7	1	0,11
N_4	9	9	7	1	2	4,08	0,47
N_5	9	9	7	1/2	1	3,09	0,35

Таблица 4

СМС	P_{1i}	P_{2i}	P_{3i}	C_i
N_1	0,03	0,07	0,51	0,1779
N_2	0,04	0,45	0,07	0,0863
N_3	0,11	0,04	0,26	0,1498
N_4	0,47	0,16	0,12	0,3418
N_5	0,35	0,29	0,04	0,2551
P_j	0,62	0,09	0,3	

Максимальному значению компонент вектора $\vec{C}$ соответствует предпочтительный вариант СМС (N_4), который характеризуется минимальной допустимой вероятностью блокировки $P_{бл}$ = 0,02, плотностью обслуживаемых абонентов N_a/S_0 = 183 аб./км^2 и количеством базовых станций N_{BTS}=18.

3. Выбор предпочтительной технологии систем мобильной связи 4-го поколения методом анализа иерархий

Для сравнительного анализа и многокритериального выбора предпочтительного варианта были рассмотрены следующие технологии мобильной связи 4-го поколения: HSPA, WiMAX и LTE [11]. Кратко приведем особенности этих технологий.

В результате развития систем мобильной связи была создана технология HSPA+ . В нисходящем канале ее отличает использование модуляции 64-QAM с SIMO (1x2) или 64-QAM с SIMO (2x2). В восходящем канале использована модуляция 64-QAM и улучшены возможности для VoIP. Поправки в соответствии с релизом 8 позволяют использовать в нисходящем канале режим MIMO (2x2) с модуляцией 64-QAM.

Системы мобильной связи с технологией WiMAX предназначены для предоставления сервисов как неподвижным, так и подвижным пользователям. Мобильный WiMAX (релиз 1.5) имеет сравнимые с HSPA+ (релиз 8) пиковые скорости в нисходящем канале при одинаковой модуляции, скорости кодирования и ширине канала. При этом у мобильного WiMAX в восходящем канале пиковая скорость выше в 2-3 раза. Мобильный WiMAX поддерживает ширину канала до 20 МГц, а также как частотное, так и временное дуплексирование. Его частотные профили планируются в диапазонах 700, 1700, 2300, 2500 и 3500 МГц. Мобильный WiMAX обеспечивает «гладкую IP - сеть» (из конца в конец).

Следующим шагом эволюции СМС 3GPP, являются системы Long Term Evolution (LTE). Их отличает использование модуляции OFDMA в нисходящем канале и SC-FDMA – в восходящем. Модуляция – 64-QAM, ширина канала – до 20 МГц, дуплексирование TDD и FDD. Применяются адаптивные антенные системы. Сетевая архитектура - полностью IP–сеть. В системе LTE применяются технологии и методы, уже применяемые в мобильном WiMAX. Системы LTE – это революционное улучшение СМС 3G. LTE представляет переход от систем CDMA к системам OFDMA, а также переход к полностью IP–системе с коммутацией пакетов. Поэтому внедрение этой технологии в существующих СМС означает необходимость обеспечения новых радиочастотных ресурсов для получения преимущества от широкополосного канала.

В табл. 5 представлены исходные значения показателей качества для разных стандартов СМС: K_1 - спектральная эффективность (нисходящий канал), K_2 - радиус действия, K_3 - скорость передачи данных. Показатели качества рассмотренных технологий СМС носят конкурирующий характер. Поэтому следует применять методы многокритериальной оптимизации для выбора предпочтительного варианта технологии СМС.

Таблица 5

Показатели качества	HSPA		WiMAX	LTE
	Релиз 7	Релиз 8	Релиз 1.5	
Спектральная эффективность, бит/Гц/с	0,87	1,75	1,59	1,57
Радиус действия, км	30	40	50	5
Скорость передачи, Мбит/с	21	35	48	75

Рассмотрим особенности применения метода анализа иерархий для выбора предпочтительного варианта технологии СМС 4-го поколения. На

рис. 2 показано иерархическое представление задачи выбора предпочтительного варианта СМС.

В табл. 6 приведены матрица парных сравнений показателей качества, а также вычисленные согласно (6) и (7) оценки компонент главного собственного вектора и вектора приоритетов показателей качества.

Далее с учетом суждений экспертов выполнены парные сравнения вариантов технологий на 3-м уровне иерархии. В частности, выполнены парные сравнения технологий СМС по отношению к выбранным показателям качества.

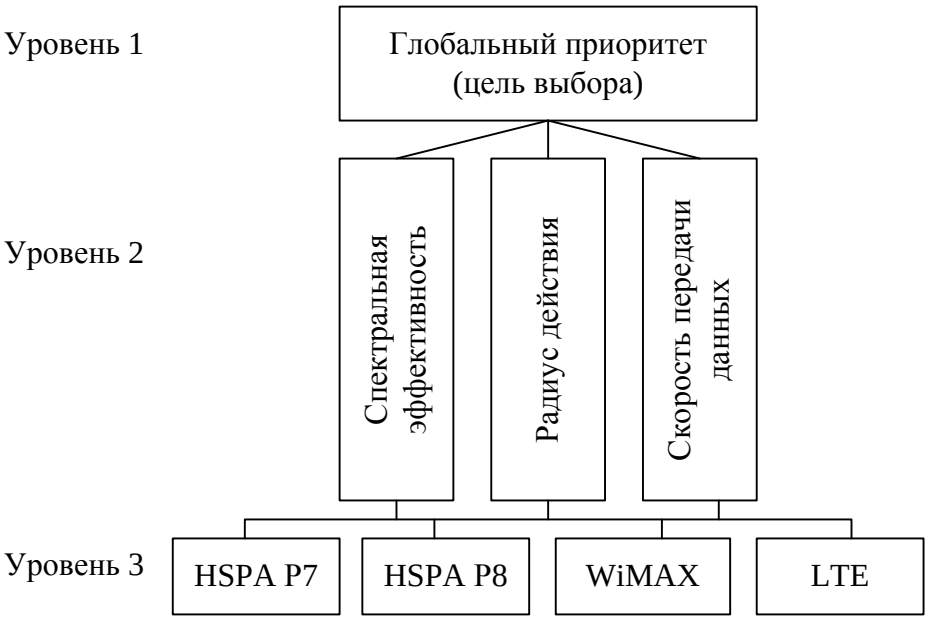


Рис. 2. Декомпозиция задачи выбора в иерархию технологий сетей мобильной связи

Таблица 6

	$К_1$	$К_2$	$К_3$	V_i	P_j
$К_1$	1	3	1/3	1	0,2584
$К_2$	1/3	1	1/5	0,4058	0,1049
$К_3$	3	5	1	2,464	0,6367

В результате обработки полученных матриц парных сравнений вычислены соответствующие главные собственные векторы и векторы приоритетов. Для примера в табл. 7, 8, 9 приведены матрицы парных сравнений вариантов технологий СМС по отношению к спектральной эффективности, радиусу действия, скорости передачи данных, а также вычисленные компоненты соответствующих главных собственных векторов и векторов приоритетов. С использованием вычисленных векторов приоритетов вычислены согласно (8) компоненты вектора глобальных приоритетов. Результаты вычислений приведены в табл. 10.

Таблица 7

	N_1	N_2	N_3	N_4	V_j	P_{1j}
N_1	1	1/5	1/5	1/5	0,299	0,057
N_2	5	1	3	3	2,59	0,4935
N_3	5	1/3	1	3	1,495	0,2849
N_4	5	1/3	1/3	1	0,863	0,1645

Таблица 8

	N_1	N_2	N_3	N_4	V_j	P_{2j}
N_1	1	1/3	1/5	5	0,76	0,1301
N_2	3	1	1/3	7	1,627	0,2785
N_3	5	3	1	7	3,201	0,5481
N_4	1/5	1/7	1/7	1	0,253	0,0433

Таблица 9

	N_1	N_2	N_3	N_4	V_j	P_{3j}
N_1	1	1/3	1/5	1/7	0,31	0,0433
N_2	3	1	1/3	1/5	0,67	0,0928
N_3	5	3	1	3	2,59	0,3593
N_4	7	5	5	1	3,64	0,5046

Таблица 10

СМС	P_{1i}	P_{2i}	P_{3i}	C_i
N_1	0,057	0,13	0,043	0,0553
N_2	0,494	0,279	0,093	0,2158
N_3	0,285	0,548	0,359	0,3586
N_4	0,165	0,043	0,505	0,3704
P_j	0,26	0,1	0,64	

Видно, что максимальному значению компоненты вектора глобальных приоритетов соответствует предпочтительный вариант технологии СМС 4-го поколения – N_4. Это технология LTE со скоростью передачи данных 75 Мбит/с, спектральной эффективностью 1,57 бит/Гц/с и радиусом действия базовых станций 5 км.

Выводы

1. Приведена методология многокритериального выбора оптимальных проектных решений с учетом совокупности показателей

качества, который включает 2 этапа. На первом этапе на множестве возможных проектных решений выделяется множество Парето-оптимальных вариантов и исключаются из дальнейшего рассмотрения безусловно худшие решения. На втором этапе из множества Парето-оптимальных вариантов выбирается единственный предпочтительный вариант с использованием дополнительной информации от экспертов.

2. Рассмотрены примеры многокритериального выбора оптимальных проектных вариантов на этапе планирования систем мобильной связи 3-го и 4-го поколений. Парето-оптимальные варианты СМС найдены методом дискретного выбора. Единственный предпочтительный вариант из множества Парето-оптимальных выбран методом анализа иерархий.

Литература

1. Подиновский В.В. Парето-оптимальные решения многокритериальных задач / В.В. Подиновский, В.Д. Ногин. – М.: Наука, 1982. – 256 с.
2. Брахтман Т.Р. Многокритериальность и выбор альтернатив в технике. – М.: Сов. радио, 1984. – 326 с.
3. Березовский Б.А., Многокритериальная оптимизация. Математические аспекты / Б.А. Березовский, Ю.М. Барышников, В.И. Борзенко, Л.М. Кепнер. – М.: Наука, 1986. – 128 с.
4. Дубов Ю.А. Многокритериальные модели формирования и выбора вариантов систем / Ю.А. Дубов, С.И. Травкин, В.Н. Якимец. – М.: Наука, 1986. – 221 с.
5. Саати Т. Аналитическое планирование. Организация систем / Т. Саати, К. Кернс. – М.: Радио и связь, 1991. – 224 с.
6. Чернoруцкий И.Г. Методы оптимизации и принятия решений. – СПб.: Издательство „Лань”, 2001. – 384 с.
7. Ногин В.Д. Принятие решений в многокритериальной среде: Колличественный подход. – М.: ФИЗМАТЛИТ, 2002. – 176 с.
8. Безрук В.М. Векторна оптимізація та статистичне моделювання в автоматизованому проектуванні систем зв’язку. – Харків: ХНУРЕ, 2002. – 164 с.
9. Чеботарёва Д.В. Многокритериальная оптимизация проектных решений при планировании сотовых сетей мобильной связи / Д.В. Чеботарёва, В.М. Безрук. – Харьков: Компания СМИТ, 2013. – 148 с.
10. Безрук В.М. Принятие оптимальных решений при планировании сетей мобильной связи с учетом совокупности показателей качества / В.М. Безрук, Д.В. Чеботарёва // Радиотехника. – Харьков, 2014. – Вып.178. – С.119 – 130.
11. Безрук В. М. Применение метода анализа иерархий при выборе средств телекоммуникаций с учетом совокупности показателей качества / В. М. Безрук, Ю. В. Скорик // Радиоэлектроника и информатика. – Харьков: ХНУРЭ, 2013. – С. 24-29.

ПІДХІД ДО ЕНЕРГОЕФЕКТИВНОГО ПЛАНУВАННЯ ЗАДАЧ У СЕРВЕРНОМУ КЛАСТЕРІ ДЛЯ ОПТИМІЗАЦІЇ ОБРОБКИ ТРАФІКА ПЕРЕДАЧІ ДАНИХ

Глоба Л.С., Гвоздецька Н.А., Прокопець В.А., Степурін О.В.

Вступ

Технології 4G LTE (Long Term Evolution) мобільного зв'язку 4 покоління на сьогоднішній день широко впроваджуються у світі. Оператори зв'язку та експерти висловлюють своє бачення концепції 5G. Вона включає в себе розвиток та поширення технології Інтернету речей (Internet of Things, IoT) та міжмашинної взаємодії M2M. Ці сервіси продукують великі масиви даних, що потребують високої швидкості обробки.

Яскравим прикладом цього може стати керування «розумним автомобілем», що самостійно відстежує ситуацію на дорозі, отримуючи дані від інших подібних автомобілів, та автоматично приймає рішення, щодо зміни швидкості або напрямку руху без безпосередньої участі водія. Така концепція вимагає гарантованого обміну даними у режимі реального часу. З мінімальними втратами пакетів, з мінімальними затримками. Перепоною на шляху до впровадження такої концепції може стати не тільки затримка при передачі даних, а й затримка при їх обробці у центрі обробки даних (ЦОД).

В зв'язку із цим, для забезпечення повноцінного функціонування систем типу M2M та IoTy режимі реального часу необхідно забезпечити високу швидкість та якість обслуговування на таких етапах згідно загальної структури системи (рис. 1):

1. Збір даних датчиками.
2. Передача даних від датчиків до віддаленого сервера (ЦОД).
3. Обробка даних віддаленим сервером.
4. Передача відповіді сервера до пристрою.

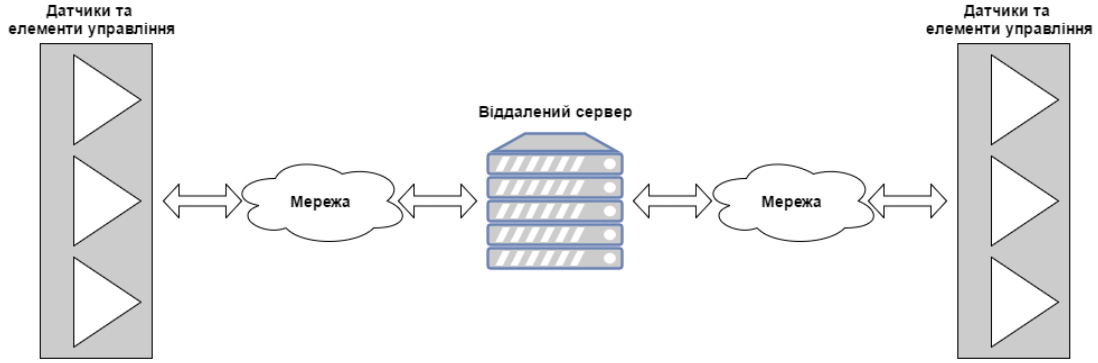


Рис. 1. Типова структура M2M системи

Забезпечення якості та швидкості обслуговування на етапі 1 здійснюється шляхом вдосконалення апаратного забезпечення пристроїв M2M та IoT. На етапах 2, 4 необхідні вимоги задовольняються шляхом

збільшення пропускної здатності мережі, пріоритезації трафіка, скорочення ділянки передачі. У роботі докладно розглянуто та проаналізовано можливі шляхи підвищення швидкості та якості обробки даних на етапі 3. В роботі запропоновано алгоритм планування задач, що має на меті збільшення швидкості обробки даних із одночасним підвищенням енергоефективності.

Високої швидкості передачі та обробки потребують також сервіси високоякісного відео у форматі 4К та 8К, сервіси віртуальної реальності.

Обробка даних потребує значних енергозатрат. Згідно джерел [1], [2] кількість енергії, що була спожита центрами обробки даних (серверами, сховищами, апаратурою зв'язку та охолодження) по всьому світу у період з 2000 до 2010 років мала такий розподіл: становила близько 70,8 ТВт у 2000 році; 152,5 ТВт – у 2005 році (спостерігалось зростання споживання на 115%); зросла до 271,8 ТВт у 2010 році. При цьому частка цієї потужності у загальній споживаній у світі становить в середньому 1,5%.

Обсяги енергії, спожитої ЦОД продовжують зростати і сьогодні. Проблема енергоспоживання серверів та центрів обробки даних (ЦОД) в цілому набуває все більшої актуальності із розповсюдженням концепцій інтернету речей (Internet of Things), машино-машинної взаємодії (M2M), збільшенням ролі інтернет-сервісів. Кількість даних, що потребують обробки безперервно зростає. До швидкості передачі та обробки даних при цьому висуваються жорсткі вимоги, що означає що економити на енергії за рахунок зменшення продуктивності роботи серверів неприпустимо.

У зв'язку з необхідністю підвищення енергоефективності та зростанням вимог до швидкості обробки інформації гостро постає проблема розробки енергоефективних алгоритмів обробки даних у ЦОД. У роботі розглянуто процес планування задач у рамках серверного кластеру, що є складовою одиницею ЦОД. Запропоноване рішення перевірено на можливість масштабування до використання у реальному ЦОД.

1. Аналіз існуючих стратегій планування задач у серверному кластері

При обробці задач у серверному кластері ключовим параметром є продуктивність кластера. Середній час обробки однієї задачі – величина, обернено пропорційна продуктивності кластера (формула 1):

$$P = \frac{m}{t_\Sigma}, \tag{1}$$

де m – кількість задач, що підлягають обробці;

t_Σ – час роботи кластера над обробкою набору з m задач.

Тоді, середній час виконання однієї задачі з набору:

$$t_{task_i} = \frac{t_\Sigma}{m}. \tag{2}$$

Виходячи з цього, можливо оперувати параметром «середній час обробки задачі» як обернено пропорційним до параметру продуктивності. Мінімізація параметру часу приведе до максимізації параметра продуктивності.

Другим параметром, за якими проводиться оптимізація обробки задач у кластері є енергоефективність. Цей параметр більшість існуючих підходів до балансування задач не враховують, що буде показано далі.

Найбільш простим та розповсюдженим алгоритмом балансування навантаження у рамках серверного кластера є алгоритм Round Robin. Згідно цього алгоритму всі задачі розподіляються по всім активним серверам за циклічним принципом. Цей алгоритм швидкий, адже не потребує знання про поточний розподіл ресурсів серед серверів. Але, оскільки задачі та вузли обробки (сервери) за даним алгоритмом не мають жодного пріоритету, алгоритм не здатний призначити задачу на обробку найбільш підходящому для цього серверу. Це є основним недоліком Round Robin.

Однією з успішно використовуваних модифікації цього алгоритму є Round Robin зі зваженими коефіцієнтами (Weighted Round Robin). На відміну від простого алгоритму Round Robin, дана модифікація використовую знання про обчислювальні потужності вузлів обробки і, відповідно до значень обчислювальних потужностей, присвоює вагові коефіцієнти серверам, що в подальшому враховується при розподілі задач між ними.

Обидва вищеописані алгоритми не беруть до уваги параметр енергоспоживання кластерного вузла.

Для підходів, що враховують параметр енергоефективності при розподілі задач необхідністю є вибір правильної енергетичної моделі вузлів кластера. Енергетична модель у даному випадку являє собою функцію (формула 3):

$$P = f(CPU), \tag{3}$$

де CPU – ступінь зайнятості центрального ядра обробки (процесора) у відсотках, P – потужність, споживана вузлом кластера.

Визначати енергетичну модель у даній роботі запропоновано індивідуально для кожного вузла кластера. Існуючі підходи до балансування, які враховують параметр енергоефективності, та енергетичні моделі, що вони використовують, наведені далі.

Серед відомих алгоритмів балансування, що враховують параметр енергоспоживання слід відмітити алгоритм CTES – Cooperative Two-Tier Energy-Aware Scheduling [3]. Автори розглядають дворівневий підхід до планування задач із регулюванням швидкості їх виконання, з метою досягнення оптимального використання процесору, замість міграції задач на інші вузли. Використовуються декілька стратегій планування з прогнозом виконання задач для оптимального призначення їх на доступні

машини. Результати моделювання, представлені в роботі, показують, що цей підхід зменшує загальне споживання енергії у серверному кластері.

Недоліком даного алгоритму є те, що автори вважали модель енергоспоживання сервера лінійною. Насправді така модель не точно відображає характер залежності споживаної потужності від завантаженості сервера. Натомість у даній роботі пропонується проведення попередньої атестації кожного сервера кластеру для отримання реальної залежності потужності споживання від завантаженості для кожного вузла кластеру.

Іншим прикладом алгоритму, що має на меті підвищення енергоефективності обробки задач у хмарних системах, EDRP – Energy and Deadline aware Resource Provisioning [4]. Автори зосередились на проблемі мінімізації затрат на хмарні системи, підвищуючи ефективність використання енергії, але гарантуючи терміни виконання користувацьких задач, визначених в умовах щодо якості обслуговування (SLA). Вони беруть до уваги два типи задач, незалежні пакетні передачі і задачі із залежностями.

Їхня модель розрахунку споживання електроенергії у момент часу t включає статичне $P_{xstatic}(t)$ і динамічне $P_{xdynamic}(t)$ енергоспоживання. Обидві характеристики розраховуються на основі відсотку завантаження процесору $Util_x(t)$, в якому враховуються тільки параметри використовуваної машини Q-t. Недоліком запропонованого алгоритму є те, що автори не враховують відмінностей між машинами, на яких виконуються задачі, і машинами, що знаходяться в режимі очікування.

У роботі [5] описано стратегію розподілу задач Min_c. Дана стратегія враховує різноманітність задач, що приходять на обробку та відповідну різноманітність ресурсів, яких вони потребують. Недоліком описаного у статті [5] алгоритму є те, що енергетична модель, яку використали автори, хоч і має нелінійний характер, близький до дійсності, проте використана модель має єдиний вигляд для всіх машин в незалежності від їхніх характеристик. Натомість у даній роботі пропонується індивідуально визначати модель енергоспоживання для кожної машини окремо.

2. Постановка задачі визначення стратегії енергоефективного балансування задач у серверному кластері

Введення основних термінів та припущень

У даній роботі автори мають на меті здійснення оптимізації процесу розподілу задач між вузлами кластера за двома критеріями: продуктивність та енергоефективність. При постановці задачі використані такі терміни та поняття:

– Комп’ютерний кластер – набір комп’ютерів, що працюють разом і можуть розглядатись як єдина система.

– Вузол – одиниця комп’ютерного кластера, представлена фізично однією ЕОМ:

$$ComputerCluster = \{N_j\},$$

де N_j - j-й вузол комп'ютерного кластера.

– FLOPS – одиниця вимірювання продуктивності комп'ютера (кількість операцій з плаваючою комою, що можуть бути виконані за секунду) [6].

– Задача – набір елементарних операцій, що повинні бути оброблені неподільно в комп'ютерному кластері:

$$Task = \{t_{task_i}\}.$$

– Робота – задача разом з вимогами до її виконання:

$$Jobs = \{job_i\} \rightarrow \{t_{task_i}, \{V_{req_i}, k_{core_{req_i}}, t_{i_{max}}\}\},$$

де V_{req_i} – об'єм оперативної пам'яті для виконання задачі;

$k_{core_{req_i}}$ – кількість ядер процесора, що вимагає задача t_{task_i} для свого виконання;

$t_{i_{max}}$ – максимальний час, за який задача t_{task_i} має бути оброблена.

Формулювання задачі в математичному вигляді

Вихідні дані:

Серверний кластер складається з n вузлів $\{N_j\}$. Кожен вузол N_j характеризується:

– V_j – об'ємом доступної оперативної пам'яті;

– $flops_j$ – продуктивністю вузла N_j, що має k_{core_j} обчислювальних ядер;

– $P_j = f_j(CPU_j)$ – функцією енергоспоживання від навантаженості процесора $f_j(CPU_j)$, що експериментально визначена для кожного вузла N_j.

Нова задача t_{task_i} приходить до системи в момент τ.

Кожна задача потребує певних значень вищеназваних параметрів:

$$\{V_{req}, k_{core_{req}}, t_{max}\},$$

$$job_i \rightarrow \{t_{task_i}, \{V_{req_i}, k_{core_{req_i}}, t_{i_{max}}\}\}.$$

Задача:

Розробити алгоритм планування задач, такий що $P_\Sigma \rightarrow \min$, $t_\Sigma \rightarrow \min$, де P_Σ – сумарна потужність спожита усім серверним кластером, t_Σ - час виконання набору із m задач.

3. Запропонований підхід до енергоефективного балансування задач у серверному кластері

Загальна характеристика запропонованого підходу

Основна ідея запропонованого підходу складається із реалізації таких основних етапів:

I. Проведення попередньої атестації серверного кластера (визначення базових параметрів вузлів серверного кластера).

II. Застосування запропонованого алгоритму балансування, що складається із таких основних кроків:

1) оцінка стану кластера в момент τ_{k-1};

2) виключення з розгляду всіх непідходящих вузлів (по оперативній пам'яті і кількості доступних ядер);

3) знаходження набору $P_\Sigma = \{P_{\Sigma_j}\}$ (за допомогою відомих функцій $P_j = f_j(CPU_j)$, що являє собою теоретичні потужності, що будуть спожиті кожним вузлом кластера у разі розміщення задачі на нього;

4) сортування набору вузлів N_j за теоретично встановленим внеском в потужність споживання всього кластера: $N_P = \{N_{P_j}\}$;

5) сортування набору вузлів по продуктивності: $N_{FLOPS} = \{N_{FLOPS_j}\}$;

6) присвоєння вузлам вагових коефіцієнтів (балів) за енергоефективність і продуктивність в залежності від їх позиції в сортованих масивах N_P та N_{FLOPS};

7) призначення задачі вузлу, що має найбільший сумарний бал.

Новим у запропонованому підході є те, що енергетичні моделі визначаються індивідуально для кожного вузла кластера у рамках процесу його попередньої атестації. Також новим є запропонований алгоритм, що враховує при балансуванні задач 2 параметри – продуктивність вузла та його енергоефективність.

Детальний опис запропонованого алгоритму

Крок 1. У момент часу τ нова задача (робота) job_i надходить до кластера. У момент часу τ_{k-1} (попередня оцінка) необхідно виміряти такі параметри кожного вузла N_j:

1) $\Delta V_{j_{avail}} = V_j - V_{j_{used}}$ - доступний об’єм оперативної пам’яті j-го вузла; де $V_{j_{used}}$ – оперативна пам’ять, зайнята на момент часу τ_{k-1};

2) $\Delta k_{core\,j_{avail}} = k_{core\,j} - k_{core\,j_{used}}$ - кількість ядер процесора, доступних на вузлі N_j; де $k_{core\,j_{used}}$ - кількість ядер вузла N_j, що зайняті іншими задачами на момент часу τ_{k-1};

3) $P_{\downarrow}\Sigma = \Sigma_{\downarrow} j \equiv P_{\downarrow} j \big| \tau_{\downarrow}(k-1)$ - сумарна потужність, що споживається кластером в момент часу τ_{k-1}.

Крок 2. Якщо доступний об'єм оперативної пам'яті менший за той, що вимагається для виконання задачі $task_i, (\Delta V_{j_{avail}} \le V_{req_i})$ вузол N_j кластера виключається із розгляду доступних для обробки даної задачі вузлів. Аналогічно, якщо кількість незадіяних у роботі ядер вузла N_j менша, ніж цього вимагає задача t_{task_i} вузол N_j виключається із розгляду доступних для обробки даної задачі вузлів.

У результаті таких операцій отримуємо масив $\{N_{avail_j}\}$ теоретично доступних для розміщення задачі t_{task_i} вузлів, такий що $\{N_{avail_j}\} \subset \{N_j\}$.

Крок 3. Для кожного j-го вузла серверного кластера залежність потужності споживання від навантаженості процесора є відомою функцією $P_j = f_j(CPU_j)$. Цю залежність було виміряно і збережено при проведенні попередньої атестації кластера.

Припустивши, що робота job_i була призначена для виконання вузлу N_j, та знаючи залежність $P_j = f_j(CPU_j)$ для цього вузла, обчислимо теоретичне значення сумарної спожитої кластером потужності у разі розміщення задачі на обробку на j-й вузол:

$$P_{\Sigma_1} \big| \tau_k \;\&\; N_j = P_{\Sigma} \big| \tau_{k-1} + \Delta P_{ji} \big| \tau_k .$$

Проведемо такі обчислення для кожного вузла кластера.

Таким чином, маємо масив значень $P_{\Sigma} = \{P_{\Sigma_j}\}$ отриманий шляхом теоретичного розміщення роботи job_i на кожний з вузлів N_j

Крок 4. Проведемо сортування масиву доступних для розміщення задачі t_{task_i} вузлів $\{N_{avail_j}\}$ за їх відомою продуктивністю $flops_j$ у порядку спадання. В результаті отримаємо масив $N_{flops_{avail}}$:

$$N_{avail_{flops}} = \{N_{flops_{j_{max}}}, \ldots N_{flops_{j_{min}}}\}.$$

Крок 5. Проведемо сортування масиву доступних для розміщення задачі t_{task_i} вузлів $\{N_{avail_j}\}$ за їх теоретично обрахованим на кроці 3 вкладом у загальне сумарне енергоспоживання всього серверного кластера $[[\Delta P]_{\downarrow} j]_{\downarrow} i \big| \tau_{\downarrow} k$. Нехай сортування буде також проведене в порядку спадання. В результаті отримаємо масив N_{avail_p}:

$$N_{avail_p} = \{N_{p_{j_{max}}}, \ldots N_{p_{j_{min}}}\}.$$

*Крок 6.*В залежності від позиції вузла із номером j у сортованому масиві $N_{avail_{flops}}$ та N_{avail_p}, він отримує два вагових коефіцієнта (дві

оцінки): mark_p і mark_{flops} відповідно. При чому mark_p рівна позиції j-го вузла у сортованому масиві N_{avail_p}, mark_{flops} – позиції цього вузла у сортованому масиві $N_{avail_{flops}}$ відповідно.

Крок 7. Сумарна оцінка кожного вузла:

$$\text{mark}_{\Sigma_j} = \text{mark}_{flops_j} + \text{mark}_{p_j}.$$

Вузол, для якого $\text{mark}_{\Sigma_j} = \max_j \text{mark}_{p_j}$ вибирається для здійснення обробки завдання t_{task_i}.

4. Експериментальна оцінка запропонованого підходу

Для оцінки роботи запропонованого підходу було проведено натурний експеримент, в ході якого було порівняно роботу запропонованого підходу із широко використовуваними алгоритмами балансування навантаження.

Загальний опис та план експерименту

Для порівняння із запропонованим підходом було використано алгоритм Round Robin, що використовується широко у реальних обчислювальних кластерах та ЦОД та є водночас простим для розуміння та використання. Суть даного алгоритму полягає в тому, що задачі розподіляються між вузлами кластера послідовно, без урахування індивідуальних параметрів вузлів.

Для проведення експерименту було організовано серверний кластер з 5 вузлів, що мав топологію зірки (рис.2).

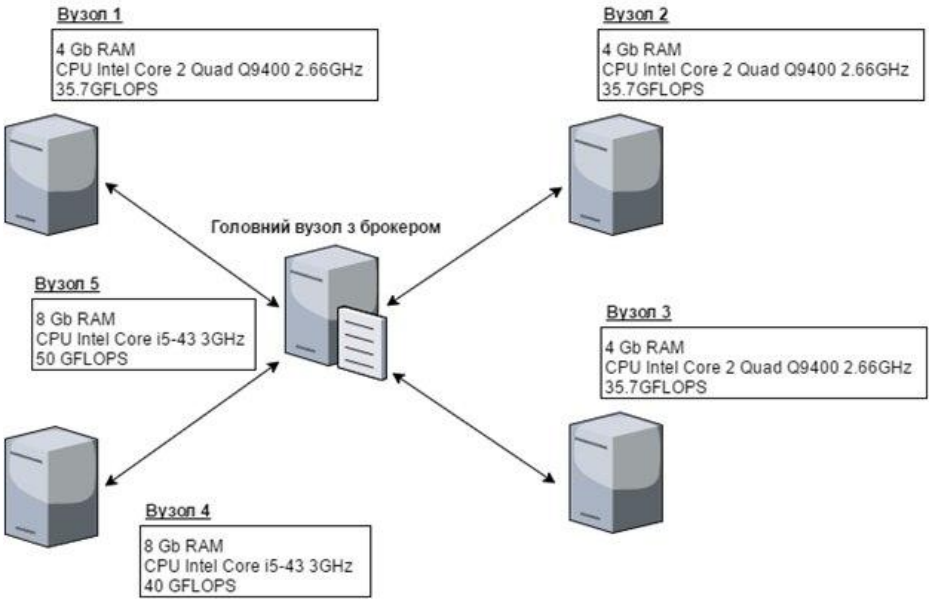


Рис. 2. Серверний кластер, який було використано у ході експерименту

Комп’ютери, що були використані у якості вузлів кластера були двох типів (за їх фізичними параметрами). Параметри кожного з вузів наведено на рис. 2.

Чотири з п’яти вузлів були рівноправними, а один був обраний центральним. На ньому розміщувався брокер (балансувальник задач), код якого був написаний на мові програмування С.

Для зняття характеристик $P_j = f_j(CPU_j)$ кожного вузла, було використано цифровий мультиметр, що здатний вимірювати струм, напругу та миттєву потужність безпосередньо між джерелом живлення (розеткою) та навантаженням.

Хід експерименту та аналіз результатів

Згідно із запропонованою методикою в першу чергу необхідно провести атестацію кожного з вузлів серверного кластера для визначення зокрема їх енергетичних моделей. Для визначення енергетичних моделей кожен вузол кластера послідовно було навантажено стресс-тестом, що завантажував процесор вузла послідовно на 25, 50, 75 та 100%. При цьому між джерелом живлення та машиною був увімкнений цифровий мультиметр, що знімав покази миттєвої потужності споживання кожну секунду. Це дозволило отримати основні точки графіка енергетичної моделі. Для уточнення характеристики моделі кожен вузол було послідовно навантажено задачами підрахунку числа π із різною кількістю знаків після коми (точністю). Це дозволило краще прослідкувати характер залежності потужності споживання від ступеня завантаженості процесора CPU. В результаті проведення попередньої атестації вузлів кластера для кожного з них була отримані характеристики, представлені на рис. 3.

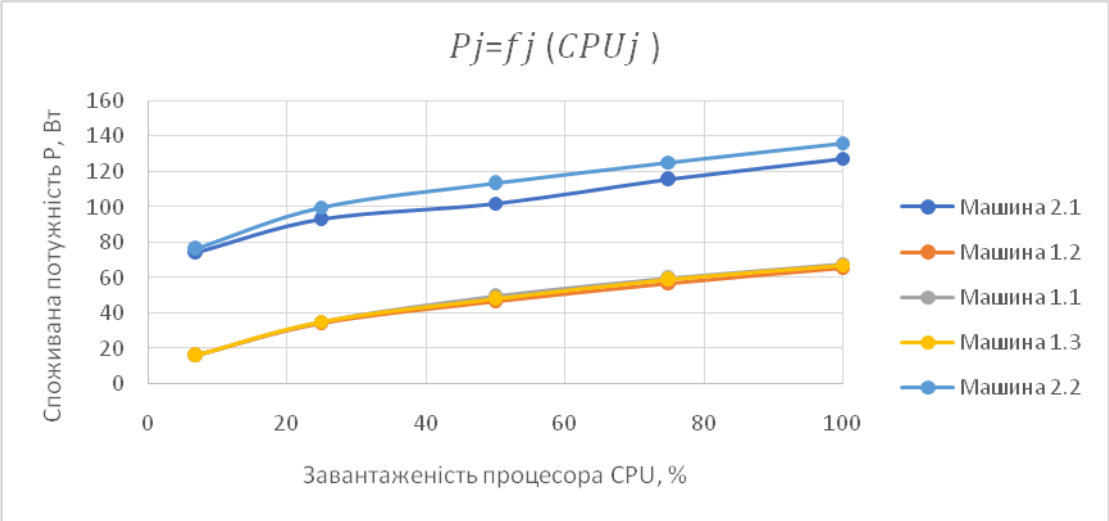


Рис. 3. Експериментально визначені енергетичні моделі для кожного з вузлів кластера

На рисунку нумерація машин складається із двох цифр, розділених крапкою. Перша з них означає номер типу машини (у експерименті було

використано 2 основних типи за фізичними характеристиками), друга – номер машини у групі.

Після проведення етапу попередньої атестації на центральний вузол кластера, що виступає брокером задач, було завантажено програмний код із логікою алгоритму Round Robin.

Кластер було навантажено пакетом задач, кожна з яких являла собою підрахунок числа π із точністю 1000, 5000, 10000, 25000, 50000, 75000, 100000 знаків після коми.

Під час експерименту було зафіксовано час обробки пакету задач та кількість спожитої кластером енергії. Для нівелювання впливу випадкових похибок було проведено 20 тотожних експериментів та статистична обробка отриманих даних. Результати, показані експериментом, наведені у таблиці 1 (значення в таблиці являють собою середні арифметичні значення 20 вимірювань. Похибка вимірювання становить не більше 1%). Спожита потужність – це сумарна потужність, що була спожита всіма вузлами кластеру за час обробки заданого пакету задач.

Після дослідження роботи алгоритму Round Robin код брокера було змінено на той, що відповідає логіці запропонованого алгоритму. Пакет задач, що був використаний у ході експерименту ідентичний попередньому. Результати порівняння запропонованого алгоритму із попередньо дослідженим Round Robin наведені у табл. 1.

Таблиця 1

Результат експериментального порівняння запропонованого алгоритму та алгоритму Round Robin

Алгоритм	**Середній час обробки однієї задачі, с**	**Сумарна спожита потужність *, Вт**	**Виграш у порівнянні із алгоритмом RoundRobin, %**	
			По часу	**По спожитій потужності**
Round Robin	14.835	376.88	0%	0%
Запропонований алгоритм	13.324	365.8	**10,2%**	**3%**

Отриманий завдяки використанню запропонованого алгоритму виграш становить 10,2% за продуктивністю (зв’язок продуктивності кластера із середнім часом обробки однієї задачі пояснено у розділі 1 роботи) та 3% за енергоефективністю.

Отриманий виграш не є значним. Це пояснюється тим, що серверний кластер, організований у експерименті, складався лише із 5 вузлів, що мали подібні фізичні характеристики. Для перевірки залежності

ефективності алгоритму від розміру кластера та параметрів його вузлів було проведено імітаційне моделювання у середовищі MATLAB, що докладно описано у розділі 5 роботи.

5. Імітаційне моделювання роботи запропонованого підходу для великих кластерів та його порівняння з іншими алгоритмами балансування навантаження

Експериментальний аналіз роботи запропонованого підходу для кластера з 5 вузлів показав, що даний підхід підвищує ефективність обробки задач за обома досліджуваними параметрами – продуктивності та енергоефективності обробки задач.

Для підтвердження припущення, що запропонований підхід проявляє себе краще при балансуванні навантаження у великих серверних кластерах, вузли яких є гетерогенними за своїми параметрами, було проведено імітаційне моделювання із використанням спеціалізованого середовища моделювання MATLAB.

План проведення моделювання

Для доведення можливості використання методу імітаційного моделювання у даному випадку та адекватної відповідності результатів моделювання реальному експериментові, у першу чергу необхідно провести моделювання роботи кластера за умов, ідентичних до проведеного натурного експерименту та зробити відповідні висновки.

При досліджені моделі було вирішено порівняти запропонований підхід із рядом відомих. Для цього, для дослідження були обрані такі алгоритми балансування:

1. Перший доступний (FIFO) – як найпростіший алгоритм, що може бути використаний для планування задач в рамках серверного кластера.
2. Round Robin (RR) – як алгоритм планування, що широко застосовується у роботі реальних ЦОД (центрів обробки даних).
3. Least Connections – як ще один широко використовуваний алгоритм, що використовує дані про поточну завантаженість вузлів кластера.

Всі вибрані алгоритми докладно описано у розділі 1 даної роботи.

Згідно цього, моделювання було проведено у відповідності з наступним планом:

1. Визначення аналітичних виразів функцій P(CPU) кожної машини.
2. Моделювання роботи алгоритму PCPB у кластері з 5 вузлів, що відповідає реально проведеному експерименту, для валідації адекватності моделювання.
3. Послідовне моделювання роботи алгоритмів FIFO, RoundRobin, LeastConnections у кластері з 5 вузлів.

4. Моделювання роботи алгоритму PCPB у кластері з 20 вузлів, для перевірки залежності ефективності алгоритму від розміру кластера.

5. Послідовне моделювання роботи алгоритмів FIFO, RoundRobin, LeastConnections у кластері з 20 вузлів.

6. Моделювання робочої доби (24 години) серверного кластера з 20 машин з використанням реального розподілу навантаження на кластер впродовж дня. Перевірка та порівняння для даних умов алгоритмів PCPB та FIFO, RoundRobin, LeastConnections.

Задачі, що використовувались при моделюванні були задані наступним чином:

– Інтервал часу між появою нових задач – величина розподілена по нормальному закону, значення якої коливаються в діапазоні 0 ... 3600 секунд.

– Об’єм оперативної пам'яті, необхідний для виконання кожної задачі: 100 Мб ... 1 Гб.

– Кількість обчислювальних ядер вузла кластера, необхідних для обробки кожної із задач: 1 ... 4.

– Обсяг кожної задачі, як число операцій з плаваючою комою, які мають бути проведені: 10 ... 500 GFLOPS.

– Кількість задач у вхідному наборі: 1750.

Пакет задач був однаковим для кожної ітерації моделювання, при цьому задачі в рамках пакету були різними за характеристиками.

Опис процесу моделювання та аналіз результатів

Визначення енергетичних моделей вузлів кластера

Згідно описаного у розділі 2 підходу у першу чергу необхідно визначити енергетичні моделі для кожного з вузлів кластера. Для цього, на основі графіка, отриманого в ході натурного експерименту, криву енергетичної моделі засобами середовища MATLAB було апроксимовано поліномом 4 ступеня вигляду:

$$P(CPU) = a * x^4 + b * x^3 + c * x^2 + d * x^1 + e * x^0.$$

Коефіцієнти a, b, c, d відповідно різнилися для кожного вузла кластера.

Моделювання кластера з 5 вузлів

Для перевірки адекватності імітаційної моделі було змодельовано кластер з 5 вузлів, параметри яких були ідентичними натурному експерименту.

– Кількість вузлів в кластері: 5.

– Об’єм оперативної пам'яті кожного вузла {4096, 4096, 4096, 8192, 8192} [Мб] відповідно.

– Кількість ядер процесора для кожного вузла {4; 4; 4; 4; 4} відповідно.

– Продуктивність кожного вузла {35,7; 35,7; 35,7; 40; 50} [GFLOPS].

– Енергетична модель P(CPU) у вигляді полінома:

$$P(CPU) = a * x^4 + b * x^3 + c * x^2 + d * x^1 + e * x^0.$$

Результати моделювання для кластера із 5 вузлів представлено у табл. 2.

Як видно з таблиці, отриманий виграш запропонованого алгоритму статистично еквівалентний експериментально визначеному виграшу, що дає підставу стверджувати, що імітаційне моделювання застосовне у даному випадку та дає результат, що відповідає дійсності у достатній мірі.

Таблиця 2

Результати моделювання роботи кластера з 5 вузлів у середовищі MATLAB

Алгоритм	**Середній час обробки однієї задачі, с**	**Сумарна спожита потужність *, Вт**	**Виграш у порівнянні із алгоритмом FIFO, %**		**Виграш у порівнянні із алгоритмом RoundRobin, %**	
			По часу	**По спожитій потужності**	**По часу**	**По спожитій потужності**
First available (FIFO)	13.146	371.61	0%	0%	12,8%	1,4%
Round Robin	14.835	376.88	-12,8%	-1,4%	0%	0%
Least Connections	13.483	375.64	-2,5%	-1,1%	9,1%	0,3%
Запропонований алгоритм	12.824	364.6	**2,4%**	**1,5%**	**13,6%**	**3,3%**

Моделювання кластера з 20 вузлів

Для перевірки ефективності запропонованого підходу у кластері із більшою кількістю вузлів, було проведено моделювання серверного кластера у складі 20 вузлів із різними фізичними параметрами. Параметри були наступними:

– Кількість вузлів в кластері: 20.

– Об’єм оперативної пам'яті кожного вузла:{4096, 4096, 4096, 4096, 4096, 16384, 8192, 8192, 8192, 8192, 8192, 8192, 4096, 4096, 8192, 4096, 8192, 4096, 4096, 8192} [Мб].

– Кількість ядер процесора для кожного вузла:{Всі 20 вузлів – по 4 ядра} [ядер].

– Продуктивність кожного вузла: {35.7, 35.7, 45.7, 45.7, 45.7, 75, 25.7, 64.7, 64.7, 45.7, 30.7, 30.7, 45.7, 45.7, 64.7, 35.7, 35.7, 45.7, 45.7, 64.7} [GFLOPS].

Послідовно було проведено моделювання роботи кожного із досліджуваних алгоритмів, як і в попередньому випадку. Результат моделювання для кластера з 20 вузлів відображено у табл. 3.

Таблиця 3

Результати моделювання роботи кластера з 20 вузлів у середовищі MATLAB

Алгоритм	**Середній час обробки однієї задачі, с**	**Сумарна спожита потужність * , Вт**	**Виграш у порівнянні із алгоритмом FIFO, %**		**Виграш у порівнянні із алгоритмом Round Robin, %**	
			По часу	**По спожитій потужності**	**По часу**	**По спожитій потужності**
First available (FIFO)	11.636	1436.1	0%	0%	25.2%	4.5%
Round Robin	15.55	1500.8	-25.2%	-4.5%	0%	0%
Least Connections	13.01	1455.9	-11.9%	-1.4%	16.38%	3%
Запропонований алгоритм	10.89	1376.3	**6.4%**	**4.2%**	**30%**	**8.3%**

Аналіз результатів, наведених у таблиці, показує, що для більших кластерів запропонований підхід демонструє вищу ефективність за обома параметрами. За параметром продуктивності виграш запропонованого алгоритму становить 30% у порівнянні із алгоритмом RoundRobin та за параметром енергоефективності – 8,3%. Цей результат дає підставу стверджувати, що запропонований підхід є доречним при використанні у порівняно великих кластерах, що містять 20 і більше вузлів із різноманітними характеристиками.

Після проведення моделювання роботи кластера з 20 вузлів, для максимального наближення умов моделювання до реальних, було зімітовано роботу кластера із 20 вузлів протягом доби.

Приклад розподілу навантаження на реальний кластер протягом доби зображено на рис. 4 [7]. Даний розподіл було використано при моделюванні.

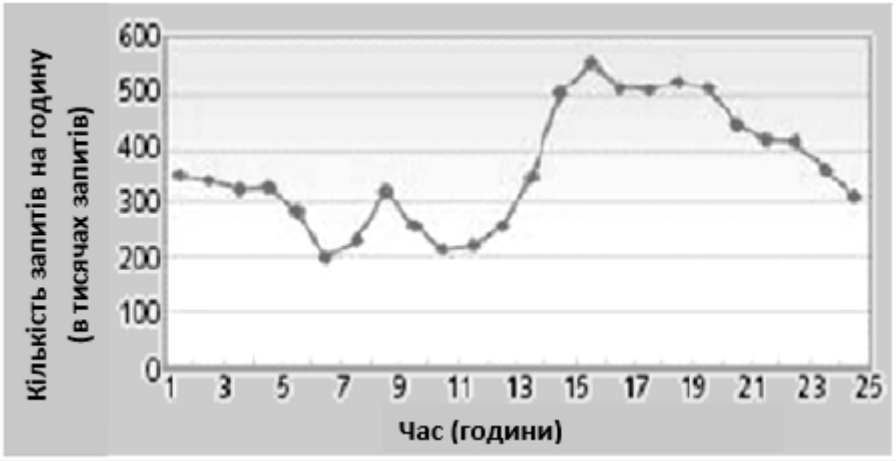


Рис. 4. Приклад розподілу навантаження на реальний серверний кластер протягом доби [7]

Модель кластера була ідентичною до моделі, описаної у розділі 5.2.2. Пакет, що підлягав обробці, складався з 68850 задач.

Результат моделювання роботи кластера з 20 вузлів впродовж доби представлений у табл. 4.

Таблиця 4

Результати моделювання роботи кластера з 20 вузлів протягом доби

Алгоритм	**Середній час обробки однієї задачі, с**	**Сумарна спожита потужність *, Вт**	**Виграш у порівнянні із алгоритмом FIFO, %**		**Виграш у порівнянні із алгоритмом RoundRobin, %**	
			По часу	**По спожитій потужності**	**По часу**	**По спожитій потужності**
First available (FIFO)	11.71	29725	0%	0%	34,58%	2,86%
Round Robin	15.76	30578	-34,58%	-2,86%	0%	0%
Least Connections	12.91	29739	-10,24%	-0,0005%	9,1%	2,74%
Запропонований алгоритм	10.07	28568	**14%**	**3,9%**	**36,1%**	**6,57%**

Як видно з таблиці, алгоритм дає виграш у продуктивності на 36,1% та у енергоефективності – на 6,57%. Це дає підставу припускати, що підхід повинен показати себе добре при роботі у реальному ЦОД.

Енергоспоживання для різних алгоритмів для даного експерименту наведено також на рис. 5. Запропонований алгоритм позначено на графіку як Proposed (від англ. *запропонований*).

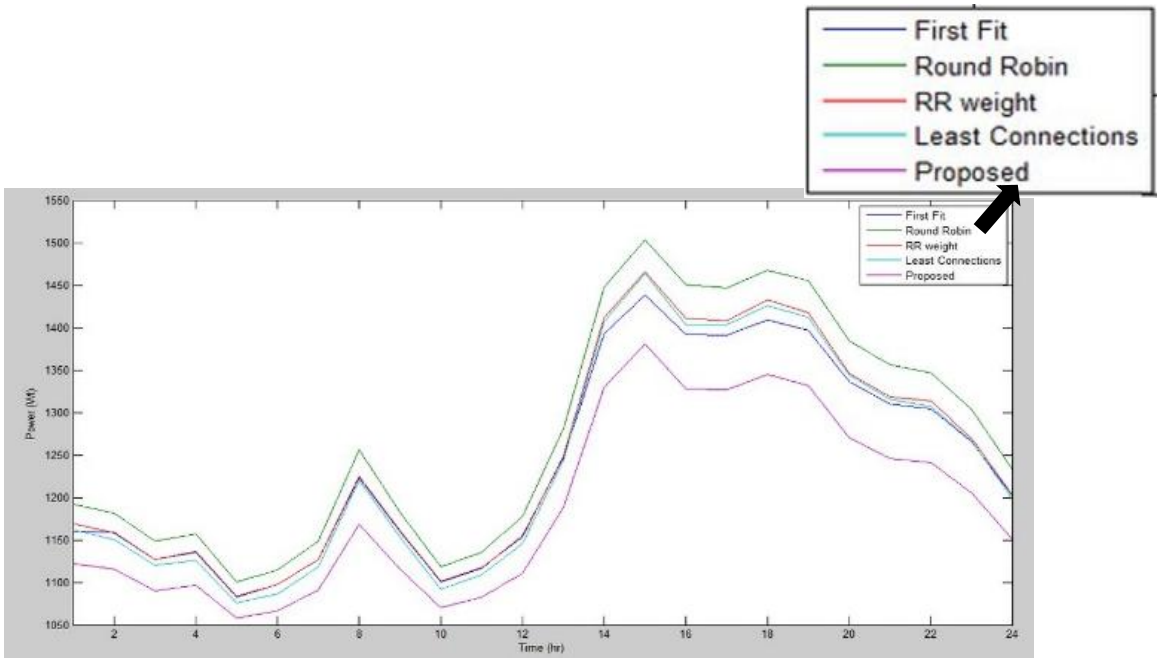


Рис. 5. Залежність споживаної кластером потужності від часу протягом доби роботи кластера (результат моделювання)

Висновки

У роботі запропоновано унікальний підхід до балансування задач у серверному кластері, що має на меті оптимізацію процесу балансування за двома критеріями – продуктивність та енергоефективність обробки задач. Запропонований підхід відрізняється тим, що передбачає використання індивідуального підходу до визначення енергетичних моделей вузлів кластера та розподіл задач з урахуванням їх енергетичних моделей.

У роботі описано процес проведення натурного експерименту у серверному кластері з 5 вузлів. Описано процес проведення імітаційного моделювання для кластера із 20 вузлів, у тому числі – роботи кластера протягом доби.

У ході досліджень виявлено, що запропонований підхід дає змогу отримати виграш у 36,1% за параметром продуктивності обробки задач та 8,3% за параметром енергоефективності у порівнянні із широко використовуваним алгоритмом балансування RoundRobin.

У роботі також запропоновані можливі майбутні модифікації до алгоритму з метою підвищення його ефективності.

Проведене дослідження показало, що запропонований у роботі підхід доцільно використовувати у промисловому масштабі у ЦОД для збільшення енергоефективності обробки задач із одночасним підвищенням продуктивності.

Запропонований підхід стане зокрема слушним при поширенні сервісів InternetofThings та M2M із розвитком технологій зв'язку 4G та 5G.

Література

1. J. G. Koomey “World wide electricity used in data centers,” Environmental Research Letters, vol. 3, no. 3, p. 034008, 2008. [Online]. Available: http://stacks.iop.org/1748-9326/3/i=3/a=034008
2. “Growth in data center electricity use 2005 to 2010,” July 2011. [Online]. Available: http://www.analyticspress.com/datacenters.html
3. S. Hosseinimotlagh, F. Khunjush, and S. Hosseinimotlagh, A Cooperative Two-Tier Energy-Aware Scheduling for Real-Time Tasks in Computing Clouds, in Proceedings of the 2014 22Nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, Washington, DC, USA, 2014, pp. 178 – 182.
4. Y. Gao, Y. Wang, S. K. Gupta, and M. Pedram, An Energy and Deadline Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems in Proceedings of the Ninth IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis, Piscataway, NJ, USA, 2013, pp. 1–10.
5. Ф. Армента-Кано, А. Черных, Х.М. Кортес-Мендоза, Р. Яхьяпур, А.Ю. Дроздов, П. Буври, Д. Клязович, А. Аветисян, С. Несмачнов Min_c: стратегия неоднородной концентрации задач для энергосберегающих компьютерных рас писаний. – Труды ИСП РАН, 2015. – том 27, вып. 6. – 355 с.
6. https://ru.wikipedia.org/wiki/FLOPS
7. http://www.osp.ru/os/2004/02/183912/

СУЧАСНИЙ СТАН ТА АСПЕКТИ ВИКОРИСТАННЯ ІНТЕРНЕТУ РЕЧЕЙ

Климаш М.М., Стрихалюк Б.М., Климаш Ю.В.

Вступ

У наш час відбувається посилення тенденції технологізації різних сфер життєдіяльності. Інформаційні та телекомунікаційні технології стали не тільки невіддільною частиною повсякденного життя сучасної людини, але і необхідною технологічною платформою для організації сучасних бізнес-процесів. Активний розвиток смартфонів, створення мобільних додатків для гаджетів вже зараз дозволяє оперативно відслідковувати, фіксувати, зберігати різні аспекти життя людини: від списку постійних контактів, послідовності виконання робочих функцій, здійснених банківських транзакцій, останніх покупок до стану фізичного та емоційного самопочуття. Однак нові інформаційні технології виводять рівень збору, агрегації і обміну накопиченою інформацією на принципово інший якісний рівень з мінімальною роллю і ступенем участі людини.

Однією з технологій, що впливає на зміну економічного, соціального та культурного ландшафту сучасного суспільства, є концепція Інтернету речей [1]. Дана концепція дозволяє не тільки об'єднувати предмети матеріального світу за допомогою Інтернету для обміну інформацією між ними, а й розвивати можливості по накопиченню, структуруванню та аналізу різної інформації про поведінку людей в різних місцях [2,3].

Дослідницькі компанії і ряд вчених протягом останніх років розглядають Інтернет речей як новий етап розвитку Інтернету, вказуючи на ті безмежні можливості по об'єднанню людей, процесів, даних і речей, які він надає пов'язані з цим технологічні рішення. Щоб скористатися потенційними перевагами для людей, суспільства і економіки, необхідно вирішити питання і проблеми, які виникають при впровадженні даної концепції. У зв’язку з цим актуальною задачею є аналіз аспектів та проблем застосування Інтернету речей.

1. Концепція Інтернету речей

У загальному випадку під Інтернетом речей (Internet of Things, IoT) розуміється сукупність різноманітних сенсорів, пристроїв, об'єднаних в мережу за допомогою будь-яких доступних каналів зв'язку, що використовують різні протоколи взаємодії між собою і єдиний протокол доступу до глобальної мережі – Інтернету.

Незважаючи на те, що термін «Інтернет речей» є порівняно новим, концепція об'єднання комп'ютерів і мереж для моніторингу та керування пристроями існує вже кілька десятиліть. Наприклад, уже в кінці 1970-х рр. здійснювалося комерційне використання систем для віддаленого моніторингу лічильників електричної мережі через телефонні лінії. У 1990-

і рр. досягнення в області безпровідної технології уможливили широке поширення корпоративних і виробничих рішень «Machine-to-Machine» (M2M) для моніторингу та управління обладнанням. Однак багато з цих ранніх рішень M2M були створені на основі закритих спеціалізованих мереж на фірмових або галузевих стандартах, а не на мережах на основі протоколу Internet Protocol (IP) і стандартів Інтернету. Ідея використання IP для підключення до Інтернету пристроїв, які не є комп'ютерами, не нова. Перший пристрій з підключенням до Інтернету – тостер з підтримкою протоколу IP, який можна було включати і вимикати через Інтернет, було представлено в 1990 році Johnom Romkey. Проте тільки в 21 столітті в зв'язку з бурхливим розвитком інформаційно-комунікаційних технологій сформувалася концепція IoT і отримала своє практичне втілення .

У 2003 році на нашій планеті проживало близько 6,3 млрд осіб, в той час до Інтернету було підключено 500 млн пристроїв. Розділивши кількість підключених пристроїв на величину населення земної кулі, ми побачимо, що на кожну людину тоді відводилось по 0,08 такого пристрою. У 2010 році в результаті стрімкого поширення смартфонів і планшетних комп'ютерів кількість підключених пристроїв зросла до 12,5 млрд, тоді як населення Землі склало 6,8 млрд чоловік.

Таким чином, вперше в історії на кожну людину стало припадати більше одного підключеного пристрою (1,84 пристрою на одну людину) (рис. 1).

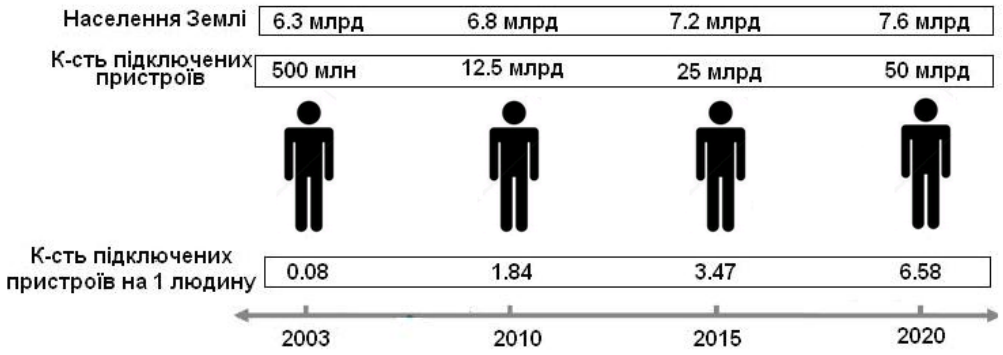


Рис. 1. Статистичні дані Cisco IBSG щодо кількості підключених пристроїв до Інтернету на людину

Заглядаючи в майбутнє, Cisco IBSG прогнозує до 2020 року – 50 млрд пристроїв. Важливо зауважити, що ці прогнози не враховують прискореного розвитку інтернет-технологій і пристроїв.

У міру збільшення числа пристроїв, підключених до Інтернету, очікується істотне збільшення трафіку. Наприклад, згідно з прогнозами Cisco, інтенсивність мережевого обміну даними між пристроями (не включаючи ПК) збільшиться з 40% в 2015 р до майже 70% в 2019 р. Cisco також прогнозує, що число «міжмашинних» підключень («M2M»)

(включаючи промислові, домашні, медичні, автомобільні та інші вертикалі IoT) збільшиться з 24% від всіх підключених пристроїв в 2015 р до 43% в 2019 р.

Сьогодні Інтернет речей складається з слабко пов'язаних між собою розрізнених мереж, кожна з яких була розгорнута для вирішення своїх специфічних завдань. Наприклад, в сучасних автомобілях працюють декілька мереж: одна керує роботою двигуна, інша – системами безпеки, і т. п. В офісних і житлових будівлях також встановлюється безліч систем для управління опаленням, вентиляцією, кондиціюванням, телефонним зв'язком, безпекою, освітленням тощо. Смартфон не вважався б «розумним», якби не наявність безлічі сенсорів, вбудованих в кожен пристрій, зазвичай це 5-9 сенсорів (в залежності від моделі). В цей перелік входять: сенсор зовнішнього освітлення, акселерометр, магнітометр, барометр, сенсор вологості, температури і т.д. Функція сенсорів набагато ширша, ніж просте забезпечення працездатності наших мобільних телефонів. Насправді вони є важливим компонентом, який «запускає» IoT. Постійно збираючи дані про своє оточення, сенсор стає основним способом отримання даних комп'ютером. Вони можуть отримувати і обробляти дані зі швидкістю і в кількостях, з якими жодна людина не зможе зрівнятися. Саме сенсори сприяли виникненню феномена, який сьогодні називається «великі масиви даних».

Перш ніж аналізувати значення Інтернету речей, потрібно зрозуміти різницю між Інтернетом і тим, що іменується «Всесвітньою павутиною» (World Wide Web, або просто «Веб»). Ці терміни часто використовуються як абсолютні синоніми, що не є коректно.

Інтернет – це фізичний рівень мереж: комутатори, маршрутизатори та інше обладнання. Його головна функція полягає у швидкій, надійній і безпечній передачі інформації з однієї точки в іншу. «Веб» – це рівень додатків, що працює поверх Інтернету. Його завдання – це створення інтерфейсу для отримання реальної користі від переданої через Інтернет інформації.

На відміну від веб-технологій, Інтернет стабільно розвивався перш за все в кількісному відношенні, майже не змінюючись якісно. Сьогодні Інтернет виконує приблизно ті ж завдання, які ставилися перед ним за часів мережі ARPANET, коли існувало кілька комунікаційних протоколів (AppleTalk, Token Ring і I).

У IoT кожна річ має свій унікальний ідентифікатор, які спільно утворюють середовище речей, здатних взаємодіяти один з одним, створюючи тимчасові або постійні мережі [4,5]. Так речі можуть брати участь в процесі їх переміщення, ділячись інформацією про поточну геопозицію, що дозволяє повністю автоматизувати процес логістики, а маючи вбудований інтелект, речі можуть змінювати свої властивості і адаптуватися до навколишнього середовища, в тому числі для зменшення

енергоспоживання. Вони можуть виявляти інші, так чи інакше пов'язані з ними речі, і налагоджувати з ними взаємодію. IoT дозволяє створювати комбінацію з інтелектуальних пристроїв, об'єднаних мережами зв'язку, і людей. Спільно вони можуть створювати найрізноманітніші системи.

IoT викликав широке поширення сенсорів температури, тиску, вібрації, освітлення, вологості і фізичних навантажень, які допомагають нам спростити та покращити наше життя. Крім того, він почав проникати в раніше недоступні сфери. Наприклад, пацієнти ковтають інтернет-пристрої, що дозволяють точно діагностувати деякі захворювання і виявляти їх причини; мікроскопічні сенсори, підключені до Інтернету, можна закріплювати на рослинах, тваринах, геологічних утвореннях тощо.

Отже, Інтернет речей передбачає, як мінімум[6]:

– наявність широкого кола пристроїв (причому не тільки «звичайних» інтернет-терміналів – персональних комп'ютерів, смартфонів і т.п.), підключених до мережі Інтернет;

– збір значного масиву даних про навколишній простір (як персональних даних, так і іншої інформації), а також обмін даних між зазначеними пристроями;

– можливість автоматичного (без безпосередньої участі людини) виконання пристроями IoT функцій (здійснення дій), які можуть мати юридичне значення та наслідки для людей.

Інтернет речей складається з безлічі сегментів і ринків. Для споживача – це технології і «розумні» прилади, такі як термостати, телевізори тощо, у промисловому секторі – автономні машини і обладнання з сенсорами, у бізнес-просторі – великі масиви даних і маркетингова аналітика. Іншими словами, Інтернет речей настільки ж різноманітний, як і світова економіка: від виробництва до споживання продуктів.

Фактори, які внесли свій вклад в розвиток IoT , представлено на рис. 2. Коротко розглянемо їх.

Повсюдне недороге і високошвидкісне з'єднання з мережею, особливо за допомогою ліцензованих і неліцензованих послуг безпровідного зв'язку і технологій, дозволяє підключити до мережі практично будь-який предмет.

Рис. 2. Фактори розвитку Інтернету речей

Протокол IP став основним глобальним мережевим стандартом, що забезпечує чітко визначену і широко використовувану платформу для програмного забезпечення та інструментів, що може бути легко і без великих витрат включено в широкий спектр пристроїв.

Хмарні обчислення, що використовують віддалені мережеві ресурси для обробки, управління і зберігання даних, дозволяють невеликим і розподіленим пристроям взаємодіяти з потужними функціями аналізу та управління на сервері [7].

Досягнення в області виробництва дозволяють застосовувати найсучасніші технології обчислень і зв'язку в об'єктах дуже малого розміру. У поєднанні з більш високою економічністю обчислень це послужило поштовхом для створення недорогих сенсорів малого розміру, на яких засновано безліч областей застосування IoT.

У міру розвитку будуть підключатися та використовуватися більш широкі засоби безпеки, аналітики та управління, і в результаті IoT отримає ще більше можливостей відкрити людству нові, великі перспективи.

2. Моделі комунікації Інтернету речей

З практичної точки зору корисно розглянути, як пристрої IoT здійснюють підключення та зв'язок відповідно до своїх технічних моделей зв'язку. У березні 2015 року Комісія з архітектури Інтернету (IAB) випустила директивний документ по архітектурі для мережевого підключення інтелектуальних об'єктів (RFC 7452), в якому визначається концептуальна основа чотирьох загальних моделей зв'язку, що використовуються пристроями IoT.

Підключення від пристрою до пристрою

Дана модель зв'язку представляє два або декілька пристроїв, що підключені і здійснюють зв'язок один з одним безпосередньо, а не через проміжний сервер додатків. Ці пристрої здійснюють зв'язок через різні типи мереж, в тому числі, мережі на основі протоколу IP або Інтернет.

Однак часто ці пристрої використовують такі протоколи, як Bluetooth, 40 Z-Wave41 або ZigBee42 для встановлення прямого зв'язку від пристрою до пристрою (рис.3).

Рис. 3. Приклад моделі зв'язку від пристрою до пристрою

Ця модель зв'язку зазвичай застосовується в таких додатках, як домашні системи автоматики, в яких використовуються пакети даних малого розміру для встановлення зв'язку між пристроями з низьким рівнем вимог в області швидкості передачі даних.

Побутові пристрої IoT, такі як лампочки, вимикачі, термостати і дверні замки, в домашній системі автоматики обмінюються малим обсягом інформації (наприклад, повідомлення про стан дверного замка або команда включення/виключення світла). Такі пристрої часто знаходяться в безпосередньому зв'язку, зазвичай вони оснащені вбудованими механізмами безпеки, але також використовують певні моделі даних для кожного пристрою, що вимагає додаткових зусиль в розробці. Це означає, що виробники пристроїв повинн вкладати кошти в розробку певних форматів даних для кожного типу пристроїв замість використання відкритої платформи для стандартних форматів.

З точки зору користувачів, це часто означає використання несумісних протоколів передачі даних, тому користувач змушений вибирати інші пристрої, що підтримують той же протокол. Наприклад, пристрої, що використовують протокол Z-Wave, несумісні з пристроями сімейства ZigBee.

Підключення від пристрою до хмари

В даній моделі зв’язку пристрій IoT підключається безпосередньо до хмарної інтернет-служби, такої як постачальник послуг оренди додатків, для обміну даними і управління трафіком повідомлень. При такому підході часто використовуються існуючі механізми зв’язку, такі як традиційні з’єднання Ethernet або Wi-Fi для встановлення з'єднання між пристроєм і мережею IP, яка, в свою чергу, підключається до хмарної служби.

Ця модель з'єднання використовується деякими популярними споживими пристроями IoT, такими як термостат Nest Labs44 і SmartTV виробництва Samsung. У випадку термостата Nest пристрій передає дані в хмарну базу даних, де ці дані можуть використовуватися для аналізу споживання електроенергії вдома. Це хмарне підключення дозволяє користувачеві отримувати віддалений доступ до свого термостата через смартфон або веб-інтерфейс, а також підтримує оновлення програмного забезпечення термостата. Аналогічно у випадку технології SmartTV виробництва Samsung, телевізор використовує підключення до Інтернету для передачі інформації про використовувані користувачем програми в Samsung для аналізу і підключення інтерактивної функції розпізнавання голосу на пристрої телевізора. У цих випадках модель підключення пристрою до хмари забезпечує додаткову цінність для кінцевого користувача за рахунок розширення стандартних функцій пристрою.

Проте проблеми інтероперабельності можуть виникнути при спробі інтеграції пристроїв різних виробників. Найчастіше використовуються

хмарні послуги та пристрій одного виробника. Якщо для зв'язку між пристроєм і хмарними службами використовуються патентовані протоколи даних, власник або користувач пристрою може користуватися лише певною хмарною службою, що обмежує його можливість користуватися послугами інших постачальників. Така ситуація позначається терміном «залежність від постачальника», яка охоплює різні аспекти відносин з постачальником, такі як володіння даними і доступ до них. У той же час користувачі зазвичай можуть бути впевнені в можливості інтеграції пристроїв, створених для певної платформи.

Підключення від пристрою до шлюзу

У випадку моделі підключення між пристроєм і шлюзом або, найчастіше, моделі підключення пристрою до шлюзу прикладного рівня (ALG), пристрій IoT підключається через службу ALG як канал для використання хмарної служби. Простіше кажучи, це означає, що прикладне програмне забезпечення функціонує на пристрої локального шлюзу, яке виконує роль посередника між пристроєм і хмарною службою та забезпечує безпеку і інші функції, такі як перетворення даних або протоколи. У призначених для користувача пристроях присутні різні варіанти цієї моделі. В багатьох випадках в якості локального шлюзу використовується смартфон з додатком для зв'язку з пристроєм і передачі даних в хмарну службу. Така модель часто використовується з популярними споживчими пристроями, такими як браслети для занять спортом. В цих пристроях відсутня функція прямого підключення до хмарної служби, тому вони часто використовують додаток смартфона для роботи в якості шлюзу підключення.

Іншим різновидом цієї моделі підключення пристроїв до шлюзу є пристрої, що виконують роль концентратора в додатках домашньої автоматики. Такі пристрої використовуються в якості локального шлюзу між окремими пристроями IoT і хмарною службою, але вони також можуть заповнювати прогалини інтероперабельності між самими пристроями. Наприклад, концентратор SmartThings – окремий пристрій шлюзу з трансиверами Z-Wave і Zigbee, встановленими для підтримки зв'язку з обома типами пристроїв. Тому концентратор буде з'єднуватись з хмарною службою SmartThings, завдяки якій користувач може отримувати доступ до пристроїв за допомогою програми смартфона і підключення до Інтернету.

Розглянута модель зв'язку використовується в тих випадках, коли інтелектуальні об'єкти вимагають інтероперабельності з пристроями, що не підтримують IP. Іноді цей підхід використовується для інтеграції пристроїв, що підтримують тільки протокол IPv6.

Підсумовуючи можна сказати, що така модель зв'язку часто використовується для інтеграції нових інтелектуальних пристроїв у традиційну систему з пристроями, які спочатку не можуть з ними

взаємодіяти. Недолік цього підходу полягає в тому, що необхідність розробки системи і шлюзу прикладного рівня збільшує складність і вартість системи в цілому.

Очікується, що в майбутньому будуть створені більш універсальні шлюзи для зниження вартості та рівня складності інфраструктури для кінцевих споживачів, підприємства і промислового застосування. Існування таких універсальних шлюзів ймовірніше в тому випадку, якщо конструкція пристрою IoT підтримуватиме універсальні протоколи Інтернету і не вимагатиме наявності шлюзу прикладного рівня для перетворення протоколів.

Модель спільного використання даних на сервері

Дана модель відповідає архітектурі, що дозволяє користувачам експортувати і аналізувати дані інтелектуальних об'єктів з хмарної служби в поєднанні з даними з інших джерел. Така архітектура підтримує можливість надання доступу для третіх сторін до завантажених даними сенсорів. Цей підхід відповідає моделі з'єднання окремих пристроїв з хмарою, що може призвести до створення вихідної бази даних, де пристрої IoT завантажують дані тільки для одного постачальника послуг оренди додатків.

Архітектура спільного використання даних на сервері дозволяє об'єднувати і аналізувати потоки даних, отриманих від одного пристрою IoT. Наприклад, корпоративний користувач, відповідальний за офіс, може бути зацікавлений в об'єднанні і аналізі даних про фактичне споживання електроенергії та інші комунальні послуги, що отримуються всіма сенсорами IoT і системами інженерного забезпечення з підключенням до Інтернет. У моделі підключення окремих пристроїв до хмарних служб дані кожного датчика або системи IoT знаходяться в окремій базі. Ефективна архітектура спільного використання даних на сервері повинна дозволяти компанії з легкістю отримувати доступ і аналізувати хмарні дані, отримані від всіх пристроїв в будівлі. Крім того, цей тип архітектури дозволяє забезпечити переносимість даних. Ефективна архітектура спільного використання даних на сервері дозволяє користувачам переміщати свої дані при перемиканні між послугами IoT, долаючи бар'єри традиційних роздільних баз даних.

Модель спільного використання даних на сервері передбачає об'єднаний підхід до хмарних послуг; в іншому випадку необхідні хмарні інтерфейси прикладного програмування (API) для забезпечення інтероперабельності, розміщених на хмарі даних з інтелектуальними пристроями.

Дана модель архітектури є підходом для забезпечення інтероперабельності між системами на базі сервера. Як вказується в IETF Journal, «стандартні протоколи можуть полегшити завдання, але їх

недостатньо для видалення вузькоспеціальних баз даних, так як для взаємодії між різними виробниками необхідна наявність загальних інформаційних моделей ».

Іншими словами, ця модель зв'язку ефективна тільки на основі архітектури системи IoT.

Проведемо підсумок розглянутих моделей комунікації.

Чотири основні моделі зв'язку демонструють стратегії розробки, що застосовуються для забезпечення зв'язку між пристроями IoT.

Крім технічних аспектів, застосування цих моделей багато в чому визначається відмінностями між патентованими і відкритими IoT пристроями в мережі. У випадку використання моделі зв'язку пристрою зі шлюзом її основною характеристикою є здатність подолання обмежень при підключенні патентованих пристроїв IoT. Це означає, що інтероперабельність пристрою і відкриті стандарти є ключовою умовою для створення і розвитку взаємопов'язаних систем IoT.

Розглянуті моделі зв'язків дозволяють краще зрозуміти можливість створення додаткової цінності для кінцевих користувачів з допомогою мережевих пристроїв. Загальна цінність пристроїв підвищується за рахунок надання користувачам більш зручного доступу до IoT пристроїв та їх даних. Наприклад, в трьох з чотирьох моделей зв'язку пристрої підключаються до служб аналізу даних на основі хмарних обчислень. За рахунок створення каналів передачі даних на хмарі користувачі і постачальники послуг можуть більш швидко і легко об'єднувати дані, проводити їх аналіз і візуалізацію, а також застосовувати технології аналітичного прогнозування, щоб скористатися додатковими перевагами даних IoT, одержаних за допомогою традиційних додатків вузькоспеціальною базою даних. Іншими словами, ефективні моделі зв'язку є важливим фактором для підвищення цінності послуг для кінцевих користувачів за рахунок можливості застосування нових способів використання інформації. Однак, незважаючи на ці переваги, також є недоліки. При виборі архітектури необхідно ретельно врахувати питання додаткових витрат для користувачів при підключенні до хмарних ресурсів, особливо в регіонах з високою вартістю послуг зв'язку.

Також слід зауважити, що ефективні моделі зв'язку IoT сприяють розвитку технічних інновацій і відкривають можливості комерційного зростання.

3. Практичне застосування IoT

На основі Інтернету речей можуть бути реалізовані «розумні» (smart) додатки в різних сферах діяльності і життя людини :

«Розумна планета» – людина зможе буквально «тримати руку на пульсі» планети: своєчасно реагувати на забруднення та інші екологічні

проблеми, а значить, ефективно розпоряджатися невідновлюваними ресурсами.

«Розумне місто» – міська інфраструктура і супутні муніципальні послуги, такі як освіта, охорона здоров'я, громадська безпека тощо, стануть більш пов'язаними і ефективними [8].

«Розумний будинок» – система буде розпізнавати конкретні ситуації, що відбуваються в будинку, і реагувати на них відповідним чином, що забезпечить мешканцям безпеку, комфорт і ресурсозбереження.

«Розумна енергетика» –надійна і якісна передача електричної енергії від джерела до приймача в потрібний час і в необхідній кількості.

«Розумний транспорт» – переміщення пасажирів з однієї точки в іншу стане зручніше, швидше і безпечніше.

«Розумна медицина» – лікарі і пацієнти зможуть отримати віддалений доступ до дорогого медичного обладнання або до електронної історії хвороби в будь-якому місці, буде реалізована система віддаленого моніторингу здоров'я, автоматизована видача лікарських препаратів хворим і багато іншого.

В останні роки в містах інтенсивно створюються інформаційні системи для автоматизації окремих сфер міського життя: безпеки міського середовища, транспорту, енергетики і ЖКП, охорони здоров'я, освіти, державного і муніципального управління та ін. Принципи і технології IoT дозволяють створити повнозв'язне інтегроване рішення, необхідне для функціонування міського середовища, що доступне всім жителям міста, співробітникам міських служб, чиновникам і керівникам різних рівнів.

Слід визнати, що Інтернет речей поки що не проник глибоко в елементи міської інфраструктури та господарства, але вже сформував сферу впливу, в рамках якої грає практично революційну роль. Це в першу чергу транспорт, енергетика та комунальні послуги, екологія, контроль злочинності, інформаційне забезпечення жителів міста тощо.

Інтелектуальні мобільні пристрої і високошвидкісні територіально розподілені мережі для доступу до них, сенсори, що вбудовуються в міське середовище, – все це забезпечує основу для створення всеосяжних міст (ubiquitous city), або u-міст, в яких об'єкти інфраструктури і люди тісно пов'язані [9]. Уряди декількох країн вже прийняли масштабні програми створення інтелектуальних міст U-City.

Найбільш ефективні U-системи (пов'язані на основі Інтернету речей) – це комунальна, транспортна, паркувальна служби, а також служба боротьби з злочинністю. Це, по суті, ключові проблеми міського життя, які можна вирішити на основі єдиної системи моніторингу та контролю. Так, в корейському місті Eunpyeong New Town ефективно працює U-система в сфері торгівлі у вигляді порталу з інформацією про магазини, кафе і т.д., а також система контролю місця розташування дітей, призначена для батьків.

За допомогою додатку Uber можна відстежити переміщення замовленої машини, виявити найближчих водіїв на онлайн-карті.

Збір інформації від автобусів, обладнаних системою GPS або ГЛОНАСС, дозволяє створювати інтерактивні табло, онлайн-ресурси і додатки, які інформують жителів про те, скільки їм доведеться чекати автобуса.

Наприклад, в Києві та Дніпрі встановлено «розумні» зупинки. Пасажири в режимі онлайн можуть відслідковувати номер і час прибуття, схеми і розклад маршрутів на інформаційних табло. Також зупинки обладнані підігрівом, розетками для підзарядки мобільних телефонів і безкоштовним Wi-Fi.

Інший цікавий приклад – «розумні» сміттєві контейнери. Сигнал про наповнення подається в централізовану систему управління, яка відстежує на карті всі сміттєзбиральні машини і включає наповнений контейнер в маршрут найближчої вантажівки. І це теж вже не фантастика: саме так працює сміттєзбиральна система в Дубліні і Барселоні.

«Розумний» будинок призначений для максимально комфортного життя людей за допомогою використання сучасних високотехнологічних засобів. Принцип роботи такої системи полягає в автоматизації всього, з чого складається житлова споруда: освітлення, кондиціонування, системи безпеки, електроенергії, опалення, водопостачання та водовідведення і т.д.[10] До основних підсистем «розумного» будинку відносяться: клімат-контроль, освітлення, мультимедіа (аудіо і відео), охоронні системи, зв'язок і інші.

У стандартному проекті «розумного» будинку можна виділити три основні підмережі: мережу мультимедійних пристроїв, мережу електроосвітлювального обладнання і сенсорну мережу. В останньому випадку це сенсори руху, світла, температури, тиску, вологості, вібрації і т.п. Таким чином, «розумний» будинок складається з програмного і апаратного забезпечення, сенсорів і провідної / безпровідної мережі.

Для автоматизації будинку смарт-вузли можуть бути інтегровані безпосередньо в побутові пристрої, наприклад в пилососи, мікрохвильові печі, холодильники тощо. Вони можуть взаємодіяти один з одним і з зовнішньою мережею через інтернет. Це дозволить кінцевим користувачам легко управляти пристроями будинку як локально, так і віддалено.

У загальному випадку реалізація такого рішення надає його власнику такі переваги:

- зниження споживання ресурсів (газ, вода, електроенергія);
- високий рівень комфорту;
- забезпечення необхідної взаємодії всіх автоматизованих систем об'єкта нерухомості, завдання різних режимів роботи;
- зниження ймовірності виникнення аварійних ситуацій;
- підвищення оперативності, простоти і зручності управління.

Ідея використовувати в Інтернеті речей таку просту, що отримала повсюдне поширення технологію, як стільниковий зв'язок, знаходить все більше застосування в усьому світі.

Розглянемо більш конкретно переваги впровадження «розумних» пристроїв в певних сферах.

Автотранспорт:

- навігація;
- безпека, попередження аварій;
- аварійне реагування;
- самодіагностика автомобіля.

Безпека:

- відеоспостереження, відстеження;
- дистанційна зброя;
- моніторинг навколишнього середовища.

Охорона здоров’я:

- мобільні лабораторії;
- віддалений моніторинг пацієнта, діагностика;
- телемедицина.

Розумне місто:

- інтелектуальний транспорт;
- «розумні» будинки;
- «розумне» водопостачання;
- «розумне» освітлення
- «розумні» парковки.

«Розумний» автобус призначений для збору і зберігання інформації про ситуацію в автобусі, а також контролю за дорожньо-транспортною ситуацією і передачею даних про адміністративно-правові порушення іншими учасниками дорожнього руху. Переваги впровадження такого рішення для міста: оперативне реагування на надзвичайні ситуації, зниження смертності, зниження завантаженості доріг, ф порушення правил дорожнього руху, збір аналітичної інформації.

«Розумна» парковка призначена для поліпшення дорожньо-транспортної ситуації в місті за рахунок надання оперативної інформації про наявність вільних місць для паркування автомобілів через додаток смартфона.

Переваги такого рішення: зниження кількості «пробок» (за статистикою 30% заторів викликано машинами, які шукають паркувальне місце), додатковий прибуток за рахунок можливостей відео аналітики (поліпшення показника фіксації порушень паркування, який зараз становить 3 з 10 випадків), додаткова вигода за рахунок впровадження плати за користування парковками.

«Розумне» освітлення призначено для зниження витрат міста на освітлення вулиць і використовуваної електроенергії, які в середньому становлять від 18% до 30% загального бюджету міста.

Переваги впровадження даного рішення: істотне зниження споживання електроенергії (~30%) при одночасному підвищенні загальної ефективності освітлення, зниження енергоспоживання на міське освітлення до 80% (Рівас-Васиамадрид, Іспанія), збільшення терміну служби ламп, можливість віддаленого управління і контролю, підвищення безпеки громадян.

І це далеко не повний перелік того, як «розумні» пристрої можуть покращити наше життя, зробити його більш комфортним і безпечним.

4. Проблеми впровадження IoT

Можливості Інтернету речей в області генерування, збору, передачі, аналізу та розподілу величезного обсягу даних в світовому масштабі дозволять людству, в підсумку, отримати нові знання, які необхідні не тільки для виживання, але і для нинішнього добробуту протягом багатьох століть.

Підтвердження цьому – включення Інтернету речей в перелік перспективних технологій в США і в число семи формуючихся національних стратегічних галузей промисловості в Китаї.

Проте широкому впровадженню Інтернету речей перешкоджає ряд проблем [11]. Розглянемо основні з них.

Проблема безпеки

Забезпечення безпеки, надійності, стійкості і стабільності додатків і послуги Інтернет, має критично важливе значення для довіри і використання Інтернету. Користувачі повинні мати високий ступінь впевненості в тому, що Інтернет і підключені до нього пристрої мають досить високий ступінь безпеки для виконання різних завдань по відношенню до допустимості ризику, пов'язаного з їх виконанням [12]. Інтернет речей нічим не відрізняється в цьому відношенні, і безпека IoT пов'язана, в основному, з довірою до середовища з боку користувачів. Якщо люди не вірять в захищеність підключених пристроїв і отриманої інформації від неприпустимого використання, це призводить до відмови від використання Інтернету. Відповідно цей фактор робить глобальний вплив на електронну комерцію, технічні інновації і практично всі інші аспекти діяльності онлайн. Забезпечення безпеки продуктів і послуг IoT має бути основним пріоритетом в даній галузі.

У міру постійного збільшення числа пристроїв, підключених до Інтернету, виникають нові потенційні вразливі місця. Недостатньо захищені пристрої можуть бути точками доступу для кібератак,

дозволяючи зловмисникам перепрограмувати пристрій або викликати його несправність.

Пристрій з недосконалою конструкцією може піддавати дані користувачів небезпеці розкрадання за рахунок недостатнього захисту потоків даних. Несправні або дефектні пристрої також можуть створювати вразливі точки.

Крім потенційних вразливих місць, суттєве збільшення кількості і типів пристроїв IoT також може сприяти збільшенню ймовірності кібератак. Кожний підключений пристрій, що не має достатнього захисту, надає потенційно негативний вплив на безпеку і стійкість Інтернету в глобальному масштабі, а не тільки локально. Наприклад, незахищений холодильник в США, заражений шкідливим програмним забезпечення, може відправляти тисячі шкідливих повідомлень електронної пошти одержувачам у всьому світі за допомогою домашнього підключення Wi-Fi.

Зростаючий рівень залежності від пристроїв IoT і інтернет-послуг, з якими вони взаємодіють, також відкриває зловмисникам можливості доступу до пристроїв. Припустимо, ми можемо відключити підключений до Інтернету телевізор, якщо він піддасться кібератаці, але ми не зможемо також просто вимкнути електролічильник або систему регулювання руху транспорту або імплантований кардіостимулятор. Саме тому безпека послуг IoT є питанням, яке потребує особливої уваги.

Розглянемо основні проблеми безпеки пристроїв IoT.

Багато систем IoT буде складатися з груп ідентичних або майже ідентичних пристроїв. Така однорідність підсилює потенційний вплив кожної уразливості, множачи його на кількість пристроїв, що мають ті ж характеристики.

Розгортання багатьох пристроїв, підключених до Інтернету речей, може здійснюватися в умовах, що ускладнюють або роблять неможливою їх модернізацію чи зміну конфігурації.

Більшість пристроїв IoT працюють таким чином, що користувач не має або майже не має уявлення про внутрішнє функціонування пристрою або створювані ним потоки даних. Це створює вразливість в області безпеки, коли користувач вважає, що пристрій IoT виконує певні функції, в той час як насправді він може виконувати небажані дії або збирати дані, які користувач не має наміру надавати.

Деякі пристрої IoT встановлюються в таких місцях, де важко або навіть неможливо забезпечити їх фізичну безпеку. Зловмисники можуть отримати прямий фізичний доступ до них. У зв'язку з цим, для забезпечення безпеки необхідні функції захисту від злому і інших інновацій.

Проблема конфіденційності

Дотримання права на недоторканність приватного життя і переваг конфіденційності є невід'ємною частиною вирішення проблеми довіри до Інтернету. Інтернет речей часто представляється масштабною мережею сенсорних пристроїв, які збирають дані про оточення і нерідко – про людей. Звичайно, ці дані можуть бути корисними для власників пристроїв, але дуже часто вони представляють інтерес для виробників і постачальників пристроїв. Збір і використання IoT-даних перетворюється на справжню проблему конфіденційності, коли уявлення людей, що знаходяться під наглядом IoT-пристроїв, про масштаб і використання даних, відрізняються від міркувань збирача даних.

При об'єднанні або зіставленні кількох потоків даних можна отримати більш точний цифровий портрет людини, ніж при використанні одного потоку IoT-даних. наприклад, підключена до Інтернету зубна щітка може записувати і передавати нешкідливі дані про те, як її власник чистить зуби. Але якщо холодильник передає дані про те, що продукти, що споживаються, а фітнес-трекер передає дані про фізичну активність, то комбінація цих потоків дозволяє отримати більш детальний і точний опис загального стану здоров'я цієї людини.

Цей тип збору даних отримує все більш широке поширення в області побутових пристроїв, таких як «розумні телевізори» і ігрові приставки. Такі пристрої оснащені функцією розпізнавання голосу і зображення, тому можуть безперервно слухати або переглядати те, що відбувається в приміщенні і активно передавати ці дані в хмарний сервіс для подальшої обробки. Людина може перебувати в оточенні подібних пристроїв, не підозрюючи про те, що її розмови або дії відстежуються, а дані записуються. Такого роду функції можуть не тільки приносити користь обізнаним користувачам, але і створювати проблеми конфіденційності тим, хто не підозрює про присутність цих пристроїв і не може контролювати використання зібраних даних.

Незалежно від того, чи відомо це користувачеві і чи згоден він з тим, що його дані збираються і аналізуються, подібні ситуації лише підкреслюють цінність персоналізованих потоків даних для компаній і організацій, які прагнуть збирати і записувати IoT-дані. Потреба в цих даних призводить до появи юридичних і нормативних проблем, пов'язаних з законами про захист і конфіденційність даних.

Якщо користувач втратить довіру до Інтернету через недотримання його прав на приватне життя в Інтернеті речей, загальна цінність Інтернету може зменшитися.

Проблема інтероперабельності

У традиційному Інтернеті інтероперабельність пристроїв являє собою ключову цінність. Найважливіша вимога до Інтернет-підключення

полягає в тому, щоб «пов'язані» системи могли «говорити однією мовою» протоколів і кодів.

Інтероперабельність настільки важлива, що перші майстерні та семінари для постачальників Інтернет-обладнання так і називалися: «Interops» (від англійського interoperability – «інтероперабельність»). Крім того, вона є ключовим моментом, на який звернено увагу всієї спільноти розробки Інтернет-стандартів, сконцентрованого навколо IETF.

В умовах повної інтероперабельності будь-який IoT пристрій зможе встановлювати зв'язок з будь-яким іншим пристроєм або системою і здійснювати обмін інформацією. Проте на практиці це більш складне явище. Взаємодія між IoT-пристроями і системами відбувається на різних рівнях і в різних шарах в рамках стека комунікаційних протоколів.

Стандартизація та прийняття протоколів, що визначають принципи зв'язку (в тому числі реальну потребу в наявності стандартів), є основною темою дискусій, що стосуються Інтернету речей.

Крім технічних аспектів, інтероперабельність значно впливає на потенційний економічний вплив IoT. Реалізація існуючих стандартів (а при необхідності – створення нових відкритих стандартів) дозволяє зменшити бар'єри для створення та впровадження нових бізнес-моделей, а також створює умови для масштабного зростання економіки.

Згідно зі звітом міжнародної консалтингової компанії McKinsey Global Institute за 2015 рік, «в середньому 40 відсотків валового продукту, який може бути створений індустрією Інтернету речей, реалізується лише завдяки інтероперабельності».

У міру розробки IoT-пристроїв в цій сфері неминуче виникають технічні, тимчасові і вартісні обмеження, що впливають на сумісність і дизайн пристроїв. Деякі пристрої мають технічні обмеження (наприклад,обмежені внутрішні ресурси обробки даних, пам'ять або витрату енергії). Крім того, виробники прагнуть знизити собівартість продукції, максимально знижуючи собівартість компонентів і розробки продукту.

У короткостроковій перспективі розробка функціоналу для забезпечення інтероперабельності пристроїв і перевірка на відповідність стандартам може виявитися менш вигідним рішенням. Іноді набагато дешевше використовувати власні протоколи і системи, щоб успішно вийти на ринок. Але з іншого боку, виробникам слід також оцінювати це з точки зору довгострокової перспективи збільшення життєвого циклу продукту за рахунок інтероперабельності.

Проблема нормативно-правового і юридичного характеру

Застосування IoT-пристроїв привело до появи цілого ряду проблем і питань нормативно-правового та юридичного характеру.

Наприклад, відправлення даних, зібраних IoT-пристроями, за межі країни може бути недопустиме. Також IoT-пристрої можуть збирати дані про фізичних осіб з однієї юрисдикції і передавати ці дані для зберігання і обробки в іншу юрисдикцію, при цьому часто спостерігається нестача або відсутність технічних бар'єрів. Це явище може швидко перерости в юридичну проблему (наприклад, якщо зібрані дані були визнані персональними або конфіденційними і підлягають захисту відповідно до законодавства декількох юрисдикцій). Складність полягає в тому, що законодавство про захист даних в тій юрисдикції, де знаходяться пристрій і суб'єкт персональних даних, може не відповідати або суперечити законам країни, де дані зберігаються і обробляються.

Подібні ситуації є проблемою транснаціональних інформаційних потоків і ставлять питання про законодавчу базу, яка повинна застосовуватися.

IoT-пристрої можуть сприяти забезпеченню правопорядку і громадської безпеки, проте необхідно ретельно обміркувати наслідки з точки зору закону і суспільства. Наприклад, у торгових точках встановлюють камери спостереження для збору відеоматеріалу і спостереження за поведінкою покупців, що корисно для збору доказів і запобігання злочинів.

Корпорація On-Star, яка є дочірньою компанією General Motors, може надавати органам правопорядку інформацію, що отримується з сенсорів, вбудованих в автомобілях, допомагаючи поліції знаходити викрадені автомобілі; крім того, вона може дистанційно блокувати двигуни транспортних засобів.

Поліцейське управління округу Нассау (штат Нью-Йорк) використовує мережу звукових датчиків ShotSpotter, щоб з точністю визначати джерело стрільби в тих районах, де вони установлені. Це приклади тих переваг, які технологія Інтернету речей пропонує органам правопорядку для боротьби зі злочинністю і забезпечення громадської безпеки.

З іншого боку, подібне використання IoT-технологій викликає занепокоєння у деяких захисників цивільних прав. Приводом служить можливість всепроникного контролю даних, недосконаліть законодавства про збереження і знищення даних, методи вторинного використання даних державними службовцями, а також можливість отримання доступу зловмисниками до цих даних [13].

Отже, оскільки функціонал IoT-пристроїв складніший від функціоналу простих автономних продуктів, то виникає досить широке коло проблем юридичного, нормативного та правового характеру.

Висновки

Інтернет речей – це не просто безліч різних пристроїв і сенсорів, об'єднаних між собою каналами зв'язку, і підключених до мережі Інтернет, це більш тісна інтеграція реального та віртуального світів, в якому «спілкування» відбувається між людьми і пристроями.

Для реалізації такого «спілкування» розглянуто чотири моделі комунікації між пристроями IoT: підключення від пристрою до пристрою, від пристрою до хмари, від пристрою до шлюзу і спільне використання даних на сервері. Крім технічних аспектів, застосування цих моделей багато в чому визначається відмінностями між патентованими і відкритими IoT пристроями в мережі. Ефективні моделі зв'язку є важливим фактором для підвищення цінності послуг для кінцевих користувачів за рахунок можливості застосування нових способів використання інформації.

Для споживачів нові продукти IoT, такі як побутова техніка з підключенням до Інтернету, компоненти домашньої автоматики і пристрої для регулювання електроенергії наближають нас до концепції «розумного» будинку, забезпечуючи більш високий рівень безпеки та енергоефективності. Інші особисті пристрої IoT, такі як пристрої для фітнесу і контролю за станом здоров'я, а також медичні пристрої з підключенням до мережі, змінюють методи надання медичних послуг. Ця технологія здатна забезпечити більш високий рівень незалежності та якості життя за розумною ціною. Такі системи IoT як транспортні засоби, підключені до єдиної мережі, інтелектуальні системи управління дорожнім рухом та вбудовані сенсори на дорогах і мостах наближають нас до ідеї «розумних міст» для зниження числа пробок, зменшення енергоспоживання тощо.

Проте, в той же час, перед Інтернетом речей стоїть ряд проблем, які можуть перешкодити скористатися його потенційними перевагами. Постійні повідомлення про злом підключених до Інтернету пристроїв, проблеми щодо особистої конфіденційності та нормативно-правового характеру вже привернули увагу громадськості. На даний момент технічні питання продовжують залишатися невирішеними, а також виникають нові складності в області політики, законодавства та подальшого розвитку IoT.

Розглянуто ключові області проблем IoT для визначення найбільш важливих питань, пов'язаних з цією технологією. Ці області включають: безпеку; конфіденційність; інтероперабельність і нормативно-правові вимоги.

Незважаючи на те, що проблеми безпеки в області інформаційних технологій не є новими, багато прикладів впровадження IoT ставлять перед нами нові унікальні проблем в області безпеки. Вирішення цих проблем і забезпечення безпеки в області продуктів і послуг IoT має бути основним пріоритетом.

Користувачі повинні бути впевнені в тому, що пристрої IoT і пов'язані з ними послуги в області даних не мають вразливих місць, особливо в міру поширення цієї технології та її інтеграції у повсякденне життя.

Потоки даних і специфіка користувачів, керованих пристроями IoT, можуть забезпечити неймовірну цінність для користувачів IoT, але проблеми конфіденційності і потенційних зловживань можуть перешкоджати повному впровадженню Інтернету речей. Це означає, що конфіденційність і повага до права на конфіденційність мають важливе значення для завоювання довіри користувачів IoT.

Незважаючи на те, що повна інтероперабельність продуктів і послуг не завжди доцільна або необхідна, користувачі можуть відмовлятися від покупки продуктів і послуг IoT при відсутності гнучкої інтеграції, високої складності володіння і можливої залежності від постачальника.

Використання загальних, відкритих і широкодоступних стандартів в якості технічних складових пристроїв і послуг IoT (таких, як IP) забезпечить широкий ряд переваг для користувачів, інновації та економічні можливості.

Використання пристроїв IoT піднімає ряд нормативно-правових питань, а також розширює коло вже існуючих проблем щодо Інтернету.

Щоб скористатися потенційними перевагами для людей, суспільства і економіки, необхідно вирішити питання і проблеми, пов'язані з IoT. Для пошуку можливості максимального використання переваг Інтернету речей при максимальному зниженні ризиків недостатньо обговорення різних точок зору, що протиставляють можливості IoT його потенційним небезпекам. Замість цього потрібна активна участь на основі наявних даних і співпраця різних сторін для пошуку найбільш ефективних шляхів вирішення проблем та впровадження IoT для покращення якості життя.

З розвитком Інтернету речей все більше пристроїв будуть підключатися до глобальної мережі, тим самим відкриваючи все нові і більш широкі перспективи та сприяючи підвищенню якості життя населення. Передбачається, що в майбутньому «речі» стануть активними учасниками бізнесу, інформаційних і соціальних процесів, де вони зможуть взаємодіяти і спілкуватися між собою, обмінюючись інформацією про навколишнє середовище, реагуючи і впливаючи на процеси, що відбуваються в навколишньому світі, без втручання людей.

Література

1. Рой Уонт, Бил Шилит, Скотт Дженсон. Механизмы Интернета вещей // Открытые системы. СУБД., 2015. – No 1. – С. 38 – 42.

2. Rose K., Eldridge S., Chapin L. The Internet of Things: An Overview. Understanding the Issues and Challenges of a More Connected World / The Internet Society (ISOC), October, 2015. – 50 P.

3. Elfrink W. The Internet of Things: Capturing the Accelerated Opportunity / Cisco Blog, October 15, 2014.

4. Koshizuka N, Sakamura K. Ubiquitous ID: standards for ubiquitous computing and the internet of things / IEEE Pervasive Comput 9(4), 2010, pp.98 – 101.

5. Barbry E.. The Internet of Things, Legal Aspects: What Will Change (Everything) /Communications & Strategies, No. 87. – Quarter, 2012 – Pp. 83 –100.

6. Atzori L., Iera A The internet of things. / Computer Networks, vol. 54, no. 15, 2010. – pp. 2787–2805.

7. Дроздов С. Eurotech, «интернет вещей» и «облако устройств» / С. Дроздов, С. Золотарев // Control Engineering Россия, 2012. – № 8(78). – С. 18 – 24.

8. Голышко, А. Строим «интеллектуальный городок» // Мобильные телекоммуникации, 2013. – №10. – С. 46 – 51.

9. Schaffers H., Komninos N., Pallot M. Smart cities and the future internet: Towards cooperation frameworks for open innovation, The future internet, 2011. – pp. 431 – 446.

10. Tabar A., Keshavarz A., Aghajan H. Smart home care network using sensor fusion and distributed vision-based reasoning, in Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks. ACM, 2006. – pp. 145 – 154.

11. Коржов, В. Опасный Интернет вещей // Открытые системы. СУБД, 2013. – №4. – С. 29 - 30.

12. Баранов А., Брыжко В. Права человека и защита персональных данных– Харьков : Фолио, 2000. – 280 с.

13. Брижко В., Швець М. Системна інформатизація правоохоронної діяльності : європейські нормативно-правові акти та підходи до упорядкування інформаційних відносин у зв'язку з автоматизованою обробкою даних : посіб. / В.Брижко, М.Швець [та ін.]. – Кн. 2. – К. : ТОВ "ПанТот", 2006. – 509 с.

НОВІТНЯ ПЕРСПЕКТИВНА АВТОМАТИЗОВАНА СИСТЕМА “KaSPer”

Колачов С.П., Гуржій П.М., Масесов М.М., Гуржій І.А., Довикоза А.П.

Вступ

Медичне забезпечення є окремим видом забезпечення Збройних Сил України (далі – ЗС України) і являє собою систему заходів щодо збереження та зміцнення здоров’я особового складу, запобігання виникненню і розповсюдженню хвороб, надання медичної допомоги військовослужбовцям, лікування і відновлення їх працездатності та боєздатності після поранень, захворювань і травм.

Основою медичного забезпечення військ у воєнний час є система лікувально-евакуаційних заходів щодо організації надання медичної допомоги пораненим, ураженим, постраждалим та хворим (далі – поранені), їх евакуації, лікування, реабілітації та призначених для цього сил і засобів медичної служби.

Лікувально-евакуаційні заходи включають розшук, збір та винесення (вивезення) поранених з поля бою або вогнищ масових санітарних втрат (далі – поле бою), надання їм необхідних видів медичної допомоги, евакуацію, лікування та медичну реабілітацію.

Суть сучасної системи лікувально-евакуаційних заходів полягає в етапному лікуванні поранених з їх евакуацією за призначенням з використанням медичних підрозділів військових частин і з’єднань, мобільних і стаціонарних військово-медичних закладів та максимальним залученням існуючої мережі цивільних закладів охорони здоров’я (рис. 1).

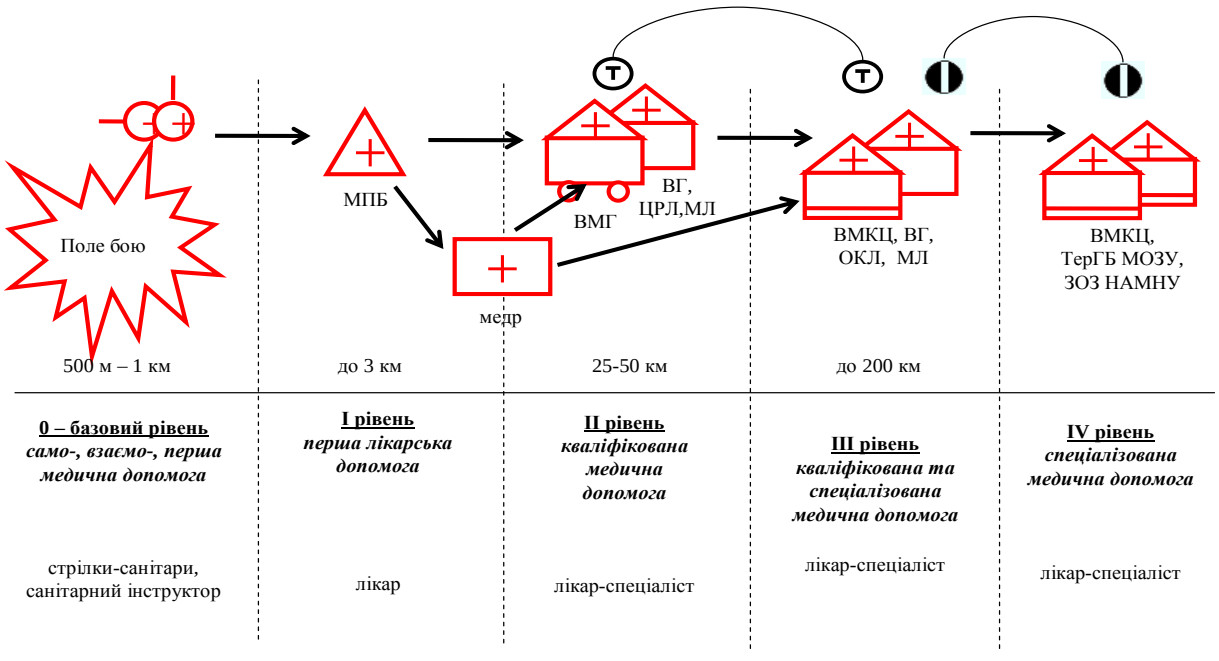


Рис. 1. Схема лікувально-евакуаційних заходів

Важливим елементом системи лікувально-евакуаційних заходів, нерозривно пов'язаним з наданням медичної допомоги та лікуванням поранених, є медична евакуація. Своєчасно та якісно проведена медична евакуація дозволяє компенсувати нестачу сил і засобів медичної служби в районі бойових дій.

Медична евакуація – це сукупність організаційних, медичних та технічних заходів щодо виносу (вивозу) поранених з поля бою та транспортування їх з одночасним медичним супроводом на етапи медичної евакуації (рівні медичного забезпечення) з метою своєчасного і повного надання необхідної медичної допомоги та лікування.

Транспортування поранених без медичного супроводу називається евакуацією поранених.

В умовах сучасної війни медична евакуація є найважливішим засобом забезпечення своєчасного надання всіх видів медичної допомоги пораненим. Вона передбачає швидку доставку їх саме на ті етапи медичної евакуації (далі – ЕМЕ), на яких найбільш раціонально за медичними показаннями та відповідно до бойової і медико-тактичної обстановки можна надати послідовно першу лікарську, кваліфіковану та спеціалізовану медичну допомогу, а також провести стаціонарне лікування до повного одужання.

Сучасний розвиток медичного забезпечення військ передбачає скорочення кількості ЕМЕ та наближення спеціалізованої медичної допомоги до поранених, що дає змогу здійснювати медичну евакуацію не послідовно через всі ЕМЕ, а минаючи деякі з них, направляти поранених безпосередньо до закладів охорони здоров'я (далі – ЗОЗ), спроможних до надання спеціалізованої медичної допомоги і лікування в повному обсязі (евакуація за призначенням).

Медична евакуація здійснюється відповідно до політики госпіталізації та евакуаційної політики, встановлених для визначеного порядку застосування військ і характеру бойових дій.

Політика госпіталізації та евакуаційна політика є ключем до визначення співвідношення спроможностей сил і засобів медичної служби на кожному ЕМЕ та засобів медичної евакуації в інтересах забезпечення оптимального лікування поранених і дозволяють збалансувати можливості щодо надання медичної допомоги з потребою у медичній евакуації.

Метою медичної евакуації є врятування життя і збереження здоров'я поранених шляхом доставки їх на відповідні ЕМЕ для своєчасного надання необхідного виду та обсягу медичної допомоги.

Основними завданнями медичної евакуації є:

– забезпечення своєчасного надання медичної допомоги пораненим;

– своєчасна доставка поранених на відповідні ЕМЕ;

– вивільнення ЕМЕ нижчого рівня для підготовки їх до переміщення або прийому нових поранених.

Медична евакуація поділяється на передову, тактичну і стратегічну.

Передова медична евакуація передбачає винос (вивіз) поранених з поля бою та їх транспортування до першого розгорнутого ЕМЕ, як правило – медичного пункту (далі – МП), медичної роти (далі – медр) або військового мобільного госпіталю (далі – ВМГ) чи стаціонарного ЗОЗ, розташованих в районі бойових дій. Вона здійснюється переважно броньованими транспортними засобами, санітарними автомобілями або гелікоптерами.

Під час здійснення цього виду медичної евакуації особлива увага має приділятися забезпеченню безпеки медичного та іншого персоналу, що залучається для її проведення.

Тактична медична евакуація – це медична евакуація поранених з передових до наступних ЕМЕ (стаціонарного військового або цивільного ЗОЗ), розташованих поза районом бойових дій у межах оперативної зони. Вона здійснюється після стабілізації стану поранених санітарними автомобілями (автобусами), залізничним, водним та авіаційним транспортом.

Стратегічна медична евакуація передбачає подальше транспортування поранених до тих закладів охорони здоров'я (військових або цивільних), де їм буде надано вичерпну медичну допомогу та проведено лікування до повного одужання. Вона здійснюється наземними і повітряними транспортними засобами.

Виконання бойових завдань медичними підрозділами та військовими частинами в зоні проведення антитерористичної операції на сході України виявило значні проблеми щодо оперативності обміну інформацією для організації екстреної евакуації поранених та медичного забезпечення Збройних Сил України в цілому. Використання передачі голосових повідомлень, відсутність впровадження автоматизованих систем не дають можливість ефективно здійснювати медичну допомогу в умовах бойових дій. Все це підтверджує **актуальність** на необхідність розробки та впровадження інфокомунікаційної системи екстреної евакуації поранених та медичного забезпечення, шифр "KaSPer".

Метою створення АС "KaSPer" є виключення передачі координат та запиту на евакуацію голосом, підвищення оперативності організації екстреної евакуації поранених та медичного забезпечення в цілому, скорочення часу виконання завдань медичними підрозділами (засобами), забезпечення точності визначення координат поранених військовослужбовців і засобів евакуації, реалізація підтримки прийняття рішення медичному персоналу.

Сутність АС "KaSPer" полягає в тому, що автоматизація, при її використанні, охоплює весь процес передачі інформації в ході здійснення евакуації та медичного забезпечення пораненого військовослужбовця і полягає у наступному.

Отримавши поранення військовослужбовець особисто (а у випадках неможливості – інший військовослужбовець) повідомляє санітара про отримання поранення шляхом застосування рятувального маяка (індивідуального рятувального засобу військовослужбовця). Інформація про отримання поранення та місцезнаходження військовослужбовця потрапляє на програмно-апаратний комплекс (далі – ПАК) санітара (стрільця-санітара) (рис. 2). Отримавши інформацію про пораненого, санітар за допомогою використання електронної карти здійснює пошук пораненого на полі бою та евакуює його до "гнізда евакуації" (рис. 3), де надає йому невідкладну медичну допомогу та надає запит на евакуацію до чергового медичного пункту батальйону.

Рис. 2. Отримання інформації санітаром про місцезнаходження пораненого

Рис. 3. Евакуація санітаром пораненького до "гнізда евакуації"

Запит на евакуацію відповідає запиту, що прийнятий та використовується у збройних силах країн-членів НАТО (за формою типу "9 ліній" стандарту НАТО STANAG 2546 — AJMedP-2) та містить наступну інформацію: координати гнізда евакуації, частота і позивний, терміновість евакуації, необхідне обладнання, кількість пацієнтів, тощо (табл. 1).

У разі неможливості передачі запиту на евакуацію в автоматизованому режимі, інформація передається голосом з використанням наявних засобів зв'язку.

В такому випадку доцільно передавати не букви англійського алфавіту (формалізовані дані), а їх позначення відповідно до міжнародного фонетичного алфавіту (наприклад, A – позначається словом "Alpha", B – "Bravo", C – "Charlie", D – "Delta", E – "Echo", тощо).

Черговий медичного пункту батальйону визначає необхідний розрахунок наряду сил та засобів для евакуації поранених до сортувального пункту бригади, причому програмне забезпечення, що встановлено на ПАК чергового медичного пункту батальйону, розраховує можливі варіанти такої евакуації та надає їх для прийняття рішення у порядку пріоритету (рис. 4).

Таблиця 1

№ з/п	Назва пункту “9 ліній”	Зміст інформації	Формалізоване позначення
1.	Місце знаходження точки забору пораненого (поранених)	Координати “гнізда”	Цифрами
2.	Частота, позивний	Частота (в МГц), налаштування радіостанції та позивний санітара	Цифрами та буквами (позивний)
3.	Кількість пацієнтів за терміновістю евакуації	Термінові(не хірургічні) 2 год.	A
		Термінові хірургічні	B
		Важливі - 4 год.	C
		Відкладені 24 год.	D
		Незначні	E
4	Необхідне додаткове обладнання	Не потрібне	A
		Підйомник	B
		Засоби для евакуації	C
		Апарат для вентиляції легень	D
5	Кількість пацієнтів	Лежачих	A
		Амбулаторних	B
6.	Безпека місця забору поранених	Відсутні ворожі війська	N
		Можливі ворожі війська	P
		Наявні ворожі війська	E
		Наявні ворожі війська (необхідний озброєний супровід) або при бойових діях	X
7	Спосіб позначення місця забору поранених	Кольором	A
		Сигнальні ракети	B
		Дим сигнал (вказати колір)	C
		Ніякого	D
		Інші	E
8	Громадянство та статус пацієнта	Український військовий	A
		Український цивільний	B
		Неукраїнський військовий	C
		Неукраїнський цивільний	D
		Ворожій військовополонений	E
		Дуже цінна мішень (необхідний озброєний супровід)	F
9	Наявність зараження місцевості (у мирний час – опис особливостей місцевості на місці забору поранених)	Радіаційне	N
		Біологічне	B
		Хімічне	C

Після вибору одного з можливих варіантів, черговий медичного пункту батальйону надає запит про готовність до евакуації на ПАК водія транспортного засобу (санітарного бронетранспортеру), а після підтвердження готовності до евакуації – віддає наказ на евакуацію з передачею інформації про координати “гнізда евакуації”, кількість, необхідне медичне обладнання.

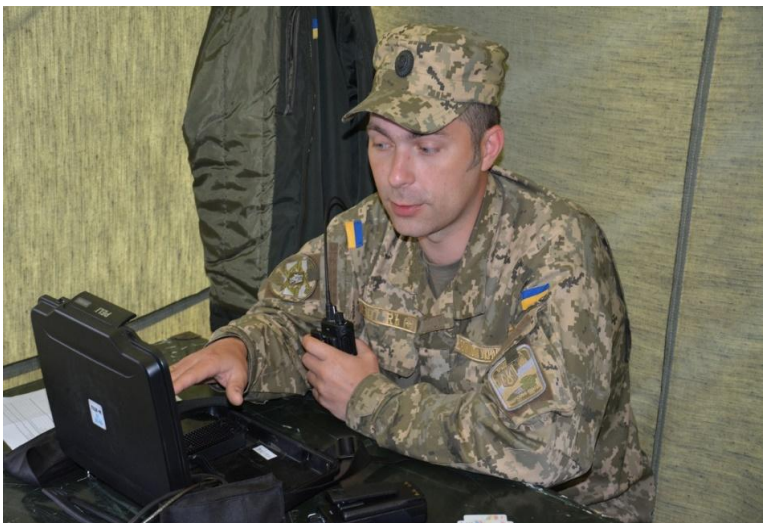


Рис. 4. Визначення необхідного розрахунку наряду сил та засобів для евакуації поранених до сортувального пункту бригади черговий медичного пункту батальйону

В процесі евакуації підтримується безперервний зв'язок з водієм: здійснюється обмін формалізованими та довільними даними, передаються доповіді про початок та закінчення евакуації, фіксуються дані про час отримання (передавання) сигналів (команд, наказів).

Після евакуації поранених військовослужбовців до пункту сортування здійснюється надання медичної допомоги та сповіщення оперативного чергового медичної роти батальйону про кількість поранених, характер отриманих ними поранень. Карта пораненого військовослужбовця має електронний вигляд та заповнюється за допомогою індивідуальної ідентифікаційної медичної картки військовослужбовця. У зазначеній картці зберігаються не тільки стандартні дані військовослужбовця (ПІБ, вік, вага, зріст, колір очей та інш.), перенесені хвороби, травми та поранення, а, що особливо важливо, наявність у організму пораненого протипоказань до деяких ліків. Передбачається, що зазначена картка (рис. 5) буде видаватися військовослужбовцю на початку служби та коректуватися на протязі проходження всієї служби.

Рис.5. Індивідуальна ідентифікаційна медична картка військовослужбовця

Доступ до персональних даних пораненого військовослужбовця

відбувається після ідентифікації та автентифікації лікаря (санітара) за допомогою персональної ідентифікаційної картки лікаря (санітара). Окрім того, одночасно здійснюється реєстрація медичного працівника у системі та відбувається фіксація всіх дій та змін які робить медпрацівник у медичній картці військовослужбовця. Картка пораненого військовослужбовця застосовується у аналогічних системах медичної евакуації поранених у збройних силах країн-членів НАТО та має назву TCCC (Tactical Combat Casualty Care – тактична допомога пораненим у бою) [2].

Отримавши ідентифікаційну картку пораненого військовослужбовця, черговий медичної роти бригади приймає рішення щодо вибору медичного закладу, в який повинен бути доставлений поранений, а також транспортного засобу, що для цього повинен бути задіяний.

Спостереження за переміщенням транспортних засобів та координація дій евакуаційних бригад відбувається за рахунок використання геоінформаційних систем.

Кожен черговий медичної роти бригади має у своєму розпорядженні актуальну базу даних про всі типи поранення, транспортні засоби та медичні заклади, які оновлюються у відповідності до визначеного регламенту, а також про всіх поранених у зоні відповідальності, військовослужбовців та їх місцеперебування.

Для підвищення оперативності, ідентифікаційна картка військовослужбовця передається до медичного закладу, куди евакуюється поранений, що надає змогу завчасно підготувати необхідне обладнання, ліки, операційні, тощо, а також викликати визначених фахівців.

При необхідності проведення наступних ЕМЕ, черговим медичної роти бригади формується запит на евакуацію поранених до військового мобільного госпіталю або цивільного закладу охорони здоров’я, що передається з використанням ПАК.

Про кількість і стан поранених (наявних та відправлених на евакуацію), а також потреби в евакуації здійснюється доповідь старшому начальнику медичної служби за формою і в терміни, визначені Табелем термінових донесень по медичній службі.

Алгоритм дій начальників (чергових) медичної служби вищих ланок управління – аналогічний роботі начальника (чергового) медичної служби військової частини і відрізняється об’ємом організації та здійснення адміністративних заходів.

Начальником (черговим) стаціонарного (мобільного) військового або цивільного закладу охорони здоров’я здійснюються заходи щодо оперативного управління силами і засобами евакуації, ведення локальних баз даних про можливості медичного забезпечення (кількість вільних ліжок, наявність ліків та лікарів, тощо), а також доповіді визначеного керівнику про стан медичного забезпечення у закладі.

Начальником ЦВМУ ЗСУ здійснюється загальне управління та координація дій силами і засобами медичного забезпечення Збройних Сил, а також взаємодія з іншими міністерствами з метою додаткового замовлення транспорту для потреб медичної евакуації. На рівні ПАК начальника ЦВМУ ЗСУ здійснюється розгортання та ведення централізованої бази даних, яка забезпечує функціонування всієї АС "KaSPer".

Порядок проходження інформації за напрямками між службовими особами, що приймають участь у процесі евакуації та медичного забезпечення поранених військовослужбовців, представлено на рис. 6.

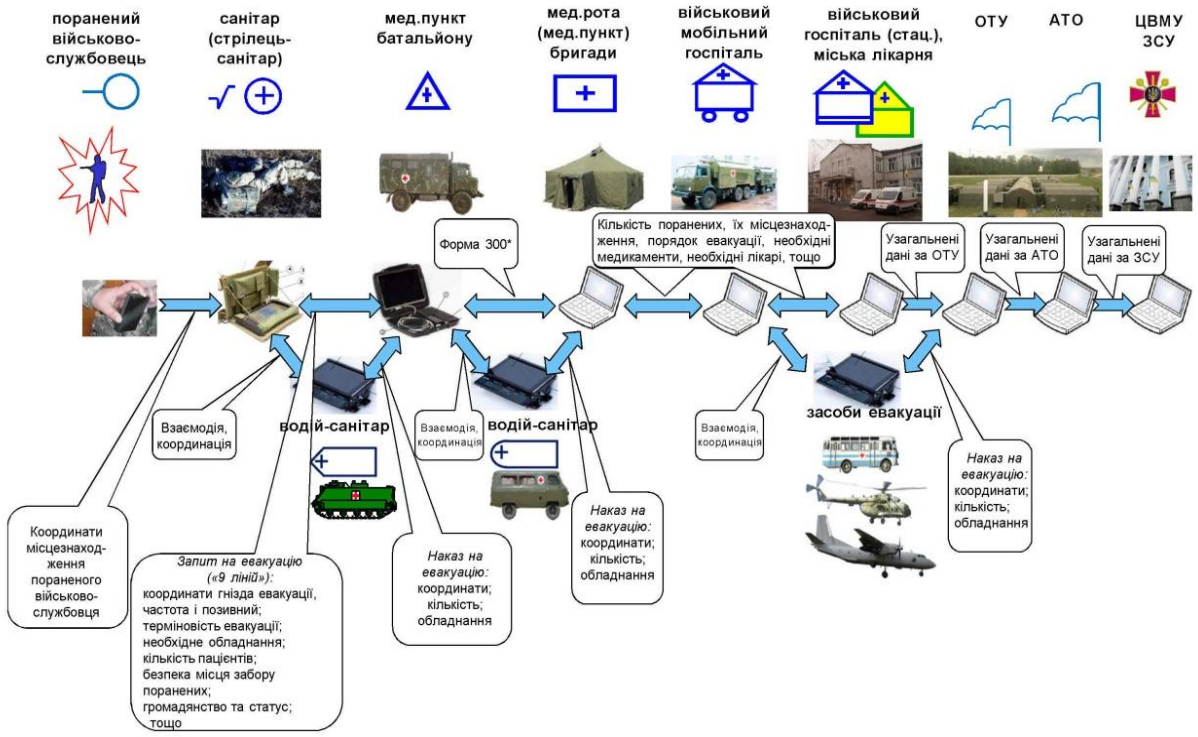


Рис. 6. Напрямки обміну інформаційними потоками при використанні АС "KaSPer"

Структурно АС "KaSPer" складається з наступних функціональних підсистем:

– підсистема оповіщення про поранених;
– підсистема виявлення поранених (опціонально);
– підсистема ідентифікації поранених;
– підсистема оперативного управління;
– геоінформаційна підсистема;
– підсистема підтримки прийняття рішення;
– інформаційно-довідкова підсистема;
– телекомунікаційна підсистема;
– підсистема захисту інформації та кібернетичної безпеки.

Зазначені підсистеми забезпечують інформаційний обмін між службовими особами на робочих місцях, обладнаних відповідними програмно-апаратними комплексами (далі – ПАК).

Підсистема оповіщення про поранених призначена для інформування службових осіб медичної служби підрозділу про необхідність здійснення евакуації поранених шляхом використання запиту на евакуацію відповідно до Військового стандарту «Спільна об'єднана доктрина з медичної евакуації», який розроблено на основі стандарту HATO STANAG 2546 - AJMedP-2 [1]..

Підсистема оповіщення про поранених дозволяє забезпечувати передачу інформації у формалізованому вигляді про: координати гнізда евакуації, частоту і позивний санітара, терміновість евакуації, необхідне обладнання, кількість пацієнтів, безпека місця забору поранених, спосіб позначення місця забору поранених, громадянство та статус поранених, опис особливостей місцевості на місці забору поранених.

Підсистема виявлення поранених (опціонально) дозволяє забезпечувати:

– використання індивідуальних рятувальних засобів військовослужбовців (рятувальних маяків)з вбудованими GPS та радіомодулями для передачі координат місця знаходження пораненого;

– пошук санітаром поранених по отриманим від рятувальних маяків координатам з допомогою відповідного програмного забезпечення, що встановлюється на ПАК санітара.

Підсистема ідентифікації пораненого призначена для встановлення (підтвердження) особи пораненого, основних медичних показників (групи крові, хронічних захворювань, непереносимості медичних препаратів, тощо) шляхом використання ідентифікаційної картки військовослужбовця.

Підсистема ідентифікації пораненого дозволяє забезпечувати:

– створення облікового запису та введення (модифікації) інформації про основні особисті дані та медичні показники військовослужбовця в ідентифікаційну картку військовослужбовця;

– зчитування введених даних з ідентифікаційної картки військовослужбовця та зручний, не двозначний перегляд цих даних на ПАК службових осіб медичної служби з допомогою програмного забезпечення.

Підсистема оперативного управління дозволяє забезпечувати:

– формування та передачу підпорядкованим силам команд, розпоряджень і прийом від них підтверджень, донесень, доповідей;

- взаємний обмін текстовою неформалізованою і формалізованою інформацією між АРМ посадових осіб органів управління медичної служби.

Геоінформаційна підсистема призначена для відображення рухомих та стаціонарних об'єктів медичного забезпечення, які беруть участь у евакуації поранених (санітарних транспортних засобів, медичних пунктів підрозділів, військових госпіталів, медичних закладів, тощо) та координації їх дій (рис. 7).

Рис. 7. Головний реанімобіль благодійного фонду АСАП "Хоттабич" обладнаний засобами відображення даних геоінформаційної підсистеми

Геоінформаційна підсистема дозволяє забезпечувати:

– надання користувачам доступу до обстановки на електронних картах місцевості відповідно до їх повноважень;

– створення єдиної картини медичного забезпечення з відображенням рухомих та стаціонарних об'єктів в зоні проведення бойових дій.

Підсистема підтримки прийняття рішення призначена для скорочення часу та підвищення обґрунтованості прийняття рішень на застосування відповідних засобів, що використовуються в системі екстреної евакуації поранених на полі бою.

Підсистема підтримки прийняття рішення дозволяє забезпечувати:

– створення та ведення довідників сил та засобів, що використовуються в АС "KaSPer", поранень та засобів, що необхідні для даних поранень, наявних медичних закладах, тощо;

– вибору на основі даних з довідників оптимальних варіантів застосування засобів екстреної евакуації та медичних закладів, куди доставити поранених.

Інформаційно-довідкова підсистема призначена для інформаційної підтримки діяльності службових осіб органів управління медичним забезпеченням за рахунок забезпечення їх довідковою інформацією з питань, пов'язаних з виконанням їх функціональних обов'язків.

Інформаційно-довідкова підсистема дозволяє забезпечувати:

– введення, пошук, збереження та оновлення інформації щодо медичного забезпечення;

– упорядкування і класифікацію інформації щодо медичного забезпечення;

– підтримку єдиної структурованої мови запитів до баз даних;

– спільний доступ користувачів до одного масиву інформації;

– доступ до даних прикладних програм підсистеми підтримки прийняття рішень;

– можливість обміну даними між прикладними програмами в об’єктно-орієнтованому вигляді;

– створення та адміністрування деревовидної структури каталогів.

Телекомунікаційна підсистема призначена для створення єдиного телекомунікаційного простору для роботи системи, взаємодії підсистем та технічних засобів шляхом використання наявних сил та засобів зв’язку.

Телекомунікаційна підсистема дозволяє забезпечувати:

– захищений голосовий зв’язок між абонентами шляхом використання засобів радіо, супутникового та проводового зв’язку;

– захищену передачу даних між ПАК службових осіб медичної служби з допомогою засобів радіо, супутникового та проводового зв’язку.

Підсистема захисту інформації є сукупністю необхідних, взаємоузгоджених організаційних та інженерно-технічних заходів, засобів і методів технічного та криптографічного захисту інформації, достатніх для запобігання навмисним чи ненавмисним спробам блокування інформації, порушенню її цілісності, конфіденційності або нав’язуванню хибної інформації.

Для забезпечення роботи підсистем між собою АС “KaSPer” включає наступні технічні засоби:

– індивідуальний рятувальний засіб військовослужбовця (рятувальний маяк);

– ПАК санітара (стрільця-санітара);

– ПАК водія санітарного бронетранспортеру;

– ПАК водія санітарного автомобілю;

– ПАК чергового медичного пункту батальйону;

– ПАК чергового лікаря сортувального пункту бригади;
– ПАК чергового медичної роти бригади;

– ПАК чергового військового медичного закладу;
– ПАК чергового цивільного медичного закладу;

– ПАК начальника медичної служби АТО (ОТУ);
– ПАК чергового Центрального військово-медичного управління ЗС України;
– ідентифікаційна картка військовослужбовця;
– ідентифікаційна картка медичного працівника (санітара, лікаря);
– засоби зв'язку.

Висновки

1. Використання АС "KaSPer" забезпечить:

– інформаційну підтримку та координацію дій засобів евакуації поранених, бригад екстреної (швидкої) медичної допомоги і медичних закладів;
– прийняття виклику екстреної медичної допомоги, його оброблення та оперативне реагування на такий виклик;
– формування інформації про місце події, характер та особливості невідкладного стану пораненого військовослужбовця, вид допомоги, тощо;
– можливість оперативної передачі відповідній бригаді інформації про виклик, характер та особливості невідкладного стану військовослужбовця;
– визначення медичного закладу, до якого бригада екстреної медичної допомоги здійснюватиме перевезення пораненого військовослужбовця у невідкладному стані, передачу цієї інформації бригаді;
– інформаційний супровід надання медичної допомоги та прийняття інформації про результат її надання на місці події, під час перевезення та прибуття пораненого військовослужбовця до медичного закладу;
– використання, збирання, оброблення, накопичення, зберігання, передачу, поширення, знищення, надання доступу до інформації про виклики екстреної медичної допомоги.

2. При розробці проекту системи використано передовий досвід та досягнення науки і техніки у галузі інформаційних технологій.

3. Прототипом АС “KaSPer” є “Тактична система сортування медичних даних надання допомоги пораненим, або тим, хто постраждав” розробки фірми Harris, але на відміну від неї має у своєму складі систему підтримки прийняття рішення для визначення наряду сил та засобів для евакуації поранених, автоматичну систему оповіщення, ведення баз даних поранених, медикаментів, обладнання та інше.

4. АС “KaSPer” з позитивними результатами пройшла тестові випробування у зоні проведення АТО, під час командно-штабних навчань на полігоні “Широкий лан”, при підготовці фахівців з тактичної медицини на базі Навчального центру “Десна” [3] та полігону ВДВ (м. Житомир). АС “KaSPer” було продемонстровано керівництву Міністерства оборони України, Генерального штабу Збройних Сил України та Президенту України – Верховному Головнокомандувачу Збройних Сил України. Окремі підсистеми АС “KaSPer” використовуються в зоні проведення АТО протягом останніх двох років.

Література:

1. STANAG 2546-AJMedP-2. (Electronic resource) / J ALLIED JOINT DOCTRINE FOR MEDICAL EVACUATION AJMedP-2 // 30 May 2011. – P. 54. – Mode of access: http://www.coemed.org/component/jifile/download/M2M0MWNhN2RmZmRhOGFhMDAxYWE0YmNlMmRlYjBkYjQ=/2546-ajmedp-2-pdf

2. The Tactical Combat Casualty Care Casualty Card (Electronic resource) // 30 April 2013. – P. 9. – Mode of access: http://www.chinookmed.com/TCCC-Change-Prop-1301-TCCC-Card.pdf

3. Новітні методики з тактичної медицини опанували в «Десні» інструктори Збройних Сил України (Electronic resource) // 19 травня 2016, 19:50. – Mode of access: https://goo.gl/UI4PB8

ТЕХНОЛОГІЇ МОНІТОРИНГУ ПОЖЕЖНОЇ БЕЗПЕКИ З БАГАТОРІВНЕВИМ ПЕРЕТВОРЕННЯМ ІНФОРМАЦІЇ

Куліца О.С., Дендаренко В.Ю., Слободянюк А.В., Третьяк В.Ф.

Вступ

Ефективність планування та реалізації заходів із профілактики пожеж істотно залежить від достовірності та оперативності моніторингової інформації, на основі якої приймаються управлінські рішення. Разом з тим для прийняття рішень із управління станом пожежної безпеки на адміністративній території в цілому вимагається інтегральна оцінка впливовості факторів та прогнозування динаміки втрат під впливом планованих керуючих впливів.

Застосування інформаційних технологій багаторівневого моніторингу при плануванні заходів із профілактики пожеж дозволяє забезпечувати інформацією процеси прийняття рішень про стан пожежної безпеки адміністративної території в цілому та якісно впливати на кількість матеріальних втрат в наслідок виникнення надзвичайних ситуацій на цих територіях.

Мета роботи: Підвищення ефективності профілактичних заходів запропоновано досягати шляхом використання в процесі планування інформації про впливовість та взаємну залежність факторів, що визначають стан пожежної безпеки об’єктів, та причин, які спричиняють пожежі впродовж останнього періоду часу.

Об’єктом дослідження є процеси перетворення інформації в технологіях моніторингу складних об’єктів.

Предмет дослідження – моделі, методи і засоби формування масивів вхідних даних в процесі моніторингу стану пожежної безпеки окремої адміністративної території.

Методи досліджень. При вирішенні поставлених завдань використовувались методи системного аналізу для дослідження процесів моніторингу з метою побудови адекватних моделей об’єктів моніторингу, методи індуктивного моделювання та методи математичної статистики для первинної обробки інформації.

При застосуванні сучасних технологій побудови інформаційних моніторингових систем виникає проблема врахування особливостей діючої системи нагляду у сфері пожежної та техногенної безпеки. Необхідно розробити спеціальні методи адаптації структури цих систем до зміни інформативності масиву вхідних даних. Одним із ефективних способів підвищити якість функціонування існуючої системи моніторингу пожежної безпеки без значних ресурсних затрат є автоматизація процесів обробки інформації. Основною ідеєю цієї роботи є дослідження процесу створення на базі існуючої системи державного нагляду у сфері пожежної

та техногенної безпеки інформаційної технології моніторингу у вигляді автоматизованої системи багаторівневого перетворення інформації.

Як відомо, під структурою системи розуміють перелік підсистем та зв'язків між ними. Оскільки поєднання підсистем дозволяє отримати систему, то це означає, що зв'язки між підсистемами повинні бути ефективними. Спосіб поєднання підсистем залежить від призначення самої інформаційної системи. Одначе відомо, що одну і ту ж систему можливо подати за допомогою кількох типів структур.

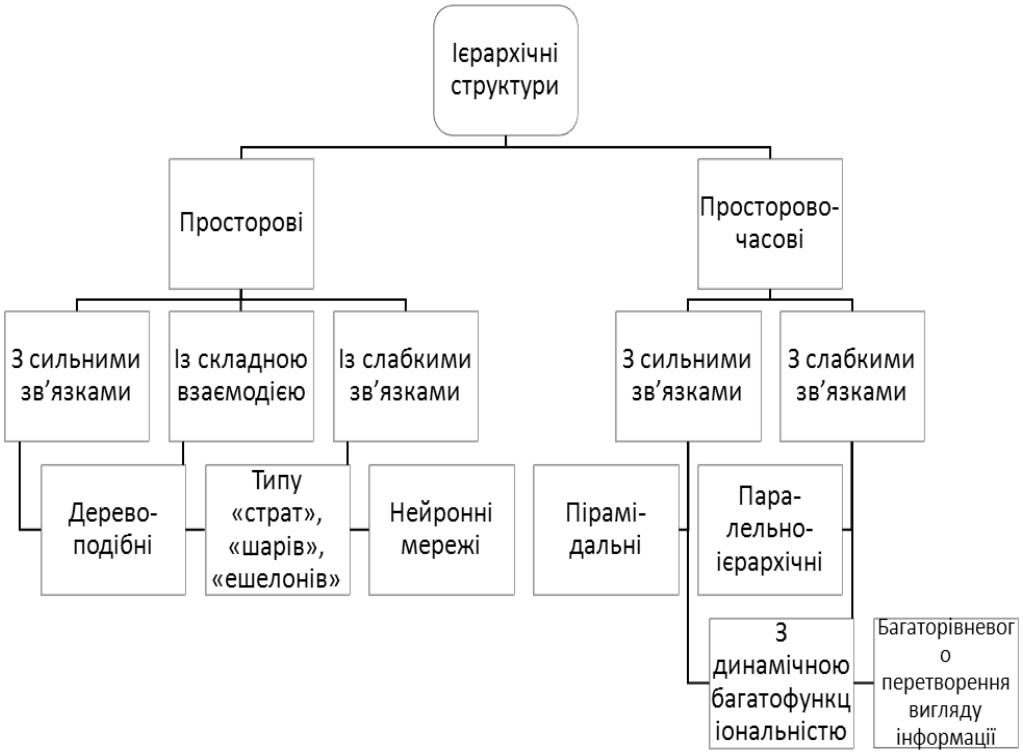


Рис. 1. Класифікація ієрархічних структур

Інформаційні системи, що мають кілька рівнів перетворення інформації подаються за допомогою *ієрархічних структур*. Вони відображають відношення підлеглості елементів, пріоритетності перетворення інформації. Спільною ознакою таких структур є наявність рівнів ієрархії [4, 5].

Оскільки система багаторівневого перетворення інформації призначена для автоматизації процедури БЕО, в якості експертів в таких системах виступають процеси синтезу моделей, які відображають окремі властивості об'єкта дослідження, та характеристики результатів їх функціонування – вихідні сигнали.

При побудові багатопараметричної моделі ставиться задача ідентифікації функціональної залежності:

$$y_i = f(x_1, x_2, \ldots, x_n), \quad (1)$$

де y_i – модельований показник (вихідний сигнал);

$x_1, x_2, \ldots, x_n$ – показники стану об'єкта або змінні моделювання (масив вхідних даних).

Традиційним методом побудови багатопараметричних моделей є множинна регресія. Функціональна залежність 1. подається у вигляді:

$$y_i = a_o + a_1x_1 + a_2x_2 + \ldots + a_nx_n. \quad (2)$$

Значення параметрів множини $A=\{a_1, a_2, \ldots, a_n\}$ визначається за методом найменших квадратів. Важливою умовою застосування цього методу є постійна дисперсія експериментально отриманих значень модельованого показника y_i і нормальність закону їх розподілу.

Оскільки на практиці таке трапляється надзвичайно рідко, то масив вхідних даних піддається попередній обробці. Закон розподілу y_i наближається до нормального в результаті усереднення значень масиву вхідних даних. Відповідно центральної граничної теореми якщо $y_1, y_2, \ldots, y_N$ – незалежні однаково розподілені випадкові величини, що мають математичне сподівання a та дисперсію σ^2, то при $N \to \infty$ закон розподілу суми випадкових чисел $\sum_{i=1}^{N} y_i$ необмежено наближається до нормального. На практиці розподіл суми $\sum_{i=1}^{N} y_i$ стає близьким до нормального, вже при N = 8 ÷ 12. При усередненні спостережень масиву вхідних втрачається частина інформації.

Уникнути цього можливо шляхом застосування індуктивного моделювання методом групового врахування аргументів (МГУА) [6, 7]. Цей метод передбачає багаторядну процедуру формування моделі оптимальної складності шляхом поетапної масової селекції множини опорних моделей за зовнішнім критерієм їх якості. Масив вхідних даних ділиться, як мінімум, на дві послідовності. Послідовність спостережень А призначення для навчання моделей. Послідовність В використовується для випробування навчених моделей та розрахунку критерію їх якості.

Початкова множина опорних моделей генерується у вигляді поліноміальних залежностей, кожна з яких поєднує попарно в своїй структурі атрибути масиву вхідних даних. Після цього кожна модель навчається шляхом визначення параметрів моделей (коефіцієнтів при атрибутах). Відповідно до значень критеріїв якості моделей відбувається їх ранжування та формування наступного ряду селекції на основі кращих із них. Формування моделі оптимальної складності завершується коли якість моделей наступного ряду гірше якості моделей ряду попереднього.

Визначення вигляду опорної моделі є однією із центральних процедур МГУА. Цей процес відбувається евристично на основі наступних підходів [1, 2]. Вважається, що повного відображення властивостей об'єкта можливо досягнути, формуючи структуру моделі на основі функціонального ряду Вольтерра:

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \ldots, \quad (3)$$

де x – елементи множини атрибутів; a – вектор параметрів.

Окремим випадком ряду Вольтерра, який реалізовує стратегію попарного поєднання атрибутів масиву вхідних даних, є поліном Колмогорова-Габора:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2 + a_4 x_1^2 + a_5 x_2^2 + a_6 x_1^2 x_2 + \\ + a_7 x_1 x_2^2 + a_8 x_1^2 x_2^2. \quad (4)$$

При використанні базових алгоритмів МГУА опорний вигляд моделі конструюють шляхом поєднання окремих елементів поліному (4).

Значимість закону розподілу показників у вхідному масиві даних нівелюється за рахунок використання індуктивними методами зовнішнього критерію якості в процесі селекції моделей. Вхідний масив даних ділиться на кілька вибірок, тільки одна з яких використовується для навчання моделей, а інші – для оцінки їх якості, селекції та формування структури індуктивної моделі оптимальної складності.

1. Стратегія досліджень

Наступні дослідження направлені на теоретичні дослідження процесів формування глобальної функції системи у вигляді технології багаторівневого перетворення інформації.

Шляхом декомпозиції глобальної функції системи отримується відповідна кількість локальних рівнів, задачі яких реалізуються шляхом синтезу локальних алгоритмів перетворення інформації (далі АПІ).

В якості локального АПІ використовуються моделі об’єктів моніторингу відповідного рівня. Досліджуються особливості синтезу моделей об’єктів моніторингу пожежної безпеки. Формалізується задача забезпечення відповідності між вимогами до моделей та можливостями засобів їх синтезу.

В процесі ієрархічного поєднання моделей об’єктів моніторингу пожежної безпеки за методом їх висхідного синтезу процесам формування масиву вхідних даних (МВД) приділяється особлива увага. Структура моделей об’єктів моніторингу наступного рівня перетворення інформації формуються автоматично синтезатором із показників МВД і якість моделей безпосередньо визначається інформативністю масиву цих показників.

Результати теоретичних досліджень дозволяють формалізувати інформаційну технологію багаторівневого моніторингу пожежної безпеки.

2. Модель процесу управління пожежною безпекою

Управління пожежною безпекою передбачає забезпечення в будь-який момент часу наперед визначеного її стану, показники якого нормовані відповідними державними актами. Процес управління реалізується шляхом моніторингу нормативних показників стану пожежної безпеки а також розробки комплексу керуючих впливів та стратегії їх застосування у вигляді плану заходів із профілактики надзвичайних ситуацій, зокрема пожеж

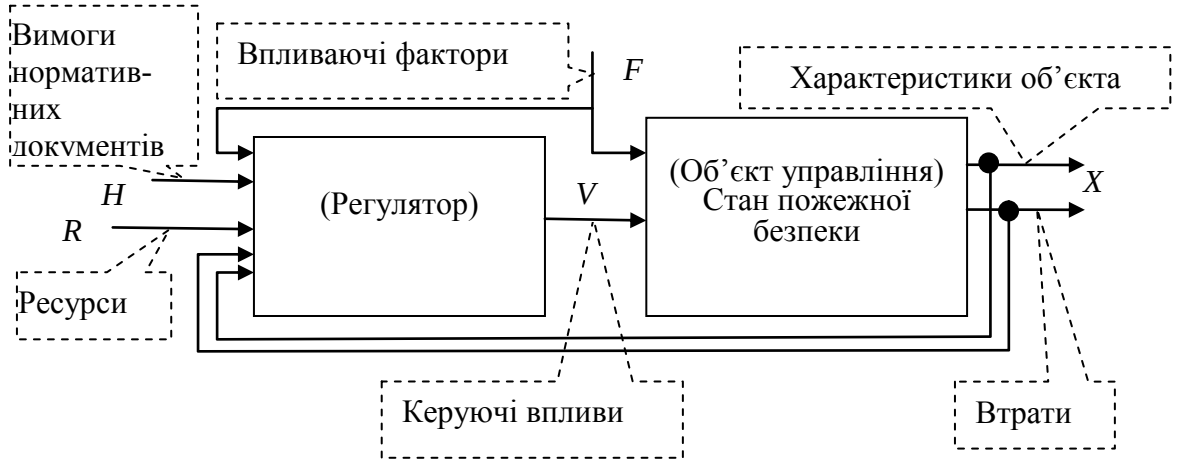


Рис. 2. Модель управління пожежною безпекою

Об'єктом управління є стан пожежної безпеки, показники якого містять чисельні характеристики нормативних показників стану об'єктів нагляду та характеристики втрат від пожеж та пов'язаних з ними подій.

Регулятор поєднує МІС, яка забезпечує інформацією процес формування керуючих впливів, та особу, що приймає рішення (далі ОПР), яка формує керуючі впливи у вигляді плану профілактичних заходів та забезпечує їх реалізацію.

Характеристики стану пожежної безпеки

$$X = \{x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n\}, \quad (5)$$

де $x_1, x_2, \ldots, x_k$ – спостережені характеристики нормативних показників стану об'єктів пожежної безпеки, що отримані внаслідок експертизи щодо пожежної безпеки і задаються нормативними документами та позначені в моделі множиною $H= \{h_1, h_2, \ldots, h_k\}$;

$x_{k+1}, \ldots, x_n$ – показники втрат внаслідок надзвичайних ситуацій, та чисельні характеристики факторів, що впливають на техногенну безпеку,

$$F = \{f_1, f_2, \ldots, f_r\}, \quad (6)$$

де $f_1, f_2, \ldots, f_r$ – чисельні характеристики відомих чинників, що зумовлюють можливість виникнення та (або) розвитку пожежі на об'єкті та впливають на показники множини X та подаються на вхід регулятора системи управління.

Регулятор перетворює вхідну інформацію у вигляді масиву вхідних даних до вигляду множини характеристик впливовості показників

$$W = \{w_1, w_2, \ldots, w_r, w_{r+1}, \ldots, w_n\}, \quad (7)$$

де w_1, w_2, …, w_n – показники впливовості зовнішніх факторів та нормативних показників стану об'єкта

та прогнозованих втрат від надзвичайних ситуацій. На основі цієї інформації формуються керуючі впливи

$$V = \{v_1, v_2, \ldots, v_m\}, \quad (8)$$

де v_1, v_2, …, v_m – перелік заходів із профілактики надзвичайних ситуацій.

3. Регулятор системи управління

Регулятор призначений забезпечувати бажаний характер роботи системи. В інформаційній системі моніторингу пожежної безпеки свої функції він реалізує шляхом періодичного моніторингу стану пожежної безпеки за нормативними показниками та причин виникнення надзвичайних ситуацій, виявляє їх індивідуальну впливовість при їх комплексній дії та реалізує отриману інформацію при організації профілактичної діяльності даного підрозділу.

Типи вхідних та вихідних сигналів. На вхід бази даних подаються показники стану об'єкта управління та характеристики втрат з причин надзвичайних ситуацій X, відомості про виділені для проведення профілактики ресурси R, значення нормативних показників пожежної безпеки Z та результати спостережень за впливаючими факторами F. Моніторингова інформаційна система обробляє ці дані та перетворює інформацію із вигляду масиву вхідних даних, який містить чисельні характеристики елементів множин X та F до вигляду (5) характеристик впливовості W елементів масиву вхідних даних на характеристики втрат.

ОПР формує план профілактичних заходів на основі отриманої інформації про зміну впливовості факторів впродовж останнього періоду часу шляхом розподілу наявних ресурсів R, пропорційно чисельним значенням елементів множини W.

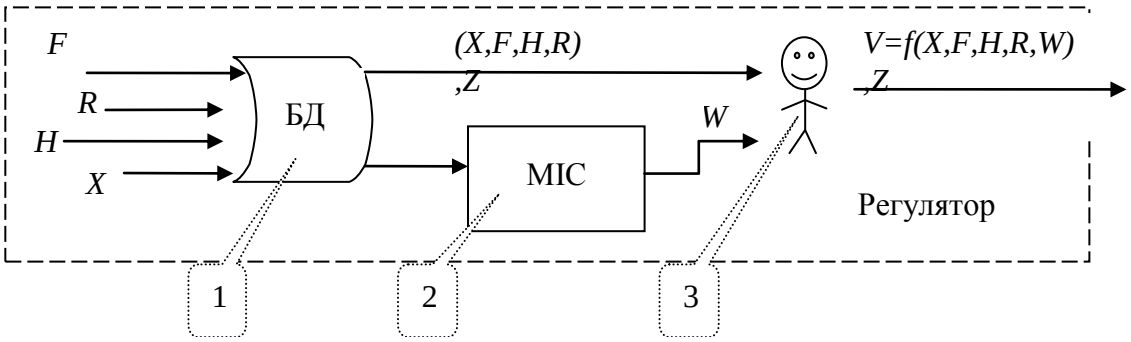


Рис. 3. Структура регулятора: 1 – база даних; 2 – моніторингова інформаційна система; 3 – особа, що приймає рішення (ОПР)

Оскільки модель процесу управління пожежною безпекою запропонована у вигляді розімкнутої системи, то формування функціональної залежності та розв'язування задачі багатопараметричної оптимізації покладається на ОПР, що формує план заходів із профілактики пожеж, і являє собою процедуру евристичну.

При цьому елементи множин X (характеристики стану об'єкта управління) та F (зовнішні впливи) визначаються із бази даних, яка формується шляхом безпосереднього нагляду за спорудами та іншими об'єктами а також реєстрації втрат внаслідок надзвичайних ситуацій. Елементи множини R (види та обсяг ресурсів, що виділяються на реалізацію профілактичних заходів) також нормовані.

Глобальною задачею системи пожежного нагляду D_n є забезпечення процесу прийняття рішень відомостями про реакцію об'єктів моніторингу на застосування керуючих впливів. Рішення буде прийматись в умовах визначеності, коли відомі наслідки реалізації кожного із n заходів із профілактики пожеж.

Глобальна функція перетворення інформації Z (x) реалізується шляхом формування багаторівневої структури автоматизованої системи моніторингу пожежної безпеки. Глобальна функція інформаційної системи протипожежного моніторингу описується відображенням множини характеристик впливаючих факторів X на множину вагових коефіцієнтів цих факторів W:

$$\pi : X \to W , \tag{9}$$

де $X = \{x_1, x_2, \ldots, x_m\}$ – множина характеристик впливаючих факторів;

$W = \{w_1, w_2, \ldots, w_n\}$ – множина вагових коефіцієнтів факторів, що впливають на характеристики пожеж та загорянь.

Складність завдання реалізації такого відображення переважає можливості наявного науково-методичного апарату. Тому для реалізації даної інформаційної системи застосовується декомпозиція глобальної функції системи. Результатом є сукупність задач, серед яких можна виділити три рівні перетворення інформації. Метою першого рівня є отримання залежності характеристик пожеж, що відбулися в окремих галузях Y, від характеристик впливаючих факторів X. Метою другого рівня перетворення інформації є визначення залежності загальних характеристик пожеж Z від характеристик пожеж за галузями, поданих елементами множини Y. Метою третього рівня перетворення інформації є визначення значимості кожного із впливаючих факторів.

Таким чином на перших двох рівнях перетворення інформації послідовно розв'язується задача ідентифікації функціональних залежностей:

$$Y = f(x_1, x_2, \ldots x_m), \tag{10}$$

$$Z = f(y_1, y_2, \ldots y_k), \tag{11}$$

де $Y = \{y_1, y_2, \ldots, y_k\}$ – множина характеристик пожеж за галузями;

$Z=\{z_1, z_2, \ldots, z_s\}$ – множина загальних характеристик пожеж.

Масив вхідних даних має вигляд матриці (12):

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} & y_{11} & y_{12} & \cdots & y_{1k} & z_{11} & z_{12} & \cdots & z_{1s} \\ x_{21} & x_{22} & \cdots & x_{2m} & y_{21} & y_{22} & \cdots & y_{2k} & z_{21} & z_{22} & \cdots & z_{2s} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{l1} & x_{l2} & \cdots & x_{lm} & y_{l1} & y_{l2} & \cdots & y_{lk} & z_{l1} & z_{l2} & \cdots & z_{ls} \end{pmatrix}, \quad (12)$$

де l – кількість спостережень за об'єктом, що містить масив вхідних даних.

На третьому рівні необхідно виявити впливовість кожного із факторів на характеристики пожеж. За результатами досліджень чутливості функціональних залежностей, отриманих на нижньому рівні, розраховуються вагові коефіцієнти впливаючих факторів, які характеризуються даними множини X [7].

Локальною задачею перетворення інформації $D_i(\gamma)$ є побудова функціональної залежності і-ї характеристики пожеж від характеристик причин, що їх викликають.

Координація локальних задач – це спосіб формування зв'язків між елементами нижнього рівня – показниками масиву вхідних даних.

Стратегія координації локальних задач перетворення інформації γ це метод висхідного синтезу елементів структури системи – моделей об'єктів моніторингу відповідного рівня [9].

Аналізуючи складові виразу (10) можна зазначити, що в процесі розв'язання наукового завдання адаптації існуючої інформаційної технології багаторівневого моніторингу до нових умов моніторингу пожежної безпеки визначеними майже залишаються всі елементи. Задано перелік глобальних та локальних задач перетворення інформації, в якості стратегії координації використовується метод висхідного синтезу елементів. Властивості глобальної функції системи та її локальних функцій перетворення інформації залежить від властивостей масиву вхідних даних X. Таким чином основним об'єктом досліджень є процес формування масивів вхідних даних для синтезу моделей об'єктів моніторингу кожного із рівнів та особливості координації елементів структури системи при зміні властивостей МВД.

Відповідно [8] технологія багаторівневого моніторингу пожежної безпеки може бути подана у вигляді агрегативної системи

Загальна структура такої системи подана на рис. 4.

Підсистема перетворення інформації (далі ППІ) призначена для реалізації глобальної функції системи — перетворення інформації із вигляду масиву вхідних даних до вигляду показників впливовості факторів та значень прогнозованих показників. Вона являє собою ієрархічне поєднання моделей об'єктів моніторингу пожежної безпеки відповідних рівнів. Синтез цих моделей та координація їх взаємодії забезпечується підсистемою управління. Масив вхідних даних (показників станів об'єктів

моніторингу) формується підсистемою збору та доставки первинних даних.

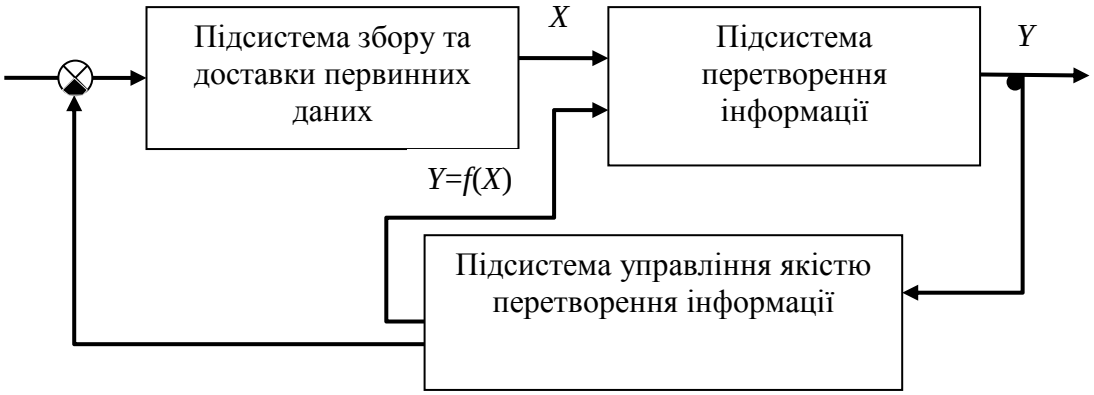


Рис. 4. Структура інформаційної системи

Структурна схема ППІ розробляється на основі положень теорії агрегатів [8]. Вона подається як послідовність алгоритмів перетворення інформації (АПІ), поєднаних в окремі страти, що організовані в агрегатовану ієрархічну структуру. Одним з основних елементів якої є агрегат-перетворювач інформації.

Функціональними вимогами до агрегату є перетворення інформації від вигляду показників стану пожежної безпеки об'єктів на певній території X до вигляду інтегральної характеристики прогнозованих втрат ресурсу певного типу, який позначається як y_i.

На підготовчому етапі функцією агрегату є формування своєї структури за допомогою синтезатора моделей.

З метою технічної реалізації структури страти, поданої на рисунку 3., необхідно створити структуру агрегату-перетворювача. Його модель описується виразом:

$$\Sigma=\{ Z, m^{\gamma}, m^{\gamma'} g;\ H, y\}, \qquad (13)$$

де Z – множина станів;

m^{γ} – алгоритм перетворення інформації, що відображає горизонтальні зв'язки елементів;

$m^{\gamma'}$ – алгоритм перетворення інформації, що формує горизонтальний зв'язок;

g – множина керуючих сигналів;

H – множина переходів;

y – вихідний сигнал, що подається на вхід вищої страти;

Структура агрегату-перетворювача подана на рис. 5.

Модель m^{γ} синтезуються на основі первинного опису, що містить вихідні сигнали попередніх страт $X = \{x_1, x_2, \ldots, x_n\}$. Вони синтезуються у вигляді функціональних залежностей (14) та (15).

$$y^{\gamma}_i = f(x_1, x_2, \ldots, x_n,),\ y_i = 0, \qquad (14)$$

$$z^{\gamma}_i = f(y_1, y_2, \ldots, y_m,), y_i = 0. \quad (15)$$

Модель типового агрегату-перетворювача подана на рис. 6 [8].

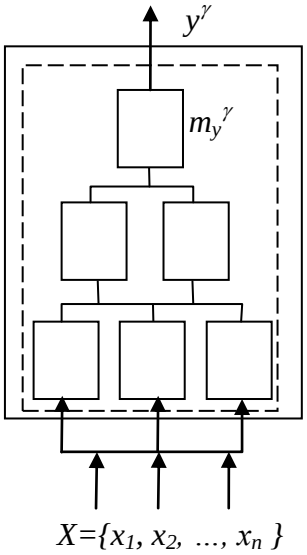


Рис. 5. Структура агрегату-перетворювача: X – масив сигналів із виходу попередньої страти; $m_{y'}$ – АПІ, що формує вихідний сигнал - горизонтальний зв'язок, m_y^{γ} –АПІ із горизонтальними зв'язками

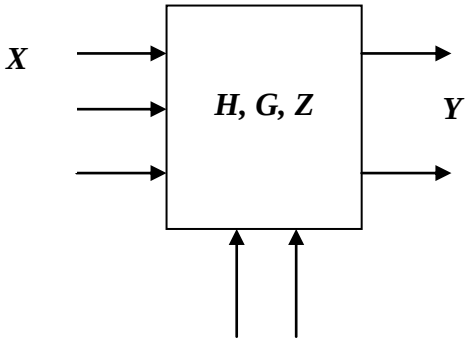


Рис. 6. Модель типового агрегату-перетворювача:
X – множина вхідних сигналів; Y – множина вихідних сигналів; M – множина алгоритмів перетворення інформації;
g – керуючий сигнал; H – оператор переходів;
G – оператор виходів; Z – множина станів агрегату;

Множина станів агрегату Z визначається множиною моделей об'єктів моніторингу даної страти M, які використовуються як АПІ.

Висновки

1. Внаслідок аналізу стану проблеми підтримки прийняття рішень в процесі управління пожежною безпекою у публікаціях вітчизняних і зарубіжних науковців встановлено, що залишаються слабо дослідженими процеси забезпечення інформацією профілактики пожеж та загорянь.

2. В результаті аналізу технології багаторівневого перетворення інформації та в ході обґрунтування вибору технології багаторівневого перетворення вигляду інформації з'ясовано, що вона дозволяє задачу отримання інтегральних характеристик стану пожежної безпеки на заданій території подати у вигляді ієрархічного поєднання локальних алгоритмів перетворення інформації.

3. В процесі управління пожежною безпекою керуючі впливи реалізуються ОПР у вигляді плану заходів із профілактики пожеж як функціональна залежність від характеристик стану пожежної безпеки, впливаючих факторів, нормативних показників та виділених ресурсів. Визначено, що ідентифікація функціональної залежності втрат ресурсів від показників стану пожежної безпеки є ключовою задачею.

Література

1. Гайдамакин, Н.А. Автоматизированные информационные системы, базы и банки данных. Вводный курс: Учеб. пособие. / Н.А. Гайдамакин. – М.: Гелиос АРВ, 2002. – 368 с.

2. Барановская, Т.П. Информационные системы и технологии в экономике. Учебник. – 2-е изд. / Т.П. Барановская, В.И. Лойко, М.И. Семенов, А.И. Трубилин; Под ред. В.И. Лойко. – М.: Финансы и статистика, 2003. – 416 с.

3. Месарович М. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Тахакара. – М.: Мир, 1973. – 344 с.

4. Александров Ю. И. Иерархическая организация поведения / Ю.И. Александров, Ю.В. Гринченко, Р.М. Хвастунов // Успехи физических наук, 1980. – Том 11. – № 4. – С. 115 - 142.

5. Саати Т. Аналитическое планирование. Оранизация систем: Пер. С англ. / Т.Саати, К. Кернс. – М.: Радио и связь, 1991. – 224 с.

6. Дендаренко В.Ю. Метод адаптивного формування структури інформаційної системи моніторингу пожежної безпеки / В.Ю. Дендаренко // Системи обробки інформації, 2010. – вип. 8 (89). – С. 174 – 178.

7. Ковзель М.О. Паралельно-ієрархічне перетворення і Q-обробка інформації для систем реального часу. Монографія / М.О. Ковзель, Л.І. Тимченко, Ю.Ф. Кутаєв, С.В. Свечніков, В.П. Кожем'яко, О.І. Стасюк, С.М. Білан, Л.В. Загоруйко. – Київ.: «КУЕТТ», 2006. – 402 с.

8. Голуб С.В. Координація взаємодій локальних агрегатів в структурі систем багаторівневого перетворення моніторингової інформації / С.В. Голуб // Вісник Східноукраїнського національного університету імені Володимира Даля, 2009. – № 6(136). – Частина 1. – С. 325 – 329.

9. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем / А.Г. Ивахненко. – К.: Наук. думка, 1981. – 296 с.

КОГНИТИВНЫЕ ТЕХНОЛОГИИ В ИНФОКОММУНИКАЦИЯХ

Никитюк Л.А.

Введение

После концепции NGN качественно новая парадигма эволюции сетей связи до сих пор не предложена. Все новации в области инфокоммуникаций сегодня определяются как post-NGN [1], под которой понимается дальнейшая активная компьютеризация средств телекоммуникаций. Пользователь при этом выступает скорее потребителем свойств среды, оснащенной smart-устройствами (умная среда), а не трафика, хотя структура последнего также существенно изменилась. Среди приложений сетей post-NGN наиболее популярны Интернет вещей (IoT), Веб вещей (WoT) и т.п. Основным же технологическим трендом, в соответствии с прогнозом развития отраслей высоких технологий, подготовленном Международным объединением фирм Deloitte, определено использование когнитивных и интеллектуальных технологий. Принципиальное отличие этих двух технологий заключается в том, что когнитивные технологии позволяют моделировать *познавательные* способности человеческого мозга для решения конкретных прикладных задач, таких как: распознавание образов (речи, сигналов, изображений и т.д.); выявление и идентификация закономерностей в массивах данных; принятие решений в условиях предсказуемо изменяющейся среды, в то время как интеллектуальные технологии предполагают *самообучение* и *адаптацию* в непредсказуемой среде. Именно эти технологии сегодня оказывают существенное влияние на процесс развития инфокоммуникаций.

Наиболее перспективными направлениями использования когнитивных технологий сегодня принято считать: когнитивное радио и беспроводные когнитивные сети, призванные обеспечивать высокое качество обслуживания мобильных пользователей и адаптивное управления частотными ресурсами; реализацию дружественного, приспосабливающегося под конкретного пользователя интерфейса. Отдельно следует отметить безусловную целесообразность применения когнитивных технологий для реорганизации системы технической эксплуатации. В эту систему оператор связи инвестирует значительные средства и по этой причине все решения, направленные на повышение ее функциональных возможностей, являются чрезвычайно актуальными и востребованными.

Одной из основных процедур, выполняемых в процессе технической эксплуатации сетей и систем связи, как известно, является мониторинг параметров их динамических характеристик. Эффективность управления сетевыми ресурсами при возникновении внештатных ситуаций в

значительной степени определяется функциональностью и качеством выполнения именно процедур мониторинга.

Когнитивные технологии в процедурах мониторинга могут быть использованы не только для своевременного обнаружения проблем в инфокоммуникационной сети, но и для прогнозирования возникновения различных внештатных ситуаций в разные периоды времени, заблаговременно оповещая о них, и тем самым обеспечивая возможность упреждения негативных последствий.

Кроме того, мониторинг параметров может быть дополнен мониторингом состояний. Принципиальным отличием мониторинга состояния от мониторинга параметров является наличие возможности получения некоторого интегрального параметра — интерпретатора измеренных параметров в терминах состояния. Это, в свою очередь, позволит повысить эффективность решений, принимаемых системой управления.

В данной работе рассматривается один из возможных подходов к реализации когнитивного мониторинга**.**

1. Анализ литературных данных и постановка задачи

Несмотря на все многообразие современных видов мониторинга, применяемых в сетях связи [2–5], все они в основном нацелены в основном на обеспечение достоверного отображения и констатацию текущего состояния объекта. Отдельные подходы к реализации прогностического мониторинга можно наблюдать в ряде научных работ [6–19].

Так, в работе [16], указывается на целесообразность реализации приведенных выше требования к процессу мониторинга, однако четкая формализация процедур прогностического мониторинга отсутствует.

В работах [6–10, 17] рассмотрено использование прогностического мониторинга в сенсорных сетях, что в итоге позволило сократить объем передаваемой служебной информации и уменьшить энергопотребление сенсоров. Однако диапазон параметров мониторинга в этом случае довольно ограничен и не может считаться информативным по отношению к инфокоммуникационным сетям.

В ряде работ [11–15, 18, 19] рассмотрена реализация процедур прогностического мониторинга с использование искусственных нейронных сетей для решения ряда практических задач в различных сферах, таких как:

– обоснование целесообразности добавления новых каналов в сетях на основе мониторинга пропускной способности соединений протокола TCP [19];

– прогноз ухудшения состояния здоровья у пациентов на основе таких данных как кардиограмма, измерение давления, определение содержания кислорода в крови и т. п. [11, 12];

– прогнозирование сбоев в доставке грузов, путем анализа и прогноза динамики бизнес-процессов и внештатных ситуаций на железной дороге [13, 14];

– мониторинг соотношения кислорода и топлива с целью прогнозирования вредных выбросов (SO_2, NO_2, CO_2) мусоросжигательного завода [15].

Однако следует отметить ряд факторов, существенно ограничивающих применение указанных методов для реализации процедур когнитивного мониторинга в инфокоммуникационных сетях, а именно:

– использование искусственных нейронных сетей предполагает процесс обучения, который является слишком затратным по времени, что создает дополнительные риски при управлении большими и сложными объектами;

– точность прогнозирования зависит от числа примеров, которые применялись во время обучения;

– высокие требования к вычислительным возможностям в аппаратной реализации.

В данной работе предлагается подход, обеспечивающий реализацию когнитивных возможностей мониторинга, основанный на использовании математических методов статистического анализа временных рядов и статистического прогнозирования, что значительно упрощает имплементацию процедур прогнозирования в существующие виды мониторинга и позволяет заблаговременно принимать решение о необходимости реконфигурации ресурсов либо реконструкции объекта мониторинга.

2. Процедуры когнитивного мониторинга состояний инфокоммуникационной сети

Состояние и поведение инфокоммуникационной сети как объекта технического обслуживания и управления отображается множеством параметров Y, мощностью m, которые описывают ее динамические характеристики. В указанном множестве могут быть выделены такие группы параметров как: параметры технического состояния, параметры качества обслуживания и параметры процессов накопления и обработки информации (контента).

Задачей когнитивного мониторинга является прогнозирование возможности возникновения внештатной ситуации на объекте мониторинга с заданным периодом упреждения, в течение которого могут быть предприняты действия, направленные на уменьшение последствий воздействия негативных факторов.

Предлагаемый подход включает последовательное прохождение фаз краткосрочного, ситуационного и долгосрочного прогнозирования [22].

Краткосрочное прогнозирование – это прогнозирование с периодом упреждения L_K, который соответствует одному шагу измерения параметра $y_k \in Y\left(k=\overline{1,m}\right)$, подлежащего мониторингу.

Ситуационное прогнозирования – прогнозирование с периодом упреждения L_C, в течение которого изменение параметра $y_k \in Y\left(k=\overline{1,m}\right)$ достигает некоторого, заданного порогового значения y_{kP}. Длительность периода L_C определяется временем, необходимым для выполнения действий по реконфигурации сетевых ресурсов.

Долгосрочное прогнозирование – прогнозирование с периодом упреждения L_D, который охватывает жизненный цикл объекта мониторинга, после которого объект необходимо реконструировать. Исходными данными для долгосрочного прогнозирования выступает выборка значений частоты F возникающих и прогнозируемых внештатных ситуаций на протяжении периода наблюдения T_F.

Краткосрочное прогнозирование позволяет определить следующее значение $y_k(i+1)$ параметра $y_k(i) \in Y\left(k=\overline{1,m}\right)$, подлежащего мониторингу, с момента накопления соответствующей репрезентативной выборки n_{min}^{K} и может быть использовано для обеспечения полноты выборки ситуационного прогнозирования n_{min}^{C},в случае, когда следующее значение не может быть получено от средств мониторинга вследствие технических сбоев. Таким образом, период упреждения для краткосрочного прогнозирования составляет $L_K = 1$.

Ситуационное прогнозирования направлено на выявление внештатной ситуации, предотвращение которой требует реконфигурации объекта мониторинга, например, вследствие резкого возрастания сетевой нагрузки. Соответствующее оповещение вырабатывается в момент времени $t_z^{кр}$, от которого начинается отсчет периода упреждения L_C, и указывает на возможность достижения наблюдаемым параметром $y_k(i) \in Y\left(k=\overline{1,m}\right)$ порогового значения и или выхода за его пределы, то есть

$$\hat{y}_k(n_{min}^{C}+L_C) \le y_{kP1} \vee \hat{y}_k(n_{min}^{C}+L_C) \ge y_{kP2}, \qquad (1)$$

где y_{kP1} и y_{kP2} – соответственно нижнее и верхнее пороговые значения параметра; n_{min}^{C} – соответствующая репрезентативная выборка.

Сигнал оповещения о необходимости реконструкции объекта мониторинга формируется в том случае, когда частота F появления внештатных ситуаций (ЧПВС) для всего множества Y параметров объекта, подлежащих мониторингу в течение периода наблюдения T_F, достигает, определенного порогового значения $F \ge F_P$, установленного на основе

экспертных оценок. В данном случае, ЧПВС выступает интерпретатором состояния объекта. Период упреждения L_D долгосрочного прогнозирования может быть определен как временной интервал, необходимый для выполнения реконструкционных задач.

Исходными данными для работы когнитивного мониторинга являются:

– перечень параметров $y_k(i) \in Y\left(k = \overline{1,m}\right)$, подлежащих мониторингу;

– периоды упреждения прогнозов соответственно L_K, L_C и L_D;

– пороговые значения $y_{kP} \in Y_P \;\forall\, y_k(i) \in Y\left(k = \overline{1,m}\right)$, где Y_P – множество пороговых значений наблюдаемых параметров, мощностью m;

– F_P – пороговое значение ЧПВС и соответствующий период наблюдения T_F.

Процедуры когнитивного мониторинга выполняются в бесконечном цикле и включают следующие шаги (рис. 1, а, б).

После накопления необходимой репрезентативной выборки значений $(n^K_{min}, n^C_{min}, n^D_{min})$ наблюдаемого параметра, выполняется прогнозирование с соответствующим периодом упреждения (L_K, L_C, L_D).

Так, с момента достижения $i = n^K_{min}$ может быть осуществлено краткосрочное прогнозирование, то есть определено прогнозируемое $\hat{y}_k(i+1)$ значение параметра $y_k(i) \in Y\left(k = \overline{1,m}\right)$, которое ожидается в следующий момент времени. Для этого серийная выборка значений $\{y_k(i)\}$, размером n^K_{min}, анализируется на наличие свойства стационарности. В зависимости от результатов проверки, определяется математическая модель описания изменения данного параметра. В случае наличия свойства стационарности $\{y_k(i)\}$ [21]:

$$y_k(i+L_K) = a_1 y(i-1+L_K) + a_2 y(i-2+L_K) + \ldots + a_p y(i-p+L_K) + e(i) - \\ - b_1 e(i-1+L_K) - b_2 e(i-2+L_K) - \ldots - b_q e(i-q+L_K), \qquad (2)$$

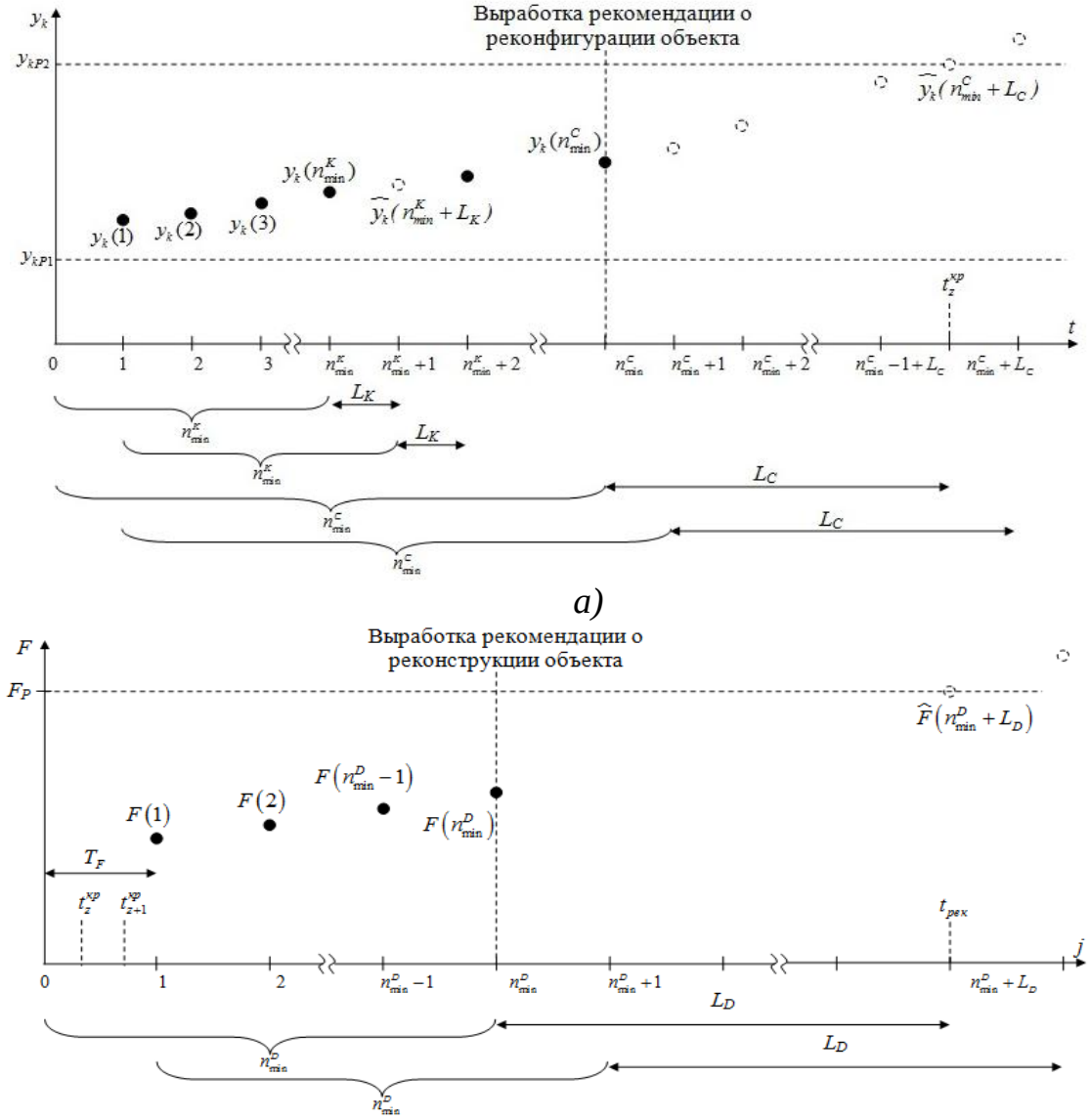
где p – порядок авторегрессии; q – порядок скользящего среднего; a_α $\left(\alpha = \overline{1,p}\right)$; b_β $\left(\beta = \overline{1,q}\right)$ – коэффициенты модели ARIMA, полученной подходом Бокса-Дженкинса [20].

В случае нестационарности $\{y_k(i)\}$:

$$\hat{y}_k(i+L_k) = \hat{T}r(i+L_k) + e(i), \qquad (3)$$

где $\hat{T}r(i+L_k)$ – прогнозируемое значение трендовой составляющей, формализованной полиномиальной моделью [21]; $e(i)$ – случайная

составляющая, с постоянной дисперсией и нулевым математическим ожиданием, последовательные значения которой являются независимыми.



а)

б)

Рис. 1. Прогнозирование: а) – краткосрочное и ситуационное; б) – долгосрочное

С течением времени прогнозируемое значение $\hat{y}_k(i+1)$ параметра $y_k(i) \in Y\left(k=\overline{1,m}\right)$ может меняться, благодаря циклическому обновлению значений $\left\{y_k(i)\right\}$.

Ситуационное прогнозирование можно начать с момента $i = n^C_{min}$, т.е. накопления соответствующей репрезентативной выборки значений наблюдаемого параметра. Его отличительной особенностью является тот факт, что осуществление прогнозирования с периодом упреждения L_C, возможно только при наличии трендовой составляющей во временном

ряду $\{y_k(i)\}$ в пределах репрезентативной выборки n_{min}^{C}. Для отслеживания тренда репрезентативная выборка должна обновляться с каждым последующим шагом мониторинга, что создает эффект «скользящего окна» во временном ряду значений наблюдаемого параметра (рис. 1).

В случае выявления трендовой составляющей, с целью экстраполяции значений в пределах периода упреждения (L_K, L_C), определяется соответствующая математическая функция для описания тренда, которая в общем случае имеет вид:

$$\begin{aligned} Tr(i) &= a_0 + a_1 \cdot i + a_2 \cdot i^2 + \cdots + a_{\lambda_\eta} \cdot i^{\lambda_\eta}; \\ Tr(j) &= a_0 + a_1 \cdot j + a_2 \cdot j^2 + \cdots + a_{\lambda_D} \cdot j^{\lambda_D}; \end{aligned} \quad (4)$$

где $\eta = K \vee C$; a_χ $\left(\chi = \overline{0,\lambda}\right)$ – статистические оценки экстраполирующего полинома степени λ.

В случае выполнения условия $y_{kP1} < \hat{y}_k(i + L_C) < y_{kP2}$ констатируется штатный режим работы объекта. В противном случае ожидается появление внештатной ситуации $(\hat{y}_k(i + L_C) \le y_{kP1} \vee \hat{y}_k(i + L_C) \ge y_{kP2} \forall i \ge n_{min}^{C})$, о чем система когнитивного мониторинга в момент $t_z^{кр}$ времени выдает соответствующее сообщение.

Долгосрочное прогнозирование выполняется после накопления соответствующей репрезентативной выборки значений ЧПВС $\{F(j)\}$, размерностью n_{min}^{D}. Здесь аналогичным образом используется эффект «скользящего окна», в пределах которого значения временного ряда $\{F(j)\}$ анализируются на наличие трендовой составляющей. При обнаружении возрастающего тренда проверяется условие:

$$F(j + L_D) \ge F_P, \quad F(j + L_D) \; \forall \; j \ge n_{min}^{D}, \quad (5)$$

где F_P – установленное пороговое значение; L_D –период упреждения, соответствующий долгосрочному прогнозированию.

Выполнение условия (5) является критерием для принятия решения о целесообразности реконструкции объекта мониторинга и свидетельствуют о том, что мероприятия по реконфигурированию сетевых ресурсов не обеспечивают переход наблюдаемого объекта в штатный режим эксплуатации.

3. Архитектура когнитивного мониторинга

На рис. 2 показана архитектура когнитивного мониторинга в системе эксплуатации инфокоммуникационной сети, отражающая взаимодействие системы когнитивного мониторинга с подсистемой средств мониторинга и системой управления.

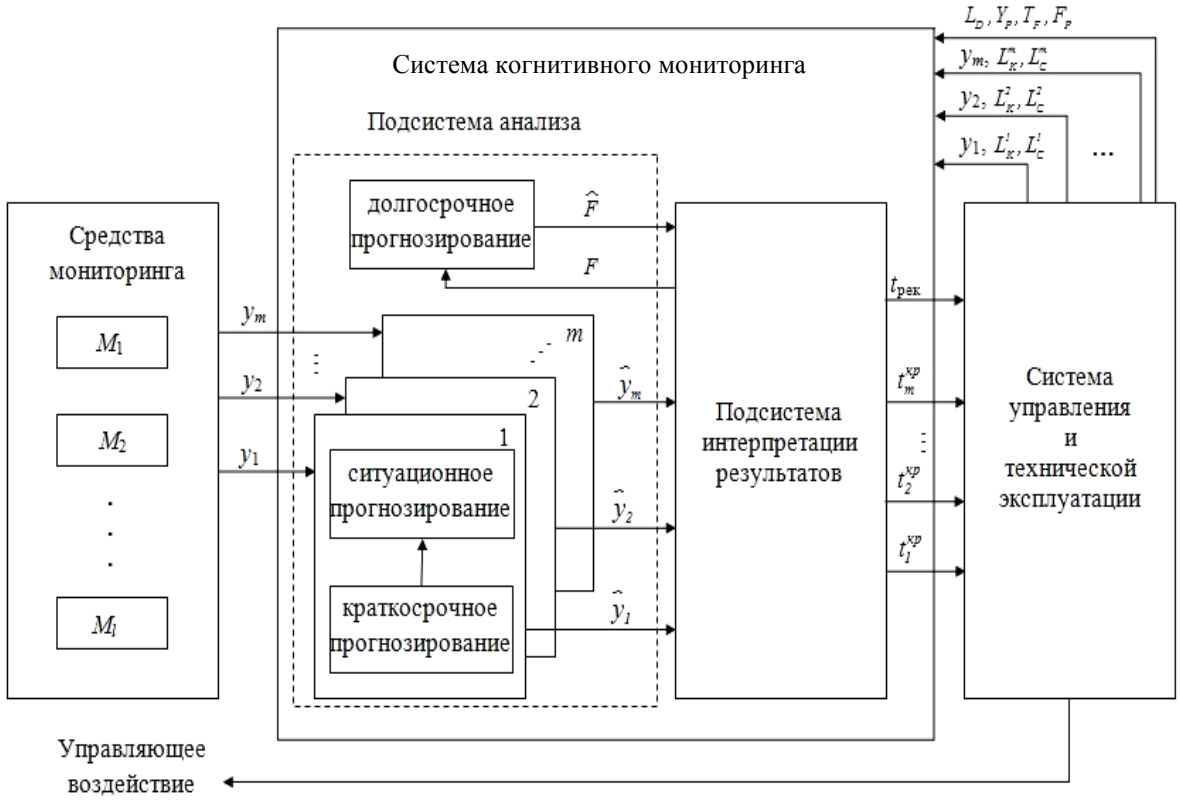


Рис. 2. Архитектура когнитивного мониторинга

Система когнитивного мониторинга состоит из двух функциональных подсистем: *подсистемы анализа* и *подсистемы интерпретации результатов* [22]. В подсистеме анализа выполняется аналитика прогнозов и выявление трендовых составляющих в изменениях параметров состояния объекта мониторинга, а в подсистеме интерпретации – формируется оценка о состоянии объекта и вырабатываются соответствующие рекомендации для системы управления и персонала технической эксплуатации.

Система управления на основе поступающих от системы когнитивного мониторинга данных, характеризующих состояние объекта и тенденции возможного его изменения, вырабатывает соответствующие команды управляющего воздействия на объект мониторинга.

Для повышения эффективности системы управления могут быть использованы интеллектуальные технологии, например, для реализации функций ЛПР (лица принимающего решения), изменения режима работы системы когнитивного мониторинга (путем изменения перечня параметров, подлежащих мониторингу, интервалов упреждения и пороговых значений для выявления внештатных ситуаций).

Выводы

1. Использование когнитивных технологий позволяет значительно повысить функциональность мониторинга динамических характеристик

инфокоммуникационной сети, обеспечивая ему прогностические возможности на основе методов теории статистического прогнозирования.

2. В последовательном прохождении этапов краткосрочного, ситуационного и долгосрочного прогнозирования в бесконечном цикле, с целью определения возрастающих трендов в изменении параметров динамических характеристик объекта мониторинга, когнитивный мониторинг способен распознавать изменения состояния объекта и, в случае необходимости, принимать решение о выдаче рекомендаций к выполнению мероприятий, упреждающих возникновение внештатных ситуаций. Таким образом, можно утверждать, что решается основная задача совершенствования системы технической эксплуатации – она становится превентивной, своевременно выполняющей мероприятия по реконфигурации ресурсов сети, вплоть до ее реконструкции.

Литература

1. Гольдштейн Б.С., Кучерявый А.Е. Сети связи пост-NGN. / Б.С. Гольдштейн, А.Е. Кучерявый А.Е. – Санкт-Петербург.: «БХВ-Петербург», 2014 – 160 с.

2. ITU-R Recommendation V.662-3 Terms and definitions [Text] / Approved 2005. – Geneva: ITU, 2005. – 19 p.

3. Рекомендация МСЭ-R BT.1790 Требования к контролю радиовещательных цепей в ходе эксплуатации [Текст]. – Женева: ITU, 2007. – 6 с.

4. ITU-T Recommendation G.8001 Terms and definitions for Ethernet frames over Transport [Текст] / Approved 2008-03-29. – Geneva: ITU, 2008. – 12 p.

5. ITU-T Recommendation I.113 Vocabulary of terms for broadband aspects of ISDN [Текст] / Approved 1997-06-20. – Geneva: ITU, 1997. – 35 p.

6. Azad A. Map based Predictive Monitoring for Wireless Sensor Networks [Электронный ресурс] / A. Ali, A. Khelil, K. Shaikh, N. Suri // Technische Universitat Darmstadt – Режим доступа: \www/ URL: http://www.gkmm.tu-darmstadt.de/publications/files/AKSS09.pdf

7. Achir, M. and Ouvry, L. Power consumption prediction in wireless sensor networks. [Электронный ресурс] / M. Achir, L. Ouvry // The Pennsylvania State University – Режим доступа: \www/ URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.60.1065&rep=rep1&type=pdf

8. Landsiedel, O. Accurate prediction of power consumption in sensor networks. [Электронный ресурс] / O. Landsiedel, K. Wehrle, S. Gotz // The Pennsylvania State University – Режим доступа: \www/ URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.6036&rep=rep1&type=pdf

9. Mini, A. F. A probabilistic approach to predict the energy consumption in wireless sensor networks. [Электронный ресурс] / A. F. Mini, B. Nath, A. F. Loureiro // The Pennsylvania State University – Режим доступа: \www/ URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.4906&rep=rep1&type=pdf

10. Wang, X. Robust forecasting for energy efficiency of wireless multimedia sensor networks. [Электронный ресурс] / X. Wang, J. Ma, L. Ding, D. Bi // MDPI AG – Режим доступа: \www/ URL: http://www.mdpi.com/1424-8220/7/11/2779

11. Clifton, L. (2014). Predictive Monitoring of Mobile Patients by Combining Clinical Observations With Data From Wearable Sensors / L. Clifton, D. A. Clifton, M. A. F. Pimentel, P. J. Watkinson, L. Tarassenko, // IEEE Journal of Biomedical and Health Informatics. – Vol.:18, Issue: 3, 722 – 730. DOI: 10.1109/JBHI.2013.2293059.

12. Moorman, J. R. Predictive monitoring for early detection of subacute potentially catastrophic illnesses in critical care / J. R. Moorman, C. E Rusin, L. Hoshik, L. E. Guin, M. T. Clark, J. B. Delos, J. Kattwinkel, D. E. Lake // Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2011. – Boston, USA August 30 – September 3, 2011. DOI: 10.1109/IEMBS.2011.6091407.

13. Metzger, A. (2014) Comparing and Combining Predictive Business Process Monitoring Techniques / A. Metzger, P. Leitner, D. Ivanovic, E. Schmieders, R. Franklin, M. Carro, S. Dustdar, K. Pohl // IEEE Transactions on Systems, Man, and Cybernetics. – Vol.:PP , Issue: 99, 1. doi: 10.1109/TSMC.2014.2347265

14. Franceschinis, M. Predictive monitoring of train wagons conditions using wireless network technologies / M. Franceschinis, F. Mauro, C. Pastrone, M. A Spirito, M. Rossi // 2013 XXIV International Symposium on Information, Communication and Automation Technologies (ICAT). – Sarajevo, Bosna i Hercegovina October 30 – November 1, 2013. DOI: 10.1109/ICAT.2013.6684032.

15. Zain, S. M Development of a neural network Predictive Emission Monitoring System for flue gas measurement / S. M Zain, Kien Kek Chua // 2011 IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA). – Penang, Malaysia March 4–6, 2011. DOI: 10.1109/CSPA.2011.5759894.

16. Бобало Ю. Я. Моніторинг об’єктів в умовах апріорної невизначеності джерел інформації. Теорія і практика [Текст] / Ю. Я. Бобало, Ю. Г. Даник, Л. О. Комарова, О. О. Лук’янов, В. М. Максимович, О. О. Писарчук, Ю. Б. Сторонський, Б. М. Стрихалюк – Львів: Коло, 2014. – 252 с.

17. Yoo, W. Network Bandwidth Utilization Forecast Model on High Bandwidth Network / Yoo, W., Sim, A. 2015 International Conference on Computing, Networking and Communications (ICNC) Garden Grove, USA. 16-19 Feb. 2015 Page(s): 494 - 498 DOI: 10.1109/ICCNC.2015.7069393

18. Zaman, F. A recommender system architecture for predictive telecom network management / Zaman, F., Hogan, G., Der Meer, S., Keeney, J., Robitzsch, S., Muntean, G.-M. Communications Magazine, IEEE(Volume:53 ,Issue: 1) Pages: 286 – 293, 2015 DOI: 10.1109/MCOM.2015.7010547

19. M. Mirza, J. Sommers, and P. Barford, “A Machine Learning Approach to TCP Throughput Prediction,”IEEE/ACM Trans. Netw., vol. 18, no. 4, pp. 1026–1039, Aug. 2010. [Online]. Available at: http: //ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5378489

20. Сажин Ю. В. Анализ временних рядов и прогнозирование [Текст] / Ю. В. Сажин, А. В. Катынь, Ю. В. Сарайкин. – Саранск: Изд-во Мордов. Ун-та, 2013. – 192 с.

21. Четыркин Е. М. Статистические методы прогнозирования. / Е. М. Четыркин. – М.: «Статистика», 1977. – 200 с.

22. Бабич Ю.О. Повышение функциональности мониторинга динамических характеристик информационно-коммуникационных сетей. / Ю.Л. Бабич, Л.А. Никитюк – Восточно-Европейский журнал передовых технологий, Vol.4, No 9(76), 2015. с. 9 – 15.

УДОСКОНАЛЕНИЙ МЕТОД ПЛАНУВАННЯ МЕРЕЖ LTE

Одарченко Р.С.

Вступ

Побудова інформаційного суспільства в Україні є одним з найактуальніших завдань сьогодення [1]. Велике значення при цьому відіграє впровадження перспективних інформаційних технологій та методів автоматизації процесів виробництва. Питання використання глобальної інформаційної мережі Інтернет є одним з пріоритетних напрямів державної політики у сфері інформатизації. В Стратегії розвитку інформаційного суспільства в Україні до пріоритетів формування сучасної інформаційної інфраструктури країни віднесено створення високошвидкісних мереж широкосмугового мобільного доступу до Інтернет на всій території України [2, 3]. Йдеться про широкосмуговий доступ на базі використання технологій мобільного зв'язку третього і четвертого поколінь 3G і 4G [4] (від англ. Generation – покоління) та 5G в майбутньому. На вирішення питань щодо впровадження 4G технологій на території України спрямовано реалізацію національного проекту “Відкритий світ”, яким передбачається створення інформаційно-комунікаційної освітньої мережі національного рівня на базі технологій радіозв'язку четвертого покоління.

Стандарт 4G обіцяє набагато більші швидкості передачі даних [5, 6]: понад 100 Мбіт/с швидкорухомим абонентам (наприклад, потягам і автомобілям) та 1 Гбіт/с абонентам з невеликою рухливістю (наприклад, пішоходам і фіксованим абонентам) згідно з міжнародною специфікацією International Mobile Telecommunications Advanced (IMT-Advanced) від 2008 року.

Особливістю розвитку мереж LTE (Long Term Evolution) є можливість їх побудови на вже розвинених мережах, як операторів GSM, так і операторів CDMA, що помітно знижує вартість розгортання мереж [6]. Складність переходу до LTE-мереж в Україні зумовлена проблемами отримання ліцензій для нового спектру частот і необхідністю спеціальних абонентських пристроїв, здатних одночасно працювати в мережах LTE і 3G.

Виходячи з вищесказаного, можна сказати, що розвиток інфраструктури широкосмугового доступу до Інтернет на всій території України на базі створення високошвидкісних мереж четвертого покоління є задачею актуальною та перспективною.

Під час організації мереж мобільного стільникового зв’язку одними з ключових залишаються задачі планування, оптимізації та максимально ефективного використання ресурсів цієї мережі.

Структура мережі LTE сильно відрізняється від мереж стандарту 3G [6, 7]. Істотні зміни зазнала і підсистема базових станцій, і підсистема комутації. Була змінена технологія передачі даних між обладнанням

користувача та базовою станцією. Також зазнали зміни і протоколи передачі даних між мережевими елементами. Для належної підтримки нових широкосмугових технологій радіодоступу в сучасних стільникових мережах повинна бути підвищена ефективність передачі інформації при зниженні вартості доставки кожного мегабайта трафіку та забезпеченні якості обслуговування (QoS), необхідного кожному типу трафіку. Таким чином, з метою оптимізації вже існуючих та побудови нових мереж 3G/4G необхідно розробляти методи, які дозволять підвищити ефективність стільникових мереж зв’язку для того, щоб вони могли відповідати ряду критеріїв:

- забезпечувати впровадження нових систем мобільного зв'язку і підтримку наявних (збереження вкладених інвестицій);
- відповідати вимогам архітектур мереж наступного покоління;
- мати ефективні засоби управління трафіком і забезпечення якості обслуговування;
- надавати зручні засоби технічного обслуговування та експлуатації.

Із розвитком стільникових мереж з’являються нові більш досконалі мережеві архітектури для передачі даних та керування. Проте залишається ряд невирішених завдань та проблемних місць, які необхідно вирішувати та усунути відповідно.

При цьому для операторів стільникового зв’язку в Україні однією з головних проблем створення та розвитку сучасних безпроводових систем зв'язку є більш ефективне планування, що дозволить з одного боку забезпечити необхідну якість обслуговування та з іншого підвищити економічну ефективність використання мережевих ресурсів.

Тому **метою даної роботи** є удосконалення методів планування мереж LTE.

Досягнення поставленої мети **передбачає розв’язання таких завдань:**

1. Проаналізувати якість обслуговування абонентів стільникових мереж в Україні з метою визначення їх ефективності та визначити вимоги до стільникових мереж нового покоління.

2. Розробити метод планування мережі LTE з урахуванням вимог до якості обслуговування та продуктивності мережі, кліматичних умов, рельєфу місцевості та характеру забудови, на основі якого розробити відповідне алгоритмічне та програмне забезпечення для оцінки радіо покриття та попереднього розрахунку вартості розгортання мережі.

Об'єкт дослідження є процес планування стільникових мереж LTE.

Предмет дослідження є методи оцінки зон радіопокриття стільникових мереж LTE.

1. Розвиток стільникових мереж в Україні

16 червня 1993 року вважається датою, коли в Україні було запроваджено стільниковий зв'язок і здійснено перший дзвінок з мобільного телефону. Першою компанією на ринку стільникового зв'язку стала компанія «Український Мобільний Зв'язок» (UMC) [1].

Основними ж віхами в історії розвитку стільникових мереж в Україні слід вважати наступні дати:

- 1997 рік – створення першої в Україні мережі стандарту GSM-900 (Global System for Mobile Communications) [2], що передбачала надання більш ширших можливостей, використання телефону у роумінгу (мережу створила компанія UMC);
- 2000 рік став переломним. Лідируючі позиції надовго перехопив Kyivstar.
- 2002 рік – UMC підключає мільйонного абонента. До цього часу «Київстар» тестово запроваджує технологію передачі даних GPRS [3].
- 2004 рік - впровадження високошвидкісної передачі даних EDGE [3].
- 2005 рік – липень Київстар реєструє 10 млн. абонентів, серпень про теж саме заявляє UMC. На ринок дуже швидко входить турецька компанія «Астеліт» зі своїм брендом Life:).

В цих умовах постійна боротьба призводить до того, що на ринку з'являються все нові більш вигідні пропозиції і акційні стартові пакети. Кожна із компаній починає анонсувати запуск мереж наступних поколінь. Кожен із найбільших стільникових операторів вже давно мали тестові закриті 3G мережі, проте із ряду причин не могли запустити їх у комерційне використання. Стільниковий оператор МТС ще в 2010 році анонсував запуск тестової мережі LTE [4]. Проте до зовсім недавнього часу не відбувалося якісного переходу (стрибку) в стільникових мережах в Україні.

Якщо відкрити карту покриття мобільного Інтернету в світі на сайті Opensignal.com і виставити мітки навпроти значків 3G і 4G, то Україна здасться білою плямою в Європі – тут є переважно лише зв'язок повільного за сучасними мірками, застарілого стандарту 2G (рис. 1).

Але від 23 лютого 2015 року, коли нарешті Нацкомісія з регулювання зв'язку продала на конкурсі 3G-ліцензії всім трьом найбільшим стільниковим операторам країни – Астеліту (торгова марка Life :)), МТС-Україна і Київстару, ситуація кардинально має змінитися на краще. Сумарна вартість ліцензій для операторів склала 8,6 млрд грн [15].

Все це означає, що подивитися онлайн-відео на великій швидкості в автомобілі швидше за все не вийде – для цього потрібен зв'язок вже четвертого покоління. Проте вдома або в офісі, якщо поруч не буде великої кількості абонентів, навіть онлайн-кінотеатр в смартфоні стане реальністю.

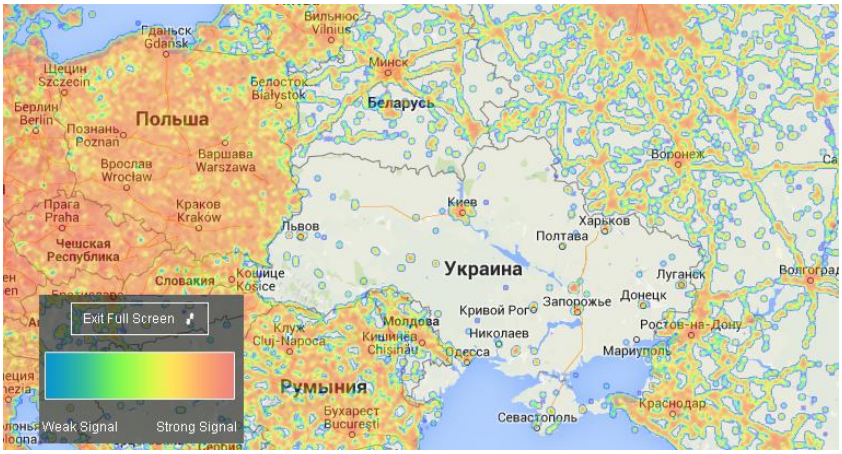


Рис. 1. Покриття стандартів покоління 3G в Україні та в сусідніх країнах

Швидкий мобільний інтернет – це не тільки зручність для користувачів-відеоманів, але і глобальна зміна якості життя в країні. Швидкий зв'язок вплине на багато сфер, починаючи з розвитку мобільної комерції і закінчуючи можливістю отримувати повноцінну онлайн-освіту в будь-який зручний час і без прив'язки до домашнього або робочого комп'ютера.

3G дасть відчутний поштовх і для розвитку економіки країни. За даними спільного дослідження компаній GSMA, Deloitte і Cisco, збільшення кількості користувачів 3G хоча б на 10% призводить до помітного прискорення темпів зростання ВВП на душу населення [16].

За оцінками Національної комісії з регулювання зв'язку, вже до 2018 року майже 40% населення України будуть користуватися новим стандартом.

Проте, слід звернути увагу на те, що поки в Україні тільки розпочинається впровадження нового стандарту, в більшості європейських держав споживачі вже перейшли на куди більш прогресивний 4G-зв'язок. Швидкість передачі даних там становить від 100 Мбіт/с до 1 Гбіт/с [7]. Більше того, у ряді країн же вже активно планується до впровадження зв'язок 5G [9].

Від пізнього впровадження 3G в країні операторам тільки один прок - на руках у абонентів вже знаходиться велика кількість терміналів, а в інтернеті вже вистачає сервісів, для яких був потрібен 3G.

Відкладений попит і грамотний маркетинг з боку операторів являється запорукою успіху стільникового зв'язку в Україні. У «Київстару» кількість користувачів 3G вже досягла 7,8 млн, у Vodafone - 7 млн, у lifecell - 2,7 млн. Якщо ці цифри здаються несерйозними, то варто згадати про те, що 3G-мережу ще навіть не у всіх операторів країни перевалила за 50% покриття абонентської бази, але при цьому кожен оператор підключив до швидкісного мобільного інтернету приблизно третину своїх абонентів.

Як буде розвиватися ситуація зі швидкістю проникнення 4G в нашій країні після видачі ліцензій поки складно передбачити. Багато чого буде залежати від того, як швидко українці оновлять свої термінали до таких, які підтримують 4G (взагалі-то напевно швидко, тому що сьогодні навіть самі бюджетні смартфони часто вміють працювати в цих мережах, а 3G простимулює покупку «розумних пристроїв» навіть в тих регіонах, де користувачі якось не поспішали це робити), і як швидко в нашу країну прийдуть сервіси, що вимагають більш швидкого підключення. Наприклад, потокове відео високої якості.

Тому дуже важливим є швидкий розвиток стільникових мереж саме в Україні, що дозволить наблизитись до розвинених країн Європи та світу. Для визначення перспектив їх розвитку необхідно дослідити реалізовані проекти стільникових мереж LTE в світі та їх основні характеристики.

2. Аналіз архітектури та основних характеристик стільникових мереж LTE

Мережі 4G на основі стандарту LTE працюють у всіх існуючих діапазонах частот, що виділені для стільникового звязку по усьому світу. У Північній Америці 700, 750, 800, 850, 1900, 1700/2100, 2500 та 2600 МГц, відповідно діапазони 4, 7, 12, 13, 17, 25, 26, 41; 2500 МГц у Південній Америці; 800, 900, 1800, 2600 МГц у Європі, відповідно діапазони 3, 7, 20; 1800 та 2600 МГц у Азії, відповідно діапазони 1, 3, 5, 7, 8, 11, 13, 40; 1800 МГц та 2300 МГц у Австралії та Новій Зеландії відповідно діапазони 3МГц та 40МГц [10].

Швидкість завантаження за стандартом 3GPP LTE в теорії досягає 326,4 Мбіт/с (download), і 172,8 Мбіт/с на віддачу (upload). Практично забезпечує швидкість передачі даних від базової станції до пристрою абонента до 100 Мбіт/с і швидкість від абонента до базової станції — до 50 Мбіт/с.

Мережа LTE складається з двох найважливіших компонентів: мережі радіодоступу E-UTRAN і базової мережі SAE (System Architecture Evolution) або EPC (Evolved Packet Core Network) [11]. (рис. 2)

Основним досягненням такої архітектури, в порівнянні з попередніми поколіннями є менші затримки при передачі як даних користувача, так і керуючої інформації у зв'язку з проходженням через менше число проміжних елементів [11].

Обмін даними в мережі EPC відбувається тільки по IP протоколу з комутацією пакетів, що суттєво відрізняє мережу LTE від мереж попередніх поколінь, в яких використовувалася комутація каналів між окремими елементами. В дану мережу входять елементи, що відповідають за управління, маршрутизацію, комутацію і зберігання різних даних, про які далі буде розказано більш докладно.

Мережа E-UTRAN, що складається з базових станцій (eNodeB) бере на себе функції радіоінтерфейсу і є сполучною ланкою між терміналами (UE) і мережею EPC.

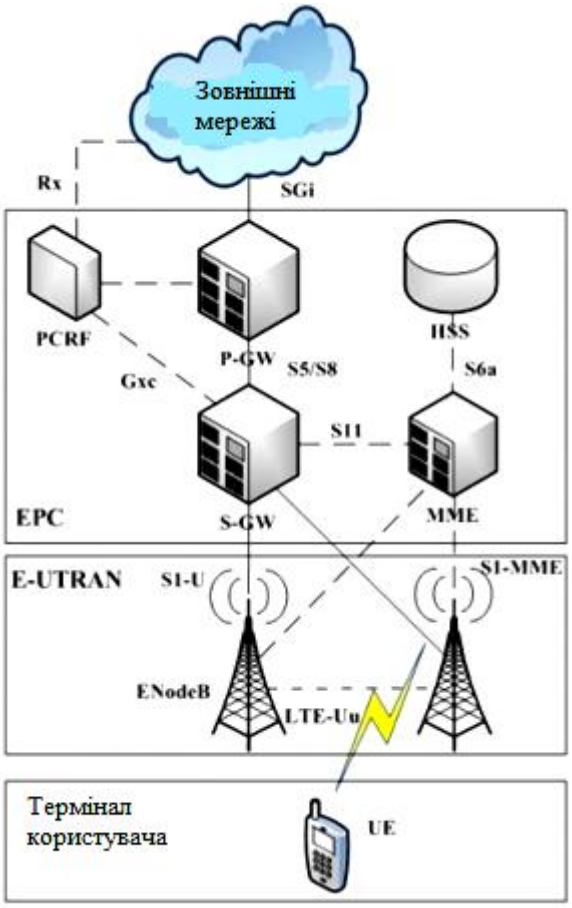


Рис. 2. Спрощена архітектура мережі LTE

Основною особливістю, що відрізняє мережу LTE від мереж інших поколінь, є те, що базові станції eNodeB можуть обмінюватися між собою інформацією по протоколу X2 і здійснювати функції управління. На відміну від стандарту GSM, де підсистема базових станцій BSS складалася з базового приймача BTS і контролера базових станцій BSC в мережі LTE в одному елементі eNodeB об'єднані функції передавача і контролера. Основні технічні характеристики стандарту LTE зведені до табл. 1 [26].

Таблиця 1

Основні технічні характеристики стандарту LTE

Характеристика		Значення
Смуга частот, МГц		1; 4; 3; 5; 10; 15; 20
Метод багатостанційного доступу	Низхідний канал	OFDM
	Висхідний канал	SC-FDMA

Продовження таблиці 1

Характеристика		Значення
Символьна швидкість, символів/с		14000
Завадостійке кодування		згортальні коди, турбокоди
Тривалість радіокадра, мс		10
Мінімальний інтервал між кадрами, мс		1
Стандартний крок між піднесними, кГц (канал “вниз”)		15
Інформаційна одиниця в каналі		ресурсний блок
Кількість піднесних на ресурсний блок (займаюча ресурсним блоком смуга)		12(180 кГц)
Циклічний префікс, мкс	Стандартний	4,7 (5,2 –перед першим символом)
	Розширений	16,7
Дуплексний режим		Частотний (FDD)
		Часовий (TDD)
Модуляція сигналу		QPSK, 16QAM, 64QAM
Максимальна ефективна випромінююча,потужність, дБм	UE	23
	eNodeB	46

3. Дослідження мережі LTE-Advanced

Вимоги до LTE-Advanced:

- високий ступінь функціональності для надання широкого діапазону високошвидкісних послуг в масштабах світового ринку з істотною економічною ефективністю і якістю;
- можливість взаємодії з іншими системами радіодоступу, включаючи повну сумісність з LTE (Rel'8);
- гармонізація і сумісність абонентських пристроїв в міжнародному масштабі;
- реалізація роумінга по всьому світу;
- підтримка ширини каналу до 40 МГц включно;
- можливість організації більш широкої смуги каналу (до 100 МГц), яка потенційно може забезпечити пікову швидкість передачі даних 3 Гбіт / с в Downlink і 1,5 Гбіт / с в Uplink [12];
- забезпечення спектральної ефективності в каналах Downlink до 15 біт/с/Гц при 4x4 MIMO і до 6,75 біт/с/Гц - при 2x4 MIMO в каналах Uplink [10].
- використання 8 передавальних антен MIMO в каналах Downlink [9].

Приводиться опис вимог 3GPP до LTE-Advanced [11]. Дані вимоги основані на вимогах ITU к IMT-Advanced системам:

1. Пікова швидкість передачі даних. Пікова швидкість передачі даних – це максимальна швидкість передачі даних, яка повинна підтримуватись з точки зору системних вимог (а не з точки зору вимог до продуктивності радіоканалу), незалежно від параметрів радіоінтерфейсу таких, як ширина каналу і конфігурація антен. Цільові значення для системи: 1 Гбіт/с в низхідному каналі і 500 Мбіт/с у висхідному каналі.

2. Затримка. Загальна затримка передачі сигнального трафіку повинна бути істотно зменшена в порівнянні з EPC-Rel 8 (LTE). Загальна затримка передачі сигнального трафіку включає в себе час передачі на ділянці радіоінтерфейсу (RAN) і опорної мережі (CN) в умовах малого навантаження (виключаючи час передачі на S1 інтерфейсі, тобто ділянці між eNB і MME). Цільовий час, необхідний на перемикання мобільної станції з холостого стану (Idle) в активний, має становити менше 50 мс. А час перемикання зі стану очікування (dormant state) в активний стан має бути менший 10 мс (виключаючи затримку, пов'язану з процедурою періодичної передачі/прийому, DRX) [13].

Система повинна бути здатна підтримати до 300 активних користувачів без використання DRX режиму (аналог режиму Sleep Mode в IEEE 802.16) при ширині каналу в 5 МГц . З використанням режиму DRX система повинна підтримувати таку ж кількість RRC з'єднань, як і в Rel.8, а саме16000 [11].

При передачі даних, призначених для користувача, повинні досягатися менші затримки в порівнянні з Rel.8, особливо в ситуаціях, коли мобільній станції ще не виділено ресурс для передачі даних і коли мобільній станції потрібно синхронізуватися і отримати ресурс для передачі .

3. Пікова спектральна ефективність. Пікова спектральна ефективність - це максимальна швидкість передачі даних (передбачається передача даних без помилок), нормована на ширину каналу всього сектора, коли весь наявний ресурс виділяється одній мобільній станції. Цільові значення для пікової спектральної ефективності при низхідній передачі 30 біт/с/Гц, а при висхідній передачі 14 біт/с/Гц [11].

4. Середня спектральна ефективність. Середня спектральна ефективність визначається як загальна пропускна здатність всіх користувачів (тобто кількість успішно переданих біт за певний проміжок часу), нормована на загальну ширину каналу сектора і поділена на кількість секторів. Середня спектральна ефективність вимірюється в біт/с/Гц/сектор.

Система повинна забезпечувати якомога більше високе значення середньої спектральної ефективності при розумній складності самої системи . Для прикладу наведемо максимальні цільові значення для середньої спектральної ефективності. Для низхідного каналу це значення дорівнює 3.7 біт/с/Гц/сектор (при конфігурації 4x4, тобто 4 передавальні і

4 прийомні антени), а для висхідного каналу - 2.0 біт/с/Гц/сектор (при конфігурації 2х4).

5. Спектральна ефективність на кордоні сектора. Спектральна ефективність для мобільної станції , що знаходиться на кордоні сектора, визначається як значення інтегральної функції розподілу, нормованої до пропускної спроможності, в точці 5 %. Система повинна забезпечувати максимально можливе значення спектральної ефективності для користувачів, що знаходяться на кордоні сектора , при дотриманні розумної складності самої системи . Для прикладу наведемо максимальні цільові значення для спектральної ефективності, характерної для кордону сектора. Для низхідного каналу це значення дорівнює 0.12 біт/с/Гц/сектор/користувач (при конфігурації 4х4), а для висхідного каналу - 0.07 біт/с/Гц/ сектор/користувач (при конфігурації 2х4). Значення наведені для випадку, коли в одному секторі знаходиться 10 користувачів.

6. *Кількість VoIP дзвінків.* Кількість одночасно підтримуваних VoIP дзвінків повинно бути збільшено в порівнянні зі значеннями, зазначеними в [2], для всіх можливих конфігурацій.

7. Мобільність. Система повинна підтримувати роботу з мобільними користувачами, які можуть рухатися зі швидкістю до 350 км/год (або навіть до 500 км/рік, залежно від використовуваних частот). Продуктивність системи повинна бути поліпшена при роботі з користувачами, які переміщуються зі швидкістю від 0 до 10 км/год. Для більш мобільних користувачів (переміщаються з більш високими швидкостями) продуктивність системи як мінімум не повинна бути гірше, ніж в Rel.8.

8. Частотні діапазони. До наявних частотних діапазонів також додаються наступні:

- 450-470 МГц;
- 698-862 МГц;
- 790-862 МГц;
- 2.3-2.4 ГГц;
- 3.4-4.2 ГГц;
- 4.4-4.99 ГГц.

Нова система (LTE - A) повинна підтримувати роботу з різними розмірами частотних діапазонів, в тому числі і з більш широкими діапазонами (наприклад до 100 МГц), ніж зазначені в Rel.8, для того, щоб забезпечити більш високу продуктивність і цільову пікову спектральную ефективність. Також повинна бути можливість як роботи в режимі частотного (FDD), так і в режимі часового (TDD) дуплексу.

4. Побудова початкового наближення мережі LTE

Завдання побудови початкового наближення мережі LTE можна сформулювати наступним чином: при заданій смузі частот потрібно

визначити просторові параметри мережі (кількість базових станцій та розміри їх зон обслуговування) за умови, що пікові швидкості передачі даних по лінії «вниз» і лінії «вгору» максимальні, а число базових станцій в складі мережі не перевищує допустимого значення.

Іншим варіантом може бути рішення задачі мінімізації числа базових станцій в складі мережі при заданих значеннях пікових швидкостей передачі даних по лінії «вниз» і лінії «вгору».

Складність завдань побудови початкового наближення мережі LTE не дозволяє знайти пряме рішення. При побудові початкового наближення мережі з ортогональним частотним поділом каналів будемо вважати, що:

1) щільність абонентів на території обслуговування мережі постійна, а розподіл абонентів по території рівномірний;

2) розміри всіх стільників мережі одні й ті ж;

3) морфоструктура місцевості однотипна (відкрита місцевість, приміський район або міська забудова).

Для підвищення точності побудови початкового наближення мережі всю територію обслуговування необхідно умовно розбити на фрагменти, де сформульовані вище допущення можна вважати прийнятними.

Виходячи з сформульованих вище обмежень, мережа має регулярну однорідну структуру, тобто вузли eNB видалені між собою на однакову відстань, технічні характеристики і кількість приймально-передавачів, а також висоти підвісу антен, азимути і кути нахилу однакові для всіх eNB.

Для побудови мережі початкового наближення потрібно досить великий набір вихідних даних, достовірність яких може істотно вплинути на адекватність прийнятого рішення. На цьому етапі проводиться оцінка бюджету втрат – показника, що характеризує допустимі втрати в радіолінії для заданого стандарту стільникового мобільного зв'язку.

Таким чином, вхідними даними до первинного планування мережі LTE будуть параметри вказані в табл. 2.

Таблиця 2

Вихідні дані для планування мережі

№ з./п.	Параметр	Одиниці виміру / варіанти
1.	Площа території, на якій необхідно забезпечити покриття	$м^2$
2.	Характер забудови	- Місто - Передмістя - Сільська місцевість
3.	Затримка в каналі	с
4.	Імовірність бітової помилки	Не вищче p_{bit}
5.	Необхідна пропускна здатність на одного абонента	Мбіт/с
6.	Кількість абонентів	

5. Методика оцінки бюджета втрат і зони покриття

Максимально допустимі втрати при поширенні в каналі рівні:

$$L = P_{TX} + G_{TX} - P_{RX} - B_{BODY} + G_{RX} - B_{fid} - IM - L_{slow} - L_{мет} - L_{ф}, \quad (1)$$

де P_{TX} – потужність передавача;

G_{TX} – коефіцієнт підсилення передавальної антени;

P_{RX} – чутливість приймача;

B_{BODY} – втрати в тілі абонента;

G_{RX} – коефіцієнт підсилення приймальної антени;

B_{fid} – втрати у фідері, IM – запас по інтерференції;

L_{slow} – запас на повільні завмирання , береться рівним 10,3 дБ.

$L_{ф}$ − врати сигналу у фідерних лініях. При відсутності фідера (коли прийомопередавачі об'єднані з антеною у вигляді моноблока) необхідно враховувати конструктивні особливості пристрою з'єднання.

$L_{мет}$ – втрати, обумовлені поглинаннями в атмосферних газах, гідро метеорах, тумані тощо, дБ.

Таким чином, $L_{мет}$ визначається наступною формулою:

$$L_{мет} = L_{тум} + L_{гідромет} + L_{аг}, \quad (2)$$

де $L_{тум}$ – втрати потужності радіосигналу в тумані, дБ;

$L_{гідромет}$ – втрати потужності радіосигналу під час опадів, дБ;

$L_{аг}$ – величина затухань радіосигналу в атмосферних газах.

Математичний апарат для розрахунку $L_{тум}$, $L_{гідромет}$ та $L_{аг}$ наступний.

Втрати у газах атмосфери $L_{Г}$, дБ, визначаються за формулою

$$L_{Г} = (\gamma_{O_2} + \gamma_{H_20}) \cdot l,$$

де γ_{O_2}, γ_{H_20} − погонні втрати (дБ/км) у кисні та водяних парах атмосфери при температурі повітря 15 ^{0}C та відносній вологості 100 % (абсолютна вологість становить 13,4 г/м^3):

$$\gamma_{O_2} = \left(7{,}19 \cdot 10^{-3} + \frac{6{,}09}{f^2 + 0{,}227} + \frac{4{,}81}{(f - 57)^2 + 1{,}5} \right) f^2 \cdot 10^{-3},$$

$$\gamma_{H_20} = \left(0{,}078 + \frac{3{,}6}{(f - 22{,}2)^2 + 8{,}5} + \frac{10{,}6}{(f - 183{,}3)^2 + 9} + \frac{8{,}9}{(f - 325{,}4)^2 + 26{,}3} \right) \times$$

$$\times f^2 \cdot 13{,}4 \cdot 10^{-4},$$

де f – робоча частота, ГГц.

Додаткові втрати $L_{ДОД}$, які складаються з втрат у антенних обтікачах та від перепаду висот приймальної та передавальної антен можна прийняти рівними 1,5 дБ.

Запас по інтерференції IM (дБ) визначається наступним чином:

$$IM = P_{ПР} - P_{ПОР}(10^{-3}),$$

Як показано в [19], чутливість приймача P_{RX} може бути представлена наступним чином:

$$P_{RX} = 10 \cdot \lg\left((E_b / N_0) \cdot k \cdot T \cdot R\right),$$

де (E_b / N_0) – відношення сигнал/шум в цифрових систмемах зв'язку – це відношення енергії сигналу на 1 біт до щільності потужності шумів на 1 герц;

R – швидкість передавання даних;

k = 1,23×10^{-23} Дж/к – стала Больцмана,

T- температура в Кельвінах (абсолютна температура).

Таким чином вираз () можна представити у вигляді:

$$L = P_{TX} + G_{TX} - 10 \cdot \lg\left((E_b / N_0) \cdot k \cdot T \cdot R\right) - B_{BODY} + G_{RX} - B_{fid} -$$
$$- IM - L_{slow} - L_{мет} - L_{ф}.$$

Для визначення співвідношень сигнал/шум, при яких не перевищується імовірність виникнення бітових помилок, проведемо моделювання для імовірності бітової помилки в обох підканалах для різних типів модуляції, яка використовується в обладнанні мереж LTE.

Імовірність бітової помилки BER_0 у гаусівському каналі дорівнює [97]

$$BER_0(\eta) = 0.5[1 - \Phi(\sqrt{a\eta})], \quad (3)$$

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)dt,$$

де α=2 и α=1 для бінарної и квадратурної фазових модуляцій, відповідно.

ВСШ в і-му власному підканалі $\eta_i = \beta_i \rho_0 \lambda_i$. Враховуючи нормування щільності імовірності і, вводячи параметр $\rho_i = \beta_i \alpha \rho_0$, отримаємо

$$BER_i = \frac{1}{2} - \frac{1}{2} \int_0^\infty f_i(\lambda) \Phi(\sqrt{\rho_i \lambda}) d\lambda. \quad (4)$$

Враховуємо, що імовірність бітової помилки P_b для BPSK та QPSK визначається виразом [5]

$$P_b = Q\left(\sqrt{2\gamma_b}\right), \quad (5)$$

де $Q(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_x^\infty \exp(-\frac{u^2}{2}) du$ – таблична функція, значення котрої надане у [21];

γ_b – відношення енергії біта E_b до спектральної щільності шуму N_0.

Для гаусівського каналу і прийому за допомогою узгоджених фільтрів імовірність бітової помилки при модуляції M-QAM, де $M=2^k$ і k – парне, визначається наступним чином [24]

$$BER = \frac{2 \cdot (1 - L^{-1})}{\log_2(L)} \cdot Q\left[\sqrt{\frac{3 \cdot \log_2(L)}{L^2 - 1} \cdot \frac{2 \cdot E_b}{N_0}}\right], \quad (6)$$

де $L=\sqrt{K}$ представляє кількість рівнів амплітуди в одному вимірі.

Підставляючи сюди вирази (3) – (4) і проводячи необхідні обчислення в результаті для імовірності бітової помилки в сильному (першому) и слабкому (другому) власних каналах MIMO – системи з довільним числом M передавальних антен будемо мати наступні графічні залежності отримані під час моделювання (рис. 3 та рис. 4).

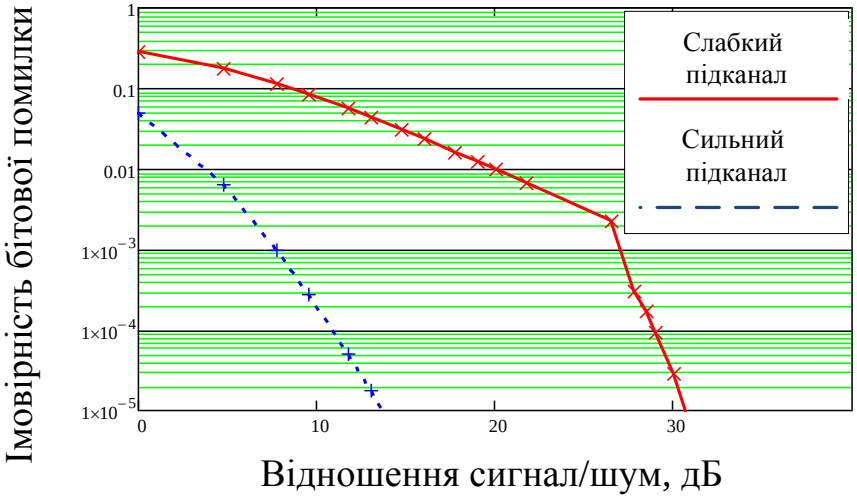


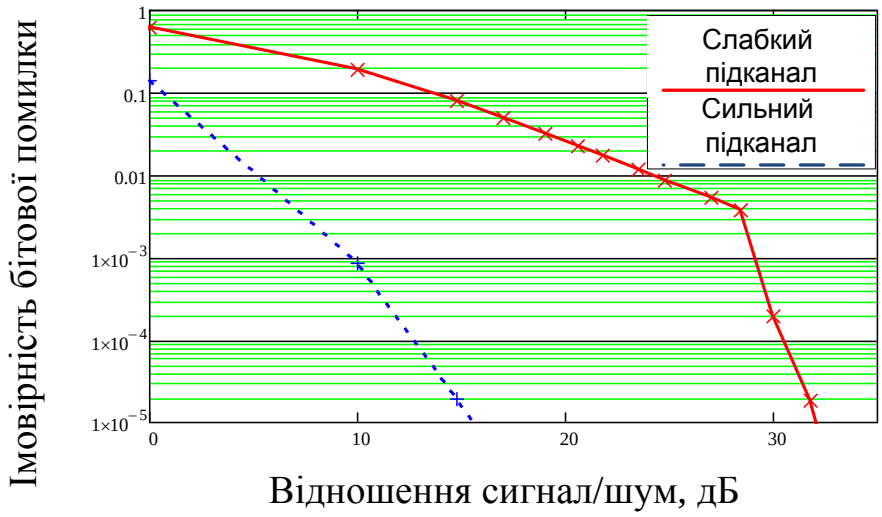
Рис. 3. Логарифмічна залежність імовірності бітової помилки від відношення сигнал/шум в слабкому та сильному власних підканалах для модуляції QPSK

Для оцінки величини L на трасі розповсюдження радіохвиль, керуючись результатами проведених досліджень, можна розробити наступні рекомендації щодо використання моделей розповсюдження радіохвиль для різних діапазонів частот (табл. 3).

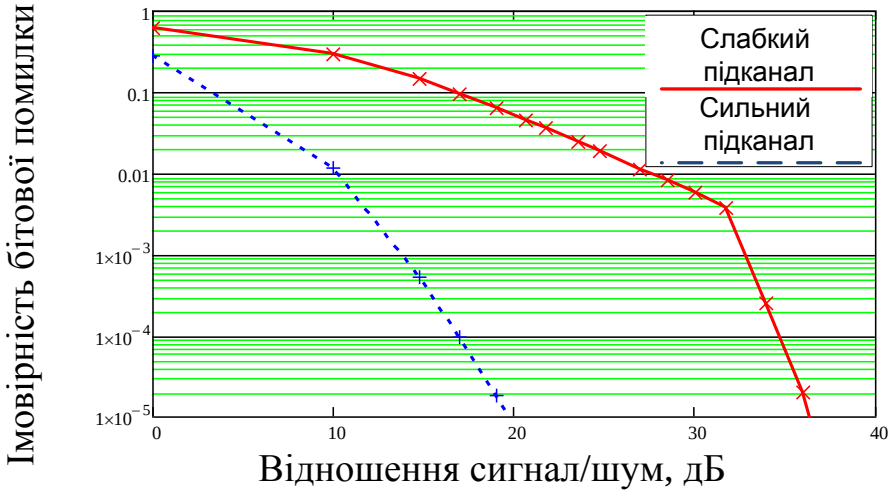
Таблиця 3

Рекомендації щодо використання емпіричних моделей розповсюдження радіохвиль для різних діапазонів частот

Частотний діапазон	**Технологія**	**Рекомендована модель розповсюдження радіохвиль**
До 2 ГГц	LTE, WiMAX, GSM, UMTS тощо	Модель Хата
2,3 ГГц; 2,5 ГГц; 2,6 ГГц	LTE, WiMAX	Модель Хата Cost 231
3,5 ГГц	LTE, WiMAX	SUI (для умов LOS) та Ericsson (для умов NLOS)
5 ГГц; 5,8 ГГц	WiMAX	SUI (для умов LOS) та Ericsson (для умов NLOS)



а)



б)

Рис. 4. Логарифмічна залежність імовірності бітової помилки від відношення сигнал/шум в слабкому та сильному власних підканалах для модуляції QAM-16 (а) та QAM-64 (б)

Таким чином, можна виразити максимальну дальність зв'язку для різних частотних діапазонів, що використовуються в LTE.

Наприклад, до 2 ГГц (за допомогою моделі Хата-Окамури):

$$\lg(d) = (L - 69{,}55 + 26{,}16\lg(f) + 13{,}82\lg(h_b) + \alpha(h_m) + K)/(44{,}5 - 6{,}55\lg(h_b)),$$

де h_b = (30–200) – висота антени базової станції, м; h_m = (1–10) – висота антени мобільної станції, м; d = (1–20) – відстань між антеною базової станції й антеною мобільної станції, км; f = (150–1500) – основна частота, МГц. Елемент $\alpha(h_m)$ – коефіцієнт виправлення висоти, що залежить від навколишнього середовища. Він дорівнює нулю для $h_m = 0$. Коефіцієнт K використовується як виправлення для формул сільської місцевості, передмістя й відкритих зон.

Результати розрахунку бюджету втрат в системах LTE показують, що збільшення смуги частот каналу призводить до зменшення допустимих втрат розповсюдження і для збереження енергетичного балансу між каналами «вгору» і «вниз» необхідно обмежувати кількість ресурсних блоків, що припадають на абонентську станцію:

– при смузі частот каналу 10 МГц допустимі втрати знаходяться в межах (125,8 – 148 дБ) і доцільно обмежувати кількість ресурсних блоків, що виділяються абонентській станції, до 4;

– при ширині смуги частот каналу 15 МГц допустимі втрати знаходяться в межах 121,8 – 144,2 дБ і доцільно обмежувати кількість ресурсних блоків, що виділяються абонентської станції, до 8;

– при ширині смуги частот каналу 20 МГц допустимі втрати знаходяться в межах 117,8 – 139,9 дБ і доцільно обмежувати кількість ресурсних блоків, що виділяються абонентської станції, до 16.

Результати моделювання показують, що для міських умов:

1. Радіус зони покриття базових станцій LTE в діапазоні 2300 – 2400 МГц при робочій смузі 10, 15 і 20 МГц зменшується і становить, відповідно:

– при модуляції QPSK 1/3 – 3.3, 2.8 і 2.4 км;

– при модуляції 16QAM ½ – 1.8, 1.6 і 1.45 км;

– при модуляції 64QAM ¾ – 1, 0.8 та 0.7 км.

2. При збільшенні смуги частот з 10 до 20 МГц площа зони покриття базових станцій LTE зменшується:

– при модуляції QPSK 1/3 з 28 до 15 кв.км;

– при модуляції 16QAM ½ з 8 до 5 кв.км;

– при модуляції 64QAM ¾ з 2.6 до 1.3 кв.км.

Відповідні модуляційно-кодують схеми вибираються, виходячи з необхідної якості послуги.

Оцінимо залежність дальності зв’язку від частоти при різних видах забудови (рис. 5).

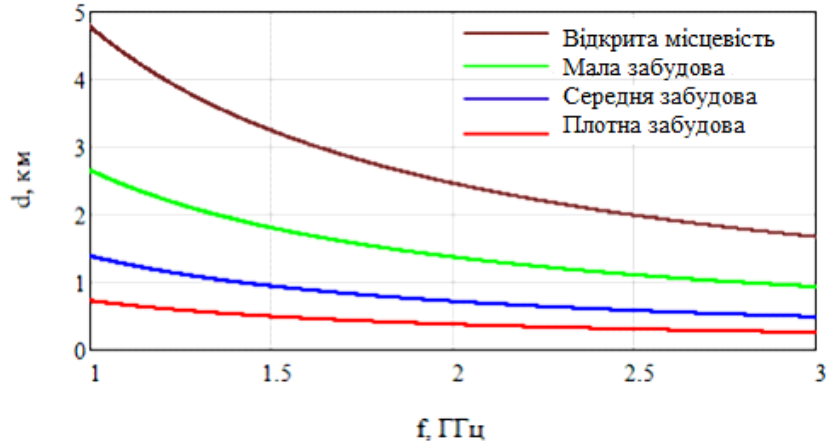


Рис. 5. Залежність дальності зв’язку від частоти для різних типів забудови

На основі використання удосконаленому в даній роботі методу було розроблене спеціалізоване програмне забезпечення (веб-додаток), яке може бути використане для оцінки зон радіо покриття та кількості БС при первинному плануванні мережі стільникового зв'язку (рис. 6).

Кількість абонентів 895895
Верхня частота діапазону частот (МГц) 2600
Нижня частота діапазону частот (МГц) 2500
Середня спектральна ефективність для мережі LTE лінія ul и dl відповідно 1x2 ▾ 2×2 ▾
Ширина каналу (МГц) 10
Кількість секторів БС 3
Полоса частот одного радіоканалу (кГц) 180
Розмірність кластера 3
Значення допустимого навантаження в секторі одного стільника 54
Кількість каналів трафіку в одному радіоканалі 1 ▾
Середнє за усіма видами трафіку абонентське навантаження від одного абонента (0,04...0,2 Ерл) 0.2
середній трафік одного абонента за місяць (Гбайт/міс) 20
коефіцієнт місцевості 2
Час найбільшого навантаження на день 7
Кількість активних абонентів у мережі 716716
Висота приймаючої антени в метрах 1.5
відстань від базової станції до приймаючої антени 0.63
Висота підйому антени базової станції 30
Радіус соти 770
Тип місцевості Місцевість з високою щільністю перешкод ▾
Корегуюча на затінення 9

а)

б)

Рис. 6. Вигляд працюючого розробленого програмного забезпечення (а – вікно введення вхідних даних для планування; б – розташування БС, як результат роботи програмного забезпечення)

Висновки

Проведений аналіз якості обслуговування абонентів стільникових мереж в Україні з метою визначення їх ефективності. Також було визначено вимоги до стільникових мереж нового покоління як в світі, так і в Україні зокрема. Це дозволило визначити напрями, за якими необхідно проводити підвищення ефективності стільникових мереж.

Тому було удосконалено метод планування мережі LTE за рахунок послідовного визначення стратегії розвитку оператора стільникового зв'язку, найбільш важливих показників якості обслуговування, виборі обладнання із множини доступних альтернатив, виборі послуг для абонентів, оцінці зон радіо покриття із врахуванням особливостей рельєфу та кліматичних умов, корегування розташування базових станцій.

Розроблений метод надає змогу операторам стільникового зв'язку проводити більш точну оцінку зон радіо покриття, вибір ключових показників якості обслуговування та розрахунок капітальних витрат, що дозволяє оцінити доцільність побудови варіанту мережі стільникового зв'язку. На основі даного удосконаленого методу розроблено навчально-інженерне програмне забезпечення (web-додаток) для оцінки зон радіо покриття базових станцій мережі LTE.

Література

1. Мобільний зв'язок в Україні [електронний ресурс] – електронні текстові дані – режим доступу: http://uateka.com/uk/article/society/1227/
2. Єрохін В.Ф., Гиндич Б.А., Кувшинов О. В. Аналіз і прогноз розвитку систем мобільного зв'язку загального користування / Збірник наукових праць ВІТІ НТУУ „КПІ" № 1, 2011. – С. 54 – 64.
3. Сети и Стандарты Мобильной Связи в Украине [електронний ресурс] – електронні текстові дані – режим доступу: http://blog.jammer.su/2012/07/seti-standarty-mobilnoj-svjazi-ukraina/
4. Первую тестовую сеть LTE в Украине построит МТС [електронний ресурс] – електронні текстові дані – режим доступу: http://podrobnosti.ua/667728-pervuju-testovuju-set-lte-v-ukraine-postroit-mts.html
5. Вишневский В. М. Энциклопедия WiMAX: Путь к 4G / В. М. Вишневский , С. Л. Портной, И. В. Шахнович – М.: Техносфера, 2009. – 472 с.
6. GSA Evolution to LTE report [електронний ресурс] – електронні текстові дані – режим доступу: http://www.gsacom.com/downloads/pdf/GSA_Evolution_to_LTE_report_060514.php4
7. Ericsson Mobility Report [електронний ресурс] – електронні текстові дані – режим доступу: http://www.ericsson.com/mobility-report
8. Understanding 5G [електронний ресурс] – електронні текстові дані – режим доступу: http://www.arnitsu.com
9. Перспективи та рекомендації по впровадженню стільникового зв'язку 4-го покоління / В. В. Ткаченко, Р. С. Одарченко, В. С. Повхліб, Т. Р. Андрійченко // Проблеми навігації та управління рухом : Всеукраїнська науково-практична конференція молодих учених і студентів; м. Київ, 21–22 листопада 2011 р.: тези доповідей / редкол. М. С. Кулик та ін. – К. : НАУ, 2011. – С. 122.

10. Тихвинский В. О. Сети мобильной связи LTE: технологии и архитектура / В. О. Тихвинский, С. В. Терентьев, А. Б. Юрчук. – М. : Эко-Трендз, 2010. – 284 с.

11. Ткаченко В. В. Вітчизняні перспективи розвитку технології LTE / В. В. Ткаченко, І. О. Дударчук, К. В. Дружиніна // Проблеми навігації та управління рухом : Всеукраїнська науково-практична конференція молодих учених і студентів; м. Київ, 23–24 листопада 2010 р. : тези доповідей / редкол. : М.С. Кулик та ін. – К.: НАУ, 2010. – С. 105.

12. В. Скрипин Глава НКРСИ: 4G в Украине появится не раньше 2016 года [електронний ресурс] – електронні текстові дані – режим доступу: http://itc.ua/news/glava-nkrsi-4g-v-ukraine-poyavitsya-ne-ranshe-2016-goda/

13. Київстар, МТС і Астеліт купили 3G-ліцензії на загальну суму 8,77 млрд грн. [електронний ресурс] – електронні текстові дані – режим доступу: http://www.business.ua/articles/media_view_ukr/Ki%D1%97vstar_MTS_%D1%96_Astel%D1%96t_kupili_Gl%D1%96cenz%D1%96%D1%97_na_zagalnu_sumu__mlrd_grn-89129/

14. Кінець епохи неоліту. Як 3G-інтернет змінить якість життя в Україні [електронний ресурс] – електронні текстові дані – режим доступу: http://news.finance.ua/ua/news/~/345778

15. Mashups: The new breed of Web app [електронний ресурс]– електронні текстові дані – режим доступу: http://www.ibm.com/developerworks/xml/library/x-mashups/index.html

16. «Київстар» будує 3G-мережу з можливістю розгорнути на ній LTE [електронний ресурс] – електронні текстові дані – режим доступу: http://www.kyivstar.ua/mk/press_center_new/news/?id=50144

17. Одарченко Р.С., Конахович Г.Ф., Ткаченко В.В. Методика вибору проектного рішення для розгортання захищеної мережі LTE // Захист інформації, 2014. – № 1 (16). – С. 63 – 68.

18. Одарченко Р.С. Стратегії розвитку операторів стільникового зв'язку в Україні // Наукоємні технології, 2016. – Вип. 2 (Том 26). – С. 141 – 148.

19. Одарченко Р.С. Обгрунтування основних вимог до систем безпеки стільникових мереж 5-го покоління // Безпека інформації, 2015. – Том 21, № 3. – С. 5 – 8.

20. Ткаліч О.П., Одарченко Р.С., Устинов О.Ю., Колодинський Д.О. Розрахунок зони покриття бездротової мережі Wi-Fi для визначення місцезнаходження абонентів в аеропорту // Проблеми інформатизації та управління, 2015. – Том 2, Вип. 50. – С. 117 – 122.

21. Ткаліч О.П., Одарченко Р.С., Устинов О.Ю., Колодинський Д.О. Оцінка адекватності моделей розповсюдження для їх використання під час визначення місцезнаходження абонентів // Наукоємні технології, 2015. – 2 (26). – С. 159 – 165.

22. Одарченко Р.С. Програмне забезпечення для попередньої оцінки вартості мережі LTE / Проблеми інформатизації та управління, 2015. – 3(51)' С. 1 – 6.

23. Solomentsev O., Zaliskyi M., Odarchenko R., Gnatyuk S. Research of energy characteristics of QAM modulation techniques for modern broadband radio systems // International Conference on Electronics and Information Technology, EIT 2016 - Conference Proceedings.

24. Р. С. Одарченко, А. О. Абакумова, Н. В. Дика Дослідження вимог до стільникових мереж нового покоління та можливості їх розгортання в Україні // Проблеми інформатизації та управління. Том 2, № 54 (2016). – С. 52 – 59.

25. Р. С. Одарченко, Н. В. Дика Дослідження архітектури мереж стільникового зв'язку в Україні та можливостей їх переходу до мереж LTE // Наукоємні технології № 3 (31), 2016. – С. 291 – 298.

ИССЛЕДОВАНИЕ СВОЙСТВ ПОМЕХОУСТОЙЧИВЫХ КОДОВ КЛАССА LDPC

Урывский Л.А., Осипчук С.А.

Введение

Актуальность. Помехоустойчивое кодирование является неотъемлемым элементом инфокоммуникационных технологий. Инструменты повышения помехоустойчивости информационных потоков всегда связаны с необходимостью привлечения дополнительных телекоммуникационных ресурсов: частотных, временных, или энергетических. Поэтому наибольший интерес представляют те классы помехоустойчивых кодов, которые предполагают наиболее рациональное использование телекоммуникационных ресурсов, предоставляемых для передачи информации. К таким кодам следует отнести коды с низкой плотность проверок на четность, LDPC [1-5]. Предложенные еще в 1960 году профессором Р. Галлагером [1], эти коды только через 50 лет нашли свое место в действующих инфокоммуникационных системах. Такой промежуток времени на реализацию кодов связан с тем, что поиск наиболее эффективных структур кодов в общем классе LDPC требует значительных вычислительных ресурсов, мощностей которых до некоторого времени было недостаточно. Вместе с тем, коды LDPC, будучи не аналитическими, а эвристическими, дают существенный выигрыш по сравнению с известными кодами Боуза-Чоудхури-Хоквингема (БЧХ) [6] и Рида-Соломона (РС) в скорости обработки в случае соизмеримых длин кодов и скоростей кодирования. Этот же фактор обуславливает существование кодов LDPC с длиной блока в десятки раз больше длин аналитических кодов, и, как следствие, с гораздо большими значениями скорости кодирования при эквивалентной исправляющей способности.

Актуальность исследования заключается в определении возможности исправления ошибок сверхдлинными LDPC кодами, которые применяются сегодня в современных телекоммуникационных стандартах, и определение места этих кодов среди известных блочных кодов путем сравнения их характеристик.

Цель представленных материалов заключается в раскрытии содержания процедуры поиска кода LDPC с наиболее эффективной структурой с точки зрения его корректирующих свойств, а также раскрытия алгоритмов формирования и декодирования информационных последовательностей с использованием синтезированных LDPC кодов.

В работе решены следующие *задачи*:

– обоснование и выбор критериев оценки исправляющей способности помехоустойчивых кодов;

– исследование свойств помехоустойчивых кодов класса LDPC;

– аналитическое описание исправляющей способности LDPC

кодов;

– сравнение эффективности LDPC и БЧХ кодов по показателю соотношения длины кода и требуемой скорости кодирования.

1. Обоснование и выбор критериев оценки исправляющей способности помехоустойчивых кодов

Одним из методов улучшения помехоустойчивости сигналов при работе в одном и том же диапазоне частот является помехоустойчивое кодирование, что позволяет улучшать помехоустойчивость системы передачи ценой перераспределения временного ресурса, а именно – снижением скорости передачи информации путем введения избыточной информации в передаваемое сообщение.

В общем случае помехоустойчивое кодирование можно описать такими параметрами, как число гарантированно исправляемых битовых ошибок t на кодовое слово, состоящее из n бит, где n = k + r, k – число информационных бит в кодовом слове, r – число избыточных (проверочных) бит кодового слова, а также скорость помехоустойчивого кодирования $r_k = k / n$.

Предельные возможности блочных кодов с длиной блока n бит, выраженные через максимально достижимую скорость кодирования r_k при заданном значении кодового расстояния d для данного кода, определяются границей Плоткина. Необходимые условия для существования помехоустойчивых кодов с заданными корректирующими способностями определяет граница Варшамова-Гильберта [7].

Граница Плоткина определяется следующими условиями: если длина кодового блока $n \geq 2d-1$, то число проверочных символов $r = n - k$, необходимых для того, чтобы минимальное расстояние линейного кода достигало значения d, равно наименее $2d - 2 - \log_2 d$. Как следствие,

$$r_k \leq 1 - \frac{2d - 2 - \log_2 d}{n}. \quad (1)$$

Другая форма записи для границы Плоткина:

$$k \leq n(2d - 2 - \log_2 n). \quad (2)$$

Граница Плоткина чаще применяется для кодов с большими значениями длины кодового слова n >> 1.

Достаточные условия для существования кодов с заданными корректирующими способностями определяет граница Варшамова-Гильберта, которая сводится к утверждению того, что существует (n,k)-код, с минимальным расстоянием, по наименьшей мере, d, который удовлетворяет следующему неравенству:

$$\sum_{i=0}^{d-2} C_{n-1}^{i} \geq 2^{n-k}. \quad (3)$$

Таким образом, критерии Плоткина и Варшамова-Гильберта

позволяют сравнивать разные блочные коды в одних координатах $r_k = f(d/2n)$ для оценки корректирующих способностей кодов.

2. Исследование свойств помехоустойчивых кодов класса LDPC

Исследуемые LDPC коды представлены моделью в стандартизированном программном обеспечении MatLab. Длина кодового слова сверхдлинного LDPC кода составляет 64800 бит и код может иметь скорости в диапазоне $r_k = 0{,}25…0{,}9$, т.е. как в стандарте DVB-S2 [8].

Исследование сверхдлинных LDPC кодов проведено на предмет потенциальной возможности исправления битовых ошибок в кодовом слове, которое поступает на декодер приемной стороны при передаче информации в современных телекоммуникационных системах. Исходными условиями являются:

– скорость кода $r_k = 0{,}25…0{,}9$ согласно стандарта для DVB-S2;

– проверочная матрица H для каждой скорости кодирования LDPC кода;

– фиксированная длина кодового слова $n = 64800$.

Для достижения поставленной цели поставлены следующие задачи:

– воспроизвести исходную информационную последовательность и соответствующий LDPC-код в программной среде MatLab с последующим введением ошибочных символов в закодированную последовательность, имитируя действие помех, для возможности проведения численных экспериментов над поступающим в декодер кодовым словом с заданным числом ошибочных бит;

– на основании статистических численных экспериментов произвести анализ возможности исправления битовых ошибок в кодовых словах LDPC кода со скоростью кодирования $r_k = 0{,}25…0{,}9$ и разной плотностью ошибок (частотой следования ошибок в кодовой последовательности);

– определить степень соответствия сверхдлинных LDPC кодов условиям достижения границ Плоткина и Варшамова-Гильберта;

– провести сравнение характеристик известных блочных кодов с большой длиной кодового слова ($n>1000$) и сверхдлинных LDPC кодов.

Проведенный численный эксперимент основан на имитации введения определенного числа битовых ошибок в принимаемое кодовое слово, с некоторой управляемой частотой (плотностью) ошибок и проведении необходимого числа итераций для декодирования переданной информации согласно методу жесткого декодирования. Из-за больших размеров реальных проверочных матриц (в работе рассмотрены коды, сгенерированные с помощью матриц размером $64800 \cdot 64800 \cdot r_k$), нужно выполнять огромное число вычислений, и сложность прямых методов декодирования высока. В связи с этим используются итерационные методы декодирования, такие как метод жесткого декодирования [9].

На рис. 1 представлен алгоритм численного эксперимента. Согласно описанному алгоритму на рис. 1, приведена зависимость числа исправленных ошибок C_{err} от числа ошибок в кодовом слове N_{err} на рис. 2 для кода со скоростью кодирования r_k, число итераций декодирования 10, 30 и 50. Важно также отметить, что на рис. 2-3 представлены только для LDPC кода со скоростью кодирования $r_k=1/2$. На первом этапе в принятое кодовое слово последовательно внесено 1000, 2000, …, 6000 рассредоточенных по длине кодового слова битовых ошибок, и определено число исправленных ошибок путем сравнения декодированных информационных бит с передаваемой информационной последовательностью. Как показано на рис. 1, если $C_{err}=N_{err}$, то информация абсолютно восстановлена и является идентичной передаваемым информационным битам. В случае $C_{err}>N_{err}>0$ – информация частично восстановлена. Если же $C_{err}<0$ – информация не восстановлена, и при декодировании внесены еще и дополнительные ошибки, кроме принятых ошибок.

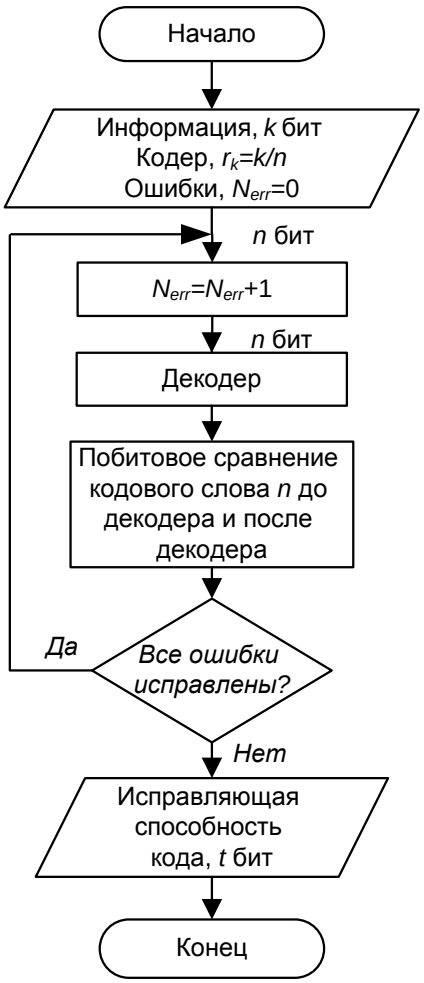


Рис. 1. Алгоритм численного эксперимента

На рис. 2 штрихпунктирная линия – число итераций декодирования 10; сплошная и пунктирная линии – число итераций соответственно 30 и

50. Как видно с рис. 2, линии для 30 и 50 итераций декодирования практически совпадают.

Как показали эксперименты, декодер значительно хуже декодирует непрерывные блоки ошибок по сравнению с рассредоточенными битовыми ошибками при одинаковом числе итераций для одного и того же кода.

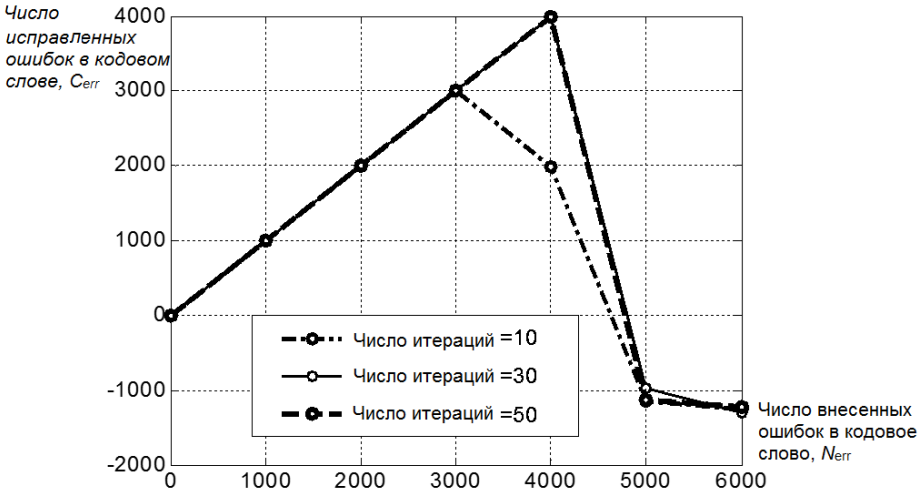


Рис. 2. Число исправленных ошибок C_{err} *от числа внесенных ошибок в кодовое слово* N_{err}

Анализ числа итераций декодирования до 100 показал, что увеличение итераций декодирования выше 50 практически не улучшает результат декодирования. С точки зрения рационального использования вычислительных ресурсов, целесообразно использовать 10…50 итераций для декодирования. Поэтому в численных экспериментах выбрано и использовано число итераций декодирования 30, 40 и 50 по критериям оптимального использования вычислительного ресурса и достоверности принятой информации.

Из полученного результата рис. 2 видно, что LDPC декодер исправляет от 4000 до 5000 битовых ошибок в кодовом слове длиной 64800 бит при скорости кодирования r_k=1/2.

С рис. 3 видно, что код исправляет от 4800 до 4810 ошибок в принятом кодовом слове. Машинное время, затраченное на процесс декодирования, растет пропорционально с увеличением числа итераций. Следовательно, 30 итераций достаточно для декодирования информации для экономии машинного ресурса и получения практически такой же достоверности декодирования. Таким образом, для LDPC кода с длиной блока n=64800 бит при скорости кодирования r_k=1/2 с рассредоточенными битовыми ошибками, числом итераций декодирования 50, число исправляемых битовых ошибок равно 4804 бит. Важно отметить, что разрежение плотности ошибок положительно сказывается на увеличении числа исправленных ошибок при тех же условиях.

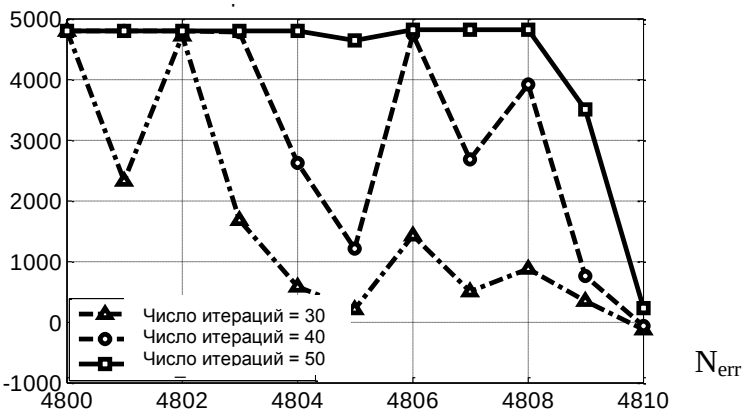


Рис. 3. Число исправленных ошибок C_{err} *от числа внесенных ошибок в кодовое слово* N_{err} *при* N_{err}=4800,4801,…,4810

Таким образом, пример исследования, представленный выше, показал, что увеличение числа итераций декодирования для исправления большего числа битовых ошибок все же дает незначительный положительный результат.

3. Аналитическое описание исправляющей способности LDPC кодов

На основании результатов экспериментальных исследований установлены статистические зависимости максимального числа исправляемых бит t в кодовом слове LDPC кода от двух параметров: длины кодового слова n и скорости помехоустойчивого кодирования r_k. Для полученных результатов предложены аналитические выражения для случая линейной, экспоненциальной и полиномиальной аппроксимаций. Обоснована наиболее точная аналитическая формула по критерию минимального расхождения с экспериментальными результатами.

LDPC коды являются линейными блочными кодами с циклическими процедурами декодирования. В связи с особенностями формирования и декодирования LDPC кодов, в отличие от других известных помехоустойчивых кодов, например, БЧХ, для LDPC кодов не существует определенной точной аналитической модели для определения их исправляющих способностей t, бит как функция длины кодового слова n и скорости кодирования r_k. Таким образом, существует задача определения зависимости исправляющих способностей t, бит как функция длины кодового слова n и скорости кодирования r_k LDPC кода, путем анализа множественных экспериментов с проверочными матрицами LDPC кодов и определения исправляющей способности LDPC кода. Получение такой зависимости позволяет указывать число исправляемых бит на длине кодового слова с заданной скоростью кодирования без поиска проверочной матрицы H с наилучшими параметрами. Многие работы в области LDPC кодирования посвящены экспериментальному исследованию

исправляющей способности LDPC кодов, но в то же время не сделана попытка обобщить полученные результаты в аналитическое выражение. Представленные результаты являются развитием положений и посвящены аналитическому описанию исправляющих способностей регулярных LDPC кодов, а именно – поиску аналитического выражения для определения числа исправляемых бит t в кодовом слове LDPC кода как функции двух параметров: длины кодового слова n и скорости помехоустойчивого кодирования r_k.

Задача заключается в поиске аналитической зависимости исправляющей способности LDPC кода t с определенной длиной кодового слова n и скоростью кодирования r_k на основе проведения множественных численных экспериментов с проверочными матрицами LDPC кода.

Входные данные: наборы проверочных матриц H, сгенерированных случайно, с параметрами: n, W_r, W_c.

Промежуточные данные: исправляющая способность t каждой проверочной матрицы H.

Выходные данные: аналитические зависимости исправляющей способности LDPC кодов t бит на длину кодового слова n как функция длины кодового слова n и скорости кода r_k:

$$t_{бит} = f\,(n,r_k). \qquad (4)$$

На основании множественных экспериментов с генерированием случайных проверочных матриц LDPC кода H, наилучшие характеристики полученных LDPC кодов представлены в координатах $r_k = f\,(d/2n)$ (рис. 4).

Под наилучшими характеристиками имеется в виду максимальное число исправляемых ошибок t на длину кодового слова n. В данной работе исследованы проверочные матрицы для длины кода n: 50, 100, 200, 500, 1000. На рис. 4 сплошной линией обозначена линия Плоткина, которая показывает границу существования блочных кодов в координатах $r_k=f(d/2n)$. Код не может существовать выше этой границы, и чем ближе точка кода находится к этой границе, тем лучше код по таким критериям как скорость кодирования и максимальное число исправляемых битовых ошибок на длину кодового слова. Построенная граница Плоткина на рис. 4 справедлива для больших длин кодового слова $n>>1$, порядка сотен и тысяч бит на кодовое слово.

Для LDPC кода $d = 2t+2$. Разделив обе части на 2n, получим

$$\frac{d}{2n} = \frac{2t+2}{2n} = \frac{t+1}{n}, \text{ откуда } r_k = f\,(\,(t+1)/n\,) \text{ и } t = f\,(n,r_k).$$

Полиномиальная аппроксимация. Каждой представленной на рис. 4 линии $r_k=f(d/2n)$ для разных n можно поставить в соответствие линию, описывающую полиномиальную аппроксимацию значений $r_k = f\,((t+1)/n)$: $y = k_1x^2 + k_2x + k_3$, где k_1, k_2, k_3 – коэффициенты функции (рис. 5), $x = (t+1)/n$ – аргумент. В табл. 1 представлены коэффициенты функций

линейной аппроксимации для каждого значения n.

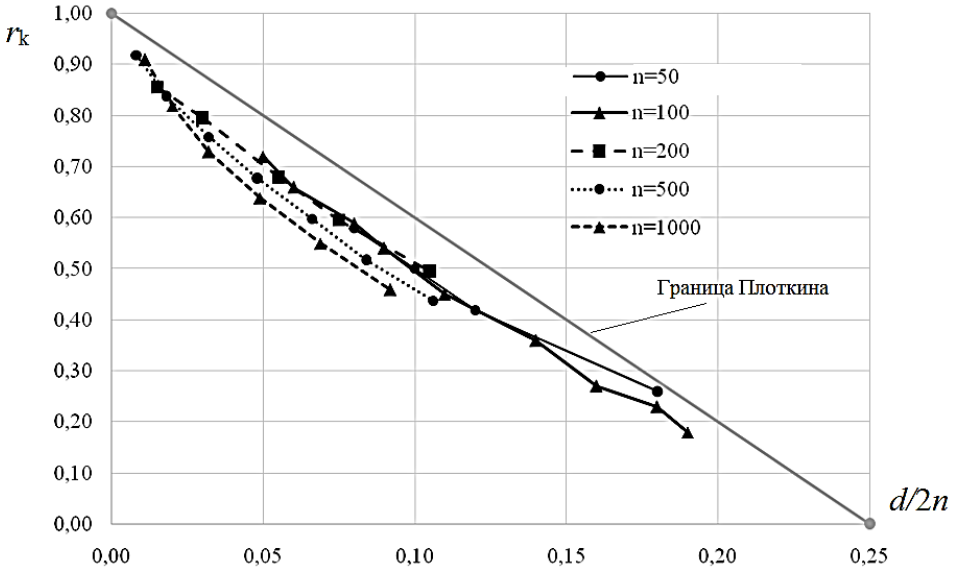


Рис. 4. Экспериментально полученные зависимости для LDPC кодов $r_k=f(d/2n)$

Таким образом,

$$r_k = k_1\left(\frac{t+1}{n}\right)^2 + k_2\left(\frac{t+1}{n}\right) + k_3, \quad (5)$$

$$t = \frac{-(2k_1 + nk_2) - \sqrt{(2k_1 + nk_2)^2 - 4k_1(k_1 + nk_2 + n^2k_3 - n^2r_k}}{2k_1}, \quad (6)$$

$$k_1 = 11{,}887e^{0{,}0011n},\ k_2 = -0{,}719\ln n - 3{,}4317, \quad (7)$$

$$k_3 = -0{,}022\ln n + 1{,}1336.$$

Таблица 1

Коэффициенты функций полиномиальной аппроксимации

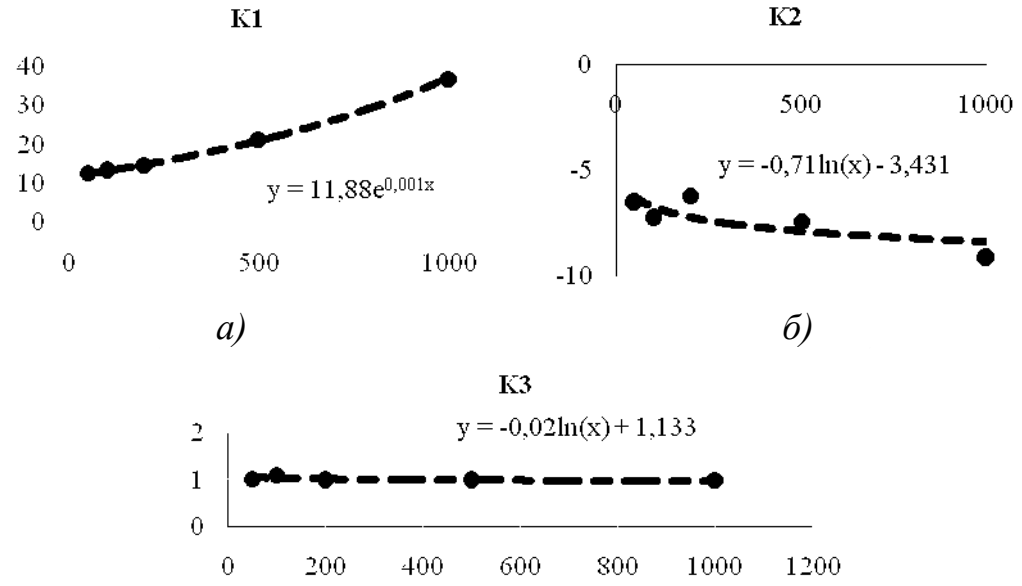
n	**k_1**	**k_2**	**k_3**
50	12.707	-6.5193	1.0215
100	1199	-7.2674	1.0983
200	14.746	-6.1859	0.98
500	21.377	-7.4009	0.9861
1000	36.906	-9.1433	0.9948

Линейная аппроксимация. Аналогично методике, представленной для полиномиальной аппроксимации, выражение для линейной аппроксимации $t = f(n,r_k)$ принимает вид:

$$t = \frac{n(r_k - 0{,}0342\ln n - 0{,}7101)}{-0{,}731\ln n - 0{,}2957}. \quad (8)$$

Экспоненциальная аппроксимация. Аналогично методике, представленной для полиномиальной аппроксимации, выражение для экспоненциальной аппроксимации $t = f(n,r_k)$ принимает вид:

$$t = \frac{n}{0{,}0004n - 8{,}2657} \ln \frac{r_k}{1{,}5354 - 0{,}085 \ln n} - 1. \qquad (9)$$



а) *б)*

в)

Рис. 5. Аппроксимация коэффициентов k_1, k_2, k_3 для полиномиальной функции $r_k = k_1\left(\frac{t+1}{n}\right)^2 + k_2\left(\frac{t+1}{n}\right) + k_3$

Сравнение значений исправляющей способности *t*, рассчитываемых на основе полученных аналитических выражений, и экспериментальных значений исправляющей способности LDPC кода на основе имитационного моделирования, показывает следующие ***результаты***:

– аналитическое выражение $t = f(n, r_k)$ для экспоненциальной аппроксимации наиболее точно описывает зависимость $t = f(n, r_k)$ из трех полученных формул. Отклонение от экспериментального значения исправляющей способности может находиться в пределах $\Delta t_{ош} = |t_{эксп} - t_{аналит}| = 0 \ldots 3$ бит;

– аналитическое выражение $t=f(n, r_k)$ для полиномиальной аппроксимации также достаточно точно описывает зависимость $t=f(n, r_k)$, но в отдельных единичных точках отклонение от экспериментального результата превышает отклонение, полученное при экспоненциальной аппроксимации и может достигать $\Delta t_{ош} = |t_{эксп} - t_{аналит}| = 3 \ldots 6$ бит;

– аналитическое выражение $t = f(n, r_k)$ для линейной аппроксимации наименее точно согласуется с экспериментальными результатами по критерию соответствия рассчитанных значений исправляющей способности t и полученных экспериментально. В отдельных случаях расходимость с экспериментальными результатами достигает

$\Delta t_{ош} = |t_{эксп} - t_{аналит}| = 10$ бит при большой длине кодового слова (n=1000) и при большой скорости кодирования, что является ожидаемым результатом в связи с очевидно нелинейной экспериментальной характеристикой $r_k = f(d/2n)$.

4. Сравнение эффективности LDPC и БЧХ кодов по показателю соотношения длины кода и требуемой скорости кодирования

Проведен качественный анализ, а также дана количественная оценка способностей исправления битовых ошибок LDPC кодами с длиной кодового слова n=1000 и БЧХ кодами с длиной блока n=1023 бит. Определены скорости кодирования LDPC и БЧХ кодов при известном отношении сигнал/шум в канале с белым шумом, при которых использование данных кодов оптимально с заданным видом модуляции для удовлетворения требуемой достоверности приема информации.

Известно, что одними из наилучших блочных кодов по критерию удовлетворения условия Варшамова-Гильберта являются коды БЧХ с большой длиной блока (n>1000) [6]. В табл. 2 представлены параметры некоторых кодов БЧХ с длиной кодового слова n=1000.

Целью исследования является сравнение кодов LDPC и БЧХ с одинаковой длиной блока в случае применения помехоустойчивого кодирования с заданным методом модуляции для достижения требуемой вероятности битовой ошибки при известных параметрах канала.

Входными известными данными для задачи являются:

– параметры канала: отношение сигнал/шум, 0…14 дБ;

– сигнальная конструкция: метод модуляции, QPSK;

– длина кодового слова для помехоустойчивого кодирования: *n*=1000 для LDPC и *n*=1023 для БЧХ;

– требование конечного пользователя к достоверности принимаемой информации: 10^{-6}.

Выходными данными являются: скорость помехоустойчивого кодирования для кодов LDPC и БЧХ.

Таблица 2

Известные параметры БЧХ кода с длиной кодового слова n=1023

k	t	d	r_k=k/n	d/2n
1013	1	3	0.990225	0.001466
973	5	11	0.951124	0.005376
923	10	21	0.902248	0.010264
. . .				
101	175	351	0.098729	0.171554
56	191	383	0.054741	0.187195
11	255	511	0.010753	0.249756

Требуется: при известных указанных входных параметрах, определить скорости помехоустойчивого кодирования $r_{k\text{-}LDPC}$ для LDPC и $r_{k\text{-}БЧХ}$ БЧХ кодов, при которых будет достигнута требуемая достоверность информации на приемной стороне.

Итак, основной задачей является поиск помехоустойчивого кода с наибольшими значениями скорости кодирования r_k и наибольшими значениями кодового расстояния d, что относится к основной проблеме теории кодирования. Для достижения цели определены следующие этапы решения поставленной задачи:

– разработка и реализация процедуры поиска минимального кодового расстояния LDPC кода с заданной длиной кодового слова и параметрами проверочной матрицы;

– определение положения LDPC кодов и БЧХ кодов в координатах $r_k=f(d/2n)$;

– определение максимальной скорости помехоустойчивого кодирования, при которой можно обеспечить требуемую достоверность приема информации.

БЧХ коды характеризуются возможностью построения кода с заранее определёнными корректирующими свойствами, а именно – минимальным кодовым расстоянием. Для любых значений m и t существует двоичный код БЧХ длины $2^m–1$, исправляющий все комбинации из t или меньше ошибок и содержащий не более чем mt проверочных символов. Таким образом, длина БЧХ кода не может быть выбрана произвольным образом и зависит от параметра m; длина кодового слова БЧХ кода всегда имеет нечетное значение. Как следует из примера, БЧХ код с длиной блока n=1023 можно формировать с шагом скорости кодирования 0,01.

При этом скорость кодирования убывает с ростом исправляющей способности кода по линейному закону:

$$r_k = 1 – 0{,}009 / t, \text{ или} \qquad (10)$$

$$t = \frac{1 – r_k}{0{,}0098}. \qquad (11)$$

Погрешность соотношений (10), (11) не превышает 2,2%.

Регулярные LDPC коды чаще показывают лучшие характеристики, чем нерегулярные LDPC коды [10-11]. В гауссовом канале лучшие характеристики проявляют именно регулярные LDPC коды. Вместе с этим, существуют условия, при которых лучшие характеристики имеют нерегулярные LDPC коды [12]. Таким образом, как регулярные, так и нерегулярные LDPC коды, имеют право на существование в теории и практике помехоустойчивого кодирования.

Алгоритм поиска кодового расстояния регулярного LDPC кода реализован на языке Java и результаты кодового расстояния LDPC кода с длиной блока n=1000 получены после численных экспериментов на

кластере суперкомпьютерных вычислений НТУУ «КПИ». Значения параметров проверочной матрицы H и кодового расстояния *d* представлены в табл. 3.

Таблица 3

Полученные параметры LDPC кодов

n	Wr	Wc	k	r_k	d	t
1000	100	10	909	0.91	22	10
		20	819	0.82	40	19
		30	729	0.73	64	31
		40	639	0.64	98	48
		50	549	0.55	138	68

Цель исследования достигается на основе определения скорости помехоустойчивого кодирования r_{k_LDPC} для LDPC и $r_{k\_БЧХ}$ для БЧХ кодов, при которых обеспечивается требуемая достоверность информации на приемной стороне при известных входных параметрах: отношения сигнал/шум h^2, метод модуляции, длина кодового слова помехоустойчивого кода, требуемая вероятность битовой ошибки на приемной стороне. Сравнив полученные значения скоростей кодирования, необходимо определить, какой метод помехоустойчивого кодирования предпочтителен по критерию r_{k_max}.

На рис. 6 построены точки положения LDPC кодов и БЧХ кодов в координатах $r_k=f(d/2n)$. В отличие от границы Плоткина, граница ВГ означает, что всегда существует код, который имеет значение скорости кодирования, которое находится на соответствующей кривой, а возможно, и выше этой кривой. Как показано на рис. 6, LDPC код с длиной блока 64800 удовлетворяет этому утверждению и лежит выше границы ВГ. Это характеризирует LDPC код с длиной кодового слова 64800 как наилучший реализуемый код, поскольку его параметры находятся между границами Плоткина и ВГ [13-14]. Из рис. 6 видно, что LDPC код характеризуется значительно лучшим исправлением ошибок, чем БЧХ код [15-17]. В первую очередь, это обосновано намного большим значением длины кодового слова.

Из рис. 6 видно, что LDPC коды находятся несколько ближе к границе Шеннона, чем БЧХ коды. Это свойство проявляется в большей степени с уменьшением скорости помехоустойчивого кода при $r_k<0,7$. В этой области LDPC код предпочтительнее БЧХ кода.

Неоспоримым преимуществом LDPC кода является возможность наращивать длину кодового слова до десятков тысяч бит.

Это поясняется относительно простыми методами кодирования и декодирования информации. Вместе с этим, преимуществом БЧХ кода является возможность аналитически определять и выбирать параметры кода (например, скорость кодирования с точностью до 0,01) для

удовлетворения требований к его исправляющей способности.

LDPC коды имеют лучшую относительную исправляющую способность, а именно – бит на длину блока, по сравнению с линейными блочными кодами БЧХ, до скорости кодирования 0,84, а выше скорости кодирования 0,84 БЧХ коды имеют незначительно большее относительное значение исправленных ошибок на длину кодового слова.

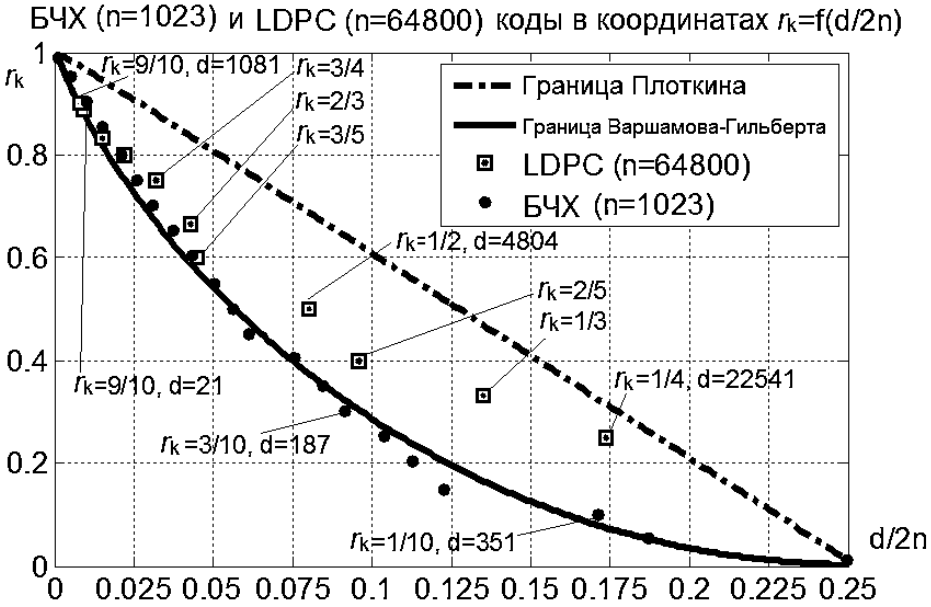


Рис. 6. Положение LDPC кодов и кодов БЧХ в координатах r_k=f(d/2n)

Выводы

1. Исследование кодов LPDC состоит из двух принципиально различающихся этапов.

Первый этап связан с синтезом оптимальных LDPC структур при заданной длине блока кода, выбираемой на основе стохастического алгоритма, как правило, более чем из миллиона возможных комбинаций. Итогом первого этапа являются аналитические зависимости, связывающие длину LDPC кода с наибольшей достижимой скоростью кодирования при заданной исправляющей способности кода, представленной в виде линий в пространстве между границей Плоткина и границей Варшамова-Гильберта.

Для применения LDPC кодирования на практике, сначала необходимо найти или иметь ранее найденную проверочную матрицу. Эта матрица должна обладать известными исправляющими способностями, и храниться в памяти кодера и декодера.

Второй этап сводится к исследованию помехоустойчивости кодированных последовательностей с помощью синтезированных LDPC кодов с известными параметрами. При этом отдельно исследуются LDPC коды с длиной блока до 2000 символов, и сверхдлинные LDPC кодов с длиной блока более 60000 символов.

Итогом второго этапа исследования является получение нового научного результата, связанного с тем, что синтезированные параметры LDPC кодов обеспечивают большую скорость кодирования (лучшую информационную эффективность), чем известные коды БЧХ при длине блока кода до 1000 символов на 12-20%, с учетом того, что ни один из известных кодов не может быть реализован как сверхдлинный по показателям быстродействия схем обработки.

2. На основании статистических численных экспериментов произведен анализ возможностей исправления битовых ошибок в кодовых словах сверхдлинных LDPC кодов со скоростью кодирования r_k=0,25…0,9 и разной плотностью ошибок. На основании численного эксперимента показано, что сверхдлинные LDPC коды лучше самых длинных БЧХ кодов по критерию достижения потенциальных границ исправляющих способностей блочных кодов (достижения границы Плоткина).

3. Как показали эксперименты, декодер значительно хуже декодирует непрерывные блоки ошибок по сравнению с рассредоточенными битовыми ошибками при одинаковом числе итераций для одного и того же кода. При наличии одного и того же числа ошибок в кодовом слове разрежение плотности ошибок положительно сказывается на увеличении числа исправленных ошибок при неизменных энергетических условиях в канале связи, что свидетельствует о целесообразности использования на передающей стороне устройств связи схем перемежения как ступени, дополняющей LDPC кодирование.

4. Экспериментальным путем получены показатели исправляющей способности LDPC кодов с длиной кодового слова n = 50, …, 1000 в виде числа исправляемых кодом ошибок t в зависимости от двух параметров кода: длины кодового слова n и скорости кодирования r_k.

Предложены аналитические соотношения между числом исправляемых ошибок t в блоке из n символов и скоростью кодирования r_k для характеристики исправляющей способности регулярных LDPC кодов.

Приведены три варианта аналитических выражений для аппроксимации исправляющей способности LDPC кода: линейная, экспоненциальная и полиномиальная. По результатам сравнения значений исправляющей способности LDPC кодов, полученных экспериментально и вычисленных аналитически по трем выведенным зависимостям, можно рекомендовать *экспоненциальную аппроксимацию* как наиболее точную для расчета исправляющей способности LDPC кода по имеющимся параметрам: длина кода и скорость кодирования.

Литература

1. Галлагер Р. Коды с малой плотностью проверок на четность / Р. Галлагер. – М.: Мир, 1966. – 144 с.

2. MacKay D. Good error-correcting codes based on very sparse matrices // IEEE Trans. Inf. Theory, vol. 45, no. 2. – 1999. – pp. 339 – 431.

3.MacKay D., Neal R. Near Shannon limit performance of low density parity check codes // Electron. Lett., 1996. – vol. 32, no. 18. – pp. 1645 ─ 1646.

4. Burshtein D., Krivelevich M., Litsyn S., Miller G. Upper bounds on the rate of LDPC codes // IEEE Trans. Inform. Theory, 2002. – 48, no. 9. – pp. 2437 ─ 2449.

5.Ohtsuki T. LDPC codes in communications and broadcasting // IEIC Trans. Commun., 2007. – vol. 90-B, no. 3. – pp. 440 ─ 453.

6.Кларк Дж. Кодирование с исправлением ошибок в системах цифровой связи / Дж. Кларк, Дж. Кейн. – М.: Радио и Связь, 1987. – 195 с.

7.Питерсон У. Коды, исправляющие ошибки / У. Питерсон, Э. Уэлдон. – М.: Мир, 1976. – 594 с.

8.Digital Video Broadcasting (DVB): Second generation framing structure, channel coding and modulation systems (DVB-S2) / ETSI EN 302 307 V1.2.1 // European Standard (Telecommunications series), 2009. – 78 p.

9.Ryan W.E. An Introduction to LDPC Codes, in CRC Handbook for Coding and Signal Processing for Recoding Systems / W.E. Ryan. – CRC Press. – 2004.

10.Luby M. Improved low-density parity-check codes using irregular graphs and belief propagation / Luby M., Mitzenmacher M., Shokrollahi A., Spielman D. – SRC Technical Note, 1998. – 9 p.

11.Miyamoto S. Sufficient conditions for a regular LDPC code better than an irregular LDPC code // Miyamoto S., Kasai K., Sakaniva K. – IEICE Trans. Fundamentals, 2007. – vol. E90–A, no. 2 – pp. 531 – 534.

12.Tian T. Construction of irregular LDPC codes with low error floors // Tian T., Jones C. – Communications, ICC ’03, IEEE International Conference, 2003. – vol. 5. – pp. 3125 – 3129.

13.Hou J. Performances Analysis and Code Optimization of Low-Density Parity-Check Codes on Rayleigh fading // J. Hou, P. H. Siegel, L. B. Milstein. – IEEE J. Selec. Areas in Comm., 2001. – vol. 19, no. 5. – pp. 924 – 934.

14. Uryvsky L. Code Rate of LDPC and Traditional Antinoise Codes Comparison // Uryvsky L., Osypchuk S. – Восьма МНТК «Проблеми телекомунікацій», ІТС НТУУ «КПІ», м.Київ, 2014. – с. 508 ─ 510.

15.Урывский Л.А. Методика оценки исправляющей способности сверхдлинных LDPC кодов // Л. А. Урывский, С. А. Осипчук. – Третя міжнародна науково-практична конференція (МНПК) молодих вчених «Інфокомунікації – сучасність та майбутнє», м. Одеса, 17-18 жовтня 2013 року, збірник тез, частина 3. – с. 91 ─ 95.

16.Урывский Л.А. Сравнение скорости кодирования LDPC и БЧХ кодов // Урывский Л. А., Осипчук С.А. – Материалы III МНПК «Академическая наука – проблемы и достижения», 20-21 февраля 2014 г., г. Москва. – с.203 ─ 205.

17.Uryvsky L. Comparative analysis of LDPC and БЧХ codes error-correcting capabilities / Uryvsky L., Osypchuk S. – Information and Telecommunication Sciences, Volume 5, Number 1, 2014. – pp. 5 ─ 9.

КИБЕР-СОЦИАЛЬНЫЙ КОМПЬЮТИНГ

Хаханов В.И., Соклакова Т.И., Чумаченко С.В., Литвинова Е.И.

Введение

Мотивация предлагаемых исследований определяется желанием многих конструктивных граждан сделать государство нравственно цивилизованным, технологически кибер-культурным, инвестиционно привлекательным и экономически эффективным. Для этого рассматриваются кибер-системные компоненты, существенные для эволюционного создания привлекательной государственности [1-5]: 1) структура кибер-социального компьютинга; 2) метрика социальной значимости; 3) государственность и коррупция – две стороны одной медали; 4) объединение и пересечение интересов в управлении социумом; метрика интеллекта социальной группы; 5) гармонический геном развития сообщества;. 6) кибер социальный компьютинг – нравственное будущее человечества.

Цель данной публикации – показать технологии совершенствования кибер-управления сообществом на основе использования облачных сервисов и точного цифрового мониторинга мнения каждого гражданина.

Задачи: 1) состояние управления социальными группами в развивающихся государствах; 2) технологии, модели и методы кибер-социального компьютинга; 3) кибер культура, измерение интеллекта социальной группы и кибер-социальное управление; 4) метрические кибер-отношения и будущее, как кибер-государственность.

Структура кибер-социального компьютинга. Современная киберкультура накладывает жесткие требования к структуре кибер-социального компьютинга, которая должна включать следующие компоненты по убыванию их значимости: 1) оцифрованные горизонтальные и вертикальные отношения в социальной группе, определяемые законодательством, культурой, историей и традициями; 2) облачное управление процессами и явлениями на основе точного цифрового мониторинга; 3) кадры, которые за счет справедливых отношений и компетентного управления, становятся производительной силой, создающей продукцию, сервисы и финансовый успех; 4) электронная инфраструктура социальной группы, предоставляющая комфортные условия для творческого, производительного труда и отдыха коллектива в формате 24/7; 5) направление движения или roadmap, которое формирует стратегические рыночно-ориентированные цели и задачи, связанные с качеством продукции, жизни сотрудников и сохранением экосистемы; 6) финансовые ресурсы, метрическое кибер-управление которыми создает глобально успешную компанию или быстрое банкротство на сегменте рынка; 7) производственные процессы, как основа

кибер-социального компьютинга, имеющая целью создание качественных продуктов и сервисов для продажи на рынке.

Семь пунктов формируют качество социальной системы, предприятия или государства. Компания Ernst&Young провела международное исследование уровня коррупции 2017 в сфере бизнеса. Украина заняла почетное первое место с результатом 88%. [6]. Чего не хватает стране, чтобы стать успешной и победить коррупцию. Системно в государстве отсутствуют всего лишь два первых пункта: нравственные отношения и компетентное управление. Их создает политическая элита (500 чиновников), которая должна иметь знания об отношениях и управлении в развитых странах: США, Германия, Англия.

Метрика социальной значимости. Технологическое совершенство следует обращать в социальную значимость для получения заслуженного морального и материального вознаграждения: 1) нравственность действий и поступков; 2) уважение культурных, языковых, исторических ценностей и традиций всех народов; 3) образованность, компетентность и знание специальных технологий; 4) выдающиеся результаты при выполнении служебных обязанностей; 5) заслуги перед социумом за волонтерскую общественно полезную деятельность.

Государственность и коррупция – две стороны одной медали. Государственность предполагает отчуждение части средств граждан через налогообложение для последующего предоставления им сервисов, связанных с инфраструктурным обслуживанием, охраной здоровья, защитой частной собственности, чести и достоинства граждан. Налоги граждан не доходят до поставщиков упомянутых сервисов. Они рассасываются в кабинетах государственных чиновников. В этом случае производители сервисов для населения создают частные компании, полиции, войска, клиники, суды, прокуратуры, которые продают уже за вторые деньги (первыми были налоги) от населения, необходимые для жизнедеятельности сервисы. Население имеет по факту двойное налогообложение. Выход тривиальный: создавать облачные кибер сервисы для распределения бюджетных средств, минуя государственных чиновников. Другой выход: уменьшать налогообложение граждан и численность государственных структур до нуля за счет создания частных коммерческих организаций. Коррупция существует до тех пор, пока налоги от граждан будут аккумулироваться на счетах привилегированной корпорации, которая называется государством. При этом всегда найдутся покупатели и чиновники, которые будут за деньги продавать не принадлежащий им товар – государственные финансы и ресурсы.

Объединение и пересечение интересов в управлении социумом. Две примитивных логических операции. Первая объединяет людей путем учета интересов, культур, историй, языков, традиций, которые ставятся в равные условия, независимо от количества носителей указанных понятий.

Государство, управляемое политической элитой с доктриной объединения, становится привлекательным и конкурентоспособным на рынке. На каждом углу и в сознании каждого гражданина США написано – “United we stand”. Результат такого позитивного зомбирования – социальное, экономическое и духовное процветание страны. Вторая операция расчленяет сообщество на отдельные и креативно слабые социальные группы, ненавидящие друг друга, за счет зомбирующей, по сути нацистской, пропаганды неравных отношений власти к языкам, историям, культурам, традициям. Такое управление со стороны политической элиты отвлекает население от критики невежественных и некомпетентных руководителей, но приводит к уничтожению граждан и государства. “Развитую цивилизацию не покорить извне, пока она сама себя не уничтожит” (У. Дюрант).

Закон аддитивности интеллекта. Интеллект нации или социальной группы определяется метрическим расстоянием между всеми членами сообщества:

$$I = \bigoplus_{i=1}^{n} P_i .$$

Здесь выполняется функция симметрической разности между любыми участниками (компонентами) социума. Для теоретико-множественного алфавита Кантора A={0, 1, X={0,1}, Æ} данная операция представлена таблицей истинности, иллюстрирующей взаимодействие интеллектов двух человек:

⊕	0	1	X	∅
0	∅	X	1	0
1	X	∅	0	1
X	1	0	∅	X
∅	0	1	X	∅

Исключительное достоинство данной функции в том, что она аддитивно объединяет все различия и превращает в пустоту все одинаковости. Парадоксально звучит, но уровень интеллекта 500 одинаково мыслящих людей сворачивается в пустое множество. Различные по компетентностям эксперты аддитивно создают интеллектуальную мощность, равную 500. Иначе, неаддитивность данного закона в отношении двух одинаковых индивидуумов утверждает, что 1+1=0. Расстояние между самим собой справедливо равно пустому множеству или нулю. Если же две персоны полностью различны, то их интеллекты суммируются: 1+1=2. Таким образом, несколько различных экспертов существенней для социальной системы, чем сотня одинаково мыслящих специалистов. Даже не зная данного закона, опытные эксперты-руководители успешных компаний принимают правильные решения при приеме на работу специалистов, которые не пересекаются по своим компетенциям. Ведущие университеты мира не собирают экспертные

советы из двух десятков ученых, чтобы оценить качество научного исследования – вполне достаточно трех ведущих специалистов в конкретной области. Три человека создают надежную систему в случае непредвиденного технического отказа одного из них и не более того. Как правило, в регионе или стране не существует более трех экспертов для конкретного объекта исследования. Теоретически качество экспертизы двух десятков ученых намного хуже, чем трех, по причине возможного демократического доминирования некомпетентно одинакового большинства при принятии решения на основе тайного голосования. Особенно негативно это проявляется при выборах руководителей всех уровней. Некомпетентное большинство, как правило, делает выбор в пользу коррумпированного кандидата. При принятии социально значимого решения необходимо следовать аксиоме: один эксперт доказательно более прав, чем десятки некомпетентных статистов. Уход от всегда невежественной демократии – путь к качеству социальных решений, предоставляемых экспертами. Позитивным примером использования приведенного закона является 200-летняя политическая культура США, выраженная в доктрине: “All roads of talented people lead to the United States”. Из предложенного закона аддитивности интеллекта следует важный, математически обоснованный, вывод для любой страны, которая желает процветания: «Могущество государства в толерантном объединении талантливых людей, различных национальностей, языков, культур, историй, религий и традиций». Сколько нужно миллиардов долларов, чтобы стать Германией? Ноль. Следует только поменять конституцию, согласно последнему тезису. Все остальное народ сделает сам, если не будут мешать чиновники.

Гармонический геном развития социальной группы. Синусоида является идеальной и фактической моделью развития общества во времени: всегда изменение, депрессия и расцвет, спады и подъемы. Природная мудрость такой модели заключается в потенциальной возможности изменять общественные отношения на более прогрессивные путем поиска и сравнения существующих культур развития человечества для принятия верного уклада социальных отношений. Отсталость страны наделена оптимизмом ее трансформирования в передовую при создании команды из нравственных и компетентных топ-менеджеров от развитых стран. Самый короткий во времени путь. Невежественность сообщества проявляется в неприятии чужих менеджеров, которые приперлись в наш коррумпированный монастырь со своим уставом честных отношений. Пусть даже это будут Нобелевские лауреаты по экономике. Петр Великий не брезговал опытом чужестранных талантов и выигрывал во времени становления нового технологического уклада. Америка 200 лет собирает лучших за счет создания идеальных условий для творчества и живет в шоколаде.

Компьютинг – отрасль знаний, которая занимается развитием теории и практики надежного метрического управления виртуальными, физическими и социальными процессами и явлениями на основе использования компьютерных центров и сетей, больших данных и цифрового мониторинга киберфизического пространства с помощью интеллектуальных поисково-аналитических сервисов, персональных гаджетов и умных датчиков.

Internet of Things – глобальный масштабируемый технологический уклад облачного метрического human-free управления виртуальными, физическими и социальными процессами и явлениями на основе использования платформ компьютерного обслуживания, центров больших данных и электронной инфраструктуры для цифрового мониторинга киберфизического пространства с помощью интеллектуальных поисково-аналитических сервисов, персональных гаджетов и умных датчиков в целях обеспечения качества жизни и сохранения экологии планеты (рис.1).

Эффективность социальной системы (на примере университета) определяется тремя чисто экономическими оценками уровней: потребления, экспорта и инвестиций (рис. 2). Университеты, которые позиционируются лидерами научного и образовательного рынка, метрически оценены высоким качеством управления и кадров, нравственными отношениями в коллективе, законами, историей, культурой и традициями. Инвестиции являются следствием, а не причиной эффективной и стабильной работы социальной системы – деньги любят тишину и устойчивость экономических, политических и творческих отношений. Корни коррупции лежат не в персонале, а в отношениях между людьми, определяемых законодательством, которое сознательно допускает субъективное распределение государственных средств и позиций политической элитой государства.

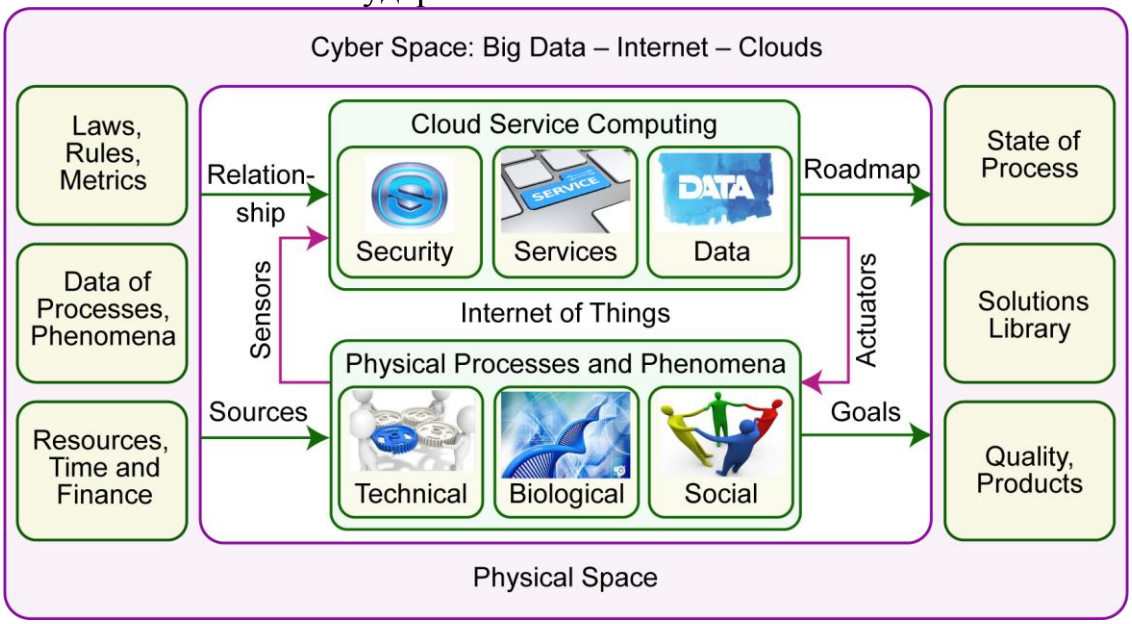


Рис. 1. IoT компьютинг

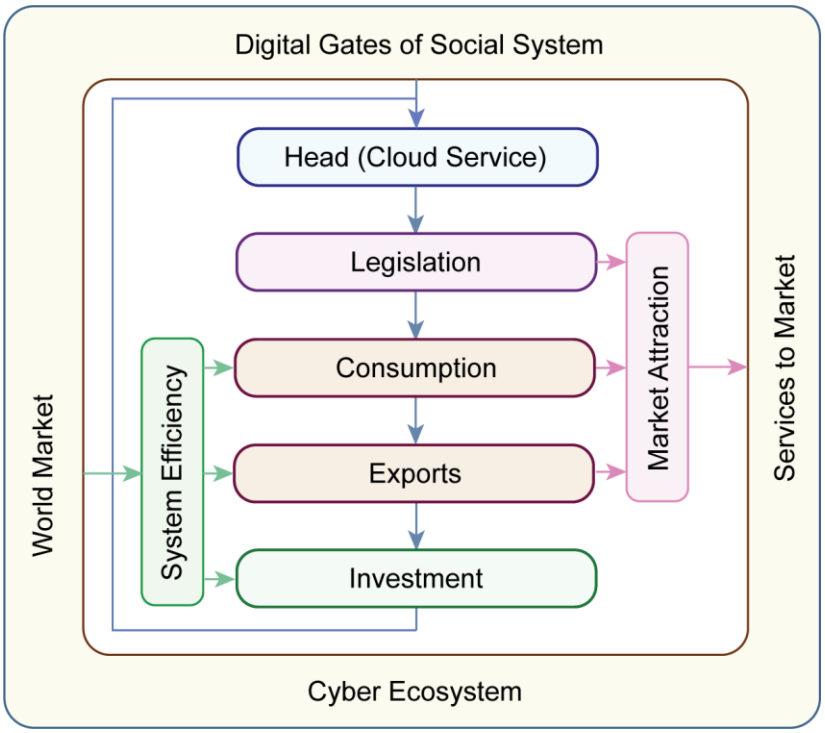


Рис. 2. Эффективность социальной системы

Решение проблемы – *кибер социальный компьютинг*, как исчерпывающий мониторинг общественного мнения и интеллектуальный анализ больших данных в целях точного облачного управления финансовыми и человеческими ресурсами путем использования е-инфраструктуры масштабируемой социальной группы на основе оцифрованных метрических отношений, определяемых действующим законодательством.

Кибер-демократия – метрическая культура социально-технологических отношений, формируемая экспертами, которая нравственно объединяет социальные группы и умную киберфизическую инфраструктуру для исчерпывающего цифрового мониторинга общественного мнения и облачного управления социальными процессами и явлениями в целях сохранения экологии планеты и достижения высокого качества жизни. Более популярно кибер-демократия представляет собой интеграцию исчерпывающей демократической дискуссии и точного экспертного мониторинга социальных проблем с облачным метрическим кибер-управлением обществом на основе оцифрованного законодательства (рис. 3).

1. Кибер-социальный компьютинг

Особое значение в 21 веке приобретает кибер-социальное правление (cyber-social governance), способное уничтожить революции, войны, социальные катаклизмы, терроризм, преступность, коррупцию, безнравственность и несправедливость со стороны политической элиты.

Одной из главных причин всех коллизий является наличие критической массы некомпетентных, несостоятельных руководителей и необразованной политической элиты на локальных территориях. Цель таких людей – купить «любовь» некомпетентного в вопросах юриспруденции и экономики электората для их последующего легитимного ограбления.

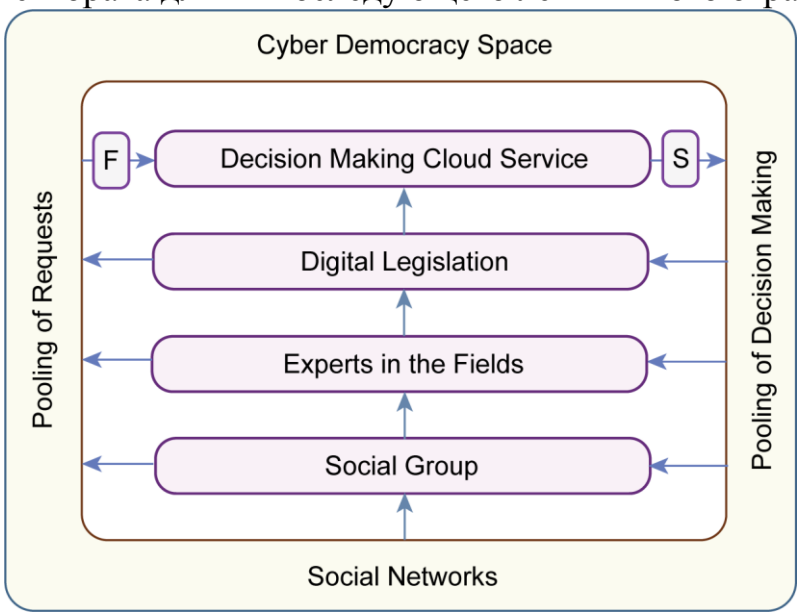


Рис. 3. Структурная схема кибердемократии

При этом о качестве жизни людей и рыночной привлекательности государства, как системы, мало кто из руководителей умеет думать. Их миссия – оставить свой, как правило, негативный или кровавый след в истории. Для народа этот след практически всегда выливается в ухудшение уровня жизни по анекдотической аксиоме – каждый следующий правитель хуже, чем все предыдущие. Следовательно, необходимо метрически выбирать компетентного руководителя группой экспертов и ограничивать его возможное негативное влияние на социальные процессы. Для этого можно предложить десять заповедей кибер-социального управления:

1. Оцифровывание социальных отношений, конституции, законодательства и подзаконных актов.
2. Электронный документооборот для государственных учреждений и частных компаний на основе первичной аутентификации. Внедрение системы электронных денег на основе создания e-infrastructure.
3. Создание электронной инфраструктуры для государственных учреждений и частных компаний.
4. Облачные сервисы для умного мониторинга и точного управления всеми социальными процессами и явлениями.
5. Внедрение облачных сервисов государственности: кибер-демократия, кибер-управление и кибер-парламент.

6. Легализация био-аутентификации по первичным признакам (отпечатки пальцев, радужная оболочка глаза, ДНК).

7. Внедрение облачных сервисов электронного голосования и электронных выборов руководителей всех уровней государственной власти.

8. Метрическая цифровая оценка физических и социальных процессов и явлений при принятии управленческих решений.

9. Распределение финансов, ресурсов и назначение кадров на основе конкурсного оценивания претендентов.

10. Онлайн-доступ к облачным сервисам государственного управления страной, городом, районом в формате 24/7.

Вред от демократии для развития человечества – всегда был, есть, но не будет в образованном культурном сообществе. Демократия в развивающихся странах есть безнравственная диктатура всегда невежественного большинства над образованным и компетентным меньшинством. Принцип Парето декларирует, что в социальной группе 20 процентов людей создают 80 процентов продукции, верно и обратное, что остальные 80 процентов граждан создают 20 процентов продукции. Вполне естественно, что 20 процентов лучших должны получать 80 процентов моральных и финансовых стимулов. Однако 80 процентов худших имеют мощное оружие – демократию, которая, в лучшем случае, бескровно побеждает конструктивное меньшинство и уничтожает государственное предприятие. Она же инициирует революции и майданы, чтобы уже физически уничтожать лучшую часть социальной группы. В результате деструкции креативных людей общество деградирует на десятки лет в прошлое. Чтобы взамен уничтоженных ученых воспитать новое поколение креативных людей, необходимо 30-40 лет. Следует учитывать, что диктатура демократии всегда уничтожает самых сильных и лучших в генетическом плане людей, а размножаться далее будут люди из 80-процентной генетически несовершенной части социума. Революции, диктаторские режимы, Первая и Вторая мировые войны унесли десятки миллионов конструктивных граждан с лучшими генами, повысив концентрацию некреативных людей до 90 и более процентов. Авторитарное правление нравственного, компетентного и креативного лидера в данной ситуации является более совершенной формой, чем демократия необразованного большинства. Такой позитивный диктатор всегда относится к социуму, как к собственной системе, которая должна быть конкурентоспособной во внешнем окружении. Поэтому он создает систему уже капиталистических отношений, где 20 процентов лучших получают 80 процентов всех благ, что обеспечивает мировое признание социальной группы, высокий уровень жизни креативных граждан. Естественно, что в экономически эффективной системе создается механизм социальной защиты 80 процентов нетворческой части населения,

которая рассматривается в качестве мощного потенциала для дальнейшего развития конструктивизма социальной группы. Не факт, что принцип Парето должен всегда иметь разделение 20/80 процентов. Существуют мощные компании (Synopsis, Apple, NASA, Cambridge, Harvard, Stanford), где эти цифры актива и пассива меняются местами, как 80/20. Что же мешает социальной системе стать успешной по метрике 80/20? Ответ на вопрос находится в истории. В постсоветских странах социализм развратил граждан уравниловкой доходов, когда все люди, активные и пассивные, получали одинаковую зарплату. Страдали лучшие люди и система, пассивная часть (80 процентов) общества гарантированно наслаждалась убогим существованием. Сегодня ничего не изменилось. Мы делаем вид, что работаем, топ-чиновники делают вид, что нам платят. Самое страшное, что из данного исторического тупика развития общества нет легитимного демократического выхода. Иллюстрация на примере университета. Исходные данные: работающих активно – 20 процентов, работающих пассивно – 80 процентов. Выборы ректора из двух кандидатур: 1) первый предлагает оставить социалистические отношения, равенство зарплат в соответствии с должностями и независимо от результатов деятельности ученых. Это приводит к гарантированной деградации университета; 2) второй предлагает внедрить нравственные отношения между сотрудниками, моральные и материальные вознаграждения по метрическому оцениванию достижений каждого ученого и сотрудника. Это приводит к созданию экономически и научно состоятельного вуза. Результаты голосования: в пяти выборных кампаниях в течение трех лет побеждает первый кандидат в ректоры. Вывод – сотрудники университета хотят иметь руководителя, который будет лоялен к их недостаткам, некомпетентности, взяточничеству и пассивной деятельности. Большинство сотрудников и ректор не ориентированы на создание экономически эффективной системы, хотя бы потому, что десятки лет университет гарантированно получал государственный бюджет без ответственности за их целевое расходование и качество произведенной научно-образовательной продукции. Науки нет, образования нет, но есть фэйковый университет. Выход – назначение ректора Президентом или Министром по метрическим параметрам оценивания компетентности каждого кандидата как: менеджера, бизнесмена, профессора, ученого, общественного деятеля. Масштабируемость выборных кампаний распространяется на все уровни социальной значимости и ответственности, вплоть до избрания президента страны. Кибердемократия предлагает метрическое оценивание компетенций кандидатов, затем осуществляются выборы одного из двух, набравших максимальное число баллов, путем голосования в среде ограниченного числа экспертов по вопросам управления. Самой близкой к

предложенной метрике кибердемократии является система выборов президента США.

Как можно устранить влияние негативного руководителя? Создать cyber governance, предназначенное для human-free распределения ресурсов по метрическим результатам деятельности людей и организаций. В этом случае уместной становится cyber democracy, ориентированная на human-free moral and metric monitoring and governance of the social group. Как ни странно, эффективность государства также оценивается принципом Парето. Если соотношение частной собственности к государственной равно 20/80, то это – развивающиеся страны. Чем ближе уровень частной собственности к 100 процентам, тем выше экономика страны, качественнее жизнь граждан, ниже уровень коррупции, стремящийся к нулю. Ни один человек в мире не будет воровать сам у себя. Будущее человечества определяется триадой кибер социальной нравственной власти: Cyber Money, Cyber Democracy, Cyber Governance и Cyber Parliament, которые в 21 веке превратятся в справедливые киберсервисы управления, стирающие границы государств с карты планеты во благо каждого гражданина Земли.

Инновационные параллели кибер (физического – социального) компьютинга. MAT (Memory – Address – Transaction) Computing. Компьютер (квантовый) будущего есть память и адресные транзакции. Память организуется на любой форме существования материи. В памяти реализуются механизмы управления и исполнения. Адресные транзакции создают все процессы и алгоритмы. Доставить одну инструкцию из control unit в память эффективнее, чем передать огромные объемы информации в АЛУ и обратно. Очевидно бутылочное горлышко, существующее 70 лет. Данные (big data) следует обрабатывать там, где они существуют. Это – ближайшее будущее кибер-физического компьютинга. Кибер-социальный компьютинг. Полная аналогия. Мегаполисы засасывают миллионы людей по утрам через бутылочное горлышко трафика ужасных дорог для их экзекуции в офисах, а вечером выплевывают их обратно по домам, в близкие и далекие деревни. Гигантские материальные, временные и финансовые затраты на бензин, аренду офисов и перемещение в реальном пространстве, стрессы в трафике, загрязнение атмосферы. Решение проблемы – кибер физический мониторинг и управление. Доставлять инструктивные воздействия через гаджеты сотрудникам в их личные виртуальные кабинеты, благодаря узаконенной online организации и созданию максимального комфорта для работы сотрудников, инвариантных к географической точке позиционирования человека (дом, отель, путешествия, отдых). Условие оплаты труда – своевременное и качественное выполнение задания. Образование и наука должны исполняться в режиме online. On-site встречи, лекции и семинары рассматриваются как роскошь живого общения. Сегодня существуют все компоненты е-инфраструктуры для создания кибер социального

компьютинга: облачный сервис управления, edge gadgets пользователей и умные вещи для реализации интеракций. Единственное, чего нет – законодательства, легализующего мульти-миллиардную сверх-инновацию в масштабах страны и планеты.

2. Киберкультура и государство

Состоятельность государства в мире определяется качеством законодательства, экономикой, способностью создавать на экспорт продукцию и сервисы, уровнем жизни граждан, а также компетентностью правительства. Экономическая и научно-образовательная отсталость страны является следствием безнравственности и несостоятельности правящей элиты, генерирующей антигуманное законодательство, что приводит к деградации и распаду государства. Цели постсоветских чиновников не имеют отношения к благосостоянию и сплочению народа. Их проблема – сохранить власть и украсть за ограниченный срок правления больше денег. Какова технология удержания власти для некомпетентных в управлении экономикой руководителей? Существует несколько беспроигрышных методов оболванивания своих народов, живущих в нищете, по сравнению с окружающими странами.

1. Создание образа внешнего врага. В наших бедах виноваты соседи. Их нужно уничтожить, тогда мы будем богатыми и счастливыми. Но если соседи более сильные, то наше инфантилизм – жаловаться мировому сообществу, признавая свою некомпетентность и неэффективность политических решений. Козырная карта патриотизма (волонтерства) всегда является неубиенной. К ней обращаются во всех случаях, когда у невежественного в управлении руководителя нет денег достойно оплачивать работу своих граждан. В 99 процентах случаев «настоящими» патриотами являются воинствующие бездельники, неспособные к производительному труду.

2. Создание образа внутреннего врага. Конституционно ввести дискриминирующие отношения между существующими на территории страны социальными группами. Целью такой политики является разделение народов по метрике неравных отношений официальной власти к языку, культуре, истории, религии, национальностям. Разделяй и властвуй. Конституционно декларируется доминирование титульной нации, все остальные – фоновые граждане. Результат – патологическая ненависть, проходящая через народы, социальные группы и семьи. Это отвлекает людей от конституционного права – переизбрать некомпетентное правительство и созидательным трудом уничтожить нищету, которая является прямым следствием невежества правящей элиты. Низкий уровень жизни 80 процентов граждан приводит к росту преступности, разрушению инфраструктуры, экономики, политической системы и государства. Слепая вера во власть – есть заклятый враг правды

(Альберт Эйнштейн). Наци(онали)зм – враг науки и образования.

Увы, но экономика страны не поднимется за счет внедрения в сознание общества деструктивных доктрин наличия внутренних и внешних врагов. Такая политика стимулирует низменные и звериные инстинкты зависти, разрушения, убийства, парализует волю граждан к созидательному труду и неспособна испечь даже булку хлеба. Как решить проблему отсталости страны? 1. Объединять народы путем создания нравственных отношений (конституции) равенства языковых, исторических, религиозных и национальных культур. 2. Метрически формировать правительство компетентных руководителей и парламентариев. 3. Дружественно сотрудничать с внешним окружением страны, соревнуясь в экономической состоятельности и качестве жизни своих граждан. 4. Не изгонять из страны тех, кто умнее чиновников, а создавать благоприятные условия для творчества и массового приглашения извне талантливых людей. 5. Создавать политически и экономически стабильные условия, инвариантные к смене правительств, президентов и парламентариев, для привлечения внешних и внутренних инвестиций в виде капиталов и компетентных кадров.

Какова роль и методы нарождающейся киберкультуры в становлении экономически состоятельного государства и тотального устранения коррупции? Далее представлены пять тривиальных тезисов, со слабой надеждой на готовность общественности к их восприятию.

1. Оцифровывание всех объектов и явлений на территории страны, включая граждан, конституцию, законы, деньги и документацию.

2. Метрическое оценивание всех процессов и явлений для последующего точного мониторинга и адекватного управления кадрами и ресурсами. Метрика – способ измерения расстояния в киберфизическом пространстве между процессами или явлениями путем сравнения их параметров.

3. Создание электронной инфраструктуры цифрового мониторинга и облачного управления умным социумом, страной, городом, домом, организацией, транспортом, финансами, наукой и образованием (рис. 4). Умный (Smart) – определение процесса или явления, связанное с сетевым взаимодействием адресуемых системных компонентов во времени и пространстве между собой и окружающей средой на основе технологий самообучения для достижения поставленных целей. E-Infrastructure – совокупность взаимосвязанных облачных сервисов, центров больших данных, компьютерных устройств, систем и сетей, а также законов, стандартов и средств аутентификации, кибербезопасности, телекоммуникации, тестирования и ремонта, обеспечивающих масштабируемый IoT-компьютинг для надежного мониторинга и устойчивого управления процессами и явлениями в целях повышения качества жизни человека и сохранения экологии планеты.

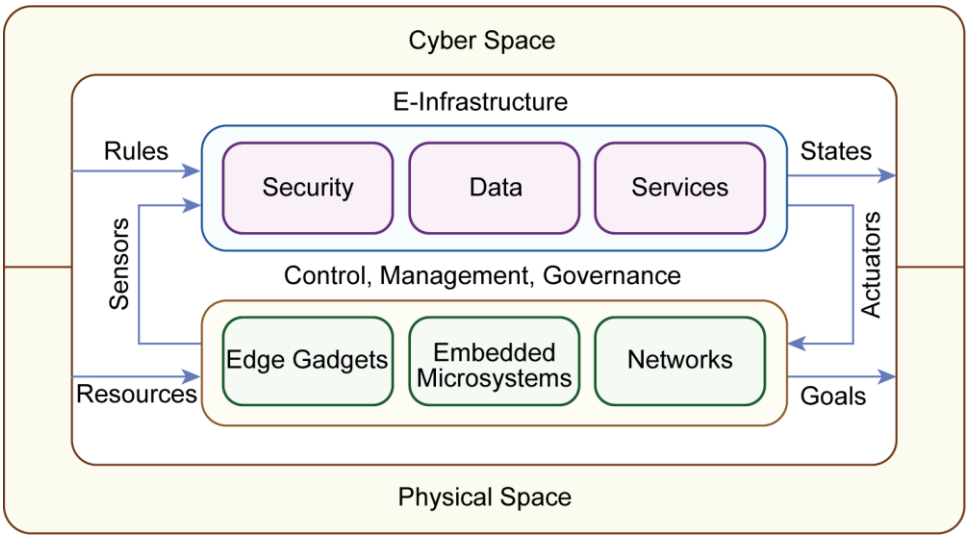


Рис. 4. Электронная инфраструктура

4. Внедрение электронного документооборота на основе аутентификации личности, и цифровой подписи во всех сферах человеческой деятельности.

5. Внедрение электронных выборов и электронного голосования по всем вопросам компетентностного оценивания и принятия решений.

Все пять пунктов достаточно легко реализовать в масштабах государства при наличии политической воли и минимального уровня киберкультуры у правящей элиты. Особое значение для воспитания киберкультуры граждан имеет электронный документооборот, который не следует примитивно понимать как сервис автоматизированного изготовления, приема-передачи, хранения и утилизации документов.

Электронный документооборот (E-Document Circulation – EDC) – киберфизическая компьютинговая система точного цифрового мониторинга и оперативного киберуправления социальными группами, использующая средства (e-infrastucture) аутентификации индивидуума, edge gadgets, cloud services в целях online решения социальных проблем без использования бумажных носителей для повышения качества жизни граждан и сохранения экосистемы планеты.

Основой электронного документооборота являются легитимные интеллектуальные транзакции потоков оцифрованных документов (сенсорных сигналов и регуляторных воздействий) в умной логически рассредоточенной сети данных, предназначенные для реализации безбумажных отношений с внешним миром, прямого мониторинга и непосредственного управления научно-образовательными процессами и подразделениями университета. Оцифрованные документы (доступные для понимания компьютером и человеком) выполняют роль цифровых сенсоров и актюаторов в замкнутой кибер-системе (например, Smart Cyber University). Это означает возможность генерирования цифровых отчетов и управление системой с помощью цифровых документов, понятных кибер-

системе, в том числе и без участия человека. Электронный документооборот зачастую ассоциируется с транзакциями электронных копий бумажных носителей информации для визуального восприятия человеком, но не киберсистемой, что было бы инновационно в 1990 году.

Киберфизическая система электронного документооборота (рис. 5) содержит:

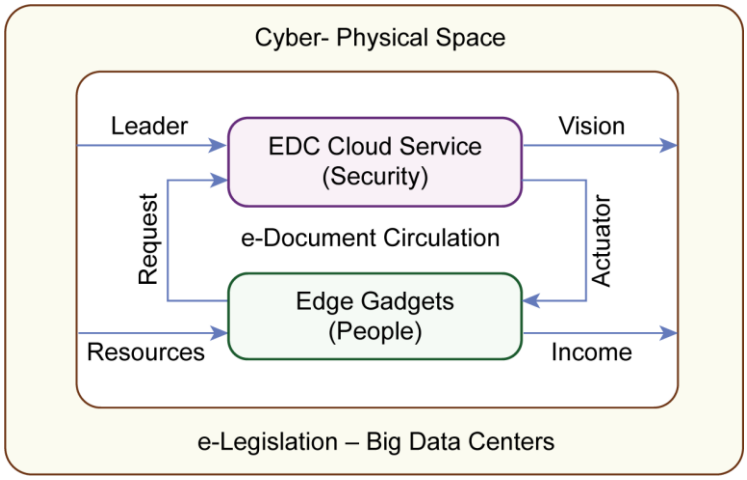


Рис. 5. Структура электронного документооборота

1) облачный сервис мониторинга, хранения и интеллектуального анализа документов в целях управления социальными группами, благодаря online генерированию документов по запросу пользователей; 2) конечные гаджеты интерфейсной связи пользователей с облачными сервисами, предоставляющими электронные документы; 3) комплексные средства первичной аутентификации защиты облачной системы от несанкционированного доступа и подделки электронных документов на основе доверительной идентификации пользователя; 4) масштабируемую социальную группу, которая создается в рамках предприятия, университета, организации, территории, государства, получающая облачные сервисы, обеспечивающие высокое качество жизни каждого пользователя.

Принципы или аксиомы электронного документооборота направлены на обеспечение высокого качества обслуживания, жизни человека и сохранение лесов для будущих поколений: 1) online цифровой мониторинг, метрическое оценивание и прогнозирование деятельности каждого гражданина, социальной группы для выработки адекватных управляющих воздействий на основе существующей электронной документации; 2) online создание конструктивного управляющего воздействия в форме цифрового документа, как результата метрического оценивания процесса или явления по запросу пользователя; 3) создание электронной инфраструктуры для комплексной защиты электронного документооборота от подделки и несанкционированного доступа к документам на основе средств первичной аутентификации: электронная

цифровая подпись, сканирование отпечатков пальцев, анализ ДНК, ручная подпись документа на сенсорном экране; 4) исключение всех вторичных материальных носителей аутентификации: электронных карт, паспортов, пропусков, водительских прав, справок, свидетельств, дипломов, аттестатов, усложняющих жизнь человека и засоряющих экосистему планеты; 5) надежность электронного документа, за счет дублирования данных и его виртуализации в киберфизическом пространстве, на порядок выше материального, который можно украсть, уничтожить, потерять; 6) online доступность электронного документа или облачного сервиса, формирующего его – 24/7, из любой точки земного шара, делает EDC особенно привлекательным для всех жителей планеты; 7) тотальное исключение из физического обращения металлических и бумажных денежных знаков, которые обладают только чудовищными недостатками, связанными с вырубкой лучших лесов, загрязнением биосферы земли, баснословными расходами на изготовление хранение и транспортирование миллионов тонн денежных знаков, но главное – с распространением инфекционных болезней по всей планете. Взамен – чистые электронные деньги и миллиарды долларов экономии. Это есть задача номер один для финансовой и политической элиты человечества; 8) приведение конституций и законодательств государств к электронным цифровым документам прямого мониторинга и исполнения в формате: “факт – оценка – действие”, путем отказа от декларативных документов, требующих сотни подзаконных актов и разъяснений. Все документы, благодаря доступности облачных сервисов, должны быть прямого действия: статья закона определяет действия гражданина; 9) абсолютная прозрачность и открытость всех видов электронного документооборота для каждого человека, не противоречащая законодательству страны, уставу предприятия, организации, социальной группы; 10) персонализация ответственности за легитимность подписанного электронного документа, которая исключает ссылки на коллективные решения некомпетентных экспертов.

Эволюционирование электронного документооборота приведет уничтожению существующего декларативного формата документов. Взамен появится интеллектуальная компьютинговая система мониторинга, управления и прогнозирования деятельности каждого человека на основе накопления опыта принятия решений. Каждый документ-программа, будет представлять собой оцифрованную триаду “факт – оценка – действие”, понятную для генерирования компьютером адекватных актюаторных воздействий, рассматриваемых в качестве ответа на запрос пользователя.

3. Метрические кибер-отношения – основа управления

Нет измерения – нет управления – нет науки и образования – нет экономики – нет государства. Метрическое оценивание процессов и

явлений для адекватного распределения моральных и материальных средств является нравственной основой справедливых социальных отношений в компании, организации и стране. Данный тезис позитивно не воспринимается общественностью, поскольку он дифференцирует людей на 20 процентов креативно работающих, создающих 80 процентов продукции. Остальные 80 процентов качественно выполняют функциональные обязанности. Естественно, что моральные и материальные стимулы между двумя указанными группами людей в развитых странах и лучших компаниях планеты с высоким уровнем технологической культуры распределяются в обратной пропорции, близкой к 80/20. Но и в этом случае каждый из сотрудников, входящих в 80 процентов, имеет высокий уровень доходов. Экономически эффективная система способна обеспечить достойные зарплаты. Однако демократия пассивного большинства не в состоянии принять конструктивную метрику оценивания деятельности людей. Поэтому необходимы непопулярные жесткие регуляторные воздействия со стороны руководства, владеющего кибер культурой системного управления компанией, организацией, страной. Таким образом, только креативные метрические отношения в социуме способны создать конкурентоспособную на рынке структуру, которая масштабируется в страну, компанию, организацию или университет. Здесь интересен 26-летний опыт управления высшей школой со стороны министерства, которое, год за годом, создавало метрические отношения для оценивания ученых, подразделений и университетов. Результат удивляет своей несостоятельностью, связанной с созданием раздутой метрики, содержащей 107 виртуальных пунктов, в мутной воде которых потонули реально-эффективные показатели. Достаточно сказать, что для формирования сводных таблиц метрического оценивания значительная часть коллектива университета «творчески» работает в течение целого месяца, не производя при этом никакой общественно полезной продукции. В масштабах страны – это миллионы потерянных долларов. Но и это еще не все. Умелые университетские чиновники и руководители подразделений научились писать фэйковые бумажные отчеты по виртуальным показателям, которые побеждают реальные научные достижения конструктивных ученых. Также много лет существуют странные штрафные санкции для сотрудников, которые имеют нулевые результаты деятельности по отдельным показателям из упомянутой 107-пунктовой метрики. Их цель – наказать ученого (кафедру) за невыполнение отдельных пунктов метрики, вместо того, чтобы наградить его за достижения в других видах деятельности. Например, если ученый получит Нобелевскую премию, а во многих других показателях он будет иметь нули, то штрафные санкции сделают его аутсайдером рейтингового оценивания с последующим моральным и материальным наказанием за

невыполнение функциональных обязанностей. Как следует исправить существующую систему рейтингового оценивания, разрушающую науку в высшей школе? Уменьшить метрику показателей до десятка, оставив признанные на международном рынке критерии научно-образовательной и волонтерской экстра активности ученого и коллектива:

1. R&D-гранты, коммерческие и государственные, образовательные и производственные проекты.
2. Наукометрия: индекс Хирша, цитируемость, количество публикаций.
3. Монографии, учебные пособия, УМК, МООС.
4. Диссертации, защищенные под руководством ученого.
5. Патенты, рыночные научно-образовательные продукты и сервисы.
6. Участие студентов во всех видах экстра деятельности.
7. Награды, звания, премии и дипломы социального признания заслуг.
8. Организация и проведение конференций, выставок и семинаров.
9. Издание журналов, фильмов и трудов конференций.
10. Договора с предприятиями и университетами.

Первичными и образующими университет являются только первые семь пунктов, которые существенно влияют на международный рейтинг, имидж, финансово-экономическое состояние вуза и качество жизни сотрудников. Только их необходимо принять к реальному исполнению, чтобы всему коллективу жить стало лучше, жить стало веселей.

Ликвидация коррупции. Структура коррупции, как системного государственного атрибута, содержит три взаимосвязанных компонента: чиновники, граждане (компании) и ресурсы. Ресурсами может быть любой государственный (значит ничейный) товар, который можно продать, включая финансы, природные ископаемые, услуги и должности. Чиновники – государственные служащие, торгующие не принадлежащий им товар. Устранение любого компонента из триады ликвидирует коррупцию как государственное явление. Самый затратный и безнаравственный путь – ликвидировать всех граждан в государстве, коррупция исчезнет, потому что не будет покупателей и государства. Второе решение проблемы – устранить чиновников путем замены их функций неподкупными облачными сервисами. Третий путь ликвидации коррупции – не накапливать государственные ресурсы для их последующего перераспределения между гражданами. Денежные потоки от граждан должны непосредственно попадать уже частным компаниям, которые предоставляют “государственные” товары и услуги. Данный вариант также не предусматривает наличия чиновников, как государственных служащих. Они переходят в частные компании, создающие и поддерживающие инфраструктуру страны, сервисы охраны

здоровья, образования, юридической защиты прав и собственности граждан. Таким образом, структура финансовых потоков и услуг в коррумпированном государстве (рис. 6), нежно трансформируется в эффективную систему путем замены армии чиновников на неподкупный облачный сервис легитимного и открытого распределения ресурсов на основе метрического мониторинга социальных проектов. Это означает конец коррупции. Новая модель стремится к нулевому числу чиновников, превращая коррумпированную по сути государственность в мощную территориальную корпорацию, свободную от управленческой избыточности.

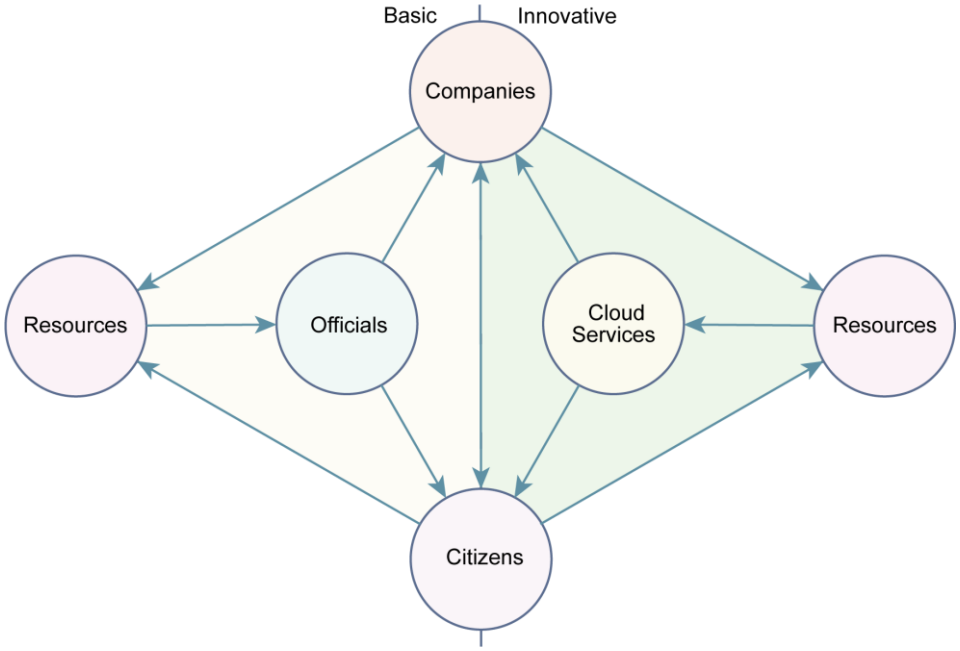


Рис. 6. Трансформация государственности в социальную систему без коррупции

Следует заметить, что государственная монополия создает для граждан безальтернативность некондиционного сервисного обслуживания. Взамен на рынок приходят качественные мультиверсные услуги от частных компаний, созданые, в том числе, на основе армии компетентных чиновников, которые заботятся о своей репутации среди граждан, влияющей на прибыль и качество жизни сотрудников компании. Открытым остается вопрос, как инициировать ликвидацию избыточного звена, создающего коррупцию и негативно влияющего на эффективность социальной системы, если это звено – власть.

Выводы

Предложена киберкультура социального компьютинга, направленная на нравственное решение вопросов метрического управления обществом на основе исчерпывающего цифрового мониторинга, в целях устранения

коррупции, обеспечения высокого качества жизни граждан и сохранения экологии планеты. Сформулирован закон аддитивности интеллекта нации или социальной группы, определяемый метрическим расстоянием между компетенциями членов сообщества.

Рассмотрены перспективные направления создания умных: e-инфраструктур, государств, городов, университетов, компаний и домов, имеющих высокий уровень капитализации в мире. Предложены электронные технологии безбумажного документооборота, сохраняющие сотни миллионов долларов и чистую экологию в масштабах государства. Рассмотрены вопросы human-free управления социальными группами на основе создания Cyber Democracy, Cyber Governance и Cyber Parliament, уничтожающих коррупцию в масштабах государства.

Уровень капитализации предложенных кибер социальных инноваций в масштабах даже слабого развивающегося государства составляет миллиарды долларов. Но еще более значимым возможным достижением является создание нравственных отношений в обществе и сохранение экологии для будущих поколений. Кроме того, кибер-культура, кибер-финансы и кибер-технологии уничтожают криминалитет, связанный с воровством материальных денег и автомобилей, коррупцией и взяточничеством, квартирными кражами и бандитизмом.

В настоящее время наблюдается устойчивое доминирование инновационной киберкультуры, включающей Нано-Аддитивные, Био-Информационные, Кибер-Физические, Социально-Когнитивные технологии, на рынке науки, образования, индустрии, транспорта и обеспечения качества жизни. В связи с этим перспективными являются следующие направления инноваций:

1. Создание облачного сервиса мониторинга и управления научно-образовательными процессами “Smart Cyber University”.
2. Разработка умного цифрового мониторинга автомобилей и streetlight-free облачного управления транспортом.
3. Создание e-infrastructure и облачных сервисов мониторинга и human-free управления социальными группами.
4. Разработка треугольных 2D-3D топологий для создания оптимальных инфраструктур киберфизического пространства: triangle-driven города, компьютинговые архитектуры, киберэкосистема планеты.
5. Создание e-инфраструктуры для аутентификации человека по первичным признакам для банкинга, поездок, доступа в помещения, исключающей паспорта, электронные карты, металлические, пластиковые ключи и все виды бумажных документов.
6. Разработка интегральной e-инфраструктуры дома и квартиры для online цифрового мониторинга и облачного управления всеми бытовыми приборами и службами.

7. Создание е-инфраструктуры online электронного голосования на основе аутентификации человека по первичным признакам.

8. Создание е-инфраструктуры для чистых электронных денег путем тотального исключения из физического обращения металлических и бумажных денежных знаков, что принесет миллиарды долларов экономии в масштабах планеты.

9. Создание электронной кибер государственности, когда гражданин планеты будет иметь право не только выбирать руководителей, но и государство, куда он будет платить налоги для получения нравственных и качественных сервисов правовой и социальной защиты.

Муниципальные сервисы от медицины и правоохранительных органов уже перешли в разряд платных. Что касается защиты от врагов, то, судя по управленческим действиям, большего врага для народа в некоторых государствах, чем собственное правительство, не существует. Государственные институции дискредитируют себя своей некомпетентностью, поэтому рынок услуг создает эффективные приватные параллельные структуры с функциональностями: суда, полиции, армии, здравоохранения, науки, образования. Финансовые потоки непосредственно попадают от потребителя сервисов до их исполнителя, минуя государство, как посредника. Не существует принципиальных ограничений трансформировать государственность в эффективную частную территориальную корпоративность, нравственно и сервисно привлекательную для своих граждан.

Computing is the Union of a Binary. Компьютерная система всегда состоит из двух взаимодействующих компонентов, выполняющих функции управления и исполнения для достижения заданной цели функционирования в киберфизическом пространстве. Любая другая система может быть представлена компьютинговым взаимодействием двух компонентов. Цель не может быть достигнута наличием только одного компонента или разрывом связей, что приводит к разрушению системы. Любое изменение структурных отношений или возникновение различных целей у двух компонентов системы приводит к коллизиям и ее самоликвидации. Примерами компьютинговых систем могут выступать:

1. Энергия и материя, взаимодействие которых создает компьютинговую пару, где целью является гармоническое изменение материально-энергетической субстанции во времени и пространстве.

2. Пространство и время, взаимодействие которых создает компьютинговую пару, где целью является гармоническое изменение пространственно-временной субстанции для создания материально-энергетического многообразия.

3. Душа (мозг) и тело, взаимодействие которых создает компьютинговую пару, где целью является выживание и познание мира.

4. Мужчина и женщина, взаимодействие которых создает компьютинговую пару, где целью является выживание и продолжение человеческого рода.

5. Электричество и кремниевый кристалл, взаимодействие которых создает компьютинговую пару, где целью является создание функциональностей: хранения информации или вычислительных процедур.

6. Человек и компьютер, взаимодействие которых создает компьютинговую пару, где целью является улучшение экологии и качества жизни человека.

7. Человечество и компьютинг (интернет), взаимодействие которых создает компьютинговую пару, где целью является бессмертие и экспансия человечества во Вселенной.

8. Человек и аватар, как его цифровая копия, взаимодействие которых создает компьютинговую пару, где целью является кибер-информационное бессмертие, востребованное для будущих поколений.

9. Социальная группа (государство) и руководящая элита, взаимодействие которых создает компьютинговую пару, где общей целью является качество жизни граждан и рыночная состоятельность сообщества. Наличие различных целей у руководства и народа (коллектива) приводит к социальным коллизиям, революциям и бунтам, а затем к самоликвидации сообщества или государства. Парадоксально, но факт, вузовский менеджмент развивающихся стран игнорирует теорию системного управления. Несостоятельность управления высшей школы заключается в отсутствии строгого кадрового разделения между ректоратом (исполнительная власть) и ученым советом (законодательная власть), что возведено в ранг закона. Это создает, как правило, коррумпированную систему управления вузом, когда ректорат своим большинством фактически диктует свою волю остаткам ученого совета при принятии невежественных или некомпетентных решений, разрушающих науку и образование.

10. Человечество и Бог, взаимодействие которых создает компьютинговую пару, где целью является нравственное и созидательное совершенство людей в гармонии с природой.

Космологические и киберфизические и киберсоциальные браки материи и энергии, пространства и времени, сущности и формы, управления и исполнения бесконечны в своем многообразии, формирующем Вселенную. Естественно, компьютинговые пары взаимодействуют с другими марьяжами, что создает конкурентную игру на выживание по существующим физическим законам.

Марьяж материи и энергии образует системное равноправное взаимодействие, где каждый из компонентов предоставляет друг другу то,

что имеет. В результате получается память-геном, которая содержит информацию о прошлом, настоящем и будущем жизненного цикла пары.

Примитив-система, содержащая два взаимодействующих и взаимно дополняющих компонента, является самой простой, надежной и жизнеспособной.

Для путешествия в пространстве и времени необходимо создавать примитивные пары материально-энергетических субстанций, формирующие память для хранения информации о структуре и алгоритмах жизненного цикла системы в целях ее воспроизведения в любой точке Вселенной с приемлемыми условиями. Всего-то нужно передать на любое расстояние информацию, как структурировать имеющуюся под рукой подходящую материю и вдохнуть в нее энергию алгоритмической жизни.

Литература

1. Gaol F. L., Hutagalung F. D. Social Interactions and Networking in Cyber Society. Springer International Publishing, 2017.
2. Meiselwitz G. Social Computing and Social Media. Springer International Publishing, 2016.
3. Koch F., Koster A., Tiago P. Social Computing in Digital Education. Springer International Publishing, .2016.
4. Barnaghi P., Sheth A., Singh V., Hauswirth. M. Physical-Cyber-Social Computing: Looking Back, Looking Forward. In: IEEE Internet Computing, 2015. – vol. 19, no. 3, May-June, 2015.
5. Hahanov V. I., Litvinova E. I., Chumachenko S. V., Mishchenko A. S. Cyber social system – Smart Cyber University. In: Radioelektronic and Computer Systems, 2016. – J 5 (79): 187 – 194.
6. http://fakty.ictv.ua/ru/ukraine/20170410-rejtyng-biznes-koruptsiyi-ukrayina

ЧАСТЬ 2
ОБРАБОТКА ИНФОРМАЦИИ В ИНФОКОММУНИКАЦИЯХ, ИТЕЛЛЕКТУАЛИЗИРОВАНОЕ КОДИРОВАНИЕ И МНОГОМЕРНЫЕ СИНТАКСИЧЕСКИЕ ПРЕОБРАЗОВАНИЯ

МЕТОД ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ ІНФОКОМУНІКАЦІЙНИХ СИСТЕМ НА ОСНОВІ ТРАНСФОРМУВАННЯ ТА КОДУВАННЯ ВІДЕОДАНИХ

Баранник В.В., Кривонос В.М., Леках А.А.

Вступ

Останнє десятиріччя характеризується стрімким розвитком бездротових інфокомунікаційних технологій передачі даних. Ефективне функціонування таких систем визначається якістю передачі та обробки інформації і оцінюється таким показником як продуктивність, що задається формулою

$$A_t = U(Pt) / T_\Sigma ,$$

де T_Σ − сумарний час обробки та передачі відеоданих обсягом $U(Pt)$.

Продуктивність бездротових мереж нового покоління допускає реалізацію відеоінформаційного обміну. Однак відеоінформаційний сектор стрімко розвивається і досягає 70%, від загального обсягу трафіку. Необхідна швидкість передачі відеопотоку за одиницю часу, структура якого представлена на рис. 1, визначається за формулою (1), і залежно від розміру зображення представлена на рис. 2.

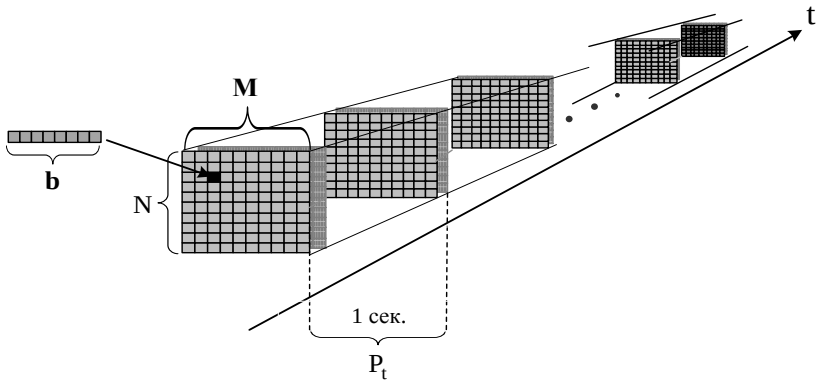


Рис.1. Структура відеопотоку

$$U(P_t) = P_t \times M \times N \times b \ \ (\text{біт}), \qquad (1)$$

де $U(P_t)$ − обсяг відеопотоку за 1 сек. з частотою P_t кадрів/с; P_t − частота кадру; $M \times N$ − розмір зображення; b − глибина оцифрування зображення.

У той же час, оцінка часу передачі нестислих відеоданих з врахуванням існуючої продуктивності бездротових технологій, показала, що час передачі може досягати десятки секунд. Звідки можна заключити, що потрібно підвищувати продуктивність існуючих бездротових технологій. Потрібно забезпечити виконання співвідношення $A_t \to \max$ для заданих обсягів відеоданих відповідних відеоінформаційних сервісів, що описується обмеженням $(U)_н \le (U)$.

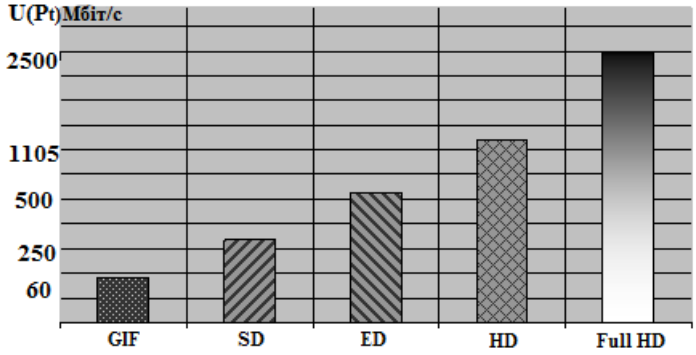


Рис.2 Необхідна середня $V_п(P_t)$ *швидкість передачі відеопотоку за 1сек. у залежності від якості зображення*

Це можна досягти за рахунок зниження сумарного часу доставки даних, як результат зниження обсягів даних, що передаються, тобто

$$T_\Sigma \to \min, \qquad T_\Sigma = T_{обр} + T_п,$$

де T_Σ − сумарний час обробки та передачі відеоданих; $(U)_н$ − необхідний обсяг відеоданих.

Отже зниження обсягів відеоданих для підвищення продуктивності функціонування інфокомунікаційних систем із заданою якістю відеосервісу, *є актуальною науково-прикладною задачею.*

Сформульовану задачу пропонується вирішувати на базі розвитку технології стиснення, з подальшою інтеграцією в інфокомунікаційні системи. Одним з найбільш розповсюджених та популярних методів стиснення є метод JPEG, основні етапи роботи якого показані на рис.3 [1].

Оцінка часу передачі зображень стислих з використанням методу JPEG проводиться за формулою:

$$T_п = \frac{U(P_t)}{k_{ст} \times V_п} \text{ (сек)},$$

де $k_{ст}$ – коефіцієнт стиснення; $V_п$ – швидкість передачі зображення по каналу зв'язку.

Аналіз компресійних характеристик методу JPEG виявив, що передача відеоданих у реальному часі можлива тільки в режимі значних втрат якості.

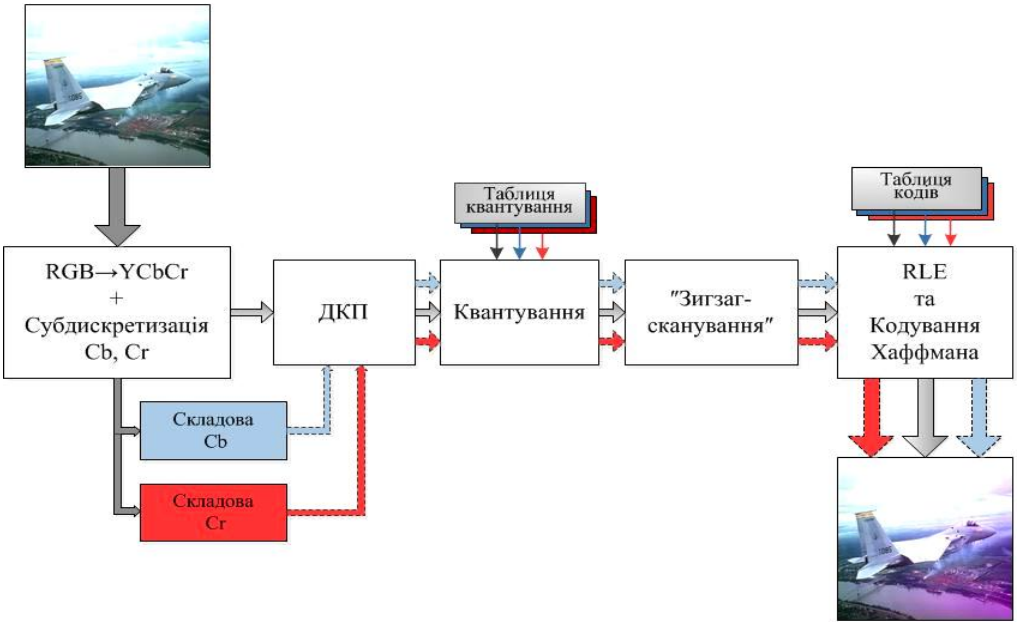


Рис.3. Блок-схема роботи алгоритму стиснення у JPEG

Для вдосконалення технології JPEG, пропонується використовувати підхід, заснований на обробці компонентного представлення трансформанти. Це дозволить враховувати такі закономірності як: переконцентрація енергії; наявність ланцюжків нульових елементів. У той же час, існуючі підходи пов'язані з такими недоліками як: в одному випадку, це додаткова передача динамічних таблиць статистичних кодів, а в другому випадку, це низька адаптивність стаціонарних статистичних таблиць до змінних характеристик зображення. Звідси виникає необхідність у розробці методу кодування відеоданих на основі компонентної структури трансформанти для зниження їх обсягів у інфокомунікаційних системах.

1. Метод стиску трансформованих відеоданих для засобів інфокомунікацій

Пропонується розділяти трансформанту після її лінеаризації на складові: вектор значущих субсмуг та вектор масштабуючих компонент рис. 4.

Тут Y_m – розгорнута трансформанта; Y_{m-1} – вектор значущих субсмуг; G_{m-1} – вектор масштабуючих компонент; DC – низькочастотна компонента трансформанти.

Внаслідок чого трансформанта складатиметься з трьох складових:

$$Y_m = \{DC; Y_{m-1}; G_{m-1}\}.$$

Це дозволить виділити характерні структурні закономірності. Для вектора значущих субсмуг

$$Y_{m-1} = \{y_2, ..., y_j, ..., y_m\};$$

структурні характеристики задаються наступними закономірностями:

Перша: сусідні значущі субсмуги послідовності мають різні значення, що задається нерівністю

$$y_\xi \neq y_{\xi+1}\ ; \qquad \xi = \overline{1, m-1}.$$

Друга: виконується закономірність щодо обмеженого діапазону субсмуг, що виражено нерівністю

$$y_{min} \leq y_2, \ldots, y_j, \ldots, y_m \leq y_{max}.$$

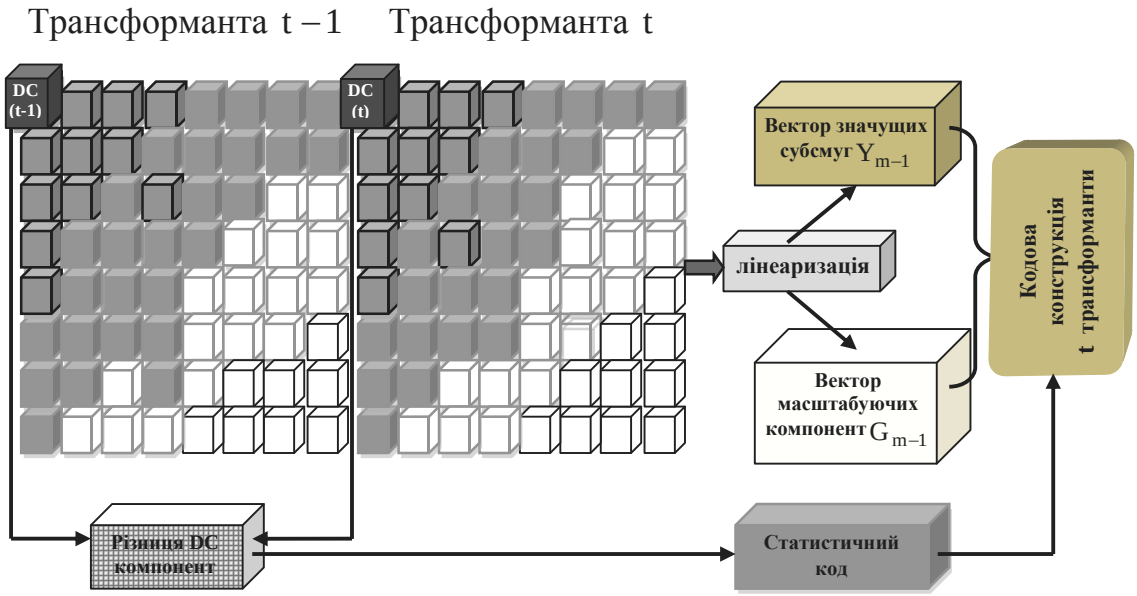


Рис.4. Структурна схема обробки компонент трансформанти

Для урахування таких закономірностей, пропонується підхід, що полягає в розгляді значень значущих субсмуг вектора, як послідовність, у якої діапазон другої компоненти визначатися нерівністю

$$y_{min} \leq y_2, \ldots, y_j, \ldots, y_m \leq y_{max};$$

для всіх інших компонент визначається виразом

$$w(y)_j = y_{max} - y_{min};\ \text{де}\ j = \overline{3, m}.$$

У результаті запропонованих перетворень для трансформант формується, послідовність значущих субсмуг, значення яких задовольняють умовам

$$y_2 \leq w(y)_2 = y_{max} - y_{min} + 1;\ y_j \leq w(y)_j = y_{max} - y_{min};\ j = \overline{3, m}. \quad (2)$$

Звідси значущі субсмуги вектора називатимуться як послідовність значень значущих субсмуг неоднорідного спектра дискретного косинусного перетворення (ДКП) [6].

Визначення 1. Вектор Y_{m-1}, для субсмуг якого виконуються умови (2) так, що у загальному випадку $w(y)_j \neq w(y)_v$, $j \neq v$, та $j, v = \overline{2, m}$,

називатиметься послідовністю значень значущих субсмуг неоднорідного спектру ДКП з системою основ $W(y)=\{w(y)_j\}$.

Тоді максимальна кількість розрядів, що витрачається на представлення вектора неоднорідних значень спектра ДКП трансформанти, обчислюється за формулою

$$D'^{(y)}_{m-1}=[\ell og_2 V'^{(y)}_{m-1}]+1=[\ell og_2(y_{max}-y_{min}+1)+(m-2)\ell og_2 w(y)]+1,$$

де $D'^{(y)}_{m-1}$ – максимальна кількість розрядів, яка необхідна на представлення вектору неоднорідних значень спектру ДКП, а середня кількість двійкових розрядів, що припадає на один елемент послідовності, визначається формулою

$$\overline{D'}^{(y)}_{m-1}=\frac{[\ell og_2(y_{max}-y_{min}+1)+(m-2)\ell og_2 w(y)]+1}{m-1},$$

де $\overline{D'}^{(y)}_{m-1}$ – середня кількість двійкових розрядів, що припадає на один елемент вектору Y_{m-1}.

Мінімальна кількість надмірності оцінюється на основі виразу

$$\bar{S}^{(y)}_{min}=(1-[\frac{\ell og_2(y_{max}-y_{min}+1)}{8(m-1)}+\frac{\ell og_2 w(y)}{8(m-1)}])100\%,$$

де $\bar{S}^{(y)}_{min}$ – мінімальна кількість надмірності у випадку представлення компоненти трансформанти як окремої величини відносно її представлення, як елементу вектору Y_{m-1}.

Звідки видно, що виконується нерівність

$$\ell og_2(y_{max}-y_{min})\le 8.$$

Отже, мінімальна кількість надмірності буде відмінною від нульового рівня $\bar{S}^{(y)}_{min}>0\%$.

Таким чином, обґрунтовано, що в результаті кодування значущих субсмуг частотного спектра ДКП в умовах його неоднорідності, досягається додаткове скорочення надмірності.

Для вектора масштабуючих компонент трансформанти, залежно від стратегії квантування, характерні наступні закономірності, а саме: наявність початкової l_g та останньої g_m нульової серії; компонент N_g, що знаходяться в середині вектора мають обмежений діапазон. З аналізу характеристик основних структурних параметрів вектора масштаующих компонент випливає, що в залежності від кроку квантування R вони змінюються за певними закономірностям. Початкова нульова серія зменшується, остання - збільшується, компоненти, що позиціонують між серіями нулів - зменшуються. Саме такі закономірності використовуються при кодуванні.

Метод стиснення пропонується будувати на основі роздільної обробки структурних параметрів трансформанти. Тут основними етапами є:

Перший. Обробка низькочастотної DC компоненти, яка кодується окремо від інших значущих компонент. Вона представляється у вигляді різниці, значення поточної компоненти та компоненти попередньої сусідньої трансформанти, вираз (3) яка кодуються двома частинами рис. (5).

$$\Delta DC(t) = DC(t) - DC(t-1). \qquad (3)$$

Низькочастотна DC компонента, кодується двома частинами, що задається формулами:

$$\Delta DC(t) < DC(t),\ l(\Delta DC(t)) < l(DC(t)),\ [\Delta DC(t)]_2 = [l_i]_2 \cup [d_i]_2;$$

де $\Delta DC(t)$ – значення різниці поточної $DC(t)$ та попередньої $DC(t-1)$ компоненти трансформанти; $l(\Delta DC(t))$ – довжина двійкового представлення різниці $\Delta DC(t)$ компоненти; $l(DC(t))$ – довжина двійкового представлення вихідної $DC(t)$ компоненти; $[\Delta DC(t)]_2$ – двійковий запис значення різниці $\Delta DC(t)$ компоненти; $[l_i]_2$ – двійковий запис основного коду; $[d_i]_2$ – двійковий запис додаткового коду; l_i – довжина основного коду; d_i – довжина додаткового коду.

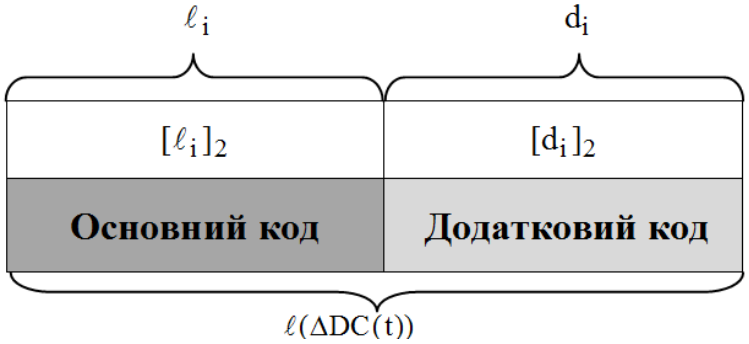


Рис.5. Структура коду низькочастотної компоненти

Другий етап. Тут відбувається кодування значущих субсмуг частотного спектра ДКП в умовах його неоднорідності. Формування кодового опису пропонується здійснювати на базі побудови кодових конструкцій для послідовності неоднорідних значень спектру ДКП. Величина $\Delta V(y)_j$ визначається як кількість допустимих послідовностей які передували вектору $\Delta Y(m-j)$, а саме:

$$\Delta V(y)_j = \begin{cases} y_j\,(w(y)-1)^{(m-j-1)} - \Delta V(y'_j = y_{j-1}), \to y_{j-1} < y_j; \\ y_j\,(w(y)-1)^{(m-j-1)}, \to y_{j-1} > y_j. \end{cases}$$

Введемо допоміжну величину μ_j, яка дорівнює

$$\mu_j = \begin{cases} y_j, & \rightarrow y_j < y_{j-1}; \\ y_j - 1, & \rightarrow y_j > y_{j-1}. \end{cases}$$

В результаті чого, співвідношення для коду $E(y)_u$ вектора значущих субсмуг трансформанти прийме вигляд

$$E(y)_u = \sum_{j=2}^{m} \mu_j \, (w(y) - 1)^{(m-j-1)}.$$

При початкових умовах, що задаються співвідношеннями

$$y'_0 = w(y) > y_2, \; y'_0 = w(y),$$

де $y_j \, (w(y) - 1)^{(m-j-1)}$ – сумарна кількість послідовностей для усіх елементів котрих, крім j-го виконуються обмеження на діапазон та неоднорідність сусідніх субсмуг; $\Delta V(y'_j = y_{j-1})$ – кількість заборонених послідовностей, які передують послідовності яка кодується $\Delta Y(m - j)$; $(w(y) - 1)^{(m-1)}$ – кількість послідовностей неоднорідного спектру ДКП [2; 9].

В результаті виключення послідовностей, що містять однорідні сусідні субсмуги, досягається усунення структурної надмірності без внесення викривлень. При цьому усунення надмірності забезпечується навіть в тих випадках, коли динамічний діапазон високочастотних компонент трансформанти прагне до динамічного діапазону низькочастотної компоненти, тобто

$$y_j \rightarrow y_1.$$

На *третьому етапі* організується кодування вектора масштабуючих компонент, який складається з трьох складових [3]

$$G_{m-1} = \{G_1^{(lg)}; G_2; g_m\},$$

де кожна складова, визначається виразами:

$$G_1^{(lg)} = \{g_1, ..., g_{l_g+1}\}, \text{ де } g_\xi = 0, \; \xi = \overline{1, l_g}, \; G_2 = \{g_{l_g+1}, ..., g_{m-1}\},$$

Для кроку квантування, відповідного необхідній якості реконструкції зображення, довжина першої серії нулів буде зростати. У зв'язку, з чим пропонується даний параметр представляти з використанням кодування довжинами серій, як показано на рис. 6.

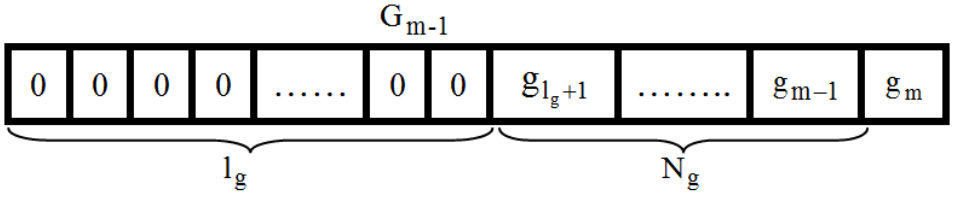


Рис.6. Структура вектору масштабуючих компонент

Перша частина коду E_1 вектора масштабуючих компонент та його довжина $l(E_1)$ буде визначатися за формулами:

$$E_1 = l_g,\ l(E_1) = [\log_2 l_g] + 1 \text{ (біт)},$$

де $G_1^{(l_g)}$ – кількість перших нульових елементів вектору G_{m-1}; G_2 – елементи вектору G_{m-1} за виключенням першої та останньої серії нулів; g_m – остання компонента вектору G_{m-1}.

Для компактного представлення, другого структурного параметра вектору маштабуючих компонент, пропонується використовувати код Бодо. Суть кодування, якого полягає у наступному:

1) визначається максимальний динамічний діапазон, за формулою:

$$W(G_2) = \max_{l_g+1 \le \xi \le m-1} \{g_\xi\} + 1.$$

2) обчислюється ціле кількість біт d_1, яка необхідна для подання чисел з виявленим динамічним діапазоном, використовуючи вираз:

$$d_1 = [\log_2 W(G_2)] + 1.$$

3) формується кодова конструкція так, що під кожну її компоненту буде виділятися рівна кількість двійкових розрядів, що визначається нерівністю:

$$l([g_\xi]_2) = d_1 \text{ (біт)}.$$

Структура компактно-представленої другої складової вектора маштабуючих компонент представлена на рис. 7.

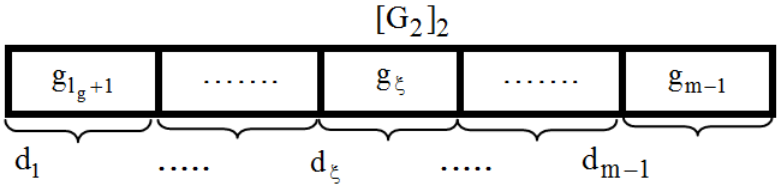


Рис.7. Структура компактно-представленого підвектора G_2 в двійковому вигляді

Таким чином, вектор масштабуючих компонент замінюється двома кодовими складовими вираз (4) та представлений на рис. 8.

$$G_{m-1} \to \{l(E_1); (d_1, ..., d_\xi, ..., d_{m-1})\}. \quad (4)$$

Граф-схема способу кодування вектора масштабуючих компонент представлено на рис. 9.

Таким чином, розроблено модель опису вектора масштабуючих компонент, який, задає масштаб значущих субсмуг в частотному спектрі ДКП. Розроблено технологію кодування масштабуючих компонент, з використанням коду Бодо та коду довжини серій, для зниження сумарних витрат на їх представлення.

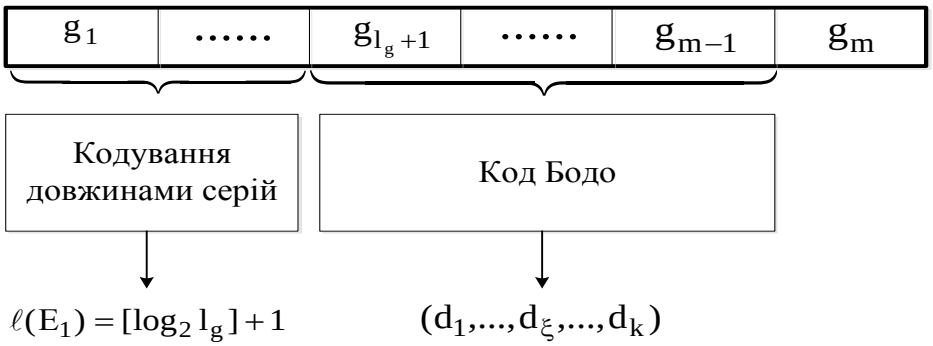


Рис.8. Структурна схема кодових складових вектору G_{m-1}

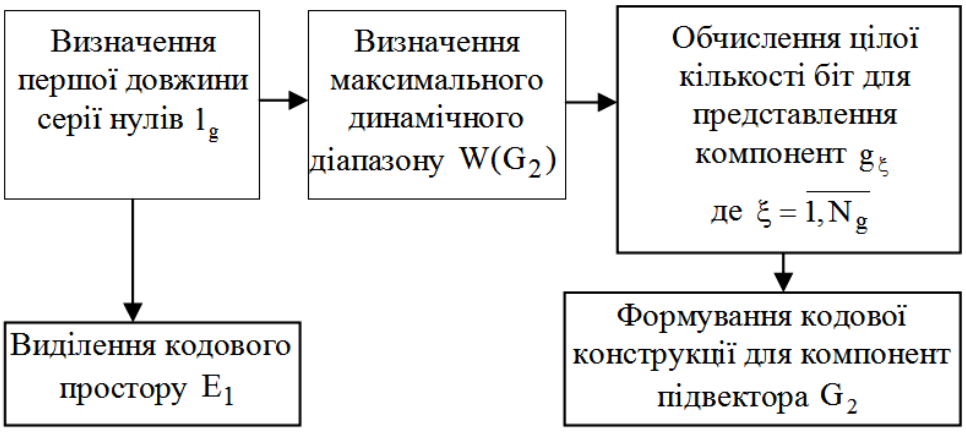


Рис.9. Граф-схема способу кодування вектора G_{m-1}

2. Метод відновлення компонент трансформанти в технології реконструкції відеоданих

Відновлення компонент трансформант включає в себе наступні етапи. Спочатку відбувається реконструкція низькочастотної DC компоненти та послідовності неоднорідних значущих субсмуг спектра ДКП, що представлено на рис. 10 [5;7;10].

На рис. 10 $E(y)_u$ – кодове значення послідовності значень субсмуг неоднорідного спектру ДКП; μ_j – допоміжна величина; $w(y)$ – діапазон субсмуг вектора Y_{m-1}; m – кількість субсмуг; j – субсмуга вектора Y_{m-1}.

На останньому етапі відбувається відновлення масштабуючих компонент рис. 11 [4].

При відновленні низькочастотної DC компоненти, кодер використовує статистичний код, який зберігається в двійковому вигляді в спеціальних кодових таблицях.

Відновлення послідовності значень субсмуг неоднорідного спектру ДКП, здійснюється в два етапи:

На першому етапі, метод відновлення значущих субсмуг полягає, в декодуванні кодового значення $E(y)_u$ послідовності значущих субсмуг неоднорідного спектра ДКП. Це дозволить отримати значення вектора

значущих субсмуг Y_{m-1}. Тут використовується інформація про діапазон субсмуг $w(y)$, кількості елементів, а також про нульовий елементі послідовності значень неоднорідного спектра ДКП, який дорівнює $y_0 = w(y)$ Відновлення елементів вектора значущих субсмуг організовується наступними діями.

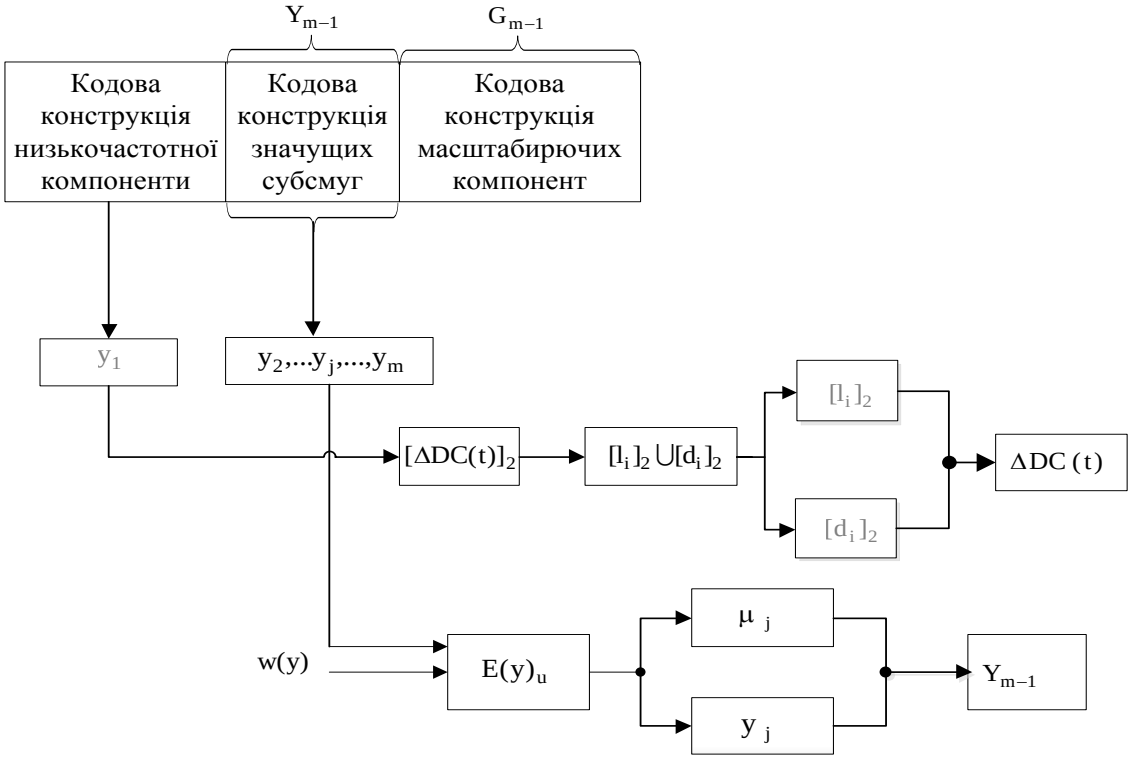


Рис.10. Загальна схема реконструкції низькочастотної DC компоненти та вектору Y_{m-1}

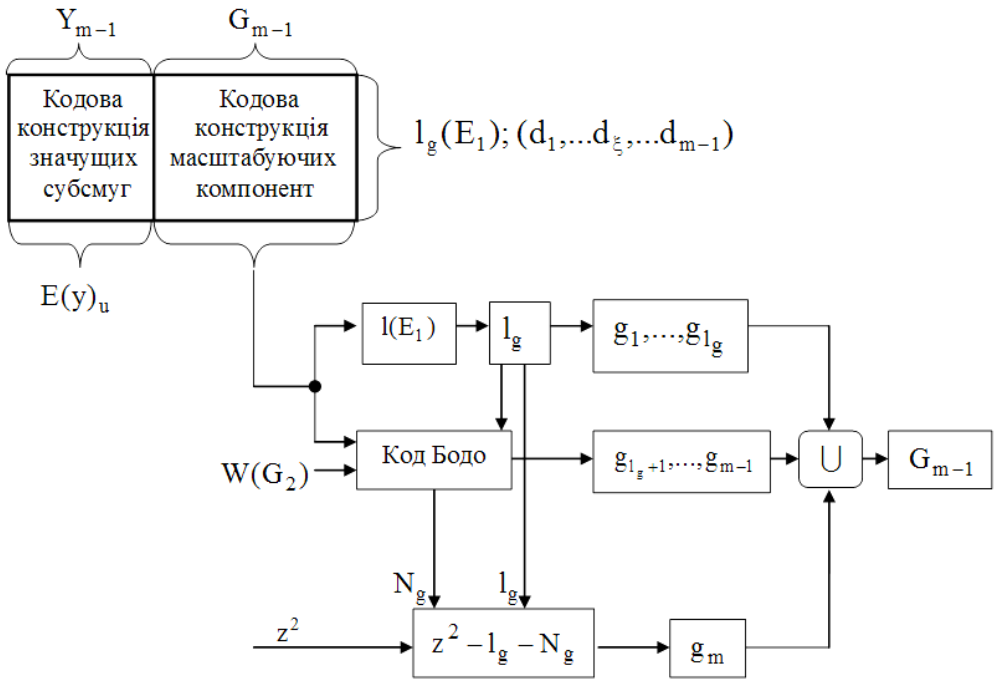


Рис.11. Структура схема відновлення компонент у векторі G_{m-1}

Відбувається відновлення допоміжних величин μ_j, на які накладається тільки одне обмеження, що визначається нерівностями:

$$\mu_j < w(y), \quad \text{якщо} \quad j=1;$$

$$\mu_j < (w(y)-1), \quad \text{якщо} \quad j=\overline{2,m}.$$

Подальше їх відновлення здійснюється за формулами:

$$\mu_1 = [E(y)_u / (w(y)-1)^{(m-1)}], \tag{5}$$

$$\mu_j = [E(y)_u / (w(y)-1)^{(m-j)}] - [E(y)_u / (w(y)-1)^{(m-j)+1}](w(y)-1), \tag{6}$$

$$j=\overline{2,m}.$$

На другому етапі, відбувається відновлення елементів y_j субсмуг вектора, на основі отриманих на попередньому етапі допоміжних величин μ_j, що задається наступною системою:

$$y_j = \begin{cases} \mu_j, & \rightarrow \mu_j < y_{j-1}; \\ \mu_j + 1, & \rightarrow \mu_j \geq y_{j-1}. \end{cases} \tag{7}$$

Об'єднавши вирази (5) - (7) отримаємо систему аналітичних співвідношень для відновлення послідовності неоднорідних значень спектра ДКП, а саме:

$$y_j = \begin{cases} [E(y)_u / (w(y)-1)^{(m-j)}] - [E(y)_u / (w(y)-1)^{(m-j)+1}](w(y)-1), \rightarrow \mu_j < y_{j-1}; \\ [E(y)_u / (w(y)-1)^{(m-j)}] - [E(y)_u / (w(y)-1)^{(m-j)+1}](w(y)-1)+1, \rightarrow \mu_j \geq y_{j-1}, \end{cases}$$

де $(w(y)-1)^{(m-j)}$ – діапазон елемента μ_j.

Тут забезпечується відновлення послідовності неоднорідних значень спектра ДКП без внесення помилок.

Реконструкція вектора масштабуючих компонент формується з трьох етапів.

Перший етап. Перша складова вектора масштабуючих компонент реконструюється в результаті зчитування десятинного числа в двійковому записі $[l_g]_2$, тобто:

$$[l_g]_2 \xrightarrow{l(E_1)} l_g,$$

або

$$l_g = \sum_{\xi=1}^{l(E_1)-1} \alpha_\xi \cdot 2^\xi, \quad [l_g]_2 = \{\alpha_1, ..., \alpha_\xi, ..., \alpha_{l(E_1)}\},$$

де α_ξ – двійковий елемент послідовності $[l_g]_2$; ξ – індекс двійкового елемента $[l_g]_2$.

Це дозволить отримати першу складову E_1 вектора маштабуючих компонент G_{m-1}, що задається наступним співвідношенням:

$$g_j = 0,\ j = \overline{1, l_g},\ E_1 = \{g_1, ..., g_j, ..., g_{lg}\}.$$

Таким чином, розроблений спосіб відновлення першої серії нульових компонент вектора масштабуючих компонент, на основі декодування двійкового коду її довжини.

Другий етап процесу відновлення здійснюється реконструкція другої структурної складової E_2 вектора масштабуючих компонент трансформанти. Для цього необхідна службова інформація про максимальне значення динамічного діапазону послідовності $g_{l_g}, ..., g_{m-1}$ (рис. 12).

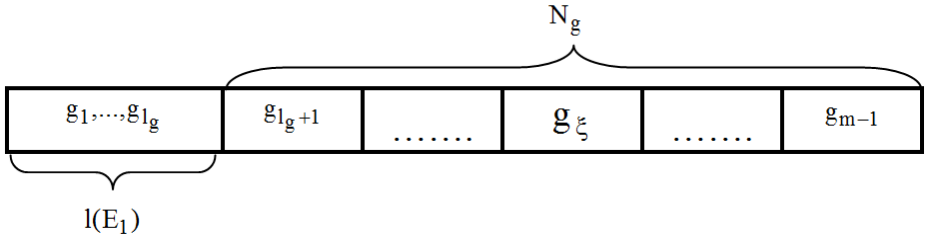


Рис.12. Структурна схема складових l_g *та* N_g *у векторі* G_{m-1}

Оскільки, простий код Бодо є рівномірним, то визначення позиції компонент другої складової, здійснюється за наступним правилом:

$$P_1(E_2) = l(E_1) + 1,$$
$$P_2(E_2) = l(E_1) + 1 + d_1,$$
$$P_\xi(E_2) = P_{\xi-1}(E_2) + d_1,$$

де P_1, P_2, P_ξ – відповідні позиції компонент, другої структурної складової вектору G_{m-1}; $l(E_1)$ – довжина першої частини коду вектора G_{m-1}; d_1 – ціла кількість біт необхідних на представлення компоненти, другої структурної складової вектора G_{m-1} кодом Бодо.

Після чого відбувається відновлення відповідного значення

$$g_\xi = \sum_{\xi=0}^{N_g-1} \alpha_\xi \cdot 2^\xi.$$

Кількість відновлених компонент, визначається на основі відомої довжини послідовності значущих субсмуг неоднорідного спектра ДКП, що задається формулою

$$N_g = (m-1) - l_g.$$

Третій етап. Структурний параметр g_m, вектора масштабуючих компонент G_{m-1}, являє собою, довжину останньої серії елементів, які мають нульові значення. Параметр g_m вектора G_{m-1} масштабуючих компонент на прийомній стороні, визначається на основі перших двох

структурних складових вектору G_{m-1}. Для цього використовується наступна умова:

$$\sum_{\xi=1}^{m-1} g_\xi = l_g + N_g .$$

З урахуванням чого, величина g_m буде обчислюватися за наступним виразом:

$$g_m = (z_1 \cdot z_2) - l_g - N_g,$$

де $z_1 \cdot z_2$ – лінійний розмір трансформанти.

Отже, знаючи розмір трансформанти та кількість компонент двох структурних параметрів l_g, N_g вектора масштабуючих компонент G_{m-1}, кодувати третю складову g_m немає необхідності (рис. 13).

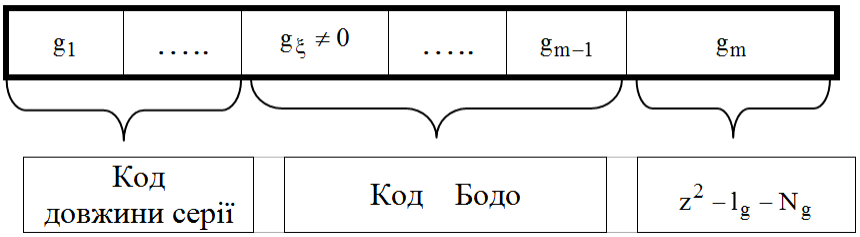


Рис.13. Структурна схема складових вектора G_{m-1}

Таким чином, розроблена технологія реконструкції зображень, стиснутих з використанням попередньої трансформації.

3. Оцінка компресійних характеристик розробленого методу стиску

Сумарна кількість розрядів, необхідна на компактне представлення всього стислого зображення, визначається виразом

$$D_{\text{ст.зоб}} = \sum_{\xi=1}^{\eta} D_{\Sigma_{\text{ст.т},\xi}} = \sum_{\xi=1}^{\eta} D^{(y)}_{m\text{-}1,\,\xi} + D^{(g)}_{m\text{-}1,\,\xi} + D_{y1,\,\xi} + D_{\text{сл},\,\xi},$$

складові виразу визначаються формулами

$$D'^{(y)}_{m\text{-}1} = [\ell\log_2(w(y)+1) + (m-2)\ell\log_2 w(y)] + 1 \text{ (біт)},$$

$$D_{\text{сл}} = D_m + D^{(y)}_w \text{ (біт)},$$

$$D^{(g)}_{m\text{-}1} = ([\log_2 l_g]+1) + ([\log_2 W(G_2)]+1) \text{ (біт)},$$

$$D_{(y1)} = l_i + d_i = l(\Delta DC(t)) \text{ (біт)}.$$

Звідки, коефіцієнт стиснення для зображення стисненого розробленим методом, визначається за формулою

$$k_{\text{ст}} = M \times N \times b / \sum_{\xi=1}^{\eta} D_{\Sigma_{\text{ст.т},\xi}} .$$

де $D_{m-1}^{(y)}$– кількість розрядів на представлення вектору Y_{m-1}; $D_{m-1}^{(g)}$ – кількість розрядів на представлення вектору G_{m-1}; $D_{(y_1)}$ – кількість розрядів на представлення компоненти y_1; $D_{сл}$ – кількість розрядів на представлення службової складової трансформанти; D_m – необхідна кількість розрядів на представлення інформації о знаках компонент трансформанти; $D_w^{(y)}$– кількість розрядів, відносно величини динамічного діапазону вектору Y_{m-1}; $D_{ст.зобр.}$ – кількість розрядів необхідних на компактне представлення усього стиснутого зображення; $\eta = M \times N / z_1 \times z_2$ – кількість трансформант в зображенні; $D_{\sum ст.т,\xi}$ – обсяг стиснутої ξ-ї трансформанти зображення; b – глибина оцифрування пікселю.

Співвідношення компресійних характеристик для різного ступеня насиченості в розробленому методі стиснення представлено на рис. 14 [8].

З аналізу рис. 14 слідує, що в залежності від ступеня насиченості та кроку квантування коефіцієнт стиснення приймає значення від 2,48 до 15,92 разів.

Залежність розподілу обсягів на представлення значимих субсмуг та масштабуючих компонент трансформант в залежності від ступеню насиченості, представлено на рис. 15.

Для слабонасичених та середньонасичених зображень найбільшу частину обсягу трансформанти склали значущі субсмуги до 66%, для сильнонасичених зображень становлять масштабуючі компоненти до 58%.

Порівняльна оцінка розробленого методу стиску з методом JPEG для середньонасичених зображень, щодо коефіцієнта стиснення наводиться на рис.16.

З аналізу випливає, що залежно від ступеня насиченості зображень та кроку квантування отриманий виграш за ступенем стиснення, складає від 5 до 20 % для зображень оброблених на базі розробленого методу.

При проведенні порівняльного аналізу, виграш за часом передачі зображень розміром 2048×1080, стиснутих розробленим методом по відношенню до метода JPEG, складає для швидкості передачі 256 Кбіт/с в середньому від 5 до 15 %, для 16 Мбіт/с - в середньому від 5 до 20%.

Розроблений метод стиснення, щодо стандарту JPEG, забезпечує скорочення кількості операцій на компресію зображення в середньому від 2 до 4 разів. Це досягається за рахунок скорочення кількості операцій складання/віднімання від 2,5 до 3 разів та операцій множення/ділення від 1,1 до 2 разів. При цьому відновлений образ зберігає високу якість насиченості деталями та переходами між фрагментами зображення.

Таким чином, для запропонованого методу стиску, заснованого на обробці компонентної структури трансформанти, досягається зниження

обсягів відеоданих та забезпечується підвищення продуктивності функціонування інфокомунікаційних систем із заданою якістю відеосервісу.

Рис.14. Зміна коефіцієнта стиснення для різних класів зображень при кроці квантування R=1 та R=5

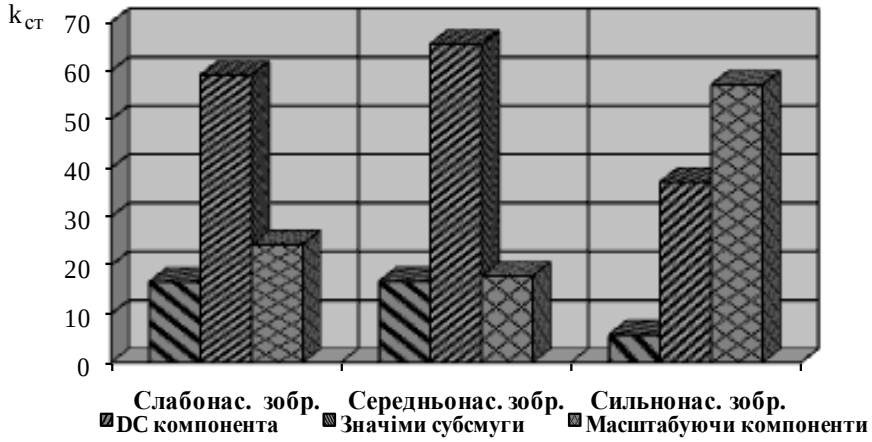


Рис.15. Процентне співвідношення кількості компонент трансформанти для різних класів зображень при кроці квантування R=3

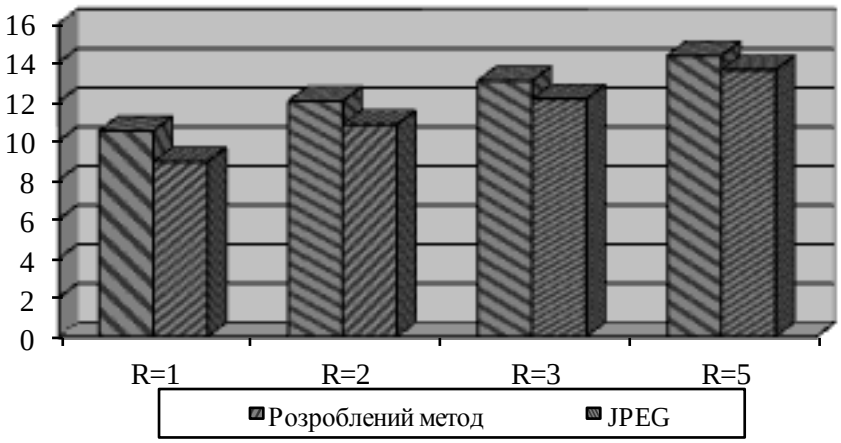


Рис.16. Співвідношення коефіцієнта стиснення розробленого методу та JPEG залежно від кроку квантування для середньоенасичених зображень

Висновки

Вирішено актуальну *науково-прикладну задачу*, яка полягає в зниженні обсягів відеоданих для підвищення продуктивності функціонування інфокомунікаційних систем із заданою якістю відеосервісу. Розроблено метод стиснення трансформованих зображень на основі обробки послідовності значущих субсмуг неоднорідного спектру ДКП.

Основними науковими результатами є:

1. Підхід для побудови технології компресії зображень з використанням попередньої трансформації, який базується на:

1) формуванні двох складових трансформанти, а саме: вектору значущих субсмуг та вектору масштабуючих компонент.

2) опису вектору значущих субсмуг трансформанти у вигляді послідовності значень субсмуг з неоднорідними сусідніми елементами. Це дозволить адаптуватися до властивостей лінеаризованих трансформант за рахунок врахування: неоднорідності значень сусідніх субсмуг спектру ДКП; обмеженого діапазону субсмуг трансформанти.

2. Модель оцінки інформативності трансформант враховуючи нерівномірності розподілу діапазонів субсмуг. На основі чого обґрунтовано, що трансформанта має структурну надмірність.

3. Обґрунтування того, що у результаті кодування неоднорідного спектру ДКП скорочується структурна надмірність, яка викликана з одногу боку когерентністю областей зображення, а з іншого боку – наявністю анізотропних властивостей зображення.

4. Метод кодування компонентної структури трансформанти який базується на формуванні кодового опису на базі побудови кодових конструкцій для послідовностей неоднорідних значень спектра ДКП.

5. Метод реконструкції вектору значущих субсмуг неоднорідного спектру ДКП, та вектору маштабуючих компонент трансформант.

Наукова новизна отриманих результатів полягає у тому, що:

1. Отримала подальший розвиток модель оцінки інформативності трансформованих зображень. Відмінна особливість моделі полягає в тому, що кількість інформації, оцінюється на основі виявлення структурних закономірностей для вектора значущих субсмуг трансформанти, а саме виключення послідовностей, які містять однорідні сусідні субсмуги трансформанти. Це дозволяє оцінити граничні межи кількості надмірності, яка буде усуватись.

2. Створено метод стиснення трансформованих відеоданих на основі кодування їх компонентного опису. Відмінна риса методу полягає в кодуванні послідовності значущих субполос неоднорідного частотного спектра дискретного косинусного перетворення (ДКП). Це забезпечує підвищення продуктивності інфокомунікаційних систем щодо обробки та передачі відеоданих.

3. Отримала подальший розвиток технологія реконструкції трансформант, яка характеризується тим, що запропоновано враховувати її поділ на вектори значущих субсмуг та масштабуючих компонент. Це дає змогу відновлювати зображення з контрольованою якістю візуального сприйняття.

4. Вперше розроблено метод реконструкції зображень на основі зворотного трансформування, який відрізняється від відомих тим, що використовується декодування значень значущих субполос послідовності неоднорідного спектру ДКП за відомим кодом і основою. Це дає змогу отримати базову інформацію про вихідне зображення без внесення втрат інформації.

Література

1. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.
2. Баранник В.В. Кодирование значимых компонент трансформант / В.В. Баранник, В.Н. Кривонос, А.В Хаханова // Радиоэлектроника и информатика. – 2012. – № 2 (57). – С. 32-35.
3. Кривонос В.Н. Метод компактного представления вектора масштабирующих компонент трансформант/ В.Н. Кривонос, Н.К. Гулак, М.В. Думанский // Сучасна спеціальна техніка, 2012. – № 3(30). – С. 28 – 33.
4. Кривонос В.М. Метод декодування вектора масштабованих компонент трансформанти / В.М. Кривонос // Наукоємні технології, 2012. – №3(15). – С. 90 – 93.
5. Кривонос В.М. Метод восстановления значимых компонент трансформант в технологии реконструкции видеоинформации / В.Н. Кривонос, А.В. Хаханова // Сучасна спеціальна техніка, 2013. – № 1(32). – С. 46 – 50.
6. Баранник В.В. Структурная модель информативности значимых компонент трансформант / В.В. Баранник, В.Н. Кривонос, А.В. Хаханова // Інформаційно-керуючі системи на залізничному транспорті, 2013. – №2(99). – С. 26 - 29.
7. Баранник В.В. Технология реконструкции кадров видеоинформационного потока в телекоммуникационной сети / В.В. Баранник, В.Н. Кривонос // Захист інформації, 2013. – № 3. – С. 196 – 203.
8. Кривонос В.Н. Метод оценки вычислительной сложности обработки изображений с выявлением значимых компонент трансформант/ Бекиров А.Э., Думанский М.В // Сучасна спеціальна техніка, 2013. – № 3(34). – С. 40 – 44.
9. Barannik V.V. Coding Tangible Component of Transforms to Provide Accessibility and integrity of Video Data / V.V. Barannik, A.V. Hahanova, V.N. Krivonos // International Symposium [″IEEE East-West Design & Test″], (Kharkov, Ukraine, September 14 – 17, 2012) / Kharkov, 2012. – P. 475 – 478.
10. Barannik V. Method Encoding Videoinformation for Increase Avaiabiity in Informatively-Telecommunication Systems / V. Barannik, A. Hahanova, V. Krivonos // XIIth International Conference [“The Experience of Designing and Application of CAD Systems in Microelectronics”], (Lviv – Polyana, Ukraine, February 19 – 23, 2013) / Lviv – Polyana, 2013. – P. 23 – 24.

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ОБРОБКИ КЛАСТЕРИЗОВАНИХ ЗОБРАЖЕНЬ З ВРАХУВАННЯМ СТУПЕНЯ СЕМАНТИЧНОЇ НАСИЧЕНОСТІ В ПРОЦЕСІ АЕРОМОНІТОРИНГУ

Бараннік В.В., Мусієнко О.П., Стасев С.Ю.

Вступ

Актуальність досліджень. Аналіз підготовки і ведення операцій, локальних війн, а також характерна риса сучасних і майбутніх конфліктів, свідчить про характер ведення бойових дій. Зростаюча динаміка і швидкість військових дій, підвищення засобів ураження збройної боротьби і маневрених можливостей військ призводить до розширення масштабу операцій, а також перетворення морського, повітряного, космічного та сухопутного простору в єдиний глобальний театр воєнних дій. Так, в першу чергу, виникає суттєва потреба в отриманні своєчасних і достовірних даних як про свої об’єкти (сили) та засоби, так і про засоби супротивника в єдиному інформаційному просторі. В теперішній час ситуація, що склалася на південному сході України, з позиції інформаційного ресурсу системи управління обумовлена необхідністю задіяти безпілотні комплекси. При цьому потрібно враховувати, що близько 80% усієї інформації отримується шляхом використання безпілотних літальних апаратів (БПЛА). Це надає можливість підвищити маневреність систем збору та обробки інформації та своєчасність її доставки.

Серед сучасних БПЛА іноземного виробництва можна виділити наступні моделі: "Pioneer", "WASP 3", "RQ - 11 Raven" - США; "BirdEye 500", Skylark - I LE - Ізраїль; "Spy Arrow" - Франція; "Орлан-10", "Форпост", "Дозор-100" - Росія. Серед українських розробок БПЛА, у тому числі і для потреб армії, можна виділити наступні моделі: А1-С "Фурія", "Spectator-M", "Patriot RV010", "Сокол-2", "Microvisor SM7", "UaViter", М-7Д "Небесний патруль", М-6-3 "Жайвір", М-10 "ОКО" та ін. Таким чином, на сьогодні у світі налічується більш ніж 100 фірм в 35 країнах, які займаються виробництвом і модернізацією безпілотних літальних систем [1]. Це свідчить про те, що безпілотні літальні комплекси стали невід'ємною частиною озброєння сучасних армій.

Завдяки сучасним БПЛА, які оснащені фотографічною, радіо- і радіотехнічною, інфрачервоною та іншою апаратурою, в процесі аеромоніторингу добуваються дані про різноманітні свої об'єкти, так і про об’єкти супротивника в різних кризових ситуаціях. Одним із затребуваних ресурсів інформації є аерофотознімки, які формуються в результаті ведення аеромоніторингу БПЛА.

Кінцевим етапом аеромоніторингу є дешифрування всього аерофотознімка та складання звіту про наземну обстановку і прийняття рішення. Основним завданням дешифрування є виявлення і розпізнавання

на аерофотознімках об'єктів місцевості, визначення їх якісних і кількісних характеристик. Виходячи з цього оперативність управління з використанням БПЛА залежить від часових витрат на обробку і передачу аерофотознімка, на зворотну обробку і дешифрування аерофотознімка. У теж час ефективність дешифрування залежить від роздільної здатності аерофотознімка і від ступеня збереження інформації про ключові ознаки дешифрування (яскравісна і контурна інформація). При цьому відбувається значний ріст інформаційної інтенсивності аерофотознімків. Передача аерофотознімків з борту БПЛА на наземний комплекс здійснюється по низькошвидкісним (обмеженим) радіолініях передачі інформації. Це призводить до різкого зростання часових затримок на доставку аерофотознімків. Тому для скорочення часу їх доставки застосовуються технології зменшення інформаційної інтенсивності, серед яких такими, що частіше застосовуються, є методи на основі JPEG-платформи [2, 3]. Більшість цих методів базуються на попередньому трансформуванні аерофотознімків, внаслідок чого відбувається скорочення як психовізуальної, так і статистичної надмірності. Вагоме зменшення інформаційної інтенсивності досягається шляхом скорочення психовізуальної надмірності. При цьому, скорочення інтенсивності досягається за рахунок втрати інформації та, як наслідок, відбувається руйнування інформації про ключові ознаки дешифрування, а також зниження достовірності отриманих аерофотознімків.

Отже, існує протиріччя, яке обумовлене тим, що з одного боку, для підвищення оперативності прийняття рішення необхідно зменшувати інформаційну інтенсивність аерофотознімків. З іншого боку, для підвищення правильності прийняття рішення в процесі дешифрування необхідно підвищувати роздільну здатність аерофотознімка та зберігати інформацію про дешифрувальні ознаки, що веде до збільшення інформаційного навантаження на канал передачі даних. Тому, підвищення оперативності доставки інформації при заданій якості дешифрування аерофотознімків з використанням безпілотних літальних апаратів аеромоніторингу*, є актуальним.*

Метою дослідження є розробка інформаційної технології щодо обробки аерофотознімків для підвищення оперативності доставки інформації в умовах забезпечення заданої якості дешифрування на основі семантичної класифікації та структурного кодування.

Основний матеріал

Технологія оцінки семантичної насиченості аерофотознімків. Висококонтрастні ділянки зображення, перепади яскравості, великі області зображення, лінійні розміри об'єктів складають найбільш інформативну частину аерофотознімка, тобто текстурні області і контури. При обробці таких областей слід враховувати безліч деталей: масштаб і роздільну

здатність (детальність) аерофотознімків, кількість і розмір об'єктів і т.д. Тут особливу увагу необхідно приділити вибору семантичної насиченості в даних аерофотознімках. Під семантичною насиченістю аерофотознімка розуміється найбільш значима інформація про контури та границі об'єктів місцевості.

Для розробки методу оцінки семантичної насиченості аерофотознімків вимагається враховувати, як особливості функціонування бортового комплексу БПЛА, так і те, що основним класом даних, що формуються на борту, є зображення різної насиченості. Тому пропонується розглядати блоки аерофотознімка по ступеню семантичної насиченості, які будуть залежати від ступеня неоднорідності структур, а саме: "слабонасичені", "середньонасичені", "сильнонасичені" [4].

Для оцінки семантичної насиченості аерофотознімків пропонується оцінювати ступінь насиченості в блоках аерофотознімка. Для цього формується ієрархія класифікації блоків за ступенем семантичної насиченості, а саме: першу групу складають блоки з відносно однорідним вмістом структури, і є висококогерентними; другу групу складають блоки з неоднорідною структурою, є низькокогерентними. Пропонуємо, класифікувати ці блоки аерофотознімка на дві загальні групи, що представлено на рис. 1.

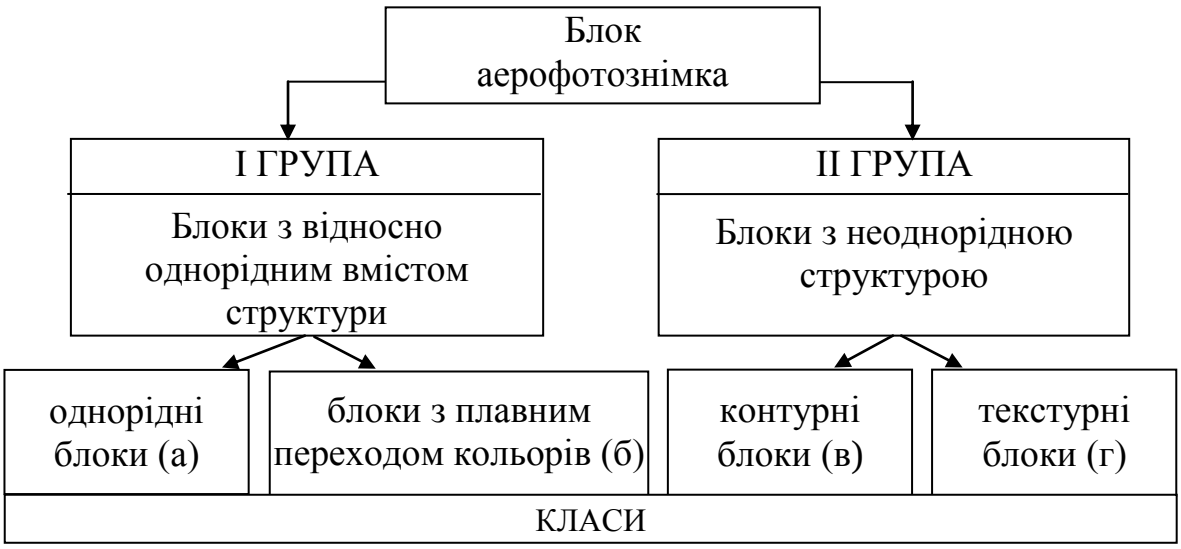
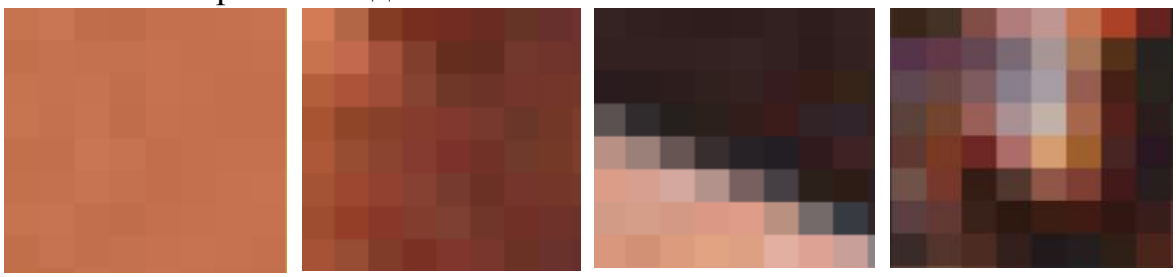


Рис. 1. Ієрархічна класифікації семантичної насиченості блоків аерофотознімка

Тут класифікуємо отримані групи блоків аерофотознімка на класи, що представлено на рис. 2. На підставі класифікації отримаємо: однорідні блоки (рис. 2 а); блоки з плавним переходом кольорів (рис. 2 б); контурні блоки (рис. 2 в); текстурні блоки (рис. 2 г).

Для підвищення ефективності класифікації фрагмента пропонується використовувати двобазисний принцип, що охоплює його просторово-часове та спектральне подання.



а б в г

Рис. 2. Приклад блоків аерофотознімку різних класів: клас 1 - однорідні блоки (однорідні блоки або блоки з плавним переходом кольорів), елементи зображення якого близькі або однакові за кольором (рис. 2 а, б); клас 2 - контурні блоки, які можна розділити на дві області з різким перепадом кольору між елементами зображенні (рис. 2 в); клас 3 - текстурні блоки, в яких є присутні різкі перепади кольорів елементу зображення на деякій локальній ділянці (рис. 2 г).

Відповідні показники для оцінки рівня насиченості фрагмента задаються наступними виразами:

– показник насиченості $P_{ДКП}$, який визначається за низькочастотними коефіцієнтами трансформанти, отриманої в результаті дискретного косинус-перетворення, і представлений виразом:

$$P_{ДКП} = [\log_2 (\prod_{\gamma=1}^{D_d} \prod_{\xi=1}^{N_y} (y_{\gamma,\xi}))], \quad (1)$$

де $y_{\gamma,\xi}$ - частотний коефіцієнт трансформанти на позиції з координатами (γ, ξ) відносно діагоналі D; D_d - кількість діагоналей в ознаковій зоні частотних коефіцієнтів; N_y - кількість частотних коефіцієнтів, які відповідають діагоналям в ознаковій зоні;

– структурний показник $P_{СТР}$, визначається по складовій яскравості вихідних блоків, що представлено виразом:

$$P_{СТР} = \left[\log_2 \left(\prod_{i=1}^{n} \left(p_{k,max} - p_{k,min} \right) \right) \right], \quad (2)$$

де $p_{k,max}$, $p_{k,min}$ - відповідно максимальне та мінімальне значення яскравості елемента строки блоку.

З урахуванням оцінки рівня семантичної насиченості фрагментів, поділ їх на три класи пропонується організовувати на базі кластерного аналізу. Структурна схема організовується на основі технології К-середніх з урахуванням двобазисного принципу оцінки рівня насиченості, що показано на рис. 3.

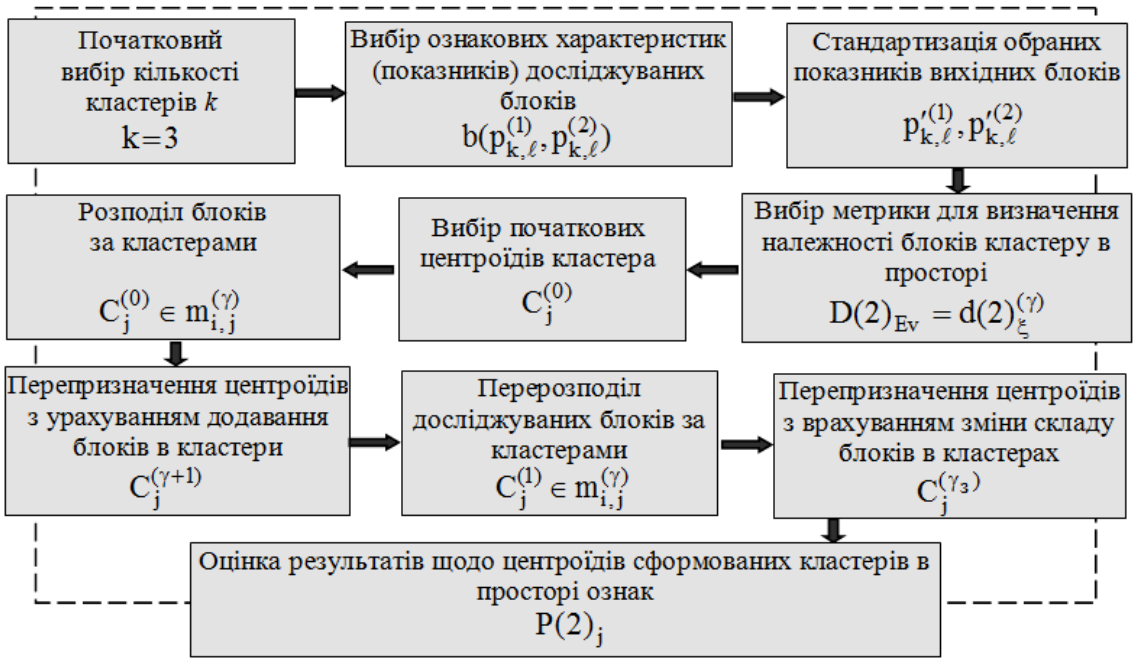


Рис. 3. Структурна блок-схема методу кластеризації блоків аерофотознімка, на основі алгоритму К-середніх

Вибір алгоритму К-середніх пов'язаний з тим, що, по-перше, цей алгоритм кластеризації простий в реалізації та має велику швидкість обробки вихідних даних (у нашому випадку, множина блоків аерофотознімка), по-друге, процес кластеризації блоків аерофотознімка може відбуватися не за однією характеристикою блоку, а за рахунок декількох ознакових характеристик (в нашому випадку показник насиченості та структурний показник блоку) [4, 5].

Тут проводиться кластеризація блоків аерофотознімка в єдиному ознаковому просторі.

В результаті процесу кластеризації блоків аерофотознімка, в ознаковому структурному просторі, сформовані кластери з розподіленими блоками за ступенем насиченості, а саме: слабонасичені, середньонасичені і сильнонасичені. Приклад кластеризації блоків аерофотознімка на три класи за ступенем семантичної насиченості на основі розробленого методу наведено на рис. 4.

Завдяки такій технології, ми зможемо отримати розподіл блоків аерофотознімків за кластерами, а також порівняти результати кластеризації блоків аерофотознімків, які отримані автоматичним способом (без участі оператора) з результатами, отриманими дешифрувальником, як експерта, ґрунтуючись на візуальному аналізі.

Таким чином, розроблений метод оцінки рівня семантичної насиченості фрагментів аерофотознімка, який дозволяє розподілити блоки за ступенем семантичної насиченості на кластери, тобто визначається приналежність блоків аерофотознімка на основі ознакових характеристик.

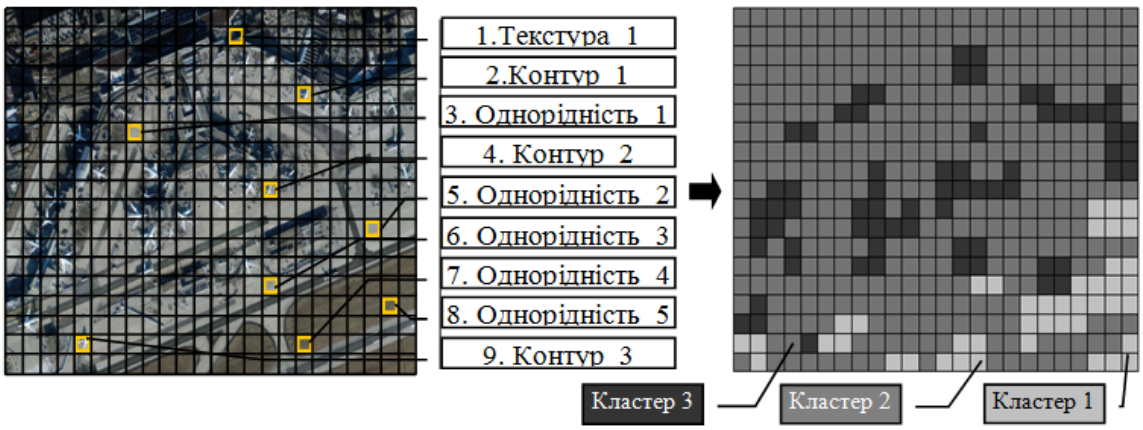


Рис. 4. Результати кластеризації блоків аерофотознімка "Аеропорт"

Поділ блоків аерофотознімків на класи семантичної насиченості дозволяє адаптувати процес скорочення надмірності відносно вимог по збереженню інформації про ключові ознаки дешифрування. Відповідно для додаткового виявлення структурних закономірностей пропонується використовувати компонентне ядро JPEG-платформи. Даний підхід заснований на виявленні послідовності $\hat{R}_{\nu_{вкр}}$ двохкомпонентних векторів $\Xi_{\chi}^{(2)} = \{q_{\chi}, b_{\chi}\}$, які базуються для лінеаризованої трансформанти ДКП відповідно як q_{χ} - довжина ланцюжка нульових компонент та b_{χ} - величини значимої компоненти [6, 8]. З особливостей, варто відзначити, що тут не враховується перший вектор, що відповідає значимій компоненті (низькочастотна компонента несе найбільшу інформацію про блок аерофотознімка) і останній вектор нульових компонент, тому що, як правило, відсутня наступна за нею значима компонента. Це представлено виразом: $\hat{R}_{\nu_{вкт-2}} = \{(q_2; b_2),\ldots,(q_{\chi}; b_{\chi}),\ldots,(q_{\nu_{вкт}-1}; b_{\nu_{вкт}-1})\}$. Після чого обробка сформованої послідовності $\hat{R}_{\nu_{вкр}}$ векторів здійснюється за двоієрархічним принципом з урахуванням виявлення структурних характеристик, як показано на рис. 5.

Тут на *першому рівні* для послідовності $\hat{R}_{\nu_{вкр}}$ двоелементних векторів $\Xi_{\chi}^{(2)}$ виявляються структурні закономірності, а саме враховується, що: по-перше компоненти вектора q_{χ} та b_{χ} мають різні значення, що представлено виразом:

$$q_{\chi} \neq b_{\chi}, \quad \chi = \overline{2, \nu_{вкт} - 1}, \tag{3}$$

а, по-друге незалежно один від одного приймають значення в межах динамічних діапазонів, згідно співвідношенню:

$$\{1 \le q_{\chi} \le \beta(q); \ \ 1 \le b_{\chi} \le \beta(b). \tag{4}$$

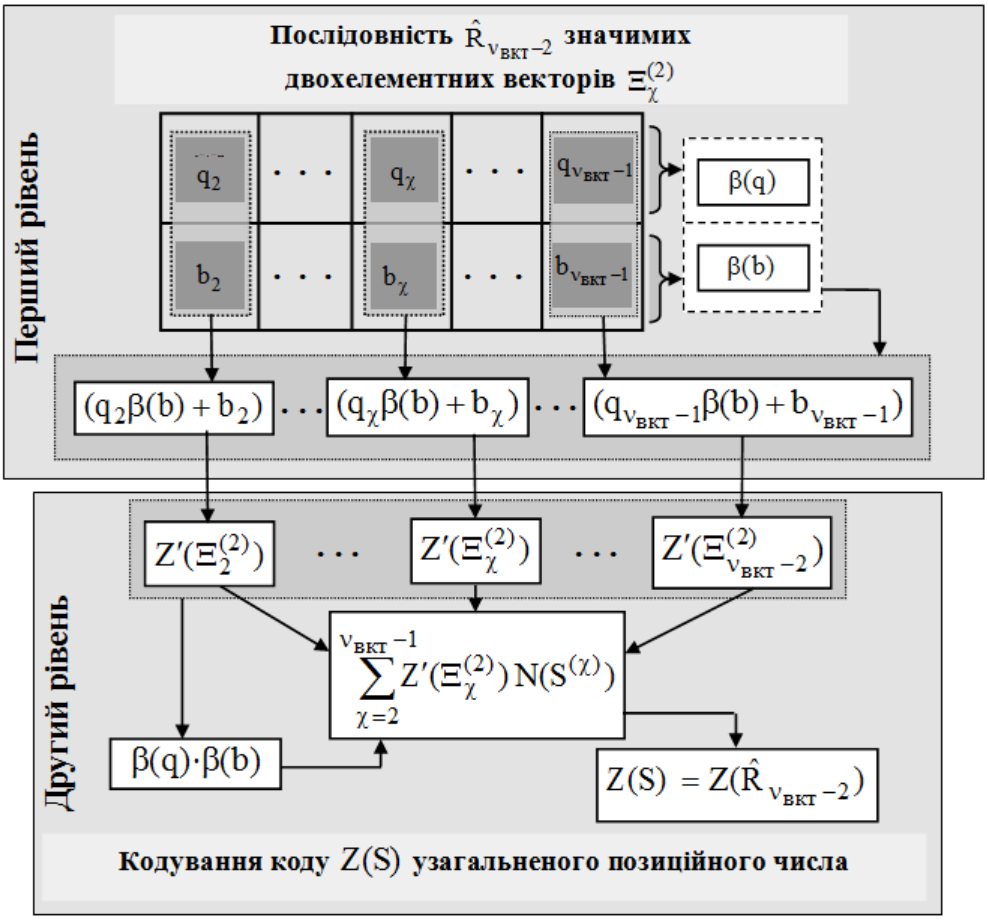


Рис. 5. Блок-схема поетапного двохієрархічного кодування послідовності значимих двоелементних векторів

Тоді з урахуванням таких властивостей, динамічний діапазон β(q) для першої компоненти q_χ вектора $\Xi_\chi^{(2)}$ буде дорівнювати:

$$\beta(q)=\max_{2\le\chi\le\nu_{ВКТ}-1}\{q_\chi\}+1,\ \text{т.к. } q_\chi\in[0;\max\{q_\chi\}], \qquad (5)$$

для другої компоненти b_χ вектора $\Xi_\chi^{(2)}$ динамічний діапазон β(b) буде визначатися як:

$$\beta(b)=\max_{2\le\chi\le\nu_{ВКТ}-1}\{b_\chi\},\ \text{т.к. } b_\chi\in[0;\max\{b_\chi\}]. \qquad (6)$$

Звідси двоелементний вектор $\Xi_\chi^{(2)}$, для компонент якого виконуються умови (5) і (6), називається двохелементним позиційним числом з нерівними сусідніми елементами.

В цьому випадку кодування таких векторів на першому рівні ієрархії представлено наступними виразами:

– значення коду $Z'(\Xi_\chi^{(2)})$ для двоелементного позиційного числа $\Xi_\chi^{(2)}$ визначається:

$$Z'(\Xi_{\chi}^{(2)}) = \eta(q_{\chi})V(q_{\chi}) + \eta(b_{\chi})V(b_{\chi}). \quad (7)$$

При цьому значення вагових коефіцієнтів $V(q_{\chi})$ та $V(b_{\chi})$ незалежно один від одного визначається такими виразами: $V(q_{\chi}) = \beta(b)$.

З огляду на те, що компонента b_{χ} є останньою в двохелементному векторі $\Xi_{\chi}^{(2)}$ компонент (відсутня наступна компонента), то ваговий коефіцієнт $V(b_{\chi})$ визначається: $V(b_{\chi}) = 1$.

– допоміжна величина η задається співвідношеннями:

$$\eta(q_{\chi}) = \begin{cases} q_{\chi}, & \to q_{\chi} < q_0; \\ q_{\chi} - 1, & \to q_{\chi} > q_0. \end{cases} \quad \eta(b_{\chi}) = \begin{cases} b_{\chi}, & \to b_{\chi} < q_{\chi}; \\ b_{\chi} - 1, & \to b_{\chi} > q_{\chi}. \end{cases} \quad (8)$$

Це дозволяє виключити послідовності, що містять рівні сусідні компоненти, тим самим досягається скорочення структурної надлишковості без внесення спотворень.

На *другому рівні* ієрархії будується кодове подання для всієї послідовності векторів. Враховуючи на те, що отримані на першому рівні ієрархії кодові значення мають обмеження, задані нерівністю:

$$Z'(\Xi_{\chi}^{(2)}) < \beta(q)\cdot\beta(b) \text{ для } \chi = \overline{2, \nu_{\text{вкт}} - 1}, \quad (9)$$

то утворені ними послідовності, називаються узагальненими позиційними числами, що представлено співвідношенням:

$$S = \{ Z'(\Xi_2^{(2)}) ; \ldots ; Z'(\Xi_{\chi}^{(2)}) ; \ldots ; Z'(\Xi_{\nu_{\text{вкт}}-1}^{(2)}) \}. \quad (10)$$

Тоді кодові значення для всієї послідовності формуються на основі виразів:

$$Z(S) = Z(\hat{R}_{\nu_{\text{вкт}}-2}) = \sum_{\chi=2}^{\nu_{\text{вкт}}-1} Z'(\Xi_{\chi}^{(2)})\, N(S^{(\chi)}), \quad (11)$$

$$N(S^{(\chi)}) = (\beta(q)\cdot\beta(b))^{\nu_{\text{вкт}}-\chi}. \quad (12)$$

Тут $Z(S)$ - кодове представлення узагальненого позиційного числа; $N(S^{(\chi)})$ - виступає як ваговий коефіцієнт χ-го елемента узагальненого позиційного числа в базисі структурних обмежень трансформанти.

Тому значення коду $Z(S)$ узагальненого позиційного числа в базисі структурних обмежень трансформанти є кодовим представленням послідовності $Z(\hat{R}_{\nu_{\text{вкт}}-2})$, тобто $Z(S) = Z(\hat{R}_{\nu_{\text{вкт}}-2})$.

В такому випадку, формується кодограма для узагальненого кодового представлення трансформанти, що показано на рис. 6.

Тут структура кодової конструкції буде складатися з двох частин, включаючи службову і інформаційну частини. Службова частина кодової конструкції містить обмеження компонент векторів $\beta(q)$ та $\beta(b)$. Інформаційна частина кодограми включає в себе кодове представлення

значення коду $Z(\hat{R}_{\nu_{вкт}-2})$ послідовності значимих двохелементних векторів.

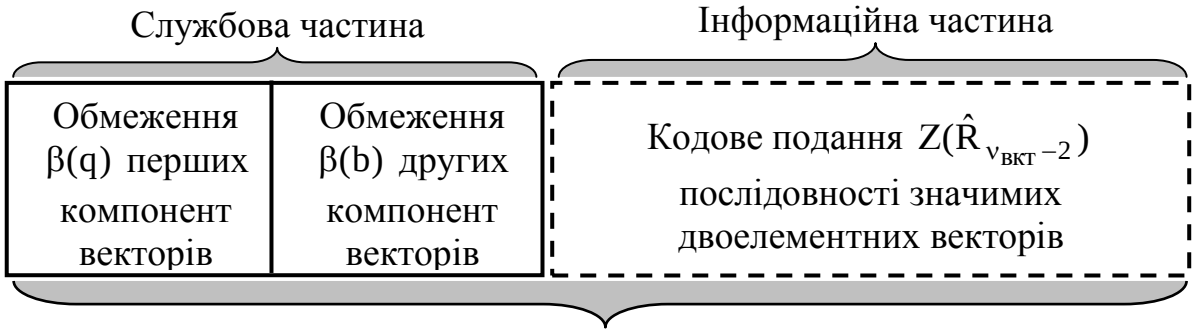


Рис.6. Структура кодограми для узагальненого позиційного числа

Такий підхід забезпечує додаткове скорочення структурної надлишковості блоків аерофотознімка в незалежності від ступеня їх семантичної насиченості.

Узагальнена схема процесу відновлення блоків аерофотознімка організується в умовах наявності змінної кількості кодограм $W(\hat{R}_{\nu_{вкр}-2})_{\omega}$, що несуть інформацію про кодове значення $Z(S)_{\ell_{\omega}}$ узагальненого позиційного числа. Ключовими етапами цього процесу є визначення кількості $Z(S)_{\ell_{\omega}}$ таких кодових слів. Тут потрібно враховувати те, що кодограми формуються за змішаним принципом, включаючи послідовність рівномірних кодових слів і кодограму $B(\hat{R}_{\nu_{вкт}-2})_{\Omega}$ нерівномірної довжини. Знаючи цю інформацію, і використовуючи технологію нерівномірного розподілу кількості векторів за кодограмами, забезпечується їх безпосереднє відновлення.

Після чого здійснюється вилучення кодових слів $W(\hat{R}_{\nu_{вкт}-2})_{\omega}$ узагальнених позиційних чисел, і декодування відповідних кодових значень $Z(S)_{\ell_{\omega}}$. Декодування проводиться за двоієрархічною схемою, спочатку в результаті узагальненого позиційного декодування відновлюються кодові значення $Z'(\Xi_{\chi}^{(2)})$ двохелементних векторів як позиційних чисел з нерівними сусідніми елементами [7-11]. Наступним кроком відбувається реконструкція самих компонент $\{q_{\chi}, b_{\chi}\}$ векторів. Після чого в результаті декомпозиції збирається двовимірний опис трансформанти. Заключним етапом процесу відновлення аерофотознімків полягає в зворотному трансформуванні та відтворенні блоків для вихідної колірної моделі уявлення.

Таким чином, розроблено метод відновлення блоків аерофотознімка на основі декодування послідовності значимих двоелементних векторів. Включає в себе такі відмінні етапи: відновлення кодової конструкції закодованого представлення блоку аерофотознімка; декодування за двоієрархічною схемою кодових значень складових послідовності значимих двоелементних векторів; відновлення значимих двоелементних векторів на основі декодування двоелементних позиційних чисел.

Оцінка ефективності розробленого методу проводиться з урахуванням побудови залежностей показника зменшення інтенсивності від ступеня насиченості аерофотознімка блоками трьох класів семантичного навантаження, що показано на рис. 7.

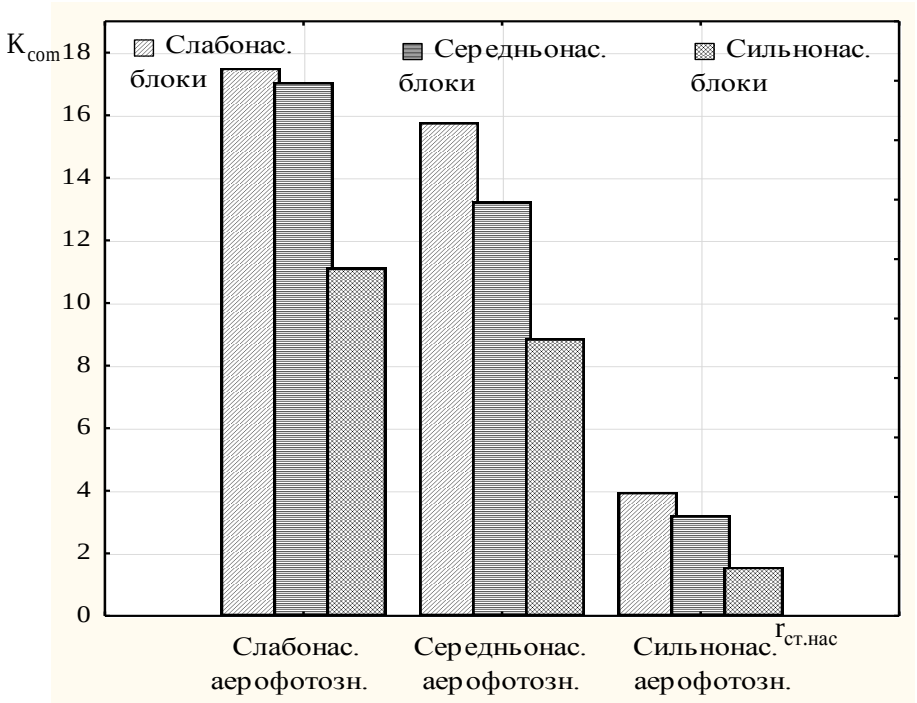


Рис.7. Залежність коефіцієнта зменшення вихідного об'єму для різних класів аерофотознімків

З аналізу рис. 7 можна зробити наступні висновки:

– в залежності від ступеня насиченості аерофотознімка блоками трьох класів коефіцієнт зменшення вихідного об'єму досягається для слабонасичених блоків 17,51 раз, тоді як найменший коефіцієнт зменшення початкового об'єму досягається для сильнонасичених блоків, який змінюється від 1,56 до 3,95 разів. В даному типі блоків використовуються невеликі значення ступеня зменшення початкового об'єму. Внаслідок чого буде збережена інформації про ключові ознаки дешифрування.

Зменшений час доставки аерофотознімків, що оброблені методом кодування, який у порівнянні з відомими методами на базі JPEG-платформи забезпечує виграш за часом доставки аерофотознімків та складає: для швидкості 256 Кбіт/с від 8 до 19%; для швидкості 512 Кбіт/с

від 13 до 27%, для швидкості 16 Мбіт/с від 15 до 30%, внаслідок чого забезпечується підвищення оперативності доставки інформації з використанням бортових комплексів повітряної розвідки

Оцінка якості збереження інформації про ключові ознаки дешифрування в аерофотознімках, які оброблені розробленим методом проведена на основі наступних показників: 1) в якості оцінки складової яскравості пропонується використовувати оцінку допустимого рівня збереження ключових ознак дешифрування з використанням показника $h_{доп}$ (ПОСШ); 2) для оцінки збереження контурній інформації (на основі методу виявлення та локалізації контурів Собела) використовувати показники збереження контурної інформації - помилки першого ε та другого μ роду.

Результати оцінок якості збереження інформації про ключові ознаки дешифрування з урахуванням виявлення контурної інформації в блоках аерофотознімків, що оброблені створеним методом кодування по показникам допустимого $h_{доп}$ рівня збереження ключових ознак дешифрування та помилок першого ε та другого μ роду дозволяють зробити наступні висновки: - виграш по значенню помилки першого ε роду в середньому збільшується на 0,029 - 0,042 (27,1 - 39,1%); значення помилки другого роду в середньому збільшується на 0,112 - 0,231 (21,21 - 34,04%); виграш щодо показника допустимого $h_{доп}$ рівня збереження ключових ознак становить від 23% до 41,5%, що і підтверджується візуальною оцінкою порівняння результатів роботи методу, що представлено на рис. 8.

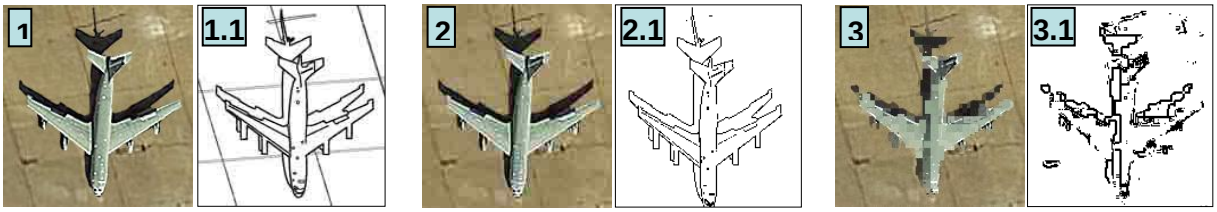


Рис. 8. 1) вихідний блок; 1.1) обробка оператором Собела; 2) оброблений блок створеним методом; 2.1) обробка оператором Собела; 3) оброблений блок методом JPEG; 3.1) обробка оператором Собела

При цьому відновлені блоки зберігають високу якість контурної інформації, що необхідно для ефективного дешифрування аерофотознімків. Таким чином, для запропонованого методу кодування, досягається зниження інформаційної інтенсивності і забезпечується підвищення оперативності доставки інформації в умовах заданої якості дешифрування аерофотознімків.

Висновки

У роботі вирішена актуальна науково-прикладна задача, а саме підвищення оперативності доставки інформації при заданій якості дешифрування аерофотознімків з використанням безпілотних літальних апаратів аеромоніторингу. Створено методи кодування і відновлення трансформованих зображень, на основі обробки послідовності двоелементних векторів. Розроблено технологію кодування послідовності значимих двоелементний векторів структурних характеристик трансформанти.

Отримано такі основні наукові результати.

1. Удосконалений метод кластеризації фрагментів аерофотознімка на основі методу К-середніх, у якому на відміну від відомого, одночасно застосовується двобазисний принцип оцінки рівня насиченості блоків аерофотознімка як в часовій, так і в спектрально-просторовій областях. Це дозволить відібрати блоки аерофотознімка, які мають найбільшу дешифрувальну насиченість;

2. Вперше розроблено метод кодування трансформованих зображень на основі усунення структурної надмірності, у якому на відміну від відомого, лінеаризована трансформанта представляється у двоієрархічному вигляді з подальшою побудовою узагальнених позиційних чисел. Це дозволить додатково усунути надмірність аерофотознімка зі збереженням інформації про ключові дешифрувальні ознаки з використанням безпілотних літальних апаратів аеромоніторингу.

3. Вперше розроблено метод обробки блоків аерофотознімка на основі реконструкції двоелементних структурних векторів лінеаризованої трансформанти, у якому на відміну від відомого, узагальнене позиційне кодування проводиться для позиційних чисел з нерівними сусідніми елементами з врахуванням ступеня насиченості блоків аерофотознімка дешифрувальними ознаками. Це дозволить зберегти необхідний рівень інформації про дешифрувальні ознаки в аерофотознімку з використанням безпілотних літальних апаратів аеромоніторингу.

4. Вперше створено інформаційну технологію обробки аерофотознімків на основі структурної обробки кластеризованих блоків, відмінною рисами якої є: 1) створення диференційованого принципу обробки блоків аерофотознімка в залежності від їх насиченості дешифрувальними ознаками; 2) трансформанти обробляються на основі структурного кодування за двоієрархічним підходом, а саме, на першому рівні формуються позиційні числа з нерівними сусідніми елементами, на другому рівні кодування здійснюється для узагальнених позиційних чисел. Це дозволить підвищити оперативність доставки аерофотознімків зі збереженням інформації про ключові ознаки дешифрування з використанням безпілотних літальних апаратів аеромоніторингу.

Література

1. Мосов С. Беспилотная разведывательная авиация стран мира: история создания, опыт боевого применения, современное состояние, перспективы развития [Текст]: монография / С. Мосов. – К.: Изд. дом. «Румб», 2008. – 160 с.

2. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2005. – 1073 с.

3. Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон. – М: Техносфера, 2004. – 368 с.

4. Бараннік В.В. Метод кластеризації фрагментів аерофотознімків у спектрально-частотному просторі // В.В. Бараннік, О.П. Мусієнко, К.С. Ялівець // Наукоємні технології, 2016. – Т. 29 (1). – С. 23 – 30.

5. Баранник В.В. Кластеризация блоков аэрофотоснимка в двухпризнаковом структурном пространстве на основе метода К-средних в системе обработки информации / В.В. Баранник, А.П. Мусиенко, А.А. Краснорутский // Радиоэлектроника и информатика, 2016. – № 2. – С. 15 – 19.

6. Бараннік В.В. Технологія кодування блоків аерофотознімку з врахуванням семантично важливої інформації для бортових комплексів повітряного моніторингу // В.В. Бараннік, О.П. Мусієнко, А.А. Леках // Наукоємні технології, 2016. – Т. 31 (3). – С. 274 – 278.

7. Мусиенко А.П. Технология декодирования блоков аэрофотоснимка на основе восстановления компонент трансформант / А.П. Мусиенко // Информационно-управляющие системы на ЖД транспорте, 2016. – №5. – С. 58 – 62.

8. Barannik V.V. Methodological Base For Transformants Representation In Nonequilibrium Positional Uneven-Diagonal Space / V.V. Barannik, A.A. Krasnorutskyi, A.P. Musienko // Science-Based Technologies, 2015. – № 3 (27). – pp. 233 – 238.

9. Barannik V. The Evaluation Method Of Coding Efficiency Of Basic Frames Of The Video Stream In Infocommunication / V.V. Barannik, A.P. Musienko// IEEE Second International Scientific-Practical Conference ["IEEE Problems of Infocommunications. Science and Technology, PICS&T'2015"], (Kharkiv, Ukraine, October 13-15, 2015) / Kharkiv, 2015. – pp. 223 – 225.

10. Barannik V. Methodological base for representation transformants in equilibrium uneven-diagonal / V.V. Barannik, Sergii Shulgin, A.P. Musienko // 2015 1st International Conference ["Advanced Information and Communication Technologies-2015 (AICT'2015)"], (Lviv, Ukraine, October 29 – November 1, 2015) / Lviv, 2015. – pp. 138 – 140.

11. Musienko A. Technology of coding of digital aerial photographs taking into account classes of a semantic saturation of blocks in system of air monitoring / A. Musienko, J. Ganjaric // VII Inter University Conference of Students, PhD Students and Young Scientists ["Engineer of XXI Century"], 08 December 2016 at the University of Bielsko-Biała (ATH) / Bielsko-Biała, Poland, 2016. – P. 215 – 220.

СТРУКТУРНО-ЭНТРОПИЙНОЕ КОДИРОВАНИЕ ДЛЯ ПОВЫШЕНИЯ ЦЕЛОСТНОСТИ ВИДЕОИНФОРМАЦИОННОГО РЕСУРСА В ИНФОКОММУНИКАЦИЯХ

Баранник В.В., Подлесный С.А., Гаврилов Д.С.

Введение

Видеоинформационные ресурсы (ВИР) в Вооруженных Силах Украины (ВСУ) используются для постоянного контроля, четкого и постоянного управления войсками в ВСУ [1]. Одним из примеров использования ВИР является применение беспилотных летательных аппаратов с целью наблюдения за территорией, выявление фактов терроризма или проведения разведки [2]. Полученная информация передается по каналам радиосвязи на командный пункт. Также областью использования ВИР в ВСУ является видеоконференцсвязь. Актуальность применения видеоконференцсвязи обусловлена необходимостью своевременного принятия решений и обсуждения определенных проблем при невозможности личной встречи [3]. При выполнении передачи видеоинформации может происходить влияние на телекоммуникационное оборудование. Это связано как наличием природных факторов, так и влиянием противника, который проводит кибератаки [4, 5]. При воздействии таких помех в телекоммуникационном оборудовании может произойти сбой, что приводит к искажению видеоинформации. В результате наличия битовой ошибки в статистическом коде происходит неверная идентификации всех последующих VLC-кодов. Этим аргументируется, что при наличии битовой ошибки в потоке кодов переменной длины влияние ошибки может сильно влиять на восстановление значений коэффициентов ДКП, то есть статистический код не является устойчивым к ошибкам [6]. Поэтому возникает проблема обеспечения целостности информации в телекоммуникационных системах.

Для борьбы с такими ошибками в существующую технологию кодирования JPEG добавляют помехоустойчивое кодирование. Принцип работы такого кодирования в том, что к передаваемой информации добавляют проверочные биты [7]. Это позволяет выявить и исправить ошибки. В результате происходит восстановленной информации. Но в такой схеме присутствуют следующие недостатки:

- применение помехоустойчивого кодирования происходит с использованием аппаратных и временных затрат. Это приводит к увеличению времени обработки;
- при добавлении дополнительных битов увеличивается объем информации. Это приводит к увеличению времени передачи видеоинформации.

Данные недостатки влияют на оперативность передачи видеоинформации, что недопустимо для использования в военной сфере.

Поэтому решение проблемы обеспечения целостности ВИР является актуальным.

Целью данной работы является разработка метода повышения целостности информации на основе существующих технологий обработки изображений, основным условием при реализации которых является сохранение временных затрат на передачу в телекоммуникационных системах.

Для достижения данной цели в работе необходимо:

1. Провести модернизацию представления ВИР в существующих технологиях обработки изображений.

2. Разработать метод выравнивающего статистического кодирования с учетом структурных характеристик обрабатываемого изображения.

3. Провести анализ категорий информационной безопасности для разработанного метода.

1. Разработка метода структурно-энтропийного кодирования линеаризированной трансформанты

Для повышения помехозащищенности видеоинформационного ресурса предлагается изменить существующую технологию статистического кодирования. Для этого **предлагается** исключить выполнение квантизации компонент трансформанты. Этим достигается следующее:

1) повышается целостность в связи с отсутствием ошибок округления при квантизации;

2) увеличивается доступность видеоинформационного ресурса вследствие уменьшения временных затрат на обработку.

Отсутствие квантизации приводит к уменьшению количества нулевых компонент в трансформанте. С другой стороны, это приводит росту информационной интенсивности. Поэтому для компенсации такого роста **предлагается:**

- формировать код $N^{(j)}$ для каждой пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты;

- в процессе формирования кода $N^{(j)}$ учитывать структурно-статистическую зависимость между элементами u_i и u_{i+1}.

Формирование информационной части $K^{(j)}$ кода $N^{(j)}$ для элементов u_i и u_{i+1} задается следующим функционалом:

$$K^{(j)} := g^{(j)} \times f_{\alpha}(u_i, u_{i+1}) + f_{\beta}(u_i, u_{i+1}), \qquad (1)$$

где j – индекс кода $N^{(j)}$ пары элементов u_i и u_{i+1} линеаризированной трансформанты, который определяется, как $j = 0{,}5(i+1)$, и изменяется в диапазоне $j = \overline{1;\, \theta/2}$;

$g^{(j)}$ – весовой коэффициент кода $N^{(j)}$, который определяется как максимум элементов u_i и u_{i+1}, т.е. $g^{(j)} = \max(u_i, u_{i+1})$;

$f_\alpha(u_i, u_{i+1})$, $f_\beta(u_i, u_{i+1})$ – функции обработки элементов u_i и u_{i+1}.

Для снижения информационной интенсивности кода $N^{(j)}$ предлагается ввести признак $n_u^{(j)}$ наличия нулевых элементов u_i и u_{i+1} согласно следующего правила

$$n_u^{(j)} = 0 \text{ при } u_i = 0 \text{ или } u_{i+1} = 0,\ n_u^{(j)} = 1 \text{ при } u_i > 0 \text{ и } u_{i+1} > 0 \quad (2)$$

и задать следующие значение функций $f_\alpha(u_i, u_{i+1})$, $f_\beta(u_i, u_{i+1})$ в зависимости от величин элементов u_i и u_{i+1} (табл. 1):

Таблица 1

Таблица состояний для функций вычисления значения $K^{(j)}$

Значение элемента u_i	Значение элемента u_{i+1}	Значение функции $f_\alpha(u_i, u_{i+1})$	Значение функции $f_\beta(u_i, u_{i+1})$
$u_i = 0$	$u_{i+1} = 0$	$f_\alpha(u_i, u_{i+1}) = 0$	$f_\beta(u_i, u_{i+1}) = 0$
$u_i > 0$	$u_{i+1} = 0$	$f_\alpha(u_i, u_{i+1}) = 0$	$f_\beta(u_i, u_{i+1}) = 0$
$u_i = 0$	$u_{i+1} > 0$	$f_\alpha(u_i, u_{i+1}) = 0$	$f_\beta(u_i, u_{i+1}) = 1$
$u_i > 0$	$u_{i+1} > 0$	$f_\alpha(u_i, u_{i+1}) = u_i - 1$	$f_\beta(u_i, u_{i+1}) = u_{i+1} - 1$

Рассмотрим значение длины $|K^{(j)}|_2$ информационной части $K^{(j)}$ кода $N^{(j)}$, сформированного согласно функционалу (1) с учетом четырех случаев значений кодируемых элементов u_i и u_{i+1} (табл. 1):

1) для нулевых значений элементов u_i и u_{i+1} весовой коэффициент $g^{(j)}$ равняется нулю, а именно:

$$g^{(j)} = 0 \text{ при } u_i = 0,\ u_{i+1} = 0.$$

Для сокращения информационной интенсивности кодового представления трансформанты предлагается для нулевых значений элементов u_i и u_{i+1} выбрать длину $|K^{(j)}|_2$ информационной части кода $N^{(j)}$ равной нулевому значению, т.е.:

$$|K^{(j)}|_2 = 0,\ [K^{(j)}]_2 = \{\} \text{ при } u_i = 0,\ u_{i+1} = 0;$$

В этом случае величина $K^{(j)}$ не передается.

2) для нулевого значения элемента u_{i+1} весовой коэффициент $g^{(j)}$ равняется u_i, а именно:

$$g^{(j)} = u_i \text{ при } u_{i+1} = 0.$$

Тогда согласно (1) значение информационной части $K^{(j)}$ кода $N^{(j)}$ получится равным нулевому значению. Длина $|K^{(j)}|_2$ соответственно равна единицы, т.е.:

$$K^{(j)} = 0,\ |K^{(j)}|_2 = 1 \text{ при } u_{i+1} = 0;$$

3) для нулевого значения элемента u_i весовой коэффициент $g^{(j)}$ равняется u_{i+1}, а именно:

$$g^{(j)} = u_{i+1} \text{ при } u_i = 0.$$

В этом случае согласно (1) значение информационной части $K^{(j)}$ кода $N^{(j)}$ будет равно единице. Длина $|K^{(j)}|_2$ соответственно будет равна единицы, т.е.:

$$K^{(j)} = 1,\ |K^{(j)}|_2 = 1 \text{ при } u_i = 0;$$

4) для ненулевых значений элементов u_i и u_{i+1} весовой коэффициент $g^{(j)}$ равняется их максимальной величине, а именно:

$$g^{(j)} = \max(u_i; u_{i+1}) \text{ при } u_i > 0,\ u_{i+1} > 0.$$

Для определения длины $|K^{(j)}|_2$ необходимо определить максимальное значение информационной части $K^{(j)}$ кода $N^{(j)}$ для заданной величины $g^{(j)}$. Анализ функционала (1) показывает, что максимальное значение $K^{(j)}_{max}$ образуется при равенстве элементов u_i и u_{i+1}, т.е.:

$$K^{(j)} = g^{(j)}(g^{(j)} - 1) + (g^{(j)} - 1) = (g^{(j)})^2 - 1 = \max \text{ при } u_i = u_{i+1} = g^{(j)}. \quad (3)$$

Полученное максимальное значение информационной части $K^{(j)}$ кода $N^{(j)}$ определяет информационную интенсивность кодового представления пары элементов u_i и u_{i+1}, а именно

$$|K^{(j)}|_2 = \log_2((g^{(j)})^2 - 1) \text{ при } u_i > 0 \text{ и } u_{i+1} > 0. \quad (4)$$

В сформированном коде $N^{(j)}$ служебной информацией $K_g^{(j)}$ является значение весового коэффициента $g^{(j)}$ кода $N^{(j)}$ и признак $n_u^{(j)}$ наличия нулевых элементов. С целью снижения информационной интенсивности служебной части кода $N^{(j)}$ предлагается значение весового коэффициента $g^{(j)}$ кодировать с учетом статистических свойств диапазона $\gamma = \log_2(g^{(j)})$ значений величины $g^{(j)}$. Для чего кодовое представление $K_g^{(j)}$ весового коэффициента $g^{(j)}$ формируется из двух частей:

1) служебной части $\Gamma_g^{(j)}$, которая характеризует диапазон γ значений величины $g^{(j)}$;

2) информационной части $I_g^{(j)}$, которая характеризует величину $g^{(j)}$ для элементов в паре $\tilde{u}_{i,\,i+1}$.

При формировании содержимого $[\Gamma_g^{(j)}]_2$ служебной части предлагается учитывать статистическую зависимость между элементами u_i и u_{i+1}. Снижение информационной интенсивности для двоичного представления разновероятностных величин реализовано в энтропийных кодов. Поэтому служебная часть является энтропийным кодом $\Gamma_g^{(j)}$, который формируется с использованием существующего аппарата кодирования низкочастотных компонент для технологий семейства JPEG (табл. 2). Это приводит к интеграции разрабатываемого метода в современные технологии обработки изображений.

Для учета структурной зависимости между элементами u_i и u_{i+1} предлагается формировать содержимое $[I_g^{(j)}]_2$ информационной части на основе содержит признака $n_u^{(j)}$ наличия нулевых элементов в паре $\tilde{u}_{i,\,i+1}$ и младших разрядов $\tilde{g}^{(j)}$ весового коэффициента $g^{(j)}$.

Поэтому в результате учета рассмотренных зависимостей разработанный метод является структурно-энтропийным кодированием. Общая длина $|K_g^{(j)}|_2$ кода весового коэффициента $g^{(j)}$ следующая (табл. 2):

Для обеспечения обработки отрицательных элементов трансформанты $Y(n, n)$ предлагается формировать матрицу знаков $Sign^{(y)}(n, n)$. Значение элемента $sign(\chi, \kappa)$ матрицы знаков формируется согласно следующему правилу:

$$sign(\chi, \kappa) = 1 \text{ при } y(\chi, \kappa) < 0,\ sign(\chi, \kappa) = 0 \text{ при } y(\chi, \kappa) \geq 0. \quad (5)$$

В результате обработки всех пар $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты образуется последовательность $\{N^{(j)}\}$, $j = \overline{1; \theta/2}$ кодов $N^{(j)}$. Здесь величина диапазона $j = \overline{1; \theta/2}$ индекса указывает, что количество кодов $N^{(j)}$ в два раза меньше, чем количество элементов вектора линеаризированной трансформанты $U(\theta)$. Каждый код $N^{(j)}$ представляет собой композицию служебной $K_g^{(j)}$ и информационной $K^{(j)}$ составляющих.

Служебная составляющая включает в себя три части, а именно:

1) энтропийный код $\Gamma_g^{(j)}$, который соответствует γ-му диапазону весового коэффициента $g^{(j)}$;

2) признак $n_u^{(j)}$ наличия нулевых элементов в паре $\tilde{u}_{i,\,i+1}$;

3) младшие разряды $\tilde{g}^{(j)}$ весового коэффициента $g^{(j)}$ кода $N^{(j)}$.

Таблица 2

Кодовая таблица двоичного представления $K_g^{(j)}$ *весового коэффициента* $g^{(j)}$

Диапазон γ значения величины $g^{(j)}$	Содержимое $[\Gamma_g^{(j)}]_2$ служебной части	Общая длина $\lvert K_g^{(j)} \rvert_2$ кода весового коэффициента $g^{(j)}$
0	010	3
1	011	4
2	100	5
3	00	5
4	101	7
5	110	8
6	1110	10
7	11110	12
8	111110	14
9	1111110	16
10	11111110	18
11	111111110	20

Содержание информационной $K^{(j)}$ составляющей кода $N^{(j)}$ зависит от наличия нулевого элемента в паре $\tilde{u}_{i,\,i+1}$. Здесь возможны следующие четыре варианта:

1) $$[K^{(j)}]_2 = \{\} \text{ при } u_i = 0,\ u_{i+1} = 0; \quad (6)$$

2) $$K^{(j)} = 0 \text{ при } u_i > 0,\ u_{i+1} = 0; \quad (7)$$

3) $$K^{(j)} = 1, \text{ при } u_i = 0,\ u_{i+1} > 0; \quad (8)$$

4) $$K^{(j)} = \max(u_i; u_{i+1})(u_i - 1) + (u_{i+1} - 1) \text{ при } u_i > 0,\ u_{i+1} > 0. \quad (9)$$

Композиция служебной $K_g^{(j)}$ и информационной $K^{(j)}$ составляющих для структурно-энтропийного кода $N^{(j)}$ представлена следующим выражением:

$$N^{(j)} = \Gamma_g^{(j)} \vee n_u^{(j)} \vee \tilde{g}^{(j)} \vee K^{(j)}. \quad (10)$$

Здесь $\Gamma_g^{(j)}$ – код диапазона γ весового коэффициента $g^{(j)}$;

$n_u^{(j)}$ – признак наличия нулевых элементов в паре $\tilde{u}_{i,\,i+1}$;

$\tilde{g}^{(j)}$ – младшие разряды весового коэффициента $g^{(j)}$;

$K^{(j)}$ – величина информационной части кода $N^{(j)}$.

Из анализа соотношений (6 – 10) следует, что длина $|N^{(j)}|_2$ каждого кода $N^{(j)}$ является переменной, так как:

– длина $|K_g^{(j)}|_2$ служебной части кода зависит от статистических характеристик кодируемых элементов;

– длина $|K^{(j)}|_2$ информационной части определяется структурной особенностью обрабатываемой пары.

Структурная схема формирования структурно-энтропийного кода $N^{(j)}$ для пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты показана на рис. 1.

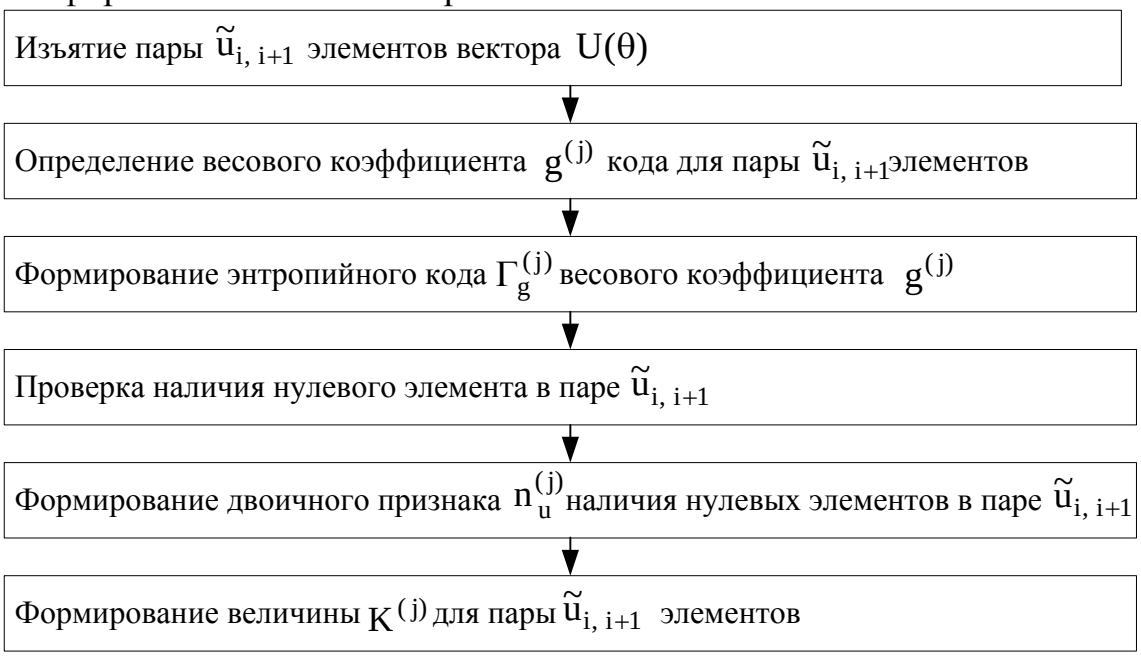


Рис. 1. Структурная схема формирования структурно-энтропийного кода $N^{(j)}$ для пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты

2. Разработка метода структурно-энтропийного декодирования линеаризированной трансформанты

Для разработки метода декодирования необходимо рассмотреть порядок восстановления структурно-энтропийного кода $N^{(j)}$. На вход декодера поступает последовательность $\{N^{(j)}\}$, $j=\overline{1;\theta/2}$ структурно-энтропийных кодов $N^{(j)}$ переменной длины, которые формируются как композиция служебной части $K_g^{(j)}$ и информационной части $K^{(j)}$. Для

восстановления структурно-энтропийного кода $N^{(j)}$ пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты сначала необходимо провести четыре последовательных этапа извлечения и восстановления его составляющих, а именно:

– три этапа обработки служебной части $K_g^{(j)}$ структурно-энтропийного кода $N^{(j)}$:

1) энтропийного кода $\Gamma_g^{(j)}$, который соответствует γ-му диапазону величины $g^{(j)}$ кода $N^{(j)}$;

2) признака $n_u^{(j)}$ наличия нулевых элементов в паре $\tilde{u}_{i,\,i+1}$;

3) младших разрядов $\tilde{g}^{(j)}$ весового коэффициента $g^{(j)}$;

– один этап обработки информационной части структурно-энтропийного кода $N^{(j)}$:

4) величины $K^{(j)}$, которая соответствует кодовому представлению пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты.

Рассмотрим подробнее каждый этап.

1. Позиционирование энтропийного кода $\Gamma_g^{(j)}$ весового коэффициента $g^{(j)}$ кода $N^{(j)}$ происходит поэлементно в соответствии с кодовой таблицей (табл. 2). Восстановление значения γ табличного диапазона весового коэффициента $g^{(j)}$ кода $N^{(j)}$ выполняется при сопоставлении двоичной последовательности $[\Gamma_g^{(j)}]_2$ соответствующей строки кодовой таблицы (рис. 2).

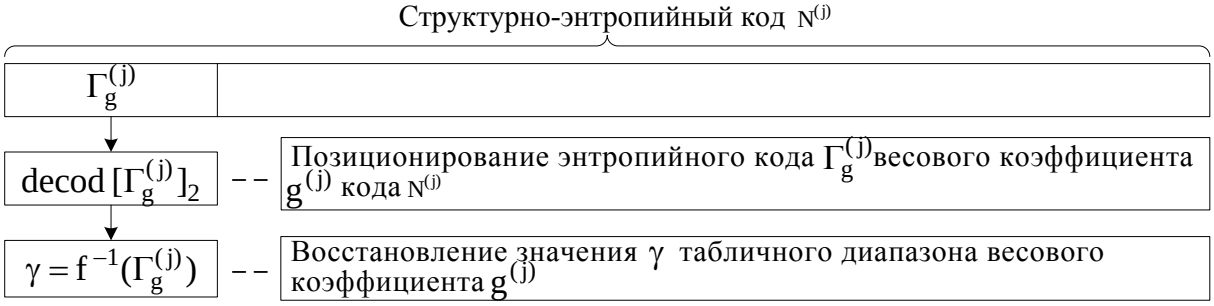


Рис. 2. Восстановление значения γ диапазона весового коэффициента $g^{(j)}$

Однако возможен случай, когда для энтропийного кода $\Gamma_g^{(j)}$, содержащего двоичную последовательность $[\Gamma_g^{(j)}]_2 = 010$, будет

определено, что кодированию подвергалась пара элементов с нулевым значением, т.е.:

$$u_i = u_{i+1} = 0 \text{ при } [\Gamma_g^{(j)}]_2 = 010. \quad (11)$$

Тогда декодирование структурно-энтропийного кода $N^{(j)}$ завершается, и дальнейшие двоичные элементы в кодовом потоке соответствуют уже следующему $(j+1)$-му структурно-энтропийному коду $N^{(j+1)}$.

2. Далее происходит извлечение двоичного признака $n_u^{(j)}$ наличия нулевых элементов в паре $\tilde{u}_{i,\,i+1}$. Восстановленный двоичный признак $n_u^{(j)}$ равняется значению двоичного элемента, расположенного в структурно-энтропийном коде $N^{(j)}$ после двоичной последовательности $[\Gamma_g^{(j)}]_2$ (рис. 3).

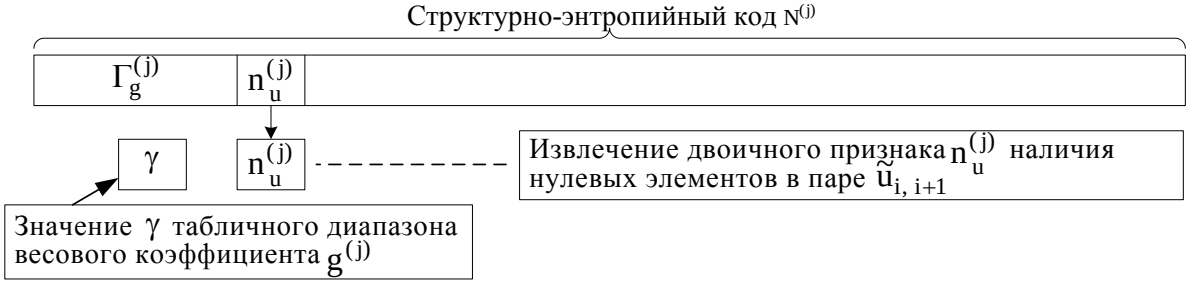


Рис. 3. Восстановление двоичного признака $n_u^{(j)}$

Восстановленное значение двоичного признака $n_u^{(j)}$ определяет структурную характеристику кодируемых элементов. Здесь единичное значение двоичного признака $n_u^{(j)}$ указывает на отсутствие нулевых элементов в паре $\tilde{u}_{i,\,i+1}$, а нулевое значение, напротив, – на наличие.

3. Окончательным этапом обработки служебной части $K^{(j)}$ является восстановление весового коэффициента $g^{(j)}$ структурно-энтропийного кода $N^{(j)}$. Для этого при позиционировании младших разрядов $\tilde{g}^{(j)}$ весового коэффициента $g^{(j)}$ применяется величина γ, восстановленная на первом этапе декодирования (рис. 4). Восстановление весового коэффициента $g^{(j)}$ заключается в дополнении к извлеченным младшим разрядам $\tilde{g}^{(j)}$ слева двоичного элемента равного единице. Это обусловлено исключением старшего разряда весового коэффициента $g^{(j)}$ в процессе формирования структурно-энтропийного кода $N^{(j)}$.

Следующим после обработки служебной части $K^{(j)}$ структурно-энтропийного кода $N^{(j)}$ является извлечение информационной части $K^{(j)}$

структурно-энтропийного кода $N^{(j)}$ и восстановление элементов в паре $\tilde{u}_{i,\,i+1}$. Для этого применяются восстановленные в процессе декодирования служебные данные.

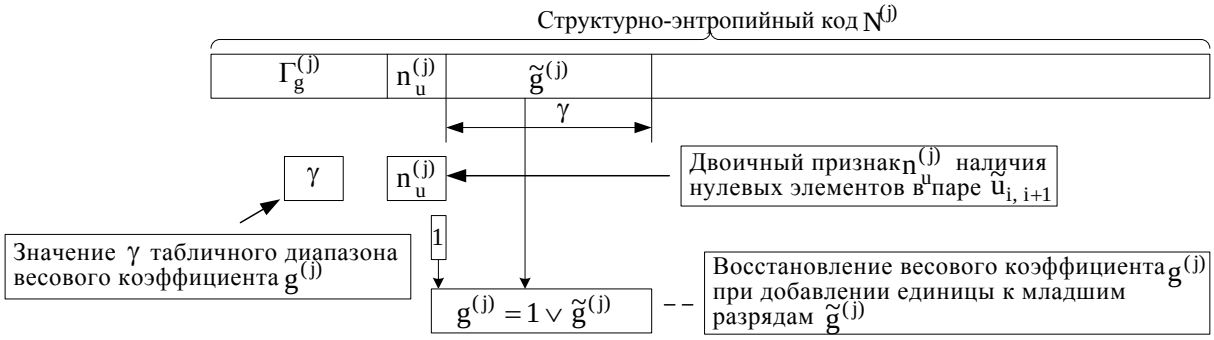


Рис. 4. Восстановление весового коэффициента $g^{(j)}$

Начальное условие позиционирования определяется двоичным признаком $n_u^{(j)}$. Это заключается в следующем:

1) для нулевого значения двоичного признака $n_u^{(j)}$ длина $[K^{(j)}]_2$ информационной части $K^{(j)}$ равняется единице (рис. 5). В этом случае восстановление элементов в паре $\tilde{u}_{i,\,i+1}$ задается значением $K^{(j)}$, а именно:

$$u_i = g^{(j)},\ u_{i+1} = 0 \text{ при } K^{(j)} = 0,\ n_u^{(j)} = 0, \qquad (12)$$

$$u_i = 0,\ u_{i+1} = g^{(j)} \text{ при } K^{(j)} = 1,\ n_u^{(j)} = 0. \qquad (13)$$

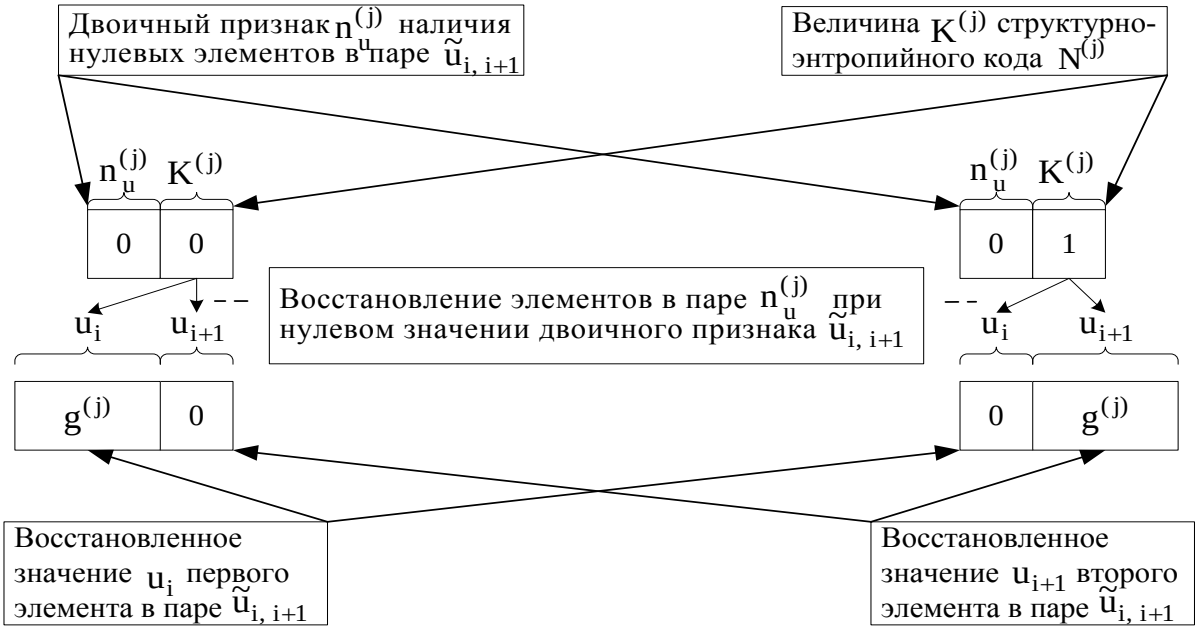


Рис. 5. Восстановление элементов в паре $\tilde{u}_{i,\,i+1}$ *при нулевом значении двоичного признака* $n_u^{(j)}$

2) при единичном значении двоичного признака $n_u^{(j)}$ длина $[K^{(j)}]_2$ информационной части $K^{(j)}$ определяется восстановленным весовым коэффициентом $g^{(j)}$ согласно следующему выражению (рис. 6):

$$|K^{(j)}|_2 = \log_2((g^{(j)})^2 - 1) \text{ при } n_u^{(j)} = 1. \quad (14)$$

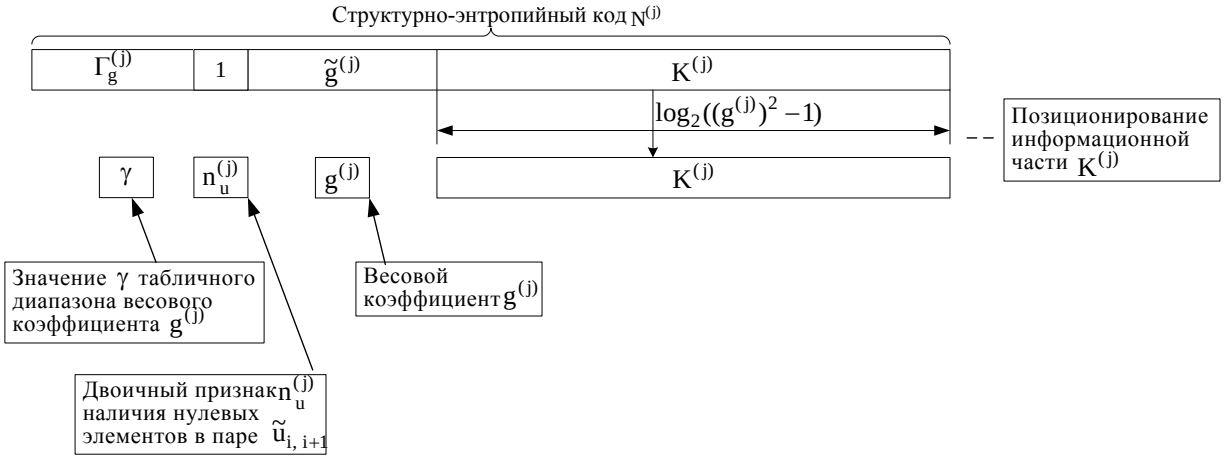


Рис. 6. Позиционирование информационной части $K^{(j)}$ при единичном значении двоичного признака $n_u^{(j)}$

Дальнейшее восстановление элементов u_i, u_{i+1} выполняется по следующим формулам:

$$u_i = K^{(j)} / g^{(j)} + 1,\ u_{i+1} = K^{(j)} \bmod g^{(j)} + 1 \text{ при } n_u^{(j)} = 1, \quad (15)$$

В результате обработки всей последовательности $\{N^{(j)}\}$, $j = \overline{1; \theta/2}$ структурно-энтропийных кодов $N^{(j)}$ будут восстановлены все пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты. Для восстановления отрицательных элементов трансформанты $Y(n, n)$ используется матрица знаков $Sign^{(y)}(n, n)$ (5). Правило изменения знака элемента трансформанты $Y(n, n)$ согласно значение элемента $sign(\chi, \kappa)$ матрицы знаков описывается следующим выражением:

$$y(\chi, \kappa) = -u_i \text{ при } sign(\chi, \kappa) = 1,\ y(\chi, \kappa) = u_i \text{ при } sign(\chi, \kappa) = 0. \quad (16)$$

Структурная схема восстановления пары $\tilde{u}_{i,\,i+1}$ элементов вектора $U(\theta)$ линеаризированной трансформанты для структурно-энтропийного кода $N^{(j)}$ показана на рис. 7.

3. Анализ информационной интенсивности для разработанного метода

Результатом кодирования вектора $U(\theta)$ элементов линеаризированной трансформанты $L(\Lambda)^{(i)}$ кодов VLC компонент любой

трансформанты является последовательность $\{N^{(i)}\}$, $i = \overline{1; \theta/2}$ структурно-энтропийных кодов $N^{(i)}$.

Рис. 7. Структурная схема восстановления пары $\tilde{u}_{i,\,i+1}$ *элементов вектора* $U(\theta)$ *линеаризированной трансформанты для структурно-энтропийного кода* $N^{(j)}$

Проведенный эксперимент показал, что при использовании разработанного метода обработки изображений происходит снижение информационной интенсивности от 10% для сильнонасыщенных изображений до 15% для средненасыщенных изображений. В результате разработано кодирование, которое осуществляется для пары $\tilde{u}_{i,\,i+1}$

элементов и учитывает такую структурную особенность, как количество нулевых компонент в паре $\tilde{u}_{i,\ i+1}$ и значение весового коэффициента $g^{(j)}$.

Формирование двоичного представления весового коэффициента $g^{(j)}$ производится с учетом статистических характеристик элементов исходного изображения. Разработанный метод называется структурно-энтропийным кодированием. Формирование кодовых конструкций с учетом структурно-статистических особенностей элементов изображения позволяет сократить интенсивность двоичного представления трансформанты с минимальным использования служебных данных. При искажении кодового представления ошибка будет локализирована в области значений весового коэффициента $g^{(j)}$. Этим достигается обеспечение целостности ВИР при заданном уровне доступности.

Выводы

По вышеизложенному материалу можно заключить, что разработана технология формирования структурно-энтропийных кодов для всех компонент трансформанты. Использование разработанной технологии приводит к снижению информационной интенсивности. Это достигается за счет:

– формирования общего кода для пары элементов вектора линеаризированной трансформанты. Здесь технология учитывает такую структурную особенность, как значение обрабатываемых элементов в паре;

– формирования энтропийного кода для максимального значения обрабатываемых элементов в паре. В данном случае учитывается статистическая особенность компонент обрабатываемой трансформанты.

Научная новизна:

1. Впервые предлагается исключить выполнение квантизации компонент. Этим достигается следующее:

– повышается целостность в связи с отсутствием ошибок округления при квантизации;

– увеличивается доступность видеоинформационного ресурса вследствие уменьшения временных затрат на обработку.

2. Впервые предлагается производить выявление компонент с нулевым значением. Это приводит к следующему:

1) отсутствию информационных затрат на передачу кодового представления пары нулевых элементов;

2) снижению информационных затрат на передачу кодового представления одного ненулевого элемента.

3. Впервые предлагается формировать код пары элементов на основе весового коэффициента. Этим достигается снижение информационной интенсивности двоичного представления пары элементов. При искажении информационной части структурно-энтропийного кода ошибка

восстановления локализируется в области значений весового коэффициента.

Этим обеспечивается условие повышения целостности видеоинформационного ресурса для заданного уровня доступности.

Литература

1. Баранник В.В. Метод локализации потери целостности информации на основе слот-технологии [Текст] / В.В. Баранник, С.А. Подлесный, Д.В. Баранник // Радиоэлектроника и информатика. – Х.: ХНУРЭ, 2015, – Вип. 4. – С. 32 – 41.

2. Баранник В.В, Метод підвищення стійкості відеоконтенту до кібернетичних атак у інфокомунікаційних системах [Текст] / В.В. Баранник, С.А. Подлесный // Безпека інформації. – К.: НАУ, 2016. – Вип. 22(2). – С. 123 – 130.

3. V.V. Barannik and S.A. Podlesny, "Analysis of the action of cyber-attacks in the video-information's resources in the information-telecommunications networks", Management Information System and Devices, 2014. – vol. 169, No 4. – pp. 16 – 22.

4. V.V. Barannik, S.A. Podlesny and S.S. Shulgin, "Methodology of impact assessment of safety cyber-attacks on video information resources in telecommunications system", Radioelektronika i informatika, 2016. – vol. 72, No 1. – pp. 61 – 64.

5. V.V. Barannik and S. A. Podlesny, "The analysis of the use of technologies of error resilient coding at influence of an error in the codeword", in 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, Lviv, Ukraine, Nov, 2016. – pp. 52 – 54.

6. V.V. Barannik, S.A. Podlesny and D.V. Barannik, "Method of location loss of integrity of information based on slot-technologies", Radioelektronika i informatika, 2015. – vol. 71, No 4. – pp. 32 – 41.

7. V.V. Barannik and S.A. Podlesny, "Basis of approach for creating technology for cyber defence of video information resources in the infocommunication space", Science-based technologies, 2016. – vol. 29, No 1. – pp. 82 – 86.

МЕТОД ИНФОРМАТИВНОГО СИНТАКСИЧЕСКОГО ПРЕДСТАВЛЕНИЯ ВИДЕОИНФОРМАЦИОННЫХ РЕСУРСОВ НА ОСНОВЕ ДВУХБАЗИСНОГО БИАДИЧЕСКОГО КОДИРОВАНИЯ

Баранник В.В., Рябуха Ю.Н.

Введение

В последнее время повышенная активность проявляется в сфере использования дистанционного мониторинга на основе бортовых комплексов, для сбора информации в системах критического управления.

Практический опыт предупреждения и локализации кризисных ситуаций показывает, что актуальным является организация процесса поддержки и принятия решений на основе видеоинформационного обеспечения с использованием дистанционных средств аэромониторинга. В тоже время здесь существует ряд объективных факторов [1; 2]: ограниченность массогабаритных и энергетических возможностей бортовых комплексов; значительная удаленность от центров принятия решений; сложный рельеф местности. В свою очередь это приводит к [1; 2] повышению задержек на обработку и передачу видеоданных с борта. Как следствие формируются угрозы нарушения категорий информационной безопасности относительно доступности и целостности видеоинформационного ресурса (ВИР) [2]. Это приводит к наличию проблемы обеспечения безопасности ВИР в системах аэромониторинга кризисных ситуаций.

Значимым подходом для решения сформулированной проблемы является создание технологий и методов эффективного синтаксического представления семантического содержания видеокадров. Для повышения эффективности синтаксического описания семантического содержания ВИР, и для повышения доступности и целостности видеоинформационного ресурса предлагается выполнять последовательность этапов дифференцированной обработки сегментов видеоснимков с введением интеллектуального анализа, а именно [2 – 4]: обнаружение и локализация семантически значимой информации в видеоизображениях; выполнение сегментного анализа видеоизображений с идентификацией семантической сложности по степени насыщенности контурами; создание адаптивной дифференцированной обработки сегментов видеоизображений с учетом идентификации степени информативности их семантического содержания.

Здесь одной из ключевых составляющих является создание метода обработки сегментов видеокадров с учетом наличия контурной информации. Полученный таким образом, сегмент разделяется на контурированные видеопоследовательности [5]. В работе [5] строится подход для обработки контурированных видеопоследовательностей (КВП), который базируется на дополнительном выявлении закономерностей $\Psi^{(1)}$,

основанных на учете локально-контурных свойств КВП сегмента видеокадра. При этом необходимо учитывать, что:

1) контурированная видеопоследовательность формируется на основе незначимой и контурной составляющих;

2) структурные характеристики для синтаксического описания незначимой и контурной составляющих КВП потенциально имеют существенные отличия.

3) незначимая и контурная составляющие КВП несут значительно отличающуюся семантическую нагрузку;

4) характеристики, используемые для описания локальных структурных закономерностей не должны снижать эффективность синтаксического представления КВП относительно стандартного позиционного подхода.

Отсюда выявление локально-контурных свойств $\Psi(\xi)^{(1)}$ КВП *предлагается* осуществлять на основе учета ограниченного локального перепада $\delta(\xi)_{i,н}^{(max)}$ как для незначимой составляющей, так и локального контурного перепада $\delta(\xi)_{i,б}^{(max)}$ для контурной составляющей, т.е.

$$F(\Psi^{(1)}): \{S(\xi)_i^{(k,\ell)}; M(\xi)_{i,кл}^{(k,\ell)}\} \to \Psi^{(1)} = \{\delta(\xi)_{i,н}^{(max)}; \delta(\xi)_{i,б}^{(max)}\};$$

$$\text{для } S(\xi)_i^{(k,\ell)} = A(\xi)_i^{(k,\ell)}.$$

Данные характеристики определяются как максимальные приращения для незначимой и контурной составляющих. В этом случае функционал $F(\Psi^{(r)})$ будет задаваться соответственно следующими выражениями [6]:

$$\delta(\xi)_{i,н}^{(max)} = \max_{2 \le j \le r(\xi)_{i,н}} \delta(\xi)_{i,н}^{(j)}; \; \delta(\xi)_{i,б}^{(max)} = \max_{r(\xi)_{i,н}+1 \le j \le r(\xi)_{i,б}} \delta(\xi)_{i,б}^{(j)},$$

где $\delta(\xi)_{i,н}^{(j)}$ - локальное приращение между смежными элементами незначимой составляющей для ξ-й видеопоследовательности, $j = \overline{1, r(\xi)}_{i,н}$, т.е. $\delta(\xi)_{i,н}^{(j)} = |a_{i,j}^{(k,\ell)} - a_{i,j-1}^{(k,\ell)}|$, $j = \overline{1, r(\xi)}_{i,н}$; $\delta(\xi)_{i,б}^{(j)}$ - локальное приращение между смежными базовыми элементами для ξ-й видеопоследовательности, $j = \overline{r(\xi)_{i,н}+2, r(\xi)}_{i,б}$, т.е $\delta(\xi)_{i,б}^{(j)} = |a_{i,j}^{(k,\ell)} - a_{i,j-1}^{(k,\ell)}|$, $j = \overline{r(\xi)_{i,н}+2, r(\xi)}_{i,б}$.

Соответственно для данного подхода создания информативного синтаксического представления видеокадров необходимо разработать метод обработки двухбазисных биадических чисел, что и является *целью исследований статьи*.

1. Разработка метода создания информативного синтаксического представления КВП на основе двухбазисного биадического кодирования

Формирование кодового представления контурированной видеопоследовательности $A(\xi)_i^{(k,\ell)}$ предлагается проводить с учетом следующих особенностей [3, 6]:

1) обеспечения эффективного синтаксического представления КВП на основе структурной информации о локальных контурных перепадах;

2) обработка незначимой составляющей КВП проводится по опорным элементам;

3) количество и позиции элементов незначимой и контурной составляющих КВП определяется соответствующей маской $M(\xi)_i^{(k,\ell)}$ контурной информации.

Для создания эффективного синтаксического представления, контурированная видеопоследовательность представляется виде двухбазисного биадического числа $A(\xi)_i'^{(k,\ell)}$, $A(\xi)_i'^{(k,\ell)} = A(\xi)_{i,б}^{(k,\ell)} \cup A(\xi)_{i,o}^{(k,\ell)}$, длиной $r(\xi)_i'$, на основе функционала выявления ограничений $F(\Psi^{(1)}) = \{F(\Psi_{i,o}^{(k,\ell)});\ F(\Psi_{i,б}^{(k,\ell)})\}$, а именно:

1) первый биадический базис $\{\Lambda(\xi)_{i,o}^{(k,\ell)};\ r(\xi)_{i,o}\}$, где функционал $F(\Psi_{i,o}^{(k,\ell)})$:

$$\lambda(\xi)_{i,j}^{(k,\ell)} = \begin{cases} H(\xi)_{i,o}, & \rightarrow\ j=1; \\ 2\,\delta(\xi)_{i,o}^{(max)}+1, & \rightarrow\ j=\overline{2,\ r(\xi)}_{i,o}; \end{cases};$$

$$H(\xi)_{i,o} = \max_{1 \le j \le r(\xi)_{i,o}} a_{i,j}^{(k,\ell)} + 1;$$

$$\delta(\xi)_{i,o}^{(max)} = \max_{2 \le j \le r(\xi)_{i,o}} (|a_{i,j}^{(k,\ell)} - a_{i,j-1}^{(k,\ell)}|);$$

$$r(\xi)_{i,o} = \left| r(\xi)_{i,н} / (\nu(\xi)_i + 1) \right|;$$

задает ограничения на элементы $a_{i,j}^{(k,\ell)}$, $j=\overline{1,\ r(\xi)}_{i,o}$ допустимых незначимых составляющих (биадических чисел $A(\xi)_{i,o}^{(k,\ell)}$, $A(\xi)_{i,o}^{(k,\ell)} = \{a_{i,1}^{(k;\ell)},\ldots,a_{i,j}^{(k;\ell)},\ \ldots,\ a_{i,r(\xi)_{i,o}}^{(k;\ell)}\}$), описываемых следующей системой формул:

$$\Psi_{i,o}^{(k,\ell)}:\begin{cases} a_{i,1}^{(k,\ell)} \le \delta(\xi)_{i,o}^{(1)} = H(\xi)_{i,o} - 1; \\ a_{i,j}^{(k,\ell)} - \delta(\xi)_{i,o}^{(max)} \le a_{i,j-1}^{(k,\ell)}, \rightarrow j = \overline{2,\ r(\xi)}_{i,o}; \\ a_{i,j}^{(k,\ell)} + \delta(\xi)_{i,o}^{(max)} \ge a_{i,j-1}^{(k,\ell)}, \rightarrow j = \overline{2,\ r(\xi)}_{i,o}; \end{cases}$$

2) второй биадический базис $\{\Lambda(\xi)_{i,б}^{(k,\ell)};\ r(\xi)_{i,б}\}$, описываемый соотношениями в соответствии с функционалом $F(\Psi_{i,б}^{(k,\ell)})$:

$$\lambda(\xi)_{i,j}^{(k,\ell)} = \begin{cases} 2\delta(\xi)_{i,б}^{(r(\xi)_{i,н}+1)} + 1, \rightarrow j = r(\xi)_{i,н} + 1; \\ 2\,\delta(\xi)_{i,б}^{(max)} + 1, \rightarrow j = \overline{r(\xi)_{i,н} + 2,\ r(\xi)}_{i}; \end{cases}$$

$$\delta(\xi)_{i,б}^{(r(\xi)_{i,н}+1)} = |a_{i,r(\xi)_{i,н}+1}^{(k,\ell)} - a_{i,r(\xi)_{i,н}}^{(k,\ell)}|;$$

$$\delta(\xi)_{i,б}^{(max)} = \max_{r(\xi)_{i,н}+1 \le j \le r(\xi)_i} (|a_{i,j}^{(k,\ell)} - a_{i,j-1}^{(k,\ell)}|);$$

$$r(\xi)_i = r(\xi)_{i,н} + r(\xi)_{i,б};$$

задает ограничения на элементы $a_{i,j}^{(k,\ell)}$, $j = \overline{r(\xi)_{i,н} + 2,\ r(\xi)}_{i}$ допустимых незначимых составляющих (биадических чисел $A(\xi)_{i,б}^{(k,\ell)}$), описываемых следующей системой формул[7]:

$$\Psi_{i,б}^{(k,\ell)}:\begin{cases} a_{i,1}^{(k,\ell)} \le \delta(\xi)_{i,б}^{(r(\xi)_{i,н}+1)}, \rightarrow j = r(\xi)_{i,н} + 1; \\ a_{i,j}^{(k,\ell)} - \delta(\xi)_{i,б}^{(max)} \le a_{i,j-1}^{(k,\ell)}, \rightarrow j = \overline{r(\xi)_{i,н} + 2,\ r(\xi)}_{i}; \\ a_{i,j}^{(k,\ell)} + \delta(\xi)_{i,н}^{(max)} \ge a_{i,j-1}^{(k,\ell)}, \rightarrow j = \overline{r(\xi)_{i,н} + 2,\ r(\xi)}_{i}. \end{cases}$$

Здесь $r(\xi)_i'$ - длина контурированной видеопоследовательности $A(\xi)_i'^{(k,\ell)}$ с интерполяцией незначимой составляющей, $r(\xi)_i' = r(\xi)_{i,o} + r(\xi)_{i,б}$; $\nu(\xi)_i$ - длина аппроксимируемого участка; $\delta(\xi)_{i,o}^{(max)}$, $\delta(\xi)_{i,б}^{(max)}$ - значения локальных контурных перепадов соответственно для незначимой и базовой (контурной) составляющей ξ-й КВП; $\delta(\xi)_{i,б}^{(r(\xi)_{i,н}+1)}$ - контурное приращение на границе между незначимой и контурной составляющей КВП, $j = r(\xi)_{i,н} + 1$; $H(\xi)_{i,н}$ - диапазон значений элементов незначимой составляющей для ξ-й КВП.

В этих условиях сформулируем и докажем следующую теорему для формирования кодового идентификатора синтаксического представления контурированной видеопоследовательности.

Теорема о кодовом значении КВП (формировании функционала $F(\Psi^{(1)})^{(1)}_{к}$*).* Кодовое значение $E(\Delta(\xi)_i; r(\xi)_i)$ для неравномерной контурированной видеопоследовательности $A'(\xi)_i^{(k,\ell)}$ с маской $M(\xi)_i^{(k,\ell)}$ по опорным элементам с учетом вектора $\Delta(\xi)_i$ локальных контурных перепадов для варианта, когда индексация элементов КВП проводится без привязки к текущей позиции в строке, т.е. индексация элементов проводится внутри ξ-й контурированной видеопоследовательности, определяется по следующему соотношению [7]:

$$E(\Delta(\xi)_i; r(\xi)_i) = \sum_{\tau=1}^{r(\xi)_{i,o}-1} a_{i,\tau}^{(k,\ell)} (\delta(\xi)_{i,н}^{(max)} + 1)^{r(\xi)_{i,o}-\tau} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}} + \sum_{\tau=r(\xi)_{i,н}+1}^{r(\xi)_i} a_{i,\tau}^{(k,\ell)} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}+r(\xi)_{i,н}-\tau}.$$

Здесь $\Delta(\xi)_i$ - вектор локальных контурных перепадов для КВП

$$\Delta(\xi)_i = \{\delta(\xi)_{i,н}^{(max)}; \delta(\xi)_{i,б}^{(max)}\}.$$

Доказательство. В соответствии с принятым лексикографическим правилом определим количество $W(a_{i,1}^{(k,\ell)}; a_{i,2}^{(k,\ell)}, \ldots, a_{i,\tau-1}^{(k,\ell)})_{r(\xi)'_i}$ двухбазисных биадических чисел длиной $r(\xi)'_i$ в условиях когда: первые $(\tau-1)$ элементов фиксированы и равны соответственно $(a_{i,1}^{(k,\ell)}; a_{i,2}^{(k,\ell)}, \ldots, a_{i,\tau-1}^{(k,\ell)})$. Тогда величина $W(a_{i,1}^{(k,\ell)}; a_{i,2}^{(k,\ell)}, \ldots, a_{i,\tau-1}^{(k,\ell)})_{r(\xi)'_i}$ находится как количество перестановок с повторениями, составленное из $(r(\xi)'_i - \tau + 1)$ элементов ДББЧ, значения которых ограничены соответствующими компонентами векторов оснований $\Lambda(\xi)_{i,o}^{(k,\ell)}$ и $\Lambda(\xi)_{i,б}^{(k,\ell)}$ [8, 9]. На основе чего получим

$$W(a_{i,1}^{(k,\ell)}; a_{i,2}^{(k,\ell)}, \ldots, a_{i,\tau-1}^{(k,\ell)})_{r(\xi)'_i} = \prod_{j=\tau}^{r(\xi)'_i} \lambda(\xi)_{i,\tau}^{(k,\ell)},$$

где $r(\xi)'_i = r(\xi)_{i,o} + r(\xi)_{i,б}$ - длина контурированной видеопоследовательности $A(\xi)_i'^{(k,\ell)}$ с интерполяцией незначимой составляющей; $\lambda(\xi)_{i,\tau}^{(k,\ell)}$ - основание τ-го элемента ДББЧ, задаваемое системами формул:

$$\lambda(\xi)_{i,j}^{(k,\ell)} = \begin{cases} H(\xi)_{i,o} = \max\limits_{1 \le j \le r(\xi)_{i,o}} a_{i,j}^{(k,\ell)} + 1, & \to j = 1; \\ 2\delta(\xi)_{i,o}^{(max)} + 1 = 2 \max\limits_{2 \le j \le r(\xi)_{i,o}} (|a_{i,j}^{(k,\ell)} - a_{i,j-1}^{(k,\ell)}|) + 1, & \to j = \overline{2,\ r(\xi)}_{i,o}; \end{cases}$$

$$\lambda(\xi)_{i,j}^{(k,\ell)} = \begin{cases} 2\delta(\xi)_{i,б}^{(r(\xi)_{i,н}+1)} + 1 = 2|a_{i,r(\xi)_{i,н}+1}^{(k,\ell)} - a_{i,r(\xi)_{i,н}}^{(k,\ell)}| + 1, \\ \to j = r(\xi)_{i,н} + 1; \\ 2\delta(\xi)_{i,б}^{(max)} + 1 = 2 \max\limits_{r(\xi)_{i,н}+1 \le j \le r(\xi)_i} (|a_{i,j}^{(k,\ell)} - a_{i,j-1}^{(k,\ell)}|) + 1, \\ \to j = \overline{r(\xi)_{i,н} + 2,\ r(\xi)}_i. \end{cases}$$

При этом учитывая второе свойства ДББЧ, весовой коэффициент текущей последовательность КВП можно разбить на два сомножителя, соответствующие весам незначимой и контурной составляющих. Учитывая данное свойство, а также выражения для оснований ДББЧ, получим

$$W(a_{i,1}^{(k,\ell)}; a_{i,2}^{(k,\ell)}, \dots, a_{i,\tau-1}^{(k,\ell)})_{r(\xi)_i'} = (\delta(\xi)_{i,o}^{(max)} + 1)^{r(\xi)_{i,o}' - \tau + 1} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}}.$$

Последовательности, удовлетворяющие перечисленным свойствам, образуют множество $\Omega(\Delta(\xi)_i; r(\xi)_i' - \tau + 1)$ двухбазисных биадических чисел. Данные последовательности будут предшествовать обрабатываемому ДББЧ, и в соответствии с лексикографическим правилом иметь меньшие порядковые номера в допустимом множестве $\Omega'(\xi)_i^{(k,\ell)}$. В тоже время с учетом четвертого свойства ДББЧ текущую последовательность КВП можно разбить на две последовательности, образуемые незначимой и контурной составляющей [8, 10]. После чего проведя суммирования по всем τ, где $\tau = \overline{1,\ r(\xi)'}$, получим:

$$E(\Delta(\xi)_i; r(\xi)_i) = \sum_{\tau=1}^{r(\xi)'} a_{i,\tau}^{(k,\ell)} W(a_{i,1}^{(k,\ell)}; a_{i,2}^{(k,\ell)}, \dots, a_{i,\tau-1}^{(k,\ell)})_{r(\xi)_i'} =$$

$$= \sum_{\tau=1}^{r(\xi)_{i,o}} a_{i,\tau}^{(k,\ell)} (\delta(\xi)_{i,o}^{(max)} + 1)^{r(\xi)_{i,o} - \tau} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}} +$$

$$+ \sum_{\tau=r(\xi)_{i,н}+1}^{r(\xi)_i} a_{i,\tau}^{(k,\ell)} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б} + r(\xi)_{i,н} - \tau}.$$

Для первого слагаемого правой части данного соотношения индексация элементов осуществляется по опорным элементам незначимой составляющей КВП. Для второго слагаемого индексация элементов проводится с учетом позиций контурной составляющей в исходном КВП до аппроксимации, т.е. относительно начиная с позиции $(r(\xi)_{i,н} + 1)$.

Теорема доказана.

На основе доказанной теоремы можно получить значения кодов $E(\Delta(\xi)_i; r(\xi)_i)$ для разных вариантов индексации элементов ДББЧ. Здесь возможны следующие основные варианты:

1. Индексация элементов ДББЧ проводится с учетом текущей j-й позиции в i-й строке, но без учета позиций опорных элементов незначимой составляющей, т.е.

$$E(\Delta(\xi)_i; r(\xi)_i) = \sum_{\tau=0}^{r(\xi)_{i,o}-1} a_{i,j+\tau}^{(k,\ell)} (\delta(\xi)_{i,o}^{(max)} + 1)^{r(\xi)_{i,o}-\tau-1} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}} + \\ + \sum_{\tau=1}^{r(\xi)_{i,б}} a_{i,r(\xi)_{i,н}+\tau}^{(k,\ell)} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}-\tau}.$$

2. Индексация организуется: с учетом текущей j-й позиции элемента в i-й строке; с учетом позиций опорных элементов в незначимой составляющей КВП. В этом случае получим такое соотношение:

$$E(\Delta(\xi)_i; r(\xi)_i) = \sum_{\tau=0}^{r(\xi)_{i,o}-1} a_{i,j+\tau\nu(\xi)_i}^{(k,\ell)} (\delta(\xi)_{i,o}^{(max)} + 1)^{r(\xi)_{i,o}-\tau-1} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}} + \\ + \sum_{\tau=1}^{r(\xi)_{i,б}} a_{i,r(\xi)_{i,н}+\tau}^{(k,\ell)} (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б}-\tau}.$$

Для свертки двух слагаемых правой части полученного соотношения в одно выражение, ***предлагается*** ввести ***признак интервала***, т.е. признак того, что позиция τ-го текущего обрабатываемого элемента не вышла за пределы незначимой последовательности, т.е. $\tau \le r(\xi)_{i,o}$. Это задается таким функционалом $\varphi(\tau; r(\xi)_{i,o})$ [6]:

$$\varphi(\tau; r(\xi)_{i,o}) = \mathrm{sign}(1 - \mathrm{sign}(\tau - r(\xi)_{i,o})) = \begin{cases} 1, \to \tau \le r(\xi)_{i,o}; \\ 0, \to \tau > r(\xi)_{i,o}. \end{cases}$$

Далее введем обозначения обратного функционала $\overline{\varphi(\tau; r(\xi)_{i,o})}$:

$$\overline{\varphi(\tau; r(\xi)_{i,o})} = 1 - \varphi(\tau; r(\xi)_{i,o}) = (1 - \mathrm{sign}(1 - \mathrm{sign}(\tau - r(\xi)_{i,o}))).$$

С учетом чего получим обобщенное выражение

$$E(\Delta(\xi)_i; r(\xi)_i) = \sum_{\tau=0}^{r(\xi)'_i} a_{i,j+\tau\nu(\xi)_i}^{(k,\ell)} {}^{\varphi(\tau; r(\xi)_{i,o}) + (\tau - r(\xi)_{i,o} + r(\xi)_{i,н})(1 - \varphi(\tau; r(\xi)_{i,o}))} \times \\ \times (\delta(\xi)_{i,н}^{(max)} + 1)^{(r(\xi)_{i,o}-\tau-1)\varphi(\tau; r(\xi)_{i,o})} \times \\ \times (\delta(\xi)_{i,б}^{(max)} + 1)^{r(\xi)_{i,б} - (\tau - r(\xi)_{i,o} + r(\xi)_{i,н})(1 - \varphi(\tau; r(\xi)_{i,o}))}.$$

В базисе формализованных множеств цепочка обработки синтаксического представления контурированной последовательности будет выглядеть следующим образом [4, 9]:

1) функциональное преобразование $\Psi^{(1)} = \Delta(\xi)_i$ относительно выявление ограничений для синтаксического представления КВП $A'(\xi)_i^{(k,\ell)}$;

2) функциональное преобразование $F(\Psi^{(1)}) = \{F(\Psi_{i,o}^{(k,\ell)}); F(\Psi_{i,б}^{(k,\ell)})\}$ относительно метода выявления множества закономерностей $\Psi^{(1)} = \Delta(\xi)_i$:

$$F(\Psi^{(1)}): \{S(\xi)_i^{(k,\ell)}; M(\xi)_{i,кл}^{(k,\ell)}\} \to \Psi^{(1)} = \{\delta(\xi)_{i,н}^{(max)}; \delta(\xi)_{i,б}^{(max)}\};$$

$$\text{для } S(\xi)_i^{(k,\ell)} = A(\xi)_i^{(k,\ell)};$$

3) функциональное преобразование $F(\Psi^{(1)})_к^{(1)}$, задающее метод кодирования (синтаксического преобразования, соответствующего семантического содержания) ВИР с учетом множества $\Psi^{(1)} = \Delta(\xi)_i$ выявленных закономерностей, задается как:

$$F(\Psi^{(1)})_к^{(1)} = f(\{\Lambda(\xi)_{i,o}^{(k,\ell)}; r(\xi)_{i,o}\}; \{\Lambda(\xi)_{i,б}^{(k,\ell)}; r(\xi)_{i,б}\});$$

$$F(\Psi^{(r)})_к^{(r)}: \{S; M_{кл}; \Psi^{(r)}\} \to W.$$

Соответственно отображение в эффективное (информативное) синтаксическое представление $W(\xi)_i^{(k,\ell)} = E(\Delta(\xi)_i; r(\xi)_i)$ формируется с учетом $S(\xi)_i^{(k,\ell)} = A'(\xi)_i^{(k,\ell)}$; $M_{кл} = M(\xi)_i^{(k,\ell)}$; $\Psi^{(1)} = \Delta(\xi)_i$ по следующему соотношению:

$$F(\Psi^{(1)})_к^{(1)}: \{A'(\xi)_i^{(k,\ell)}; M(\xi)_i^{(k,\ell)}; \Delta(\xi)_i\} \to W(\xi)_i^{(k,\ell)} = E(\Delta(\xi)_i; r(\xi)_i).$$

Значит, можно заключить, что:

1. Построена технологическая реализация режимов кодирования конструированных видеопоследовательностей, когда: индексация элементов ДББЧ проводится с учетом текущей позиции элемента в строке, но без учета позиций опорных элементов незначимой составляющей; индексация элементов организуется: с учетом текущей позиции элемента в строке с учетом позиций опорных элементов в незначимой составляющей КВП. Это обеспечивает возможность интегрирования созданного информативного представления в различных условиях построение базовой платформы обработки видеокадров.

2. Создан метод кодирования двух базисных биадических чисел с учетом свертки кодовых составляющих незначимой и контурной составляющих КВП в единое число на основе функционала, задающего признак интервала КВП, т.е. признак идентификации позиций элементов относительно незначимой и контурной составляющих.

Таким образом вопрос интеграции созданных методов и технологий кодировки ДВИР в единственный комплекс обработки находится на недостаточном уровне проработки.

Поэтому ***цель исследований*** заключается в разработке метода верификации обработки видеоинформационного ресурса на основе формирования базовых уровней построения кодовых конструкций.

2. Основная часть исследований

Верификация разработанной кодировки в систему формирования информативного синтаксического описания видеокадра с учетом их идентификации за степенью семантической информативности предусматривает процесс интегрирования, для которого нужно обеспечить:

1.Заданный уровень семантической целостности получается после реконструкции статичных ВИР. Нужно, чтобы интегрированная технология не должна снижать уровень целостности ВИР, какой устанавливается для всей системы обработки [9].

2.Необходимый уровень информативности синтаксического описания, которое отвечает требованиям относительно доступности статичных ВИР в системах аэромониторинга. Нужно обеспечить автоматическое соответствие между уровнем семантической информативности сегментов и уровнем синтаксической информативности что формируется в результате кодировки КВП.

3. Возможность обработки служебных данных, которые формируются внедряемой технологией кодировки, базовыми средствами для созданной системы обработки видеокадра. Нужно обеспечить совместимость средств обработки служебных данных в созданной системе для служебных сведений технологии что интегрируется.

Рассмотрим технологические аспекты, которые используются при обеспечении данных условий [7-9].

Обработка сегментов проводится с учетом предыдущей их интеллектуальной идентификации за степенью семантической информативности. В результате строится карта (маска) контурной информации сегмента и его семантический идентификатор. Эта информация используется на втором концептуальном этапе обработки ВИР, а именно:

– во-первых, на основе контурной маски информации проводится сегментация видеокадра на контурованые видиопоследовательности. Здесь маска обеспечивает установление взаимооднозначного позиционирования незначительной и контурной составляющих контурованных видеопоследовательностей. Следовательно, введение дополнительной служебной информации относительно позиционирования составляющих КВП не нужно.

– во-вторых, маска обеспечивает установление длины и режима аппроксимации незначительной составляющей. Это позволяет формировать синтаксическое представление КВП, плотность которого

автоматически учитывает степень насыщенности контурной информации данной области сегмента. В конечном результате этот механизм обеспечивает установление соответствия между уровнями семантической и синтаксической плотностью описания всего видеокадра.

– в-третьих, на основе информации о позициях составляющих КВП обеспечивается возможность построения базисов биадичного пространства для незначительной и контурной составляющих КВП. Это позволяет создать условия для взаимно-однозначного процесса построения информативного синтаксического представления КВП на основе двухбазисного биадичного кодировки.

– в-четвертых, на основе информации о базисах биадичных пространств незначительной и контурной составляющих обеспечивается взаимно-однозначное установление режима двоичного кодообразования для неравномерных кодограмм.

Следовательно, обеспечивается ***совместимость технологических аспектов двух концепций*** обработки статичных ВИР относительно поддержки выполнения условия доступности и целостности на уровне формирования ***информативных составляющих*** кодовых конструкций синтаксического описания.

Рассмотрим теперь особенности ***совместимости обработки служебных данных*** для двух концептуальных составляющих системы обработки статичных ВИР. Для первой концепции служебными данными являются векторы признаков наличия контурных элементов на позиции в маске. Данная контурная информация будет использоваться для дальнейшей обработки контурной маски информации для создания информативного синтаксического представления. Здесь используются методы, изложенные в работах [5].

Для второй концепции относительно наличия информации о контурных масках дополнительными служебными сведениями следующие: вектора оснований, соответственно для незначительной и контурной составляющих КВП. Обработка этой информации предусматривается в создаваемой базовой системе путем интегрирования методов обработки оснований биадичного пространства без потери информации [2].

Следовательно, на основе изложенного можно утверждать, что базовые концепции отвечают требованиям совместимости из формирования информативных и служебных частей кодовых конструкций информативного синтаксического описания видеокадров с учетом их степени семантической информативности.

Теперь разработаем структуру обобщенных кодовых конструкций синтаксического представления видеокадра с использованием двух базовых концепций обработки ВИР. Кодовые конструкции содержат четыре иерархических уровня. Первый иерархический уровень строится на

основе совокупности минимальных структурных единиц $C(\xi)_i^{(k,\ell)}$ кодового представления сегмента видеокадра. Минимальной структурной единицей информативного синтаксического описания сегмента видеокадра является кодограмма $C(\Delta(\xi);E(\xi))_i^{(k,\ell)}$ контуровaной видеопоследовательности для строки сегмента видеокадра.

Даная кодограмма содержит информационную и служебную части. Информационная часть кодограммы является неравномерной, и содержит в себе информацию о значении кода $E(\Delta(\xi)_i;r(\xi)_i)$ двухбазисного бiадичного числа, сформированного на основе контурной видеопоследовательности. Служебная часть $C(\Delta(\xi))_i^{(k,\ell)}$ включает информацию о: векторы основ $\Lambda(\xi)_{i,o}^{(k,\ell)}$, $\Lambda(\xi)_{i,б}^{(k,\ell)}$ соответственно для незначительной и контурной составляющих КВП.

Длина $V(\Delta;E)_i^{(k,\ell)}$ двоичного описания данного уровня определяется за формулой:

$$V(\Delta;E)_i^{(k,\ell)} = \sum_{\xi=1}^{\nu(i)_{квп}^{(k,\ell)}} V(\Delta(\xi);E(\xi))_i^{(k,\ell)}, \qquad (1)$$

где $V(\Delta(\xi);E(\xi))_i^{(k,\ell)}$ – длина кодового представления кодовой конструкции $C(\Delta(\xi);E(\xi))_i^{(k,\ell)}$ для ξ-й КВП, то есть

$$V(\Delta(\xi);E(\xi))_i^{(k,\ell)} = V(\xi)_{i,max}^{(k,\ell)} + V(\Delta(\xi))_i^{(k,\ell)}; \qquad (2)$$

$V(\xi)_{i,max}^{(k,\ell)}$ – максимальное количество разрядов на двоичное кодообразование $L(\xi)_i^{(k,\ell)}$ кодового значения $E(\Delta(\xi)_i;r(\xi)_i)$ для ξ- й контуровaной видеопоследовательности i-й сроки $(k;\ell)$-го сегмента видеокадра;

$V(\Delta(\xi))_i^{(k,\ell)}$ – количество разрядов на представление служебной составляющей ξ-й кодовой конструкции;

$\nu(i)_{квп}^{(k,\ell)}$ – количество КВП у i-й сроки $(k;\ell)$-го сегмента видеокадра.

На структурных единицах строятся комплексные составляющие следующего высшего уровня. Таким уровнем является уровень строк сегмента $S^{(k,\ell)}$ видеокадра.

Здесь кодовые конструкции $C(M;\Delta;E)_i^{(k,\ell)}$ содержат:

- информативную составляющую $C(\Delta;E)_i^{(k,\ell)}$, образованную на основе кодовых конструкций предыдущего уровня иерархии, то есть кодовые конструкции строк сегмента;

- служебную составляющую $C(M)_i^{(k,\ell)}$, которая содержит информацию о соответствующей сроку $M_i^{(k,\ell)}$ маски контурной информации.

Суммарная длина $V(\Delta;E)^{(k,\ell)}$ информационной части кодовых конструкций данного уровня определяется за формулой:

$$V(\Delta;E)^{(k,\ell)} = \sum_{i=1}^{\nu_{см}} V(\Delta;E)_i^{(k,\ell)},$$

или с учетом выражений (1) и (2), получим

$$V(S^{(k,\ell)}) = V(\Delta;E)^{(k,\ell)} = \sum_{i=1}^{\nu_{см}} \sum_{\xi=1}^{\nu(i)_{квп}^{(k,\ell)}} V(\Delta(\xi);E(\xi))_i^{(k,\ell)} =$$

$$= \sum_{i=1}^{\nu_{см}} \sum_{\xi=1}^{\nu(i)_{квп}^{(k,\ell)}} (V(\xi)_{i,max}^{(k,\ell)} + V(\Delta(\xi))_i^{(k,\ell)}). \quad (3)$$

Дальше с учетом кодовых посылок информативного синтаксического представления отдельных строк сегментов строится уровень кодовых конструкций $C(M;\Delta;E)^{(k,\ell)}$ сегментов видеокадра [7].

Данный уровень состоит из:

1) информационной части $C(\Delta;E)^{(k,\ell)}$, содержит кодовые конструкции $C(\Delta;E)_i^{(k,\ell)}$ синтаксического описания срок сегментов;

2) служебной части, которая включает у себя кодовое представление:

- $C(M)^{(k,\ell)}$ маски $(M)^{(k,\ell)}$ контурной информации;

- $C(\Theta)_i^{(k,\ell)}$, $C(\Theta)_j^{(k,\ell)}$ вектору $\Theta_i^{(k,\ell)}$, $\Theta_j^{(k,\ell)}$ признаков наличия контурных элементов соответственно в строках и столбцах сегментах.

Суммарная длина $V(\Delta;E)$ кодового представления уровня сегментов видеокадра находится с использованием следующего выражения [8]:

$$V(\Delta;E) = \sum_{k=1}^{N_1} \sum_{\ell=1}^{N_2} (V(M)^{(k,\ell)} + V(\Theta)_i^{(k,\ell)} + V(\Theta)_j^{(k,\ell)} + V(\Delta;E)^{(k,\ell)}), \quad (4)$$

где N_1, N_2 - количество сегментов в соответствии с направлением строк и столбцов видеокадра;

$V(M)^{(k,\ell)}$ - количество разрядов на представление маски $(k;\ell)$-го сегмента видеокадра;

$V(\Theta)_i^{(k,\ell)}$, $V(\Theta)_j^{(k,\ell)}$ - количество разрядов на представление векторов признаков наличия контурных элементов в строках и столбцах $(k;\ell)$-го сегмента.

Соответственно из отдельных кодовых конструкций сегментов формируется уровень синтаксического описания всего видеокадра. Данный уровень образует кодовую конструкцию всего информативного синтаксического описания статичного ВИР [3-5]. Уровень включает следующие составляющие:

- $C(S)$ информационную, что содержит кодовые конструкции отдельных сегментов $C(M;\Delta;E)^{(k,\ell)}$;

- $V(M)$ информационной синтаксической маски контурной информации;

- $V(\Theta)_i$, $V(\Theta)_j$ кодового описания векторов признаков наличия контурных элементов в сегментах.

Откуда общая длина $V(S)$ кодового представления уровня видеокадра оценивается с помощью такого соотношения:

$$V(S)_{\sum} = V(S)_{\text{інф}} + V(S)_{\text{сл}}, \quad (5)$$

$$V(S)_{\text{інф}} = \sum_{k=1}^{N_1}\sum_{\ell=1}^{N_2} V(\Delta;E)^{(k,\ell)} = \sum_{k=1}^{N_1}\sum_{\ell=1}^{N_2}\left(\sum_{i=1}^{\nu_{\text{см}}}\sum_{\xi=1}^{\nu(i)_{\text{квп}}^{(k,\ell)}} \left(V(\xi)_{i,\max}^{(k,\ell)} + V(\Delta(\xi))_i^{(k,\ell)}\right)\right), \quad (6)$$

где $V(S)_{\text{інф}}$ - длина информационной части данного уровня иерархии кодовых конструкций;

$V(S)_{\text{сл}}$ - суммарная длина служебной составляющей для всего видеокадра.

Данное соотношение позволяет оценить синтаксическую плотность видеокадра без учета семантической информативности.

Выводы

1. Обосновано, что кодового представления контурированной видеопоследовательности требуется проводить с учетом следующих особенностей:

1) обеспечения эффективного синтаксического представления КВП на основе структурной информации о локальных контурных перепадах;

2) обработка незначимой составляющей КВП проводится по опорным элементам;

3) количество и позиции элементов незначимой и контурной составляющих КВП определяется соответствующей маской $M(\xi)_i^{(k,\ell)}$ контурной информации.

2. Разработан метод создания информативного синтаксического представления статических видеоинформационных ресурсов. Данный метод основан на следующих концептуальных составляющих:

- композиции незначимой и контурной составляющих яркостного описания сегмента как контурированной видеопослдеовательности;

- сегментации видеокадра по контурированным видеопоследовательностям на основе информации о маски контурной информации;

- аппроксимацию контурированной видеопоследовательности двухбазисным биадическим числом с ограничениями на локально-пространственные характеристики КВП;

- технология двухбазисного биадического кодирования, обеспечивающее формирование кодового значения информативного синтаксического представления для неравномерной контурированной видеопоследовательности с маской по опорным элементам с учетом вектора локальных контурных перепадов для варианта, когда индексация элементов КВП проводится без привязки к текущей позиции в строке, т.е. индексация элементов проводится внутри контурированной видеопоследовательности.

3. Построена технологическая реализация режимов кодирования контурированных видеопоследовательностей, когда: индексация элементов ДББЧ проводится с учетом текущей позиции элемента в строке, но без учета позиций опорных элементов незначимой составляющей; индексация элементов организуется: с учетом текущей позиции элемента в строке с учетом позиций опорных элементов в незначимой составляющей КВП. Это обеспечивает возможность интегрирования созданного информативного представления в различных условиях построение базовой платформы обработки видеокадров.

4. Создан метод кодирования двухбазисных биадических чисел с учетом свертки кодовых составляющих незначимой и контурной составляющих КВП в единое число на основе функционала, задающего признак интервала КВП, т.е. признак идентификации позиций элементов относительно незначимой и контурной составляющих.

Литература

1. Кашкин, В. Б. Цифровая обработка аэрокосмических изображений [Текст] : конспект лекций / В. Б. Кашкин. – Красноярск.: ИПК СФУ, 2008. – 121 с.

2. Баранник, В. В. Методологический анализ системы аэрокосмического видеомониторинга чрезвычайных ситуаций [Текст] / В. В. Баранник, А. В. Яковенко, А. Ю. Школьник // Сучасна спеціальна техніка, 2011. – № 4 (27). – С. 12 – 22.

3. Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон. – М: Техносфера, 2004. – 368 с.

4. Красильников Н.Н. Цифровая обработка изображений. – М.: Вузовская книга, 2011. – 320 с.

5. Баранник В.В. Метод интеллектуальной обработки государственных видеоинформационных ресурсов для повышения их семантической целостности в системах мониторинга кризисных ситуаций / В.В. Баранник, Ю.Н. Рябуха // Захист інформації, 2015. - №2. – С. 32 – 40.

6. В.В. Баранник, Ю.Н. Рябуха, (2015), Метод повышения информационной безопасности в системах видеомониторинга кризисных ситуаций. – Черкассы.: ЧТУ,

2005. – 143 с.

7. Barannik, V., Krasnorutskiy, A., Ryabukha, Y.N., Okladnoy, D.E. Model intelligent processing of aerial photographs with a dedicated key features interpretation, February 2016. – pp. 736 – 738.

8. Barannik, V., Shulgin, S.S. The method of increasing accessibility of the dynamic video information resource, Lviv-Slavsko; Ukraine; 23 February 2016. – pp. 621 – 623.

9. Barannik, V., Ryabukha, Y., Krasnorutskyy, A. Method of effective syntactic description of frames using the contour information to improve the integrity of the video information resource. – Kharkiv, 2015 - 15 October 2015. – pp. 253 – 256.

10. Barannik, V., Shiryaev, A. Quadrature compression of images in polyadic space? Lviv – Slavske, 24 February 2012. – p. 422.

ON SOME FALSE ASSERTIONS IN IMAGE LOSSY COMPRESSION

Kozhemyakin R.A., Zemliachenko A.N., Abramov S.K., Lukin V.V., Vozel B.

Images have become one of the main types of information. They are widely used in various applications including medical diagnostics, remote sensing (RS), non-destructive testing, digital cameras, etc. [1, 2]. Amount of acquired images increases rapidly. Capacities of transmission lines and storage devices increase too but not so fast. This forces employing image compression where lossless coding methods have limited application due to small and uncontrolled compression ratio attained [1]. This makes lossy compression methods popular and, often, the only practical opportunity to deal with such kind of data.

On one hand, there are known standards of image compression JPEG and JPEG2000 [1, 3, 4]. The work intended on design of standard for lossy compression of multichannel RS images is close to completing too [5]. Lossy compression techniques have been practically accepted by medical imaging community [6]. On the other hand, there are quite many other than standard methods of lossy image compression, some of them outperform JPEG and JPEG2000 [7].

Most modern methods of lossy compression are based on orthogonal transforms (either discrete cosine transform (DCT) or discrete wavelet transform (DWT)) where losses (distortions) are introduced at the stage of transform coefficient quantization. Note that quantization can be done either with fixed (uniform) quantization step (QS) or with specific non-uniform (non-equal) steps. The latter approach is able to provide certain benefits, e.g., to produce better visual quality for given (fixed) compression ratio (CR) [3, 4, 7].

The main characteristic that describes a method performance (efficiency) for lossy compression is the so-called rate-distortion curve that can be represented in different forms. This can be dependence of mean square error (MSE) of introduced distortions on CR, quantization step (QS) or bpp (bits per pixel) or dependence of peak signal-to-noise ratio (PSNR) on the same parameters. Sometimes, one can use visual quality metrics instead of MSE or PSNR [7-9]. Fig. 1 shows one example how rate-distortion curves look like for different coders. The results are presented for the visual quality metric PSNR-HVS-M [9] that takes into account two important aspects of human visual system (HVS) – higher sensitivity to distortions in low spatial frequencies and masking affects of texture and other heterogeneities. Values of PSNR-HVS-M are expressed in dB and larger values correspond to better visual quality (the metric PSNR-HVS-M can be optionally used in Xvid codec, see http://www.digital-digest.com/news-62908-Xvid-130-Released.html). Note that the threshold 40 dB approximately corresponds to invisibility of distortions [10]. The curves have been obtained for the test image Lena that has a rather simple structure.

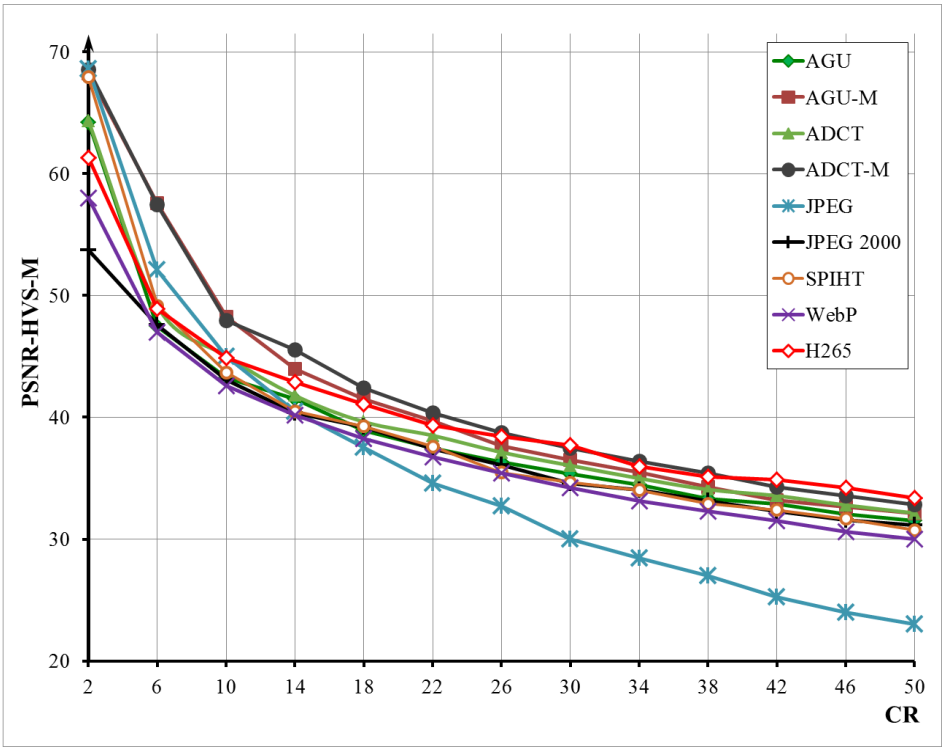


Fig. 1. Dependence of PSNR-HVS-M on CR for different coders, test image Lena

Analysis of the presented curves confirms some statements given above. In particular, it is seen that difference in PSNR-HVS-M values for different compression methods can be up to 11 dB for the same CR (see data for CR=50) – thus, the choice of a proper compression technique is important. Lossy compression with even such a large CR as 16 does not lead to degradation of visual quality for the considered image and, thus, can be thought as an excellent alternative to lossless compression. Besides, the data in Figure 1 confirm that there are compression methods that, in the sense of visual quality, sufficiently outperform both JPEG and JPEG2000. They are, in particular, H265 (available for intra-frame mode from https://hevc.hhi.fraunhofer.de), ADCT-M and AGU-M [11].

However, it is impossible to make reliable conclusions based on analysis of data for only one test image. The results for complex structure (highly textural) image Baboon are presented in Fig. 2. Their analysis shows the following. H265 occurs not as good as it is for low complexity images. JPEG with non-uniform quantization outperforms many compression techniques for CR<12 except ADCT-M and AGU-M. Distortions can be visible even for such small values of CR as 4…6. Thus, the presented examples show some problems of analyzing and comparing compression techniques. Such problems often lead to wrong conclusions and false assertions in lossy compression of images. Below we consider some of them.

In scientific publications and monographs recommended for university courses, it is often possible to meet the following statement formulated as postulates:

1) the standard JPEG2000 possesses considerably better characteristics compared to JPEG and this was the reason for accepting JPEG2000 as new standard; besides, it follows that DWT is better for compression than DCT;

2) one obvious advantage of JPEG2000 compared to JPEG is an opportunity to easily (in non-iterative manner) provide a given (desired) CR or bpp whilst it is difficult to provide a desired CR for JPEG and, in general, for any compression based on DCT [3, 7];

3) for all compression methods, based on both DWT and DCT, it seems problematic to provide a desired value of quality metric (MSE, PSNR or others) without iterations; the task is usually solved by multiple compression/decompression of a given image, metric calculation and the corresponding changing of a parameter that controls compression (PCC) – this can be bpp, QS or scaling factor (SF) [3, 7];

4) if images are corrupted by noise [10, 12], they are compressed essentially worse than the same images without noise for the same level of introduced distortions.

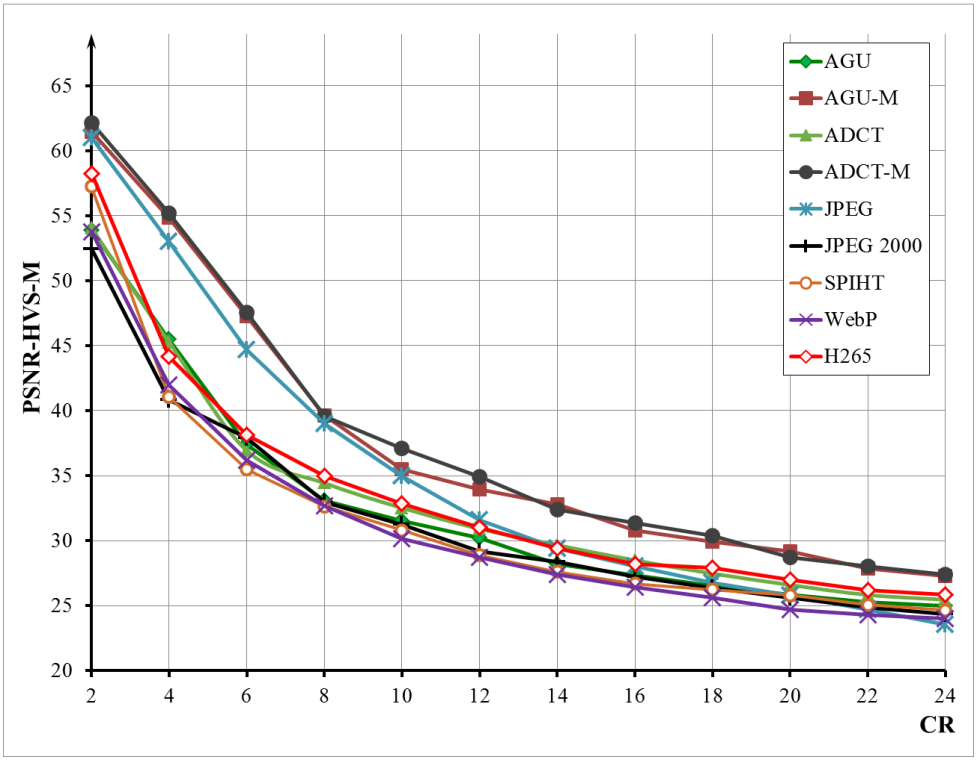


Fig. 2. Dependence of PSNR-HVS-M on CR for different coders, test image Baboon

First of all, let us consider the first statement more in details. Recall that it is possible to compare two compression methods for either the same CR (or bpp) or fixed value of a used metric. Even the examples given in Figures 1 and 2

show that JPEG with non-uniform quantization is able to perform better than JPEG2000 and SPIHT. This happens for CR<24 for the test image Lena (Fig. 1) and for CR<16 for the test image Baboon (Fig. 2). This shows that JPEG and JPEG2000 can, at least, compete for the range of compressed image quality that is of value in practice.

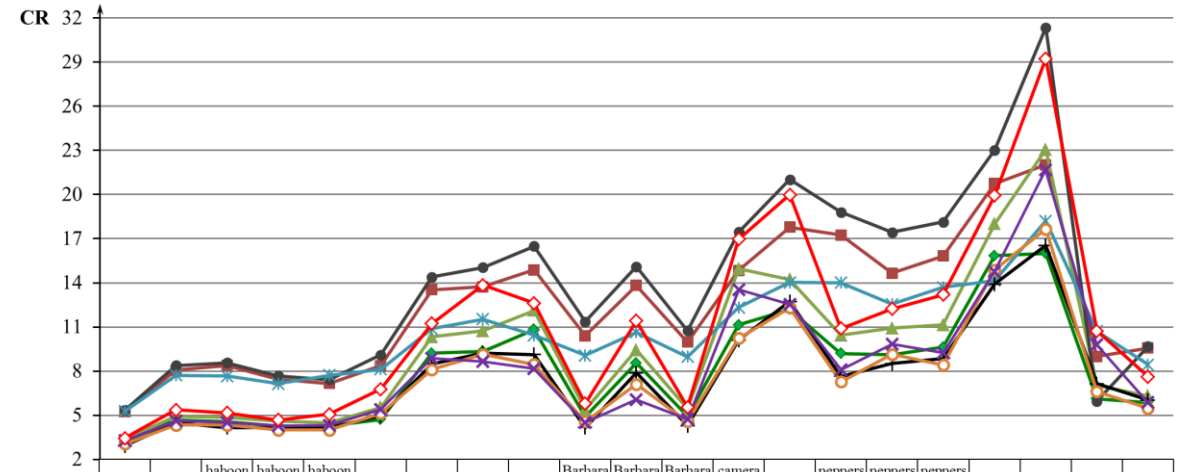
To get more data for analysis, let us present some results from [7] that allow comparing CRs for a large number of compression techniques (the same as for data in Figures 1 and 2). All test images were grayscale. In our set, there are more or less standard test images used, namely, Grass, Baboon, Airfield, Goldhill, Boat, Barbara, Cameraman, Peppers, Lenna, Pole. Grass and Baboon are examples of highly textural images; Airfield is one example of RS test image while the test image called Aviris is the second example. Besides, there are three color components of color versions of the test images Baboon, Barbara and Peppers. Considered as grayscale images, these component images exhibit considerable correlation (for a given color image) and allow comparing performance of different compression techniques. In some of these images, noise is noticeable and this allows analyzing its influence on compression parameters. Finally, one more test image called Text is included into our set. It is an example of document images for which peculiarities of lossy compression have to be studied too.

Some details have to be clarified. Firstly, there are different types of non-uniform quantization of coefficients. We have employed quantization Table recommended for intensity component of JPEG. The standard version of JPEG2000 has been used although there are special versions intended on improved visual quality. AGU and ADCT are DCT based coders that exploit uniform quantization step [13, 14]. AGU uses 32x32 pixel blocks, modern bit-plane coding and deblocking after decompression. ADCT employs adaptive partition scheme and modern coding of quantized DCT coefficients. AGU and ADCT coders have versions adapted to visual quality called AGU-M and ADCT-M [11]. Besides, we consider such popular coders as SPIHT, WebP and H.265.

Fig. 3 present the results for all aforementioned test images and compression techniques for PSNR-HVS-M equal to 40 dB. The results essentially depend upon the test image. For complex structure images (Grass, Baboon, components of color Baboon, Airfield), compression techniques "divide" into two groups where methods adapted to visual quality (JPEG, AGU-M, ADCT-M) clearly outperform other ones. For other test images, there is diversity of obtained values of CR. For some test images as Text, JPEG seems to be the best according to the considered metric. For most others, ADCT-M is the best. The results for SPIHT and JPEG2000 are among the worst for all test images.

One can argue that PSNR-HVS-M is the metric based on DCT and, due to this, DCT-based coders have got "certain benefits" in assessment of their

performance. So, let us also analyze the results based on other HVS metric, namely, MSSIM [15] which is one of the basic metrics and is not based on DCT. The results are presented in Fig. 4 and Table under it.



	grass	baboon	baboon red	baboon green	baboon blue	airfield	goldhill	boat	Barbara	Barbara red	Barbara green	Barbara blue	camera man	peppers	peppers red	peppers green	peppers blue	Lenna	pole	Text	Aviris
AGU	3.22	4.68	4.57	4.27	4.28	4.71	9.22	9.34	10.83	4.9	8.55	4.94	11.14	12.2	9.19	9.11	9.63	15.85	16	6.13	5.86
AGU-M	5.26	8.07	8.39	7.43	7.17	8.37	13.52	13.73	14.88	10.4	13.82	9.97	14.84	17.79	17.25	14.66	15.82	20.74	22.03	9	9.57
ADCT	3.36	4.94	4.85	4.6	4.49	5.54	10.32	10.74	12.13	5.37	9.44	5.44	14.95	14.24	10.43	10.92	11.15	18	23.07	7.04	6.31
ADCT-M	5.34	8.38	8.57	7.68	7.42	9.1	14.4	15.04	16.49	11.37	15.11	10.76	17.45	21.03	18.8	17.44	18.14	23.03	31.35	5.95	9.68
JPEG	5.27	7.71	7.66	7.14	7.71	8.13	10.89	11.53	10.43	9.06	10.65	8.98	12.32	14.02	14.01	12.54	13.69	14.16	18.21	10.72	8.43
JPEG2000	2.93	4.49	4.16	4.15	4.15	4.96	8.49	9.21	9.13	4.17	7.89	4.30	10.11	12.74	7.63	8.50	8.87	13.92	16.54	7.14	6.04
SPIHT	3.06	4.32	4.32	4	4	5.12	8.1	9.14	8.46	4.57	7.11	4.57	10.23	12.3	7.27	9.14	8.42	14.88	17.65	6.6	5.48
WebP	3.29	4.60	4.51	4.28	4.33	5.39	8.88	8.66	8.17	4.52	6.05	4.73	13.55	12.54	8.08	9.82	9.26	14.75	21.67	9.82	5.89
H265	3.44	5.36	5.17	4.67	5.07	6.78	11.23	13.84	12.62	5.79	11.43	5.55	16.98	20.00	10.89	12.24	13.20	19.94	29.24	10.73	7.63

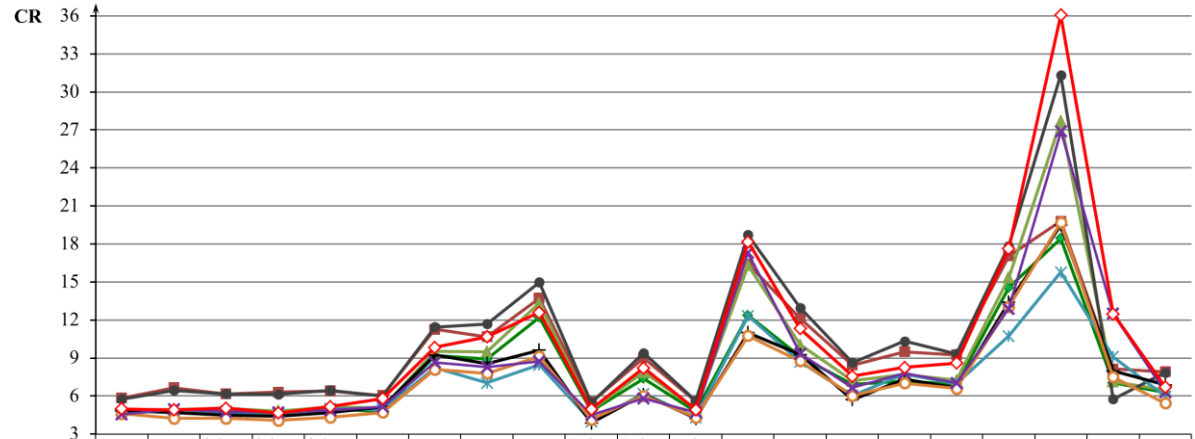
Fig. 3. CR values for different test images and used coders with the provided PSNR-HVS-M=40 dB

The metric MSSIM is within the range from 0 to 1 where the value 0.99 approximately corresponds to distortion invisibility threshold [10]. The results in Fig. 4 show the following. Firstly, they are in several senses similar to those in Fig. 3. Again, ADCT-M is among the best coders whilst SPIHT is among the worst for all test images. JPEG is at the same level as JPEG2000. H265 is good enough except the cases of texture images. ADCT-M is not good for the test image Text but it is among the best for RS test images.

More data for comparisons can be found in [7]. But even the results presented above show that JPEG provides visual quality of compressed images that are often better than for standard versions of JPEG2000 (this happens for PSNR about 32…37 dB and PSNR-HVS-M in the limits 35…40 dB). In fact, JPEG has worse performance compared to JPEG2000 for large CR values. But then introduced distortions are clearly seen and become annoying for both compression techniques. Thus, such CR values are used in practice only if providing of very large CR is of prime importance. The presented data partly explain why JPEG2000 has not found wide use in digital cameras.

It is worth noting that advantages of JPEG2000 compared to JPEG are partly explained by the use of more sophisticated methods of coding quantized transform coefficients. In recent 15 years, quite many publications appeared (for example, [16, 17]) where the authors show that more efficient coding (re-

encoding) of quantized DCT coefficients is possible. This allows reaching 10…25% larger values of CR than for standard JPEG for the same quality. Then, the corresponding modifications of JPEG become better or comparable to JPEG2000 for most practical situations. This also shows that DCT is not worse orthogonal transform than wavelets in the sense of redundancy reduction (decorrelation ability).



	grass	baboon	baboon red	baboon green	baboon blue	airfield	goldhill	boat	Barbara	Barbara red	Barbara green	Barbara blue	camera man	peppers	peppers red	peppers green	peppers blue	Lenna	pole	Text	Aviris
AGU	4.71	4.89	4.9	4.74	4.97	4.79	9.21	8.91	12.23	4.82	7.39	4.77	12.36	9.11	6.93	7.18	6.94	14.55	18.43	7.04	6.27
AGU-M	5.85	6.65	6.16	6.31	6.42	6.05	11.27	10.68	13.72	5.51	9.02	5.59	16.43	12.15	8.45	9.49	9.25	17.06	19.80	8.13	7.91
ADCT	4.92	4.94	5.01	4.82	5.15	5.06	9.52	9.49	13.33	4.89	7.89	4.99	16.33	10.05	7.20	7.72	7.25	15.34	27.69	7.15	6.47
ADCT-M	5.77	6.46	6.17	6.15	6.45	5.99	11.46	11.70	14.99	5.48	9.40	5.64	18.72	12.97	8.65	10.33	9.34	17.84	31.34	5.77	7.86
JPEG	4.58	4.96	4.71	4.60	5.01	4.90	8.22	7.05	8.46	4.00	6.19	4.33	12.32	8.71	6.05	7.76	6.97	10.73	15.78	9.10	5.99
JPEG2000	4.86	4.71	4.50	4.40	4.71	5.16	9.23	8.56	9.62	3.88	6.11	4.24	11.00	9.27	5.72	7.33	6.65	13.39	19.49	7.97	6.88
SPIHT	4.60	4.23	4.23	4.07	4.32	4.70	8.10	7.80	9.14	4.15	6.03	4.32	10.75	8.76	6.03	7.03	6.59	13.06	19.69	7.50	5.44
WebP	4.55	4.96	4.75	4.71	4.88	5.16	8.65	8.28	8.75	4.52	5.81	4.73	17.29	9.33	6.67	7.79	7.04	12.87	26.90	12.52	6.31
H265	5.00	4.90	5.06	4.67	5.17	5.80	9.85	10.67	12.62	4.97	8.23	4.93	18.17	11.33	7.58	8.26	8.60	17.64	36.07	12.49	6.72

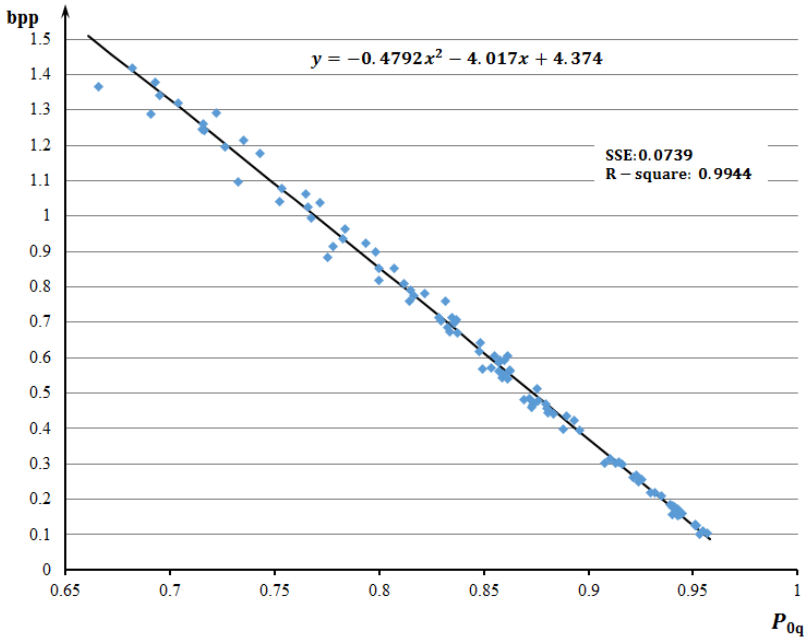
Fig. 4. CR values for different test images and used coders with the provided MSSIM=0.99

Let us come back to statement 2. Here, the situation is even more interesting. In many publications and monographs, ability of JPEG2000 and other DWT based coders to easily provide a desired CR or bpp is stressed as their main and obvious advantage (strictly saying, they provide CR not less than a desired CR_{des} or $bpp \leq bpp_{des}$ without iterations).

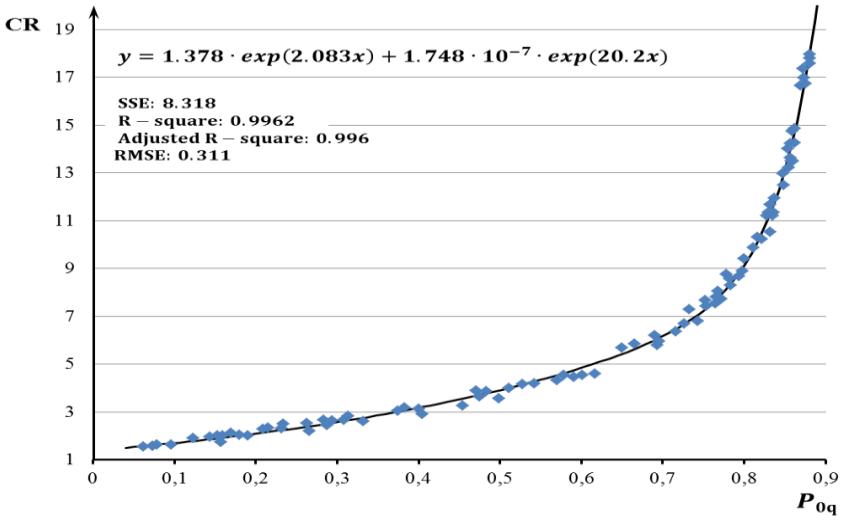
Really, this is important and valuable property for some applications. However, recently it has been shown [18] that a desired CR can be easily and quite accurately provided for JPEG. Interestingly, similar task was appealing for video and it was solved fifteen years ago [19] and practically not noticed by image processing community. The idea is mainly based on observation that there is a very strict dependence between CR and number of zeros in coded sequence. Then, the task of providing a desired CR converts to the task of providing a desired percentage P_{0q} of zeroed DCT coefficients after quantization and this is an easy algorithmic task under condition that a distribution (a set) of DCT coefficients subject to further quantization is available or type and parameters of this distribution are available or can be pre-estimated.

Dependence of CR on P_{0q} can be obtained in different ways. One of them is based on regression as shown in Fig. 5. The scatter-plot is obtained for the

coder AGU. Each point of this scatter-plot is obtained for a particular test image subject to compression with certain QS. Having a wide set of test images and QS values, a lot of points are got. Horizontal coordinate corresponds to P_{0q} determined for each particular image and vertical relates to compression parameter (bpp in Fig. 5,a and CR in Fig. 5,b).



a)



b)

Fig 5. Scatter-plots of bpp (a) and CR (b) on P_{0q} and fitted polynomials (fitting accuracy parameters are given)

Behavior of scatter-plots (compactness of points) shows that the dependences are strict and they can be got by regression (curve fitting). These dependences are obtained (for a compression technique under interest) in advance and are available at the moment of compressing a particular image. Then, for a given image, the task becomes very simple – it is necessary to find

such QS or SF that provides a desired P_{0q} after DCT coefficient quantizing. Having a set of DCT coefficients before quantization, it is easy to find a proper QS or SF. Moreover, it has been shown [20] that for advanced DCT based coders as AGU and ADCT it is not necessary to use the blocks of the same size as in these coders. It is enough to have sets of DCT coefficients for 300…500 blocks of size 8x8 pixels that are sparsely placed in an image subject to compression to carry out quite accurate prediction of bpp or CR.

The currently provided accuracy of prediction is characterized by relative error of a few percent rarely exceeding 20%. Different variants of accuracy improving are possible. Moreover, it has been shown that the described approach to bpp or CR prediction works quite well not only for compressing grayscale images but for component-wise and vector compression of multichannel images [21]. Thus, the main drawback of DCT-based compression can be avoided.

Concerning the statement 3, it is always possible to provide a desired value of a used quality metric in iterative way, by multiple compression/decompression [7, 22, 23]. As it follows from analysis of data in Figures 1 and 2, CR values for different images for the same value of a considered quality metric differ a lot. Similarly, quality of different images compressed with the same CR or bpp differs sufficiently. To show this, Fig. 6 presents the values of PSNR-HVS-M for two values of bpp (0.75, smaller values, and 1.6, larger values) for the considered test images. It is seen that for bpp=0.75 PSNR-HVS-M varies from 23 dB (very poor visual quality) to 43 dB (perfect visual quality).

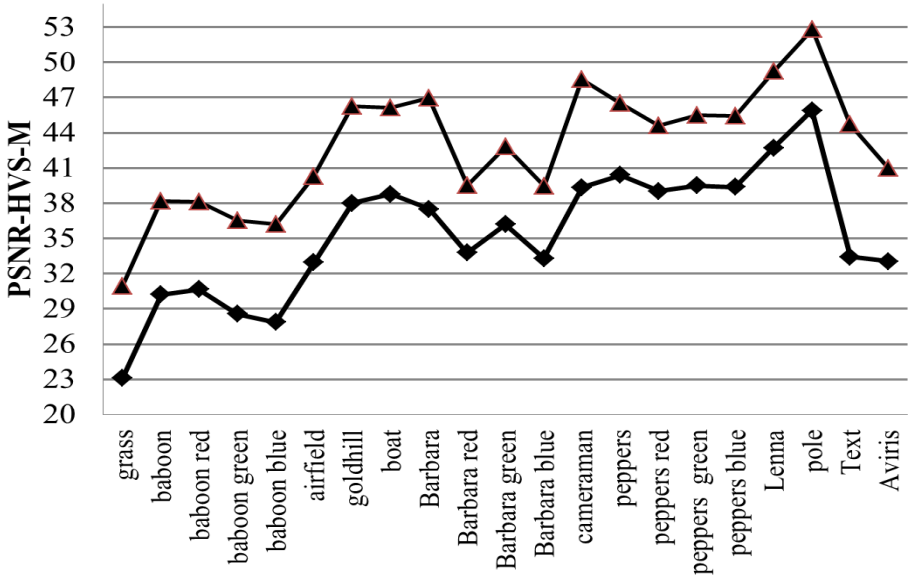
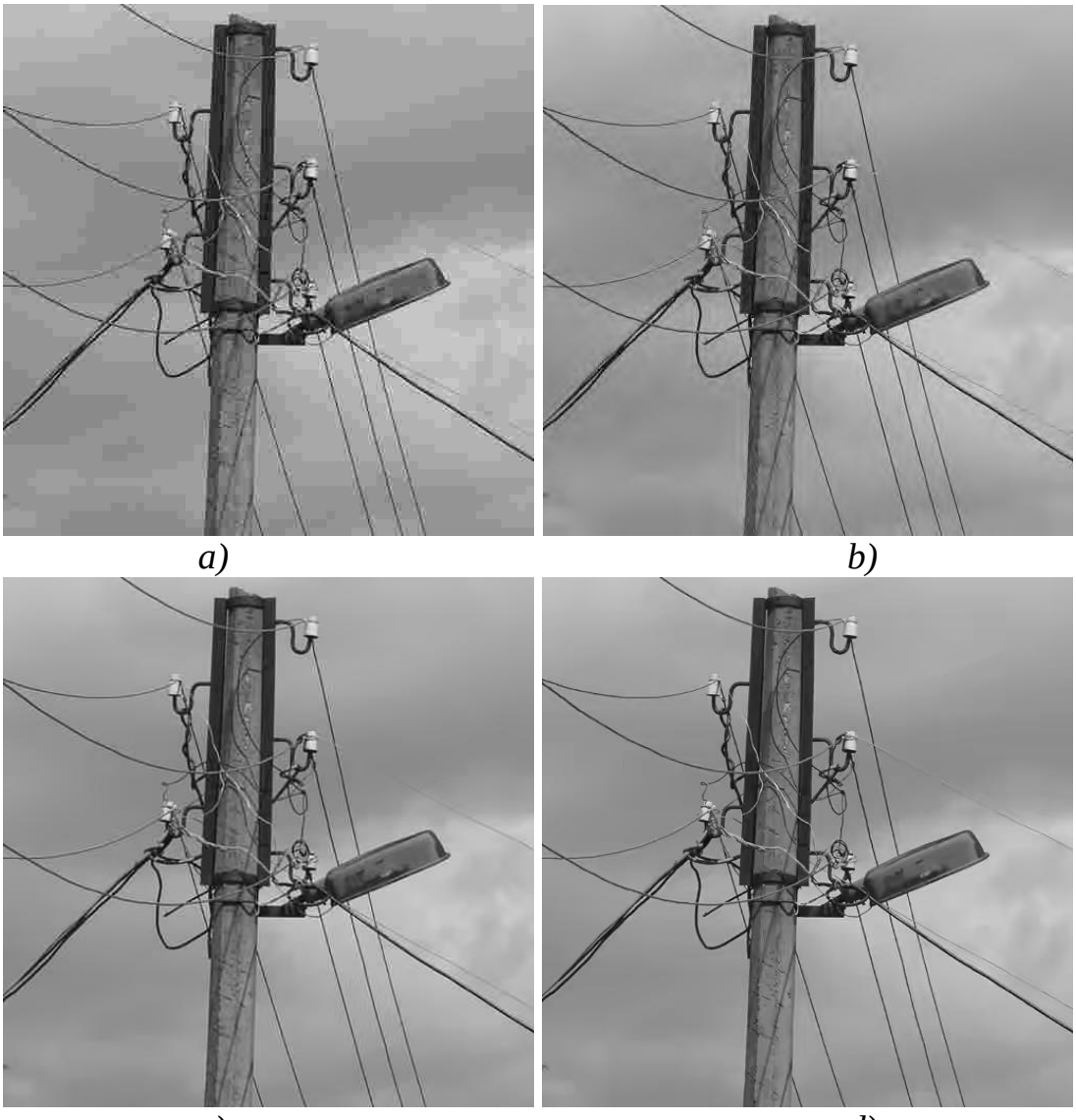


Fig. 6 Values of PSNR-HVS-M for the coder AGU-M for two fixed values of bpp: bpp=0.75 (smaller values) and bpp=1.6 (larger value)

We would like to present one more example that shows that the same CR for different coders corresponds to sufficiently different visual quality. Fig. 7

presents the test image Pole compressed with CR about 34. PSNR-HVS-M for JPEG is equal to 30.7 dB whilst for JPEG2000 this metric equals to 33.4 dB. In both cases, introduced losses are visible although they appear themselves in different manner, as blocking artifacts and ringing effects, respectively. For the coders ADCT-M and H265 the provided PSNR-HVS-M values are about 39 dB and the introduced distortions practically cannot be noticed by visual inspection.



a) *b)*

c) *d)*

Fig. 7. Results of compressing the test image Pole by JPEG(a), JPEG2000(b), ADCT-M(c) and H265(d) with provided CR about 34

There are different ways to reach compression with quality (metric value) close to a desired one with smaller number of iteration steps. They deal with adaptive choice of starting point (PCC value) and reasonable choice of other parameters of iterative procedures as step of PCC variation, accuracy of metric providing (see [7] for more details). However, number of steps remains unclear and might be too large from the viewpoint of restricted time and resources.

Meanwhile, other approaches are possible. One of them has been proposed in [24] for JPEG compression. It is strange that this work appeared in 2001 remains practically unknown by image compression community and has been cited only five times in the passed 16 years.

The approach essence is the following. The original assumption is that DCT coefficients (excluding block mean DCT coefficients) have distribution close to Laplacian with zero mean and the only parameter λ that characterizes distribution scale can be easily estimated. Then two facts are taken into account. For Laplacian distribution with known λ, it is possible to determine what QS leads to given MSE of quantization errors. Errors of DCT-coefficient quantization introduce the distortions after decompression having practically the same MSE (related to PSNR).

Thus, the approach is able to provide a desired MSE or PSNR. It also works for the case of using SF (non-equal quantization steps for different spatial frequencies). Note that PSNR is provided not absolutely accurately. Errors in PSNR can reach 0.5…0.7 dB but it is appropriate for practice. Errors in providing PSNR, to our opinion, arise due to inaccuracy of assumption on distribution of DCT coefficients. In fact, it is not Laplacian and belongs to the wider family of generalized Gaussian distributions [25].

Dependence of MSE of introduced losses on QS used as PCC for compression methods based on DCT is specific [11]. It is possible to state that if QS is of the same order as standard deviation of AC DCT coefficient distribution then MSE of introduced losses is of the order of $(QS)^2/12$. If this condition holds, this allows providing a desired PSNR easily. If QS is larger than distribution standard deviation, then MSE of introduced losses is less than $(QS)^2/12$ and it increases slower than proportionally to QS^2. Exact value of MSE depends upon several factors. For a given QS, introduced losses are smaller for simpler structure image.

It has been shown recently [26] that it is possible to predict MSE of introduced losses for JPEG without using assumption on Laplacian distribution of AC DCT coefficients. We have demonstrated that it is usually enough to have 300…500 8x8 pixel blocks chosen randomly or placed sparsely in some deterministic way for getting approximation (histogram) of AC DCT coefficient distribution. Having these data at hand, it becomes possible to estimate MSE of errors introduced by quantization in DCT coefficients and, respectively, into reconstructed image. Moreover, in this way, that is analyzing data in a limited number of blocks, it is possible to predict MSE of introduced losses and PSNR for the coders AGU and ADCT. This prediction is even more accurate and fast than the method proposed in [25]. Based on prediction mechanism, it is also possible to find QS for providing a desired PSNR. We hope that the use of modifications of the approaches in [26] will allow providing other than PSNR metrics of compressed image quality.

Concluding this discussion, we can state that providing of a desired PSNR for compression techniques in non-iterative manner is, in fact, possible. Moreover, it is possible not only for DCT-based methods but for DWT-based compression techniques as well [27].

Finally, let us consider statement 4. This is, in fact, true that images with noise are harder to compress than almost noise-free images. In this sense, we can draw attention to data for color components of the test images Barbara, Baboon, and Peppers. The test intensity (grayscale) image is compressed better (with a larger CR for the same metric, see Figures 3 and 4) than color components which are noisier.

However, one should be careful with final conclusions. The main problem is that in lossy compression of noisy images one has to apply specific criteria of compression performance. The introduced losses partly relate to information loss (distorting) and this is the negative effect whilst, at the same time, losses partly relate to noise removal and this filtering effect if obviously positive [11, 12, 28]. This filtering effect is illustrated in Fig. 8 taken from [29]. If positive effect is larger than negative, one can talk about optimal operation point (OOP).

If a compressed image according to a given metric is closer to noise–free image than original noisy image, one can talk about compression in OOP neighborhood. Parameters of compression for which closeness is maximal are associated with OOP. Thus, one can talk about QS_{OOP} or bpp_{OOP}. If OOP exists (and it is possible to predict OOP existence [30]), then it is reasonable to compress images in OOP neighborhood. Note that ways of non-iterative reaching OOP in practice have been proposed for DCT-based compression first [11]. Recently they have been also proposed for JPEG2000 and SPIHT [31].

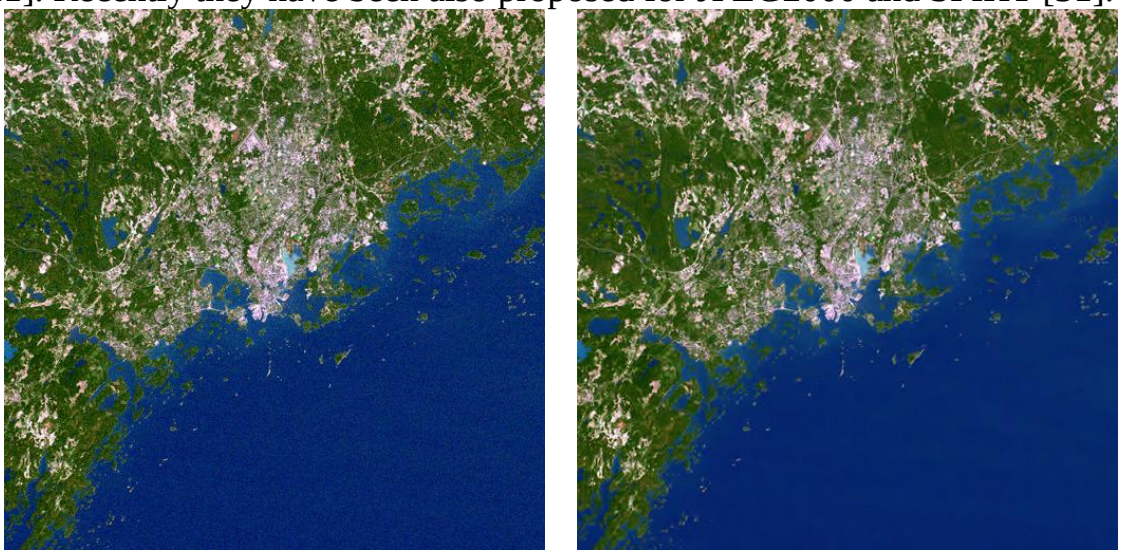


Fig. 8. Original noisy image (left) and compressed image with CR=11 (right)

Thus, traditional assertions on advantages and drawbacks of existing DCT and DWT-based methods of lossy image compression, including standards, are not always true or are valid only in certain situations.

References

1. Salomon D. Handbook of Data Compression: 5th edition / D. Salomon, G. Motta // Springer-Verlag, London. – 2009. – 1361 p. – DOI: 10.1007/978-1-84882-903-9.
2. Чобану М.К. Многомерные многоскоростные системы обработки сигналов // Техносфера. – Москва, 2009. – 480 с.
3. Wallace G.K. The JPEG still picture compression standard // Commun. ACM. – 1991. – Vol. 34. – P. 30 – 44.
4. Taubman D.S., Marcellin M.W. JPEG 2000: Image Compression Fundamentals, Standards and Practice / D.S. Taubman, M.W. Marcellin // Kluwer Academic Publishers. – 2001. – 776 p.
5. Blanes I. A Tutorial on Image Compression for Optical Space Imaging Systems / I. Blanes, E. Magli, J. Serra-Sagrista // Geoscience and Remote Sensing Magazine IEEE. – 2014. – Vol. 2(3). – P. 8 – 26.
6. Fidler A. Lossy JPEG compression: easy to compress, hard to compare / A. Fidler, B. Likar, U. Skaleric // Dentomaxillofac Radiol. – 2006. – Vol. 35. – P. 67 – 73.
7. Zemliachenko A. Still Image/Video Frame Lossy Compression Providing a Desired Visual Quality / A. Zemliachenko, N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola // Multidim. Syst. and Signal Proc. – June 2015. – 22 p.
8. Wang Z. Multi-scale structural similarity for image quality assessment / Z. Wang, E.P. Simoncelli, A.C. Bovik // Proceedings of IEEE Asilomar Conference on Signals, Systems and Computers. – 2003. – Vol. 6. – 5 p.
9. Ponomarenko N. On between-coefficient contrast masking of DCT basis functions / N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, V. Lukin // Proceedings of the Third Int. Workshop on Video Processing and Quality Metrics, USA. – 2007. – Vol. 3. – 4 p.
10. Lukin V. Lossy compression of images without visible distortions and its applications / V. Lukin, M. Zriakhov, S. Krivenko, N. Ponomarenko, Z. Miao // Proceedings of ICSP 2010, Beijing, October. – 2010. – P. 694 – 697.
11. Ponomarenko N. Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study / N. Ponomarenko, S. Krivenko, V. Lukin, K. Egiazarian // EURASIP J. on Advances in Signal Processing. – 2010. – 13 p.
12. Zemliachenko A.N. Lossy Compression of Hyperspectral Images Based on Noise Parameters Estimation and Variance Stabilizing Transform / A.N. Zemliachenko, R.A. Kozhemiakin, M.L. Uss, S.K. Abramov, N.N. Ponomarenko, V.V. Lukin, B. Vozel, K. Chehdi// SPIE J. of Applied Remote Sensing. – 2014. – Vol. 8. – 26 p.
13. Ponomarenko N.N. DCT Based High Quality Image Compression / N.N. Ponomarenko, V.V. Lukin, K. Egiazarian, J. Astola // Proceedings of 14th Scandinavian Conference on Image Analysis, Joensuu, Finland. – 2005. – P. 1177 – 1185.
14. Ponomarenko N. ADCT: A new high quality DCT based coder for lossy image compression / N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola // CD ROM Proceedings of LNLA, Switzerland. –2008. – 6 p.
15. Wang Z. Multi-scale structural similarity for image quality assessment / Z. Wang, E.P. Simoncelli, A.C. Bovik // IEEE Asilomar Conference on Signals, Systems and Computers. – 2003. - Vol. 6. – 5 p.
16. Ponomarenko N. Additional lossless compression of JPEG images / N. Ponomarenko, K. Egiazarian, V. Lukin, J. Astola // CD-ROM Proceedings of the 4th Symposium on Image and Signal Processing and Analysis, Zagreb, Croatia. – 2005. – P. 117 – 120.
17. Matsuda I. A lossless re-encoding scheme for MPEG-1 video / I. Matsuda, K. Wakabayashi, Y. Ikeda, S. Itoh // Proceedings of EUSIPCO. – 2009.

18. Осокин А.Н. Модифицированный кодер стандарта JPEG с контролем битрейта / А.Н. Осокин, Д.В. Сидоров // Интернет-журнал "Науковедение. – 2013. – № 5. – 9 с. http://naukovedenie.ru/PDF/69tvn513.pdf

19. He Z. A linear source model and a unified rate control algorithm for DCT video coding / Z. He, S.K. Mitra // IEEE Transactions, Circuits and Systems for Video Technology. – 2002. – Vol. 12(11). – P. 970 – 982.

20. Kozhemiakin R.A. An approach to prediction and providing of compression ratio for DCT-based coder applied to remote sensing images / R.A. Kozhemiakin, A.N. Zemliachenko, V.V. Lukin, S.K. Abramov, B. Vozel // Український журнал дистанційного зондування Землі. – 2016. – № 8. – С. 22 – 29.

21. Kozhemiakin R. An approach to prediction and providing of compression ratio for DCT-based coder applied to multichannel remote sensing data / R. Kozhemiakin, V. Lukin, S. Abramov, M. Simeunovic, B. Djurovic, I. Djurovic, // Telecommunications and Radio Engineering. – 2016. – Vol. 75(14). – P. 1255 – 1269.

22. Земляченко А.Н. Сжатие изображений без визуально заметных искажений / А.Н. Земляченко, В.В. Лукин // Радиоэлектронные и компьютерные системы. – 2011. – № 3. – С. 73 – 79.

23. Земляченко А.Н. Ускорение сжатия изображений с требуемым визуальным качеством / А.Н. Земляченко, О.Е. Колганова, В.В. Лукин // Радиоэлектронные и компьютерные системы. – Харьков: ХАИ, 2011. – №4(52). – С. 52 – 59.

24. Minguillon J. JPEG Standard Uniform Quantization Error Modeling with Applications to Sequential and Progressive Operation Modes / J. Minguillon, J. Pujol // Electron. Imaging. – 2001. – Vol. 10(2). – P. 475 – 485.

25. Kurkin D. Image DCT-coefficient statistics and their use in blind noise variance estimation / D. Kurkin, V. Lukin, V. Abramova, S. Abramov, B. Vozel, K. Chehdi // Proceedings of MMET 2012, Kharkov, Ukraine. – 2012. – P. 316 – 319.

26. Kozhemiakin R. Image Quality Prediction for DCT-based Compression / R. Kozhemiakin, V. Lukin, B. Vozel // Proceedings of CADSM 2017, Ukraine. – 2017. – P. 265 – 268.

27. Jiang H. Quality Prediction of DWT-Based Compression for Remote Sensing Image Using Multiscale and Multilevel Differences Assessment Metric / H. Jiang, K. Yang, T. Liu, Y. Zhang // Mathematical Problems in Engineering. – 2014. – 15 p.

28. Al-Chaykh O.K. Lossy compression of noisy images / O.K. Al-Chaykh, R.M. Mersereau // IEEE Transactions on Image Processing. – Vol. 7(12). – P. 1641 – 1652.

29. Lukin V. Challenges in Pre-processing Multichannel Remote Sensing Terrain Images. Importance of GEO initiatives and Montenegrin capacities in this area / V. Lukin, E. Bataeva // The Montenegrin Academy of sciences and arts Book. – No 119. – 2012. – P. 63 – 76.

30. Zemliachenko A. Lossy Compression of Noisy Remote Sensing Images with Prediction of Optimal Operation Point Existence and Parameters / A. Zemliachenko, S. Abramov, V. Lukin, B. Vozel, K. Chehdi // SPIE Journal on Advances in Remote Sensing. – 2015. – Vol. 9(1). – 26 p. – Doi: 10.1117/1.JRS.9.095066.

31. Zemliachenko N. Automatic Lossy Compression of Noisy Images by SPIHT or JPEG2000 in Optimal Operation Point Neighborhood / N .Zemliachenko, S.K. Abramov, V.V. Lukin, B. Vozel, K. Chehdi // Proceedings of EUVIP, Marseille, France. –2016. – 6 p.

БАГАТОФУНКЦІОНАЛЬНА ЛАЗЕРНА ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНА СИСТЕМА КОНТРОЛЮ І УПРАВЛІННЯ ЛІТАЛЬНИМ АПАРАТОМ

Коломійцев О.В.

Вступ

На сьогодення високу точність вимірювання параметрів руху (ВПР) літальних апаратів (ЛА) та невеликі ваго-габаритні характеристики забезпечують лазерні системи (ЛС), лазерні інформаційно-вимірювальні системи (ЛІВС) і суміщенні (багатофункціональні або комбіновані) системи, на яких акцентується увага усіх технологічно розвинених країн світу. Для того, щоб забезпечити високу якість інформації, яка отримується від таких систем, до них пред'являються жорсткі вимоги по забезпеченню високої точності вимірювання параметрів сигналів, що нерідко характеризується погрішностями у долі відсотків.

Висока точність ВПР ЛА ЛС (ЛІВС) обумовлена, у першу чергу, використанням у якості джерела випромінювання – лазера, який має велику несучу частоту і спектральну яскравість, монохроматичність, просторову і часову когерентність. Завдяки цьому у ЛС формуються понадвузькі діаграми спрямованості (ДС) лазерного випромінювання (ЛВ) і виходять великі коефіцієнти підсилення при порівняно малих оптичних антенах. Саме лазери генерують велетенські за потужністю і ультракороткі за тривалістю імпульси, що забезпечує високу точність ВПР ЛА і якісну передачу інформації.

Основна частина

Використання лазерних сигналів для передачі інформації дозволяє: передавати більші, ніж у радіочастотному діапазоні, об'єми інформації у одиницю часу; здолати обмеження радіочастотного спектру, виділеного для радіозв'язку; підвищити спрямованість випромінювання, що забезпечує високу скритність і завадозахищеність, завдяки малої розузгодженності ЛВ.

В процесі обробки відбитих від ЛА лазерних сигналів у ЛС, завдяки використанню отримуваних з поляризаційних матриць розсіяння (ПМР) поляризаційних ознак, можливо здійснювати розпізнавання ЛА, а за умови формування зондуючого лазерного сигналу з складною просторово-часовою структурою – формувати і обробляти його зображення, а також визначати матеріали, з яких складається ЛА.

Проте існуючі ЛС мають і ряд істотних недоліків. Вони вимірюють обмежену кількість параметрів руху ЛА (похилу дальність та кутові координати), що більшою мірою пов'язане з відсутністю методів вимірювання інших параметрів, а також здійснюють автоматичне супроводження (АС) ЛА за програмою, через відсутність методів боротьби

з флуктуаційними і динамічними помилками. Детальна інформація про розширення кількості і підвищення точності ВПР ЛА, а також – стійкості кутового АС у широкому діапазоні дальностей у відомій літературі відсутня.

Однак, концентрація енергії у ЛВ і стабільність частот, що випромінюються, а також урахування особливостей спектру ЛВ і можливостей частотно-часового методу вимірювання [1] дозволяють створити сучасну багатофункціональну ЛС, яка за своїми тактико-технічними характеристиками (ТТХ) є високоточною, стосовно ВПР ЛА, більш компактна та розвинена за своїми функціональними можливостями (ЛІВС), ніж системи радіодіапазону.

Створення багатофункціональної ЛІВС пов’язано з вирішенням наступних науково-технічних задач: електромагнітної сумісності багатьох вимірювальних (за параметрами руху ЛА) та інформаційного каналів; підвищення точності вимірювання похилої дальності (відстані) до ЛА; підвищення точності вимірювання швидкості (радіальної) ЛА; підвищення точності вимірювання кутів азимута і місця ЛА; підвищення точності вимірювання кутових швидкостей ЛА; підвищення стійкості і точності АС ЛА при сумісній обробці параметрів його руху, а також їх похідних за фільтрацією; підвищення ефективності передачі інформаційних каналів тощо.

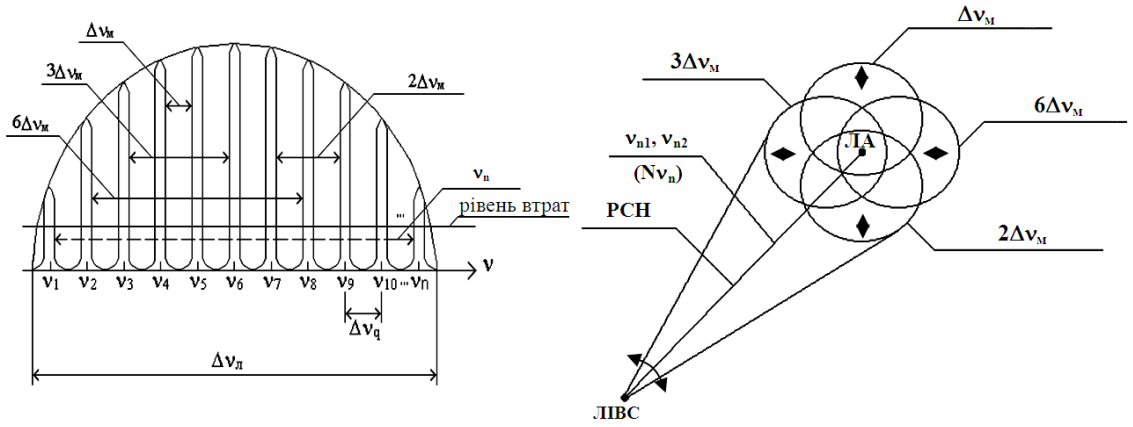
При подоланні цих проблем багатофункціональна ЛІВС отримує ряд переваг перед існуючими системами радіо і оптичного діапазонів довжин хвиль: висока точність вимірювання шести параметрів руху (траєкторні вимірювання) ЛА (також за рахунок використання їх взаємозв’язку); висока стійкість кутового АС ЛА за напрямком; висока швидкість передачі інформації на ЛА; об’єктивний контроль ЛА у денних і нічних умовах та, за необхідністю, його розпізнавання; обробка, збереження, відображення і передача інформації, що отримується під час проведення випробувань ЛА; висока економічна ефективність; висока точність геодезичної прив’язки; висока точність прив’язки до системи єдиного часу; високі надійність, мобільність тощо.

За допомогою частотної селекції каналів багатофункціональної ЛІВС, вирішуються задачі електромагнітної сумісності та кількості ВПР ЛА. В спектрі одномодового багаточастотного з синхронізацією подовжніх мод ЛВ є набір подовжніх мод (частот), які еквідістантно рознесені та достатні для селекції.

Тому, зі спектра ЛВ, з виходу лазера-передавача багатофункціональної ЛІВС, за допомогою модифікованого селектора подовжніх мод (МСПМ) [2] виділяються необхідні подовжні моди, різниця частот яких відповідає кожному з вимірювальних каналів (рис. 1а). Лазерні сигнали N інформаційних каналів для зв’язку з ЛА використовують окремі несучі подовжні моди (частоти).

Для лазера з $\lambda_0 = 532$ нм, або $\nu_0 = 6 \cdot 10^{14}$ Гц, при довжині лазера l=0,3 м, міжмодова частота складатиме $\Delta\nu = 10^9$ Гц. Смуга пропускання фільтру нижніх частот (ФНЧ) каналу АС ЛА за напрямком (АСН) не більш 10 кГц. Це означає, що при такій вузькій смузі пропускання частотна селекція достатньо висока і що система може мати високу завадостійкість, тим паче, що усі частоти міжмодових биттів мають високу стабільність.

Використання в багатофункціональній ЛІВС одномодового багаточастотного з синхронізацією подовжніх мод випромінювання єдиного лазера-передавача, ЧЧМ вимірювання, методів виділення подовжніх мод, формування лазерних сигналів, що зондують тощо дозволяє побудувати канали кутового АСН з вимірюванням шести параметрів руху ЛА, передачі команд управління на ЛА та, за необхідністю, його розпізнавання.



а) *б)*

Рис.1. а) спектр одномодового багаточастотного з синхронізацією подовжніх мод ЛВ; б) Створення рівносигнального напрямку (РСН) та сканування сумарною ДС ЛВ у невеликому куті і окремо 4-мя ДС ЛВ в ортогональних площинах

Завдяки зустрічному скануванню пар парціальних ДС ЛВ в кожній з двох ортогональних площин (рис. 2) можливо здійснити вимірювання: кутів азимута і місця ЛА, кутових швидкостей, а також радіальної швидкості – з використанням ефекту Допплера та дальності – за затримкою сигналу.

Використання в структурі багатофункціональної ЛІВС оптичного електронного модуля (ОЕМ), який складений з телевізійного і інфрачервоного каналів дозволяє здійснювати об'єктивний контроль ЛА у денний і нічний час.

Розробка структурно-сигнальної моделі багатофункціональної ЛІВС – це етап ухвалення евристичного рішення про структуру і сигнали по заданих тактико-технічних вимогах (ТТВ) та витратних показниках, користуючись досвідом побудови таких систем, науковими, технічними і

технологічними ідеями їх реалізації [3]. По суті, ця інтуїтивна структурно-сигнальна фізична побудова системи з деталізацією на рівні функціональних елементів (ФЕ) [4, 5].

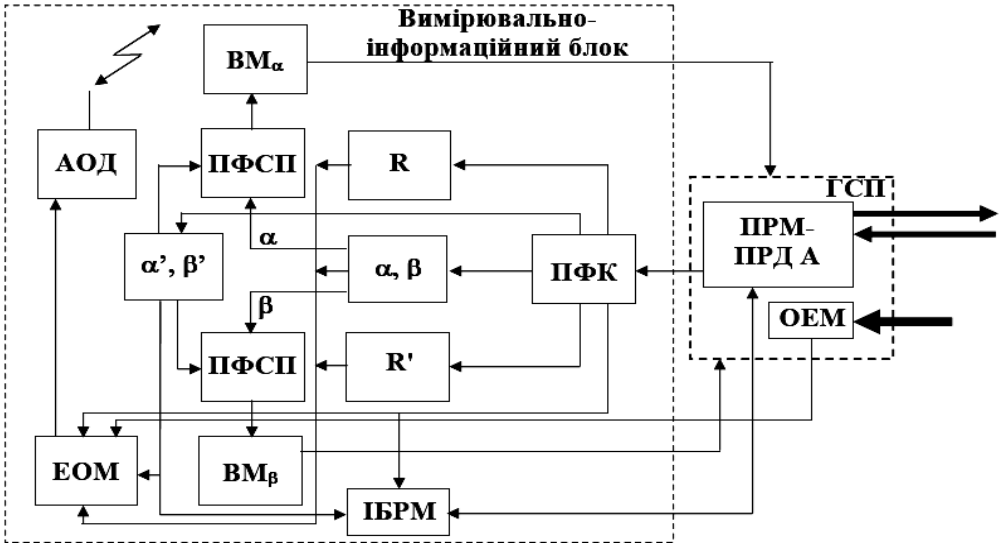


Рис. 2. Узагальнена функціональна схема багатофункціональної ЛІВС

Далі, після структурно-сигнальної побудови системи на етапі ескізного проектування, повинен слідувати етап оптимальної параметричної побудови системи, потім етап розширення діапазонів показників якості, етап отримання "кривих обміну" і етап порівняння "кривих обміну" для різних оптимальних систем, який в результаті можна називати етапом об'єктивної оптимальної побудови системи за ТТВ на множені сигналів, структур і технічних параметрів. Такий інженерний підхід до побудови структури і формування сигналів містить пропозиції використовувати для багатофункціональної ЛІВС МСПМ, розробку основної структури і сигналів вимірювальних і інформаційного каналів, які використовують нові та вдосконалені методи вимірювань для отримання усієї інформації про параметри руху ЛА, вдосконалену фільтрацію за методом Калмана-Бьюсі для підвищення стійкості АСН за кутами і аналогічно по дальності.

Основна ідея для підвищення ефективності багатофункціональної ЛІВС – це селекція подовжніх мод для створення окремих та парних оптичних сигналів, які підібрані для своїх каналів таким чином, щоб у фотоприймачі можна було селектирувати їх міжмодові биття, які фактично, на відміну від відомих систем, перетворюють сукупний оптичний сигнал у набір частот канальних міжмодових биттів (радіочастоти). Селекція каналів здійснюється канальними фільтрами, налаштованими на свої міжмодові биття.

Пари частот подовжніх мод можна вибрати таким чином, щоб кожна використовувалася тільки один раз і пари утворювали різні міжмодові биття. Це необхідно для того, щоб неможливо було отримати ті ж

міжмодові биття з іншого набору подовжніх мод.

На рис. 2 представлена узагальнена функціональна схема мобільної багатофункціональної ЛІВС, а на рис. 3 функціональна схема приймально-передавальної частки системи, де: ПРМ-ПРД А – приймально-передавальна апаратура; ОЕМ – оптико-електронний модуль; ГСП – гіростабілізірована платформа; ПФК – пристрій формування каналів; R – канал вимірювання похилої дальності до ЛА; R’ – канал вимірювання радіальної швидкості ЛА; α і β – канал вимірювання кутів азимута і місця ЛА; α’ і β’ – канал вимірювання кутових (тангенціальних) швидкостей ЛА; ПФСП – пристрій формування сигналів помилки по кутах азимута α і місця β; ВМ – виконавчі механізми по кутах азимута α і місця β; ЕОМ – електронно-обчислювальна машина; ІБРМ – інформаційний блок з розширеними можливостями та АОД – апаратура обміну даними. На ГСП розміщені приймально-передавальна апаратура (ПРМ-ПРД А) та ОЕМ.

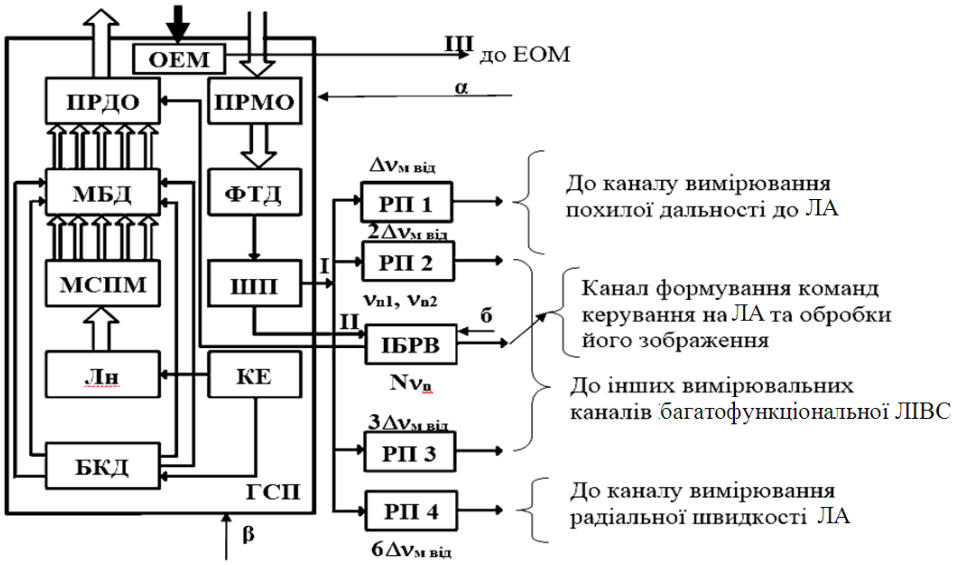


Рис. 3. Функціональна схема ПРМ-ПРД частки багатофункціональної ЛІВС

ПРМ-ПРД А складається з наступних частин (рис. 3):

– передавальної, яка включає: КЕ – керуючий елемент; лазер з накачкою (Лн), БКД – блок керування дефлекторами; МСПМ – модифікований селектор подовжніх мод; МБД – модифікований блок дефлекторів; ПРДО – передавальну оптику та ГСП – гіростабілізированну платформу;

– приймальної, яка включає: ПРМО – приймальну оптику; ФТД – фотодетектор; ШП – широкосмуговий підсилювач; ІБРМ – інформаційний блок з розширеними можливостями з б – введенням сигналу від каналу вимірювання кутових (тангенціальних) швидкостей ЛА для уточнення похибки оцінки кутів і детального його розпізнавання; РП – резонансних

підсилювачів, налаштованих на комбінації частот міжмодових биттів, що приймаються; I – вимірювальний сигнал; II – інформаційний сигнал і сигнал з просторово-часовою модуляцією поляризації та III – інформація про об'єктивний контроль ЛА.

Необхідність в такій побудові ПРМ-ПРД А системи, передусім, пов'язана з використанням спектра одномодового багаточастотного з синхронізацією подовжніх мод ЛВ, ОЕМ і АОД, а також суттю ЧЧМ вимірювання.

Режим активної синхронізації мод лазера-передавача істотно змінює спектральні характеристики випромінювання, знижує рівень флуктуацій амплітуд і частот генерованих коливань, пригнічує технічні флуктуації частот міжмодового биття, пов'язані з порушенням еквідістантності поширення мод на частотній осі. Ця обставина обумовлює ефективність використання такого режиму для побудови передавальної частини системи.

Принцип роботи багатофункціональної ЛІВС, полягає у наступному.

У ПРМ-ПРД А зі спектру одномодового багаточастотного з синхронізацією подовжніх мод ЛВ лазера-передавача (Лн) за допомогою МСПМ, виділяються необхідні моди та їх комбінації для створення:

– РСН на основі формування сумарної ДС ЛВ, завдяки 4-х парціальних ДС ЛВ, що частково перетинаються, за умови використання комбінацій подовжніх мод, "підфарбованих" різницевими частотами міжмодових биттів

$$\Delta\nu_{54}=\nu_5-\nu_4=\Delta\nu_м,\ \Delta\nu_{97}=\nu_9-\nu_7=2\Delta\nu_м,\ \Delta\nu_{63}=\nu_6-\nu_3=3\Delta\nu_м,\ \Delta\nu_{82}=\nu_8-\nu_2=6\Delta\nu_м;$$

– багатоканальної (N) передачі інформації (команд керування ЛА), за умови використання сигналу подовжніх мод (на несучих частотах ν_n);

– лазерного сигналу з просторовою модуляцією поляризації, за умови використання сигналу з двох подовжніх мод (несучих частот ν_{n1}, ν_{n2}).

Зустрічне сканування пар парціальних ДС ЛВ у кожній з двох ортогональних площин (напівперекритті ДС ЛВ) створює РСН, який проходить через ЛА. В разі необхідності виявлення ЛА у заданої точці простору, складений з частот міжмодових биттів груповий сигнал сканується у вигляді сумарної ДС ЛВ за допомогою МБД ПРМ-ПРД А, де кут та напрямок відхилення сумарної ДС ЛВ задається БКД.

Передача команд керування на ЛА здійснюється на несучих частотах (ν_n), що виділяються МСПМ, а необхідна інформація формується за допомогою ІБРМ та ЕОМ. Кількість інформаційних каналів (N) залежить від кількості мод (ν_n), які мають необхідні вихідні характеристики для використання.

За необхідністю [3], за допомогою МСПМ та ІБРМ створюється

лазерний сигнал з просторовою модуляцією поляризації шляхом створення лазерного випромінювання з двох несучих частот (ν_{n1} і ν_{n2}) у вигляді двох променів з вертикальною (ν_{n1}) та горизонтальною (ν_{n2}) поляризаціями (рис. 4).

При цьому випромінювання апертури першого і другого поляризаційних каналів в апертурної плоскості V0U рознесені на відомій відстані $\Delta\nu_q$. Різність ходу пучків до картинної плоскості ЛА X0У змінюється вдовж осі X від точки до точки. Обумовлена цім різність фаз (амплітуд) між поляризованими компонентами, що ортогональні, поля у картинної плоскості, також змінюється від точки до точки.

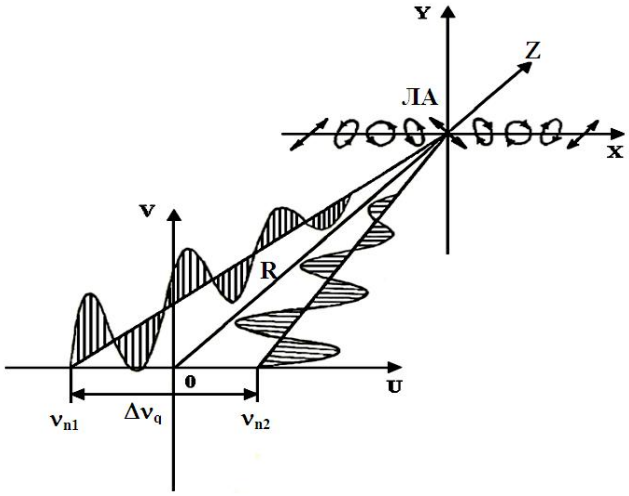


Рис. 4. Формування лазерного сигналу з просторовою модуляцією поляризації

В залежності від різності фаз (амплітуд) у картинній плоскості змінюється вигляд поляризації сумарного поля сигналу, що зондує від лінійної через еліптичну і циркулюючу до лінійної, ортогональної к начальної і т.д. Період зміни вигляду поляризації визначається базою між випромінювачами $\Delta\nu_q$ та відстанню до картинної плоскості R. Розподіл інтенсивності в реєстрованому зображенні ЛА промодульовано по гармонійному закону з коефіцієнтом модуляції та дорівнює значенню ступеня поляризації випромінювання, що відбито в даній ділянці поверхні ЛА.

Лазерні сигнали з просторовою модуляцією поляризації (ν_{n1} і ν_{n2}) від ІБРМ через ПРМ-ПРД А багатофункціональної ЛІВС разом з командами керування проходять по вздовж РСН. Відбиті від поверхні ЛА сигнали ν_{n1} і ν_{n2}, у зворотному напряму приймаються ПРМ-ПРД А, чим забезпечують його детальне розпізнавання.

При відбитті лазерного сигналу з просторовою модуляцією поляризації, що зондує, від поверхні ЛА змінюються амплітудні і фазові співвідношення між ортогонально поляризаційними компонентами,

параметри їх поляризаційні і, відповідно, комплексні коефіцієнти когерентності відбитого поля. Просторовий розподіл поляризаційних характеристик такого відбитого сигналу по зміні контрасту модуляційної структури зображення несе також інформацію про типи матеріалів у складі поверхні ЛА, їх характеристики і тощо, що приймається ПРМ-ПРД А і відображається у ЕОМ. Тому у ІБРМ також здійснюється поляризаційна обробка поля, що приймається.

Відбиті сигнали від ЛА на частотах міжмодових биттів обробляються ПРМ-ПРД А та розподіляються ПФК по вимірювальних каналах багатофункціональної ЛІВС.

В каналі вимірювання кутів азимута і місця ЛА за тривалістю зрушень періодів огинаючих пачок імпульсів частот міжмодових биттів за один повний прохід ДС ЛВ у прямому і зворотному напрямку сканування в кожній з двох ортогональних площин часо-імпульсним методом з високою точністю вимірюються кути азимута і місця та формуються сигнали похибок по двох вісях координат (α і β), які поступають на ПФСП (α і β) [3].

Середньоквадратична похибка (СКП) вимірювання кута відхилення ЛА від РСН визначається за формулою:

$$\sigma_\theta^2 = \frac{2e\theta_m^2}{q}, \qquad (1)$$

де θ_m – ширина ДС ЛВ; q – відношення сигнал/шум.

За попередніми розрахунками, на відстані R=150 км ЛА від системи погрішність визначення кутів азимута і місця ЛА складатиме $\delta_{\alpha,\beta}$<0,2", а СКП кута відхилення БР від РСН $\sigma_\theta \approx 10^{-5}$ (рад).

В каналі вимірювання кутових швидкостей ЛА за тривалістю півперіодів огинаючих пачок імпульсів частот міжмодових биттів за один прохід ДС ЛВ у одному (прямому або зворотному) напрямку сканування в кожній з двох ортогональних площин часо-імпульсним методом з високою точністю вимірюються кутові швидкості [3].

СКП вимірювання кутової швидкості ЛА визначається за формулою:

$$\sigma_{\dot\theta}^2 = \frac{e}{2}\cdot\frac{\omega_{ск}^2}{q}, \qquad (2)$$

де $\omega_{ск}$ – кутова швидкість сканування ДС ЛВ.

За попередніми розрахунками, на відстані R=150 км ЛА від системи погрішність визначення кутових швидкостей ЛА складатиме $\delta_{\alpha',\beta'}$<0,2", а СКП вимірювання кутової швидкості ЛА $\sigma_{\dot\theta} \approx 4\cdot 10^{-6}$ $\left(\frac{\text{рад}}{\text{с}}\right)$.

Вимірювальна інформація про кутові швидкості ЛА від каналу кутових швидкостей використовується у ІБРМ, де завдяки додаткової

обробці елементів поляризаційної матриці розсіяння від отриманого поляризаційного поля (суми сигналів різної поляризації) забезпечується точне значення кутових швидкостей ЛА, розширюється набір ознак його розпізнавання, підвищується ефективність та скорочується час на розпізнавання ЛА, що супроводжується [4].

В каналі вимірювання похилої дальності до ЛА, за умови використання запізнювання частот міжмодових биттів, еталонних частот, пікосекундних і "бланкуючих" імпульсів та формуванню багатошкального методу з високою точністю вимірюється похила дальність [3].

СКП вимірювання похилої дальності до ЛА визначається за формулою:

$$\sigma_R^2 = C^2 \cdot \frac{T_{\Delta\nu_м}}{q}, \qquad (3)$$

де С – швидкість світла; $T_{\Delta\nu_м}$ – час (період) випромінювання на частоті міжмодових биттів.

За попередніми розрахунками, на відстані R=150 км ЛА від системи погрішність визначення похилої дальності до ЛА складатиме δ_R<0,4 м, а СКП вимірювання похилої дальності до ЛА $\sigma_R^2 \approx 0{,}014\,(м^2)$, $\sigma_R \approx 0{,}03\,(м)$.

В каналі вимірювання радіальної швидкості ЛА, за умови використання запізнювання частот міжмодових биттів, еталонних частот, фазової автопідстройки частоти (частоти підставки) та ефекту Допплера з високою точністю вимірюється радіальна швидкість [3].

СКП вимірювання радіальної швидкості ЛА визначається за формулою:

$$\sigma_{\dot{R}}^2 = \frac{\sigma_{\Delta\nu_д}^2}{\Delta\nu_м} \cdot C, \qquad (4)$$

де $\sigma_{\Delta\nu_д}^2$ – СКП визначення частоти Допплера; $\Delta\nu_м$ – частота міжмодових биттів.

За попередніми розрахунками, на відстані R=150 км ЛА від системи погрішність визначення радіальної швидкості ЛА складатиме $\delta_{R'}$<0,1 м/с, а СКП вимірювання радіальної швидкості ЛА $\sigma_{\dot{R}} \approx 0{,}05\,(рад)$.

В ПФСП по кутах α і β, формуються сигнали похибки по кутових координатах, що корегуються прогнозованими динамічними похибками від каналу вимірювання кутових швидкостей ЛА. Отримані сигнали, через ВМ по кутах α і β, розвертають ПРМ-ПРД А та ОЕМ таким чином, щоб РСН постійно проходив через ЛА.

ГСП забезпечує дотримання просторової стабілізації платформи системи, на якої розміщена суміщена ПРМ-ПРД А та ВМ по кутах α і β.

ОЕМ постійно здійснює у денних і нічних умовах у видимому та інфрачервоному діапазонах спостереження за ЛА, яка супроводжується.

Об’єктивний контроль та інформація про зовнішньотраєкторні вимірювання ЛА (похилу дальність, радіальну швидкість, кути азимута і місця, кутові швидкості α', β') обробляється, відображається та запам’ятовується у ЕОМ.

Збереження інформації, яка оброблена під час проведення випробувань ЛА, здійснюється в пам’яті ЕОМ. Для цього використовується база даних – сукупність взаємопов’язаних даних, організованих у відповідності до схеми даних таким чином, щоб з ними міг працювати користувач.

Підвищення швидкості обробки інформації, яка поступає на ЕОМ здійснюється за рахунок використання методів та моделей паралельної часупараметризованої обробки даних.

Видача інформації, яка отримана під час проведення випробувань ЛА, споживачам та отримання додаткової інформації від них здійснюється за допомогою апаратури обміну даними за радіоканалом.

Інформація про кутові швидкості ЛА додатково використовується у каналі АСН для компенсації динамічної помилки автоматичного супроводження ЛА. Ця обставина пов'язана з тим, що при точному прогнозі кутової швидкості ЛА істотною залишиться лише флуктуаційна помилка, яку можна компенсувати використовуючи фільтрацію за методом Калмана-Бьюсі для центрованого процесу [6].

Якість фільтрації можливо підвищити завдяки формуванню зваженої оцінки кутової швидкості ЛА. Отриманий вдосконалений фільтр Калмана-Бьюсі [6], дозволяє враховувати вимірювальну інформацію про кутові швидкості ЛА, а так само усувати динамічну і флуктуаційну помилки в ФНЧ.

Таким чином, звужуючи смугу пропускання в каналі АСН, можна підвищити точність і стійкість кутового автоматичного супроводження ЛА.

Як відомо, фільтр буде стійкий, якщо перехідні в ньому процеси будуть затухаючими, тобто усі речові частини коренів (дійсних і комплексних) характеристичного рівняння системи будуть негативні. Для дослідження стійкості фільтру застосовувався критерій за Раусом-Гурвіця [6].

Побудова багатофункціональної ЛІВС також вимагає облік особливостей поширення і взаємодії світлових хвиль з атмосферними неоднородностями (хмарними, аерозольними, пиловими, турбулентними тощо) і перешкод, що виникають при цьому, а також утримання вузьких ДС ЛВ (РСН на ЛА).

Істотний вплив атмосфери на автоматичне супроводження та ВПР ЛА багатофункціональної ЛІВС призводить до деяких особливостей побудови її каналів. Тому вони повинні враховувати наступне:

– складність підсистем попарного управління 4-мя ДС ЛВ, що

сканують одна назустріч другої у кожній з двох ортогональних площин ПРМ ПРД А, до яких зазвичай відносять підсистеми стеження і узгодження оптичних осей. Ці підсистеми перевершують підсистеми функціональних каналів за числом оптичних елементів і електронних схем, складності їх компонування, вартості;

– включення до складу апаратури системи автоматичних метеорологічних приладів та ін. пристроїв для оцінки "оптичної погоди" і вимірювання оптичних характеристик атмосфери з метою зменшення впливу турбулентності атмосфери на точність ВПР БР;

– використання іншого багаторівневого принципу автоматизації процесу функціонування системи. Застосування ЕОМ, погодженою з характеристиками функціональних каналів багатофункціональної ЛІВС і підсистемою стеження за БР (наведення РСН на ЛА);

– використання просторово-часових фільтрів (матричних приймачів, активної оптики, диссекторів тощо), волоконно-оптичної і оптоелектронної техніки, інтегральних оптичних схем та ін.

Використання методів адаптивної компенсації фазових флуктуацій за допомогою пристроїв активної (адаптивної) оптики дозволить істотно зменшити вплив турбулентної атмосфери на точність ВПР ЛА. За умови поширення ЛВ в атмосфері, коли хмарні і аерозольні поля відсутні, величина послаблення оптичного сигналу з довжинами хвиль (λ=0,5 – 10,6 мкм), що потрапляють у "вікна прозорості атмосфери", істотно менше.

ФТД на λ=10,6 мкм мають високу квантову ефективність (η=30 – 50 %), проте вимагають охолодження до температури 77 – 100 К для отримання хороших шумових характеристик і не мають внутрішнього підсилення.

Фотоелектронні підсилювачі (ФЕП) і лавинні фотодіоди (ЛФД) на λ=1,06 мкм мають велике підсилення за струмом. Проте мають малий квантовий вихід (η=0,8 – 12,5 %).

Таким чином, найбільш прийнятою, в порівнянні з існуючими, є генерація лазера на довжині хвилі 0,53 мкм (або 1,54 мкм). В цьому випадку ФЕП мають досить високий квантовий вихід ($\eta\approx$25 %), а коефіцієнт корисної дії (ККД) твердотілого лазера, працюючого у режимі подвоєння частоти (λ=0,53 мкм) складає одиницю та більш. Крім того, збільшення квантового шуму у приймачі багатофункціональної ЛІВС видимого діапазону (λ=0,53 мкм) та нижчий ККД передавача у порівнянні з системою ІЧ діапазону (λ=10,6 мкм) можуть компенсуватися звуженням ДС ЛВ.

При однакових розмірах передавальних апертур коефіцієнт підсилення передавальної антени на λ=0,53 мкм буде в 400 разів більше, ніж λ=10,6 мкм.

Вужчі ДС ЛВ передавальної антени у багатофункціональній ЛІВС на

твердотілому лазері вимагають високої точності утримання та стеження (близько 1 мкм/рад), що істотно ускладнює канал АСН (утримання передавальних і приймальних антен) [6 – 8]. Використання фільтрації кутомірного сигналу ефективно вирішує задачу стійкого (без зриву) стеження за польотом ЛА, при одночасному вимірюванні його параметрів руху тощо.

Рівень сигналу (число фотоелектронів) на виході ФТД багатофункціональної ЛІВС визначається за формулою:

$$S_i = \frac{W_i \tau_i e^{-\gamma R} d^2 \prod_i T_i \prod_i T_{i\,T(гр)} \eta_i}{\Theta_i^2 R^2 \hbar \nu_i}, \qquad (5)$$

де W_i – пікова потужність випромінювання оптичного джерела; τ_i – тривалість імпульсу; e – заряд електрона; γ – повний коефіцієнт послаблення сигналу на трасі; R_a – максимально допустима відстань між ПРД і ПРМ антенами системи (обмежена необхідним відношенням сигнал/шум); d – діаметр ПРМ антени; $\prod_i T_i$ – загальний коефіцієнт передачі поєднаної оптики ПРД і ПРМ; $\prod_i T_{iT(гр)}$ – загальний коефіцієнт передачі оптики приймача грубого (точного) націлювання; η_i – квантовий вихід ФТД; Θ – ширина ДС ЛВ приймачів грубого (точного) стеження; $\hbar$ – постійна Планка; ν_i – частота випромінювання лазера-передавача.

Дисперсія погрішності вимірювання параметрів руху ЛА багатофункціональною ЛІВС визначається за формулою:

$$\min D[\Delta V_\lambda] = \min_{\left\{\frac{X_i}{X_j}\right\}} \left\{ \frac{\text{const}}{\prod_{j=1}^{n_1} X_j(Y_j)} + \sum_{i=1}^{n_2} X_i^2(Y_i) \right\} + D_c, \qquad (6)$$

де X_i – узагальнені значення технічних параметрів; X_j – функція обмеження за вартістю функціональних елементів (ФЕ); j – номер каналу; i – номер ФЕ; n_1, n_2 – число параметрів ФЕ j-ої та i-ої груп; $n = n_1 + n_2$ – число параметрів системи, що використовуються; $X_j(Y_j)$ – монотонна (сепарабельна) функція від технічного параметра, який впливає на втрати енергії сигналу (чим вона більша, тим краще); $X_i(Y_i)$ – монотонна функція нестабільності еталонів ФЕ (чим вона менша, тим вище точність); D_c – втрати точності вимірювань ПР за рахунок проходження в атмосфері.

Вартість багатофункціональної ЛІВС з зібраної маркетингової техніко-економічної статистики визначається за формулами:

$$\sum_{o=1}^{n1} C_j(X_j) \le C_{д1}, \quad \sum_{i=1}^{n2} C_i(X_i) \le C_{д2}, \quad C_{д1} + C_{д2} = C_д, \tag{7}$$

де $C_j(X_j)$, $C_i(X_i)$ – вартості ФЕ з параметрами j-ої та і-ої груп; $C_{д1}$, $C_{д2}$ –допустимі значення вартості ФЕ j-ої та і-ої груп, що визначають відповідний технічний параметр системи.

Функція обмеження за вартістю ФЕ (оптимальний алгоритм – оптимальні параметри) визначається за формулою:

$$X_{j(1)} = \frac{C_{e1(k-1)}(\overline{X}_{(k-1)})}{n_1 C'_{j(k-1)}(X_{j(k-1)})}, \tag{8}$$

де C_{e1} – економічна ефективність від застосування ФЕ:

$$C_{e1} = C_{д1} - \sum_{j=1}^{n_1} [C_j(X_{j0}) - C'_j(X_{j0}) X_{j0}], \tag{9}$$

де C'_j – похідна вартості ФЕ з j-м параметром; k – номер кроку ітерації; $\overline{X} \equiv (X_1, ...X_j, ...X_m)$ – вектор технічних параметрів.

Мінімальна дисперсія погрішності вимірювання параметрів руху ЛА системою за рахунок перешкод за методом невизначеної безлічі Лагранжа визначається за формулою:

$$\sigma_{n_1}(C) = \frac{k \prod_{j=1}^{n} C'_j(X_j)}{\left(\frac{C_{e1}}{n_1}\right)^{n_1}}. \tag{10}$$

Мінімальна дисперсія погрішності за рахунок нестабільності еталонів визначається за формулою:

$$\sigma_{n_2}(C_{e2}) = \frac{C^2_{e2}(\overline{X})}{\sum_{i=1}^{n_2} (C'_{oi}(X_{0i}))^2}. \tag{11}$$

Оптимальний алгоритм для отримання оптимальних параметрів еталонів визначається за формулою:

$$X_{n_2 \min} = \frac{C_{e2}(\overline{X}) C'_{on_2}(X_{0i})}{\sum_{i=1}^{n_2} (C'_{oi}(X_{0i}))^2}, \tag{12}$$

де $C_{e2} = \sum_{i=1}^{n_2} C_{oi} + \Delta C_{n2}$; $\Delta C_{n2} = \sum_{i=1}^{n_2} C'_{oi} X_{oi} - C_2$.

Якщо лінії середньоквадратичної регресії вартості на параметр є лінійні функції, то результат отримується з обчислення (12). А якщо ні, то формули використовуються ітераційно, послідовно покроково.

Отже, вартість багатофункціональної ЛІВС – це сума вартостей її

окремих частин і блоків (ФЕ), що дозволяє спростити задачу і представити її у вигляді блокового (сепарабельного) програмування. За допомогою вартості можливо порівняти внесок параметрів системи в показник її якості.

Висновки

Постановка і рішення задачі оптимізації параметрів (побудови) багатофункціональної ЛІВС за критерієм вартості ФЕ при обмеженнях на відношення сигнал/шум у вимірювальних і інформаційному каналах дозволяє заключити наступне:

– незважаючи на складний (пуассоновський) закон розподілу сигнальних фотонів та багатоетапність входження системи у зв'язок з ЛА, поставлена і вирішена у загальному вигляді задача оптимізації параметрів багатофункціональної ЛІВС з урахуванням значного числа каналів та ФЕ;

– у постановці задачі оптимізації параметрів багатофункціональної ЛІВС була використана уся інформація про структуру і принцип дії лінії зв'язку та у загальному вигляді – інформація про техніко-економічні показники $C_i(X_i)$ виробництва ФЕ;

– ця постановка задачі дозволяє, користуючись методом динамічного (блокового, сепарабельного) програмування, отримати рішення у вигляді квадратури (на кінцевому етапі);

– запропонований підхід дозволяє проаналізувати область оптимального рішення, а також визначити шляхи вдосконалення виробництва ФЕ за рахунок серійності та технологічності;

– отримані рішення дозволяють порівнювати за векторним показником якості аналогічні системи, а також визначати доцільність досягнення будь-якого значення показників якості;

– рішення задачі дозволяє, при заданих показниках якості системи, правильно планувати фінансування на неї;

– рішення задачі є основою для складання програм системи автоматизованого проектування технічних параметрів лазерної лінії зв'язку;

– отримані залежності показників якості від технічних параметрів системи можуть використовуватися для розрахунку характеристик ліній зв'язку у першому наближенні.

Мобільну багатофункціональну ЛІВС корисно використовувати в структурі полігонного випробувального комплексу для забезпечення проведення випробувань сучасних ЛА. Тому, окрім вище перелічених модулів і каналів до основного складу системи також входить наступна апаратура:

– геопозиціювання і метеоапаратура;

– прив'язки до системи єдиного часу;

– градуювання і калібрування тощо.

Таким чином, створення мобільної багатофункціональної ЛІВС дозволить забезпечити стійке кутове АС ЛА за рахунок використання усього цифрового вимірювального комплексу та високоточне вимірювання кутів азимута і місця, похилої дальності, радіальної і кутових (тангенціальних) швидкостей у широкому діапазоні дальностей, а також передачу інформації (команд керування) в умовах підвищеної скритності і завадазахищеності, об'єктивний контроль у денних і нічних умовах та, в разі необхідності, пошук ЛА у заданій зоні і його розпізнавання.

Література

1. Патент на корисну модель № 55645, Україна, МПК G01 S 17/42, G01 S 17/66. Частотно-часовий метод пошуку, розпізнавання та вимірювання параметрів руху літального апарату / О.В. Коломійцев – № u201005225; заяв. 29.04.2010; опубл. 27.12.2010; Бюл. № 24. – 14 с.

2. Патент на корисну модель № 43725, Україна, МПК H04 Q 1/453. Модифікований селектор подовжніх мод / О.В. Коломійцев, Г.В. Альошин, В.В. Бєлімов та ін. – № u200903693; заяв. 15.04.2009; опубл. 25.08.2009; Бюл. № 16. – 6 с.

3. Информационные технологии и системы в управлении, образовании, науке: Коллективная монография А.В. Коломийцев и др. // под ред. В.С. Пономаренко. – Х.: Цифрова друкарня, 2013. – № 1. – 278 с.

4. Информационные системы в управлении, образовании, промышленности: Коллективная монография Г.В. Алешин, А.В. Коломийцев и др. // под ред. В.С. Пономаренко. – Х.: Вид-во ТОВ "Щедра садиба плюс", 2014. – 498 с.

5. Информационные технологии и защита информации в информационно-коммуникационных системах: Коллективная монография Г.В. Алешин, А.В. Коломийцев и др. // под ред. В.С. Пономаренко. – Х.: Вид-во ТОВ "Щедра садиба плюс", 2015. – 486 с.

6. Казаков Е.Л. Адаптивная обработка сигнала однолокационных локаторов при распознавании воздушных целей: Монография Е.Л. Казаков, А.Е. Казаков О.В. Батурин, Д.Г. Васильев, А.В. Коломийцев // под ред. Е.Л. Казакова. – Х.: КП "Городская типография", 2011. – 174 с.

7. Казаков Е.Л. Распознавание целей по сигнальной информации в однопозиционных и многопозиционных локаторах: Монография Е.Л. Казаков, А.Е. Казаков, К.В. Садовый, А.В. Коломийцев // под ред. Е.Л. Казакова. – Х.: Міськдрук, 2015. – 459 с.

8. Kudriashov V. 'Experimental Evaluation of Opportunity to Improve the Resolution of the Acoustic Maps'. In: Kountchev R. and Nakamatsu K. (eds.), New Approaches in Intelligent Image Analysis, Intelligent Systems Reference Library 108, 2016. – pp. 353 – 373. Springer International Publishing Switzerland. DOI: 10.1007/978-3-319-32192-9_11, SJR: 0.154.

КОНЦЕПТУАЛЬНІ АСПЕКТИ ПО ВИРІШЕННЮ ПРОБЛЕМИ НАДАННЯ ІНФОРМАЦІЇ НА АЕРОФОТОЗНІМКУ

Краснoруцький А.О., Корольова Н.А.

Вступ

До бортового комплексу аеромоніторингу пред'являється ряд технічних вимог серед яких забезпечення: заданій інформаційній інтенсивності відеоінформації, заданій роздільній здатності для виконання певних завдань і заданої достовірності отримуваної інформації, обумовлена процесом її формування, обробки і передачі. В умовах роботи системи управління в кризовій ситуації до дистанційного відеосервісу пред'являються особливі вимоги: забезпечення необхідного рівня оперативності доставки аерофотознімка; забезпечення необхідного рівня надання інформації на отриманому аерофотознімку.

Тут існує дисбаланс: забезпечується можливість необхідного рівня оперативності доставки аерофотознімка, але з сумнівною достовірністю інформації, і навпаки: забезпечивши необхідний рівень надання інформації на доставленому аерофотознімку втрачається його оперативність доставки, що позначається на достовірності одержуваної відеомоделі аерофотознімка щодо реальних подій. Цей дисбаланс пов'язаний з особливістю цифрового аерофотознімка і особливістю сучасних технологій обробки зображень в системі надання відеопослуг.

Існуючі підходи надання відеопослуг в кризових ситуаціях є неефективними. Тому мета статті полягає в оцінці потенційної можливості в розробці концептуальних аспектів по вирішенню проблеми надання інформації на аерофотознімку в загальній системі сервісу дистанційного надання відеопослуг в процесі управління кризових ситуацій.

1. Аналіз структури аерофотознімка та його інформаційної надлишковості

Цифровий аерофотознімок відноситься до класу реалістичних зображень з функцією кореляції міжелементної залежності, прагне до дельта-функції і досить низьким коефіцієнтом кореляції між елементами зображення. Відмінною особливістю аерофотознімка є великий обсяг даних. Пояснюється це тим, що в аерофотознімку, як і в будь-якому вигляді зображень, є надмірність. В роботах [1-2] докладно описані відомі види присутньої на аерофотознімку надмірності: структурна, статистична, психовізуальна. Структурна та статистична надмірність аерофотознімка відносяться до синтаксичного типу виявлення закономірностей в синтаксичному описі аерофотознімка.

Синтаксичний опис – це цифрове представлення (опис) аерофотознімка без аналізу його семантичного змісту.

Статистична надмірність пояснюється наявністю в фотознімку статистичної залежності між різними групами елементів. В основі статистичної надмірності лежить виявлення статистичних закономірностей в синтаксичному представленні аерофотознімка. Прикладом статистичних закономірностей є нерівномірність розподілу значень елементів на аерофотознімку. Другим прикладом є кореляційні та статистичні залежності елементів на аерофотознімку. Статистичну надмірність аерофотознімка усувають методи статистичного кодування (метод Хаффмана, метод Шеннона, арифметичне кодування).

Структурна надмірність обумовлена наявністю на фотознімку структурних залежностей, які не пов'язані з законами розподілу. Прикладом структурної надлишковості є наявність ліній, геометричних фігур і т.ін. Структурну надмірність аерофотознімка усувають методи, засновані на афінних перетвореннях, виявлення довгих ліній, поліадичному кодуванні.

В основі психовізуальної надмірності лежить виявлення психовізуальних закономірностей. До психовізуальних закономірностей відносять закономірності обумовлені особливостями сприйняття аерофотознімка зоровою системою людини. Наприклад, зору людини дрібні деталі непомітні як і незначні перепади яскравості, в той же час у людського зору висока чутливість в області зеленої складової кольорового перепаду і більш висока чутливість до складової яскравості.

Людина, при аналізі будь-якого зображення, в основному оперує його контурами і загальним перепадом кольорів і, в той же час, людський зір порівняно нечутливий до малих змін. Але при цьому, психовізуального кодування як такого, не існує. Пояснення цьому лежить у тому, що немає адекватної моделі людського зору. Немає балансу (межі) де існує психовізуальна надмірність і безповоротна втрата важливої інформації.

Особливістю сучасних технологій обробки зображень заснований-них на виявленні різних закономірностей з подальшим етапом скорочення надмірності не розкривають семантичну інформацію, а відповідно і не спрямовані на її збереження.

2. Особливість надання інформації на аерофотознімку з урахуванням її дешифрування

У процесі моніторингу поверхні важлива роль відводиться дешифровщику. Дешифровщик, перш ніж сформувати з отриманого відеоматеріалу інформацію про важливі об'єкти а також передбачити подальшу їх поведінку, систематизує вихідні дані про об'єкти шляхом послідовного логічного проходження етапів від виявлення об'єкта до аналізу його стану. На ефективність роботи дешифровщика впливає як час доставки відеоматеріалу (аерофотознімка) з борту літального апарату (ЛА)

безпосередньо дешифровщику (це пов'язано зі старінням інформації), так і рівень підготовленості дешифровщика.

Дешифрування аерофотознімків – це метод дослідження території за її аерофотографічним зображенням. При цьому проводиться виявлення, розпізнавання та ідентифікація на аерофотознімках об'єктів місцевості, визначення їх якісних і кількісних характеристик.

Важкою проблемою роботи дешифровщика є завдання швидкої і безпомилкової оцінки великої кількості інформації, що у вигляді аерофотознімків надходять з літального апарату. Пов'язано це в тому числі і з тим, що для збільшення площі моніторингу за один політ необхідно встановлювати на літальний апарат кілька засобів фотографування. Це веде до того, що від моменту початку виконання завдання і до моменту отримання дешифровщиком готового аерофотознімка проходить багато часу, і, після його дешифрування отримані дані часто втрачають свою цінність. Метою дешифрування є своєчасне отримання документальних даних про місцевість і розташованих на ній об'єктів. Завданням - виявлення, правильне розпізнавання і класифікацію об'єктів, визначення їх кількісних характеристик, взаємозв'язків, стану, характеру діяльності, а також документування отриманої інформації [3].

Дешифровщик працює з аерофотознімком в умовах нестачі інформації з одного боку і надлишком її з іншою (рис.1).

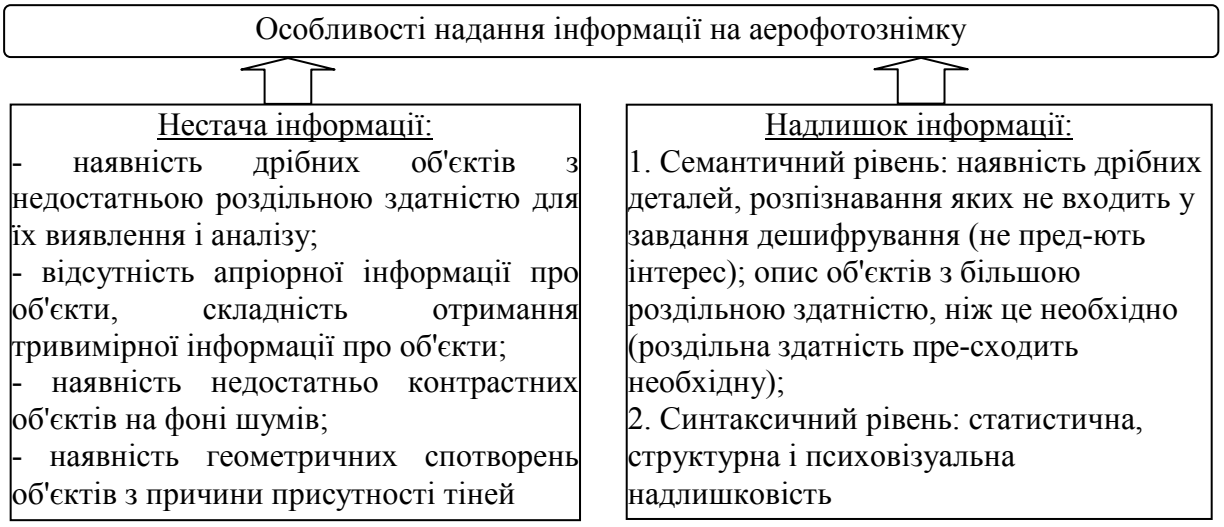


Рис. 1. Особливості надання інформації на аерофотознімку

Зауважимо, що не вся наявна інформація на аерофотознімку необхідна дешифровщику для виконання поставленого завдання по розкриттю об'єктів. Однак така інформація тут присутня і так чи інакше дешифровщик змушений з нею працювати, що відволікає його увагу і, в кінцевому підсумку, позначається на часі і якості виконання завдання. На аерофотознімку присутня, також, інформація, яка не є ключовою для виявлення об'єктів інтересу при дешифруванні. Ця інформація є

надлишкової з позиції дешифрування. Тому, з урахуванням особливостей такого складного об'єкта як аерофотознімок вводиться нове поняття дешифрувальної надмірності (рис.2).

Дешифрувальна надмірність аерофотознімка – це надлишок інформації, який пов'язаний з зображенням дрібних деталей і безлічі всіляких об'єктів, розпізнавання яких, на певних етапах, не входить у завдання дешифрування. У той же час, дешифрувальна надмірність включає і надмірність на синтаксичному рівні, тобто на рівні цифрового опису аерофотознімка. Виявлення та усунення дешифрувальної надмірності, згодом буде впливати як на оперативність доставки відеоматеріалу, так і на якість надання інформативного відеоматеріалу дешифровщику. При цьому необхідно враховувати, що кількість скасування дешифрувальної надмірності не має вплинути на зниження дешифрувальної інформативності аерофотознімка. Для розкриття цього поняття простежимо за роботою дешифровщика на різних етапах процесу дешифрування.

Рис. 2. Дешифрувальна надмірність аерофотознімка

Етап виявлення об'єктів, що полягає у сприйнятті об'єктів на аерофотознімку без визначення його сутності (тобто об'єкт в певній ділянці аерофотознімка або його немає). Дія дешифровщика тут відбувається два етапи. Перший етап полягає у виявленні безпосередньо самого об'єкта. Другий етап: визначення розташування об'єкта на аерофотознімку (в якому місці аерофотознімка знаходиться об'єкт). На цьому етапі дешифровщик стикається з труднощами виділення важливої інформації із загальної кількості інформації на аерофотознімку. Пов'язано

це з надлишком інформації на аерофотознімку або ж з потоком аерофотознімків. Надлишок інформації пов'язаний з зображенням дрібних деталей і безлічі всіляких об'єктів, розпізнавання яких, на певних етапах не входить у завдання дешифрування (відволікають увагу дешифровщика). У теж час геометричні та оптичні характеристики зображень об'єктів на аерофотознімку спотворені, порівняно з дійсними об'єктами, які підлягають дешифрування. Це веде до помилкового сприйняття об'єкта або ж, навпаки, до ігноруванню важливих об'єктів, що теж є помилкою дешифрування. Все це веде до неякісної роботи дешифровщика або ж до зростання затрат на сприйняття об'єктів (кілька проходів огляду ділянок аерофотознімка, залучення дешифровщика більш високої кваліфікації).

Методи подолання таких недоліків: зниження розмірності аерофотознімків. Це веде до зростання кількісних проходів літального апарата над певною зоною моніторингу, що, в свою чергу, веде до зростання витрат на отримання необхідного відеоматеріалу вказаної зони з борту літального апарату [4].

Етап розпізнавання об'єктів, що полягає у визначенні сутності виявлених об'єктів. На цьому етапі сприймаються і аналізуються складальні ознаки об'єктів. Таким чином, дешифровщик на цьому етапі виконує розпізнавання загального типу об'єктів (наприклад, автомобіль) і розпізнавання конкретної моделі (який автомобіль). Цей етап дешифрування може недостатньою детальністю об'єктів. Для подолання цього недоліку застосовуються методи оптичної обробки (збільшення або масштабування) певних ділянок знімка. Вимога до аерофотознімку на цьому етапі: ті ділянки аерофотознімка де є важливі об'єкти не повинні піддаватися методів стиснення з втратами інформації.

Знизити такі негативні прояви при масштабуванні ділянок аерофотознімка (збільшити розмірність об'єктів на знімку) можливо шляхом збільшення розмірності ПЗС-матриці фотокамери або ж шляхом зниження висоти польоту літального апарата над ділянкою моніторингу. Однак, у першому випадку, такий підхід веде до різкого збільшення розмірності всього аерофотознімка, що, в свою чергу, веде до зростання інформаційної інтенсивності і, як наслідок, до зростання тимчасової затримки доставки відеоматеріалу дешифровщику. Застосування другого підходу (зниження висоти польоту літального апарату) веде до підвищення кількості проходів над зоною моніторингу. Крім того зниження висоти: польоту літального апарату веде до його демаскуванню та підвищенню ймовірності його втрати.

Етап інтерпретації об'єктів на аерофотознімку. Для дешифровщика цей етап вважається найбільш трудомістким і вимагає певної майстерності, після чого приймається рішення про результати аеромоніторингу. В процесі інтерпретації виконується як детальний опис розпізнаних об'єктів, так і аналіз стану цих об'єктів на аерофотознімку. Тобто виконується

аналіз і узагальнення кількісних та якісних характеристик (ознак) об'єктів на аерофотознімку. На цьому етапі дешифровщик встановлює стан об'єктів (їх значущість) у конкретній обстановці, визначається їх попередні дії і прогнозується можливі подальші дії цих об'єктів. На цьому етапі виконується прив'язка конкретного об'єкта до мапи.

Для реалізації цих етапів дешифровщик пред'являє високі вимоги до первинного відеоматеріалу його якості з точки зору розташованих (на аерофотознімку) об'єктів. Виходячи з цього можна зробити висновок, що головне призначення дешифрування аерофотознімків як процесу це отримання «смислової» інформації про елементи місцевості та інших об'єктах, розташованих на ній. Тому в процесі обробки аерофотознімка, як на борту літального апарату, так і на землі необхідно мати такі технології обробки зображення, які з одного боку скорочують синтаксичну і семантичну складові зображення (для зниження інтенсивності відеопотоків і розвантаження каналів передачі даних), а з іншого боку - не руйнують семантичну складову аерофотознімка (забезпечують інформаційну цілісність щодо дешифрувальних ознак об'єктів інтересу) і дають дешифровщику можливість отримання «смислової» інформації про об'єкти інтересу (тим створити суттєві передумови для зниження часових витрат роботи дешифровщика).

Виходячи з цього, введемо визначення дешифрувальної інформативності (рис. 3). Дешифрувальна інформативність аерофотознімка – це інформативна частина аерофотознімка (необхідна достовірність інформації про об'єкт з урахуванням процесів обробки і передачі знімка), яку використовує дешифровщик в процесі його дешифрування і без якої неможливо якісно виконати процес дешифрування.

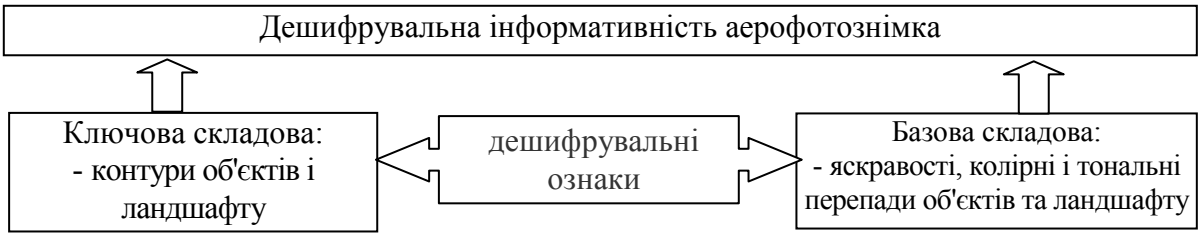


Рис. 3. Структура інформації для процесу дешифрування

Тут йдеться про якісну характеристику аерофотознімка з позиції сприйняття дешифровщиком об'єктів які представляють інтерес. Тобто об'єкт, що цікавить на аерофотознімку дешифровщик: сприймає, не сприймає, або сприймає після додаткової обробки аерофотознімка (застосування сервісних засобів). Для забезпечення скорочення дешифрувальної надмірності в умовах забезпечення заданої дешифрувальної здатності пропонується ввести таке поняття як дешифрувальне кодування (рис.4).

Дешифрувальне кодування – це отримання знань про

дешифрувальних ознак і ефективне синтаксичний опис аерофотознімків за цими знаннями.

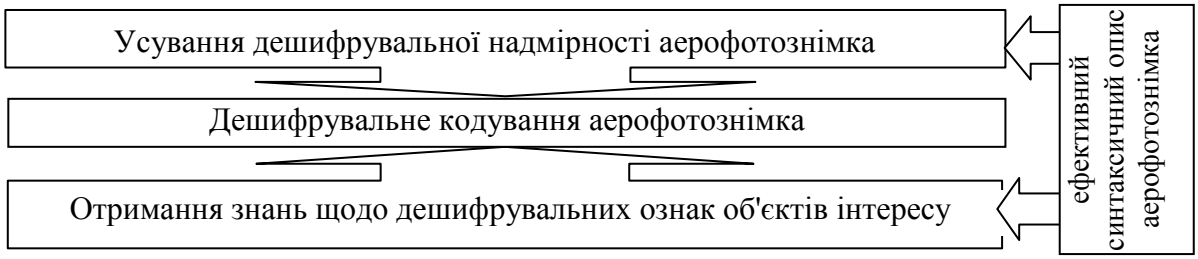


Рис. 4. Структура дешифрувального кодування

На ефективність роботи дешифровщика впливає наданий відеоматеріал, а саме дешифрувальна інформативність аерофотознімка. Дешифрувальна інформативність аерофотознімка, в свою чергу, залежить від роздільної здатності (детальності) аерофотознімка і достовірності відповідності аерофотознімка сформованого бортовими засобами реєстрації та моделі аерофотознімка, який отримує безпосередньо дешифровщик [5].

Тут виникає дисбаланс між достовірністю одержуваного аерофотознімків і оперативністю доставки: з одного боку є можливість оперативно доставити аерофотознімок з борту ЛА, з іншого боку збільшується ймовірність помилки правильного дешифрування (рис.5). Причина полягає в застосуванні технологій бортової та наземної обробки аерофотознімка з метою оперативної передачі даних по існуючих бортовим каналах зв'язку. Тому процес дешифрування диктує (накладає) наступні вимоги (обмеження) щодо методів обробки аерофотознімків на борту літального апарату:

– обмеження втрат інформації про дешифрувальних ознак об'єктів інтересу (збереження яскравості складових і контурної інформації про об'єкти) при відновленні аерофотознімків;

– мінімізація складності алгоритмів обробки (своєчасність обробки аерофотознімків).

Робимо припущення, що на основі відомих методів і технологій обробки зображень неможливо побудувати дешифрувальне кодування в умовах забезпечення одночасного скорочення дешифрувальної надмірності та забезпечення необхідної дешифрувальної здібності аерофотознімка.

Це обумовлено, з одного боку, структурною особливістю цифрового аерофотознімка, а з іншого – проблемними недоліками існуючих методів обробки зображень по відношенню до аерофотознімка.

Однак плата за такий підхід: зниження інформативності зображень (зниження якості, часткова втрата інформації) або ж зростання обчислювальної складності алгоритму стиснення (збільшення вимог до

продуктивності процесора або підвищення витрат часу на реалізацію алгоритму стиснення).

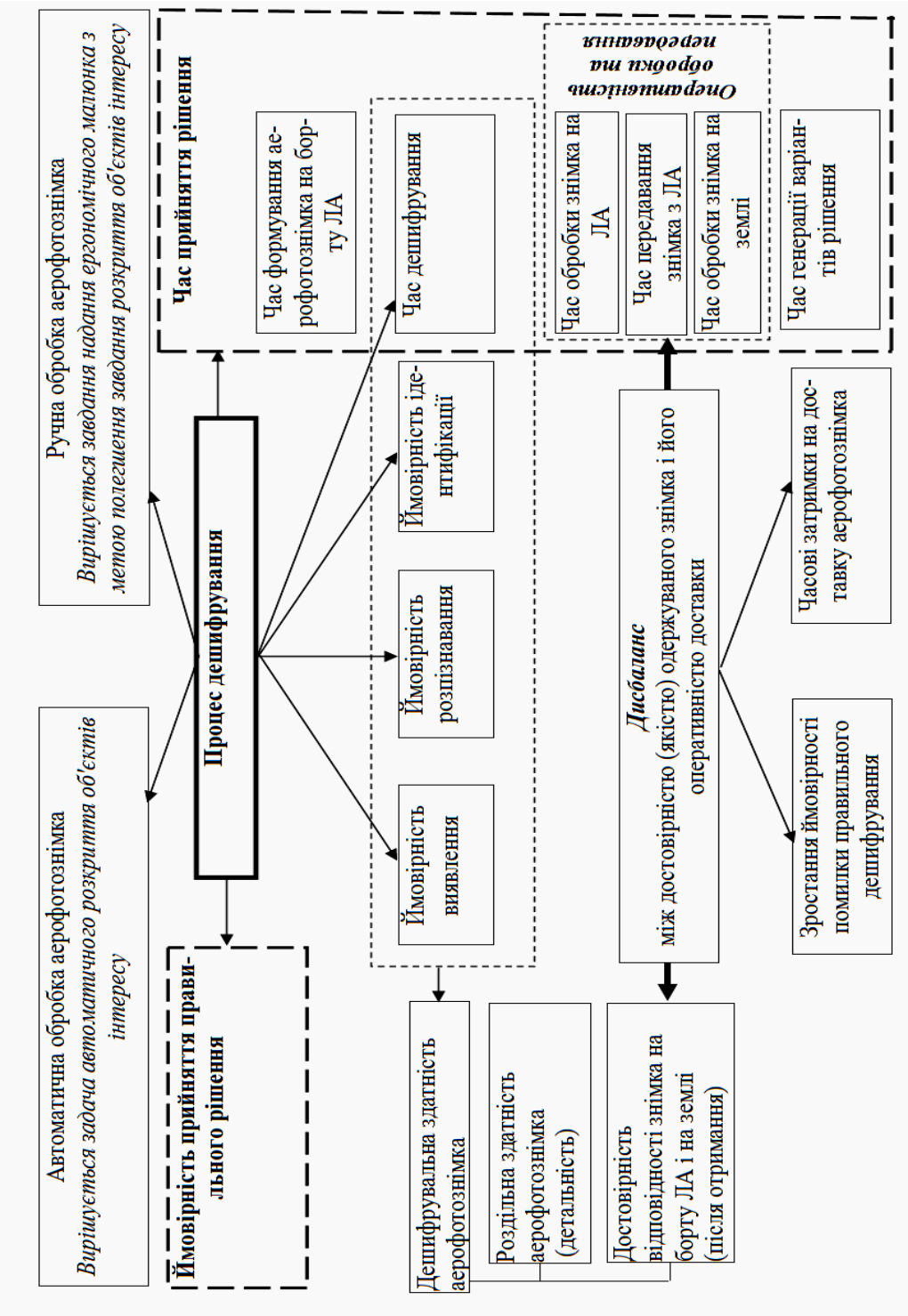


Рис. 5. Схема процесу дешифрування в умовах дефіциту часу (необхідного часу прийняття рішення) і динамічної зміни обстановки

По відношенню до аерофотознімки: перше ставить під сумнів достовірність інформації отриманого аерофотознімки, друге знижує

вірогідність одержуваної відеомоделі аерофотознімки щодо реальних подій.

3. Аналіз проблемних аспектів існуючих методів обробки зображень, спрямованих на підвищення оперативності доставки інформації

Особливістю сучасних технологій обробки зображень (компресії, фільтрації) заснованих на виявленні різних закономірностей з по-наступним етапом скорочення надмірності (статичної, структурної) не ріжуть семантичну інформацію, а відповідно і не спрямовані на її збереження (не допущені до її зміни).

Тому проводиться аналіз проблемних аспектів існуючих методів обробки зображень (компактного представлення даних), які направлені на зниження часу доставки інформації (аерофотознімки) з борта ЛА шляхом зниження інформативною інтенсивності аерофотознімки, для забезпечення необхідної ефективності дешифрування отриманої інформації.

Перший проблемний аспект відноситься до питання зниження обсягів відеоданих, що надходять в канал передачі даних з борту літального апарату (рис.6).

В даний час для зниження інформативною інтенсивності аерофотознімку, при доставці його з борту літального апарату застосовують технології компресії відеоданих.

Пов'язано це з досить таки повільною пропускною спроможністю бортових каналів передачі відеоданих. Така технологія спрямована на підвищення інформаційної щільності синтаксичного опису всього зображення на основі виявлення його інформативних складових. Однак тут присутнє наступне протиріччя. Класичні методи компресії відеоданих (на платформах JPEG, JPEG2000) будуються на спектрально-частотному поданні відеоданих (зображення) з подальшою їх квантування. В такому спектрі здійснюється концентрація значної частини енергії в невеликій кількості спектральних низькочастотних компонентах (складових спектра), які і є інформативними. Таким чином, виявлення інформативних ознак зображення здійснюється шляхом обліку психовізуальних закономірностей.

Тут не враховується нерівномірність розподілу інформативної (семантичної) складової по всьому зображенню (аерофотознімка). Крім того застосування таких методів має ряд негативних наслідків; оскільки низькочастотні спектральні компоненти несуть інформацію про структурну частини об'єктів зображення, то вони є важливим аргументом для правильної ідентифікації об'єктів.

Але ж при відновленні зображення (зворотний процес компресії) частина низькочастотних компонент або відновлюються з похибкою або ж зовсім не відновлюється. Пояснюється це розподілом помилки (шумів) по

всіх елементах зображення внаслідок квантування її спектральних компонентів.

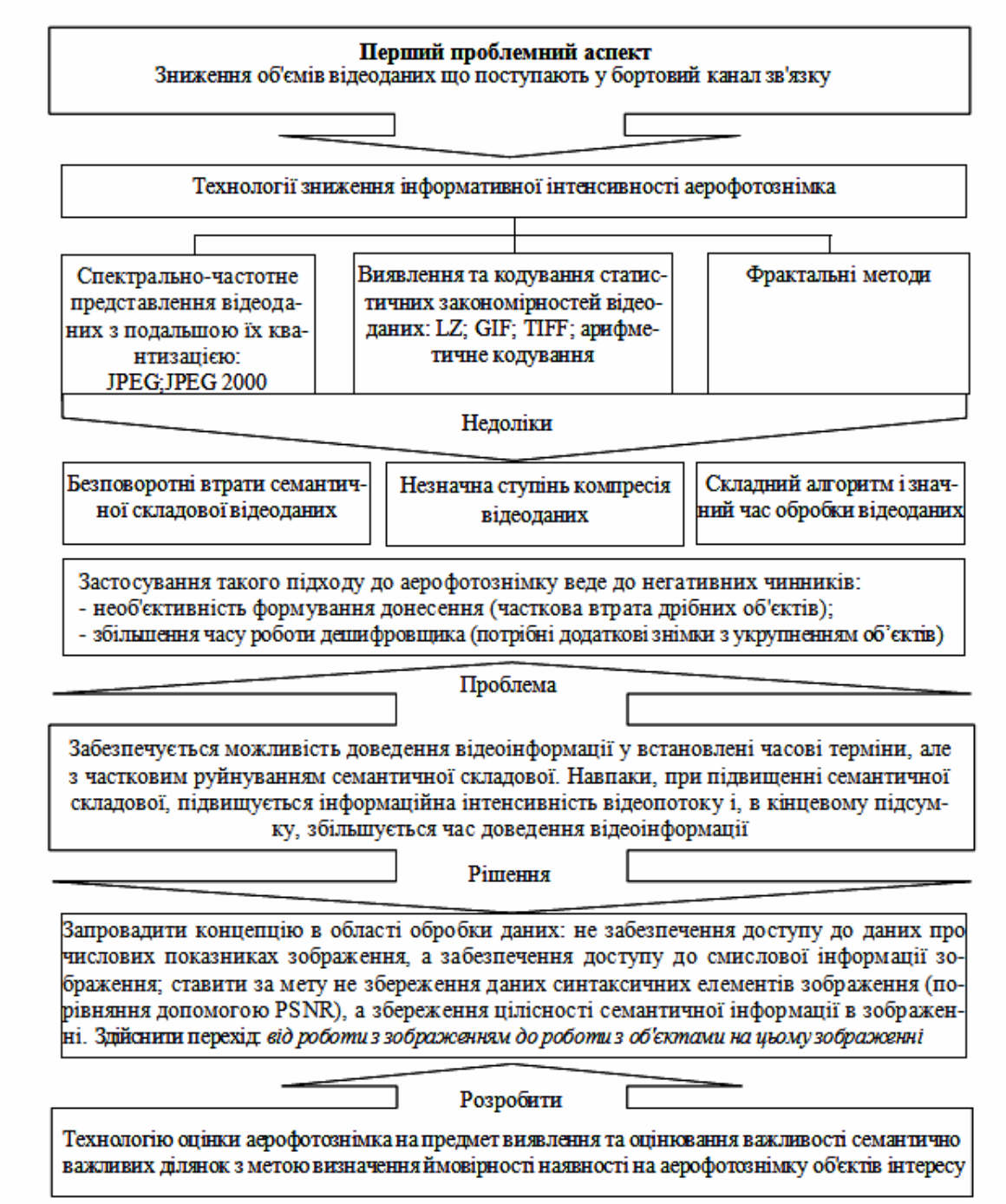


Рис. 6. Структурна схема першого проблемного аспекту (зниження обсягів відеоданих надходять в бортовий канал зв'язку)

Це веде до втрати семантично значимої інформації для дешифрування:

– немає можливості виділення інформативних ознак на фонових ділянках аерофотознімка (прим'ята трава, сліди від коліс і т.ін.). наслідком

чого є ще більш притуплення непрямих ознак дешифрування об'єктів інтересу;

– відсутній баланс психовізуальної надмірності і безповоротної втрати значущої інформації. На певному етапі це призводить до безповоротної втрати високочастотних компонент (перепаду яскравості) елементів зображення і як наслідок: істотне зниження роздільної здатності аерофотознімка;

– застосування різних видів ортогональних перетворень (в залежності від ступеня кореляції міжелементної залежності зображення), як складової частини платформи JPEG, не можуть рівномірно апроксимовані різні області аерофотознімка, такі як різкі і плавні зміни яскравості (наявність тіней, фоновий пейзаж і т.ін.). Наслідком чого є внесення додаткової перешкоди при обробці зображень, що, в свою чергу веде до безповоротних втрат семантичної складової аерофотознімка.

У той же час існують інша група методів обробки зображень, засновані на виявленні та кодування статистичних закономірностей (LZ, GIF, TIFF, арифметичне кодування). Однак такі методи характеризуються незначним ступенем компресії аерофотознімки, а в деяких випадках і навпаки, сприяють збільшенню обсягів вихідного зображення. Пояснюється це низьким коефіцієнтом кореляційної залежності елементів зображення і як наслідок невисокої ймовірністю повторення символів елементів зображення.

Третя група методів обробки зображень (фрактальні методи) характеризуються складним алгоритмом обробки даних і значними часовими витратами (близько кілька годин) на їх обробку. Обмежені продуктивні можливості бортової апаратури обробки відеоданих, через невисокі енергетичні можливості бортового генератора, накладає обмеження на реалізацію складних алгоритмів їх обробки (кількість арифметичних і логічних обчислювальних операцій).

Таким чином, існуючі групи методів обробки відеоінформації мають суттєві недоліки: забезпечується можливість доведення відеоінформації в встановлені тимчасові терміни, але з частковим руйнуванням семантичної складової. Навпаки, при підвищенні семантичної складової, підвищується інформаційна інтенсивність відеопотоку і, в кінцевому підсумку, збільшується час доведення відеоінформації. Крім того коди і конструкції, які базуються на відомих методах характеризуються підвищеною вразливістю при проходженні відеоданих по радіоканалах (за рахунок зниження надмірності інформації підвищується ймовірність спотворення цієї інформації в результаті дії перешкод присутніх в каналах зв'язку).

Практично всі ці технології мають один і той же недолік - це певний показник ступеня втрати якості (даних) в результаті когось прес з подальшим відновленням зображення. Застосування такого підходу до аерофотознімки, веде до часткової втрати дрібних (на аерофотознімки)

об'єктів, що, в кінцевому підсумку, веде або до необ'єктивності формування донесення або ж до збільшення часу роботи дешифровщіка (потрібні додаткові аерофотознімки з великою роздільною здатністю для укрупнення об'єктів на аерофотознімку).

Для вирішення першого проблемного аспекту пропонується ввести концепцію, яка полягає в принципово новому підході в області обробки даних, а саме незабезпечення доступу до даних числовим показниками даних зображення, а забезпечення доступу до смислової інформації зображення. Ставити за мету не збереження синтаксичних даних елементів зображення (порівняння за допомогою PSNR), а збереження цілісності семантичної інформації в зображенні. Здійснити перехід: від роботи з зображенням до роботи з об'єктами на цьому зображенні. Для здійснення такого підходу необхідно розробити наукомісткі і технологічні основи для оцінки аерофотознімка на предмет виявлення і оцінювання важливості семантично значущих ділянок з метою визначення ймовірності наявності об'єктів інтересу.

Другий проблемний аспект відноситься до питання дослідження семантичної складової аерофотознімка. Цей проблемний аспект пов'язаний з технологіями розпізнавання образів (рис. 7).

Існує безліч алгоритмів вирішують завдання розпізнавання образів. Однак всі ці алгоритми здійснюють тільки фільтрацію контурів

Результатом чого, після обробки таким фільтром, замість зображення з'являються тільки контури, а інформація про ландшафті аерофотознімки втрачається (застосовується для обробки рентгенівських знімків).

Такий підхід в області обробки зображень з подальшим дешифруванням не придатний, так як втрачається зв'язок об'єктів інтересу і ландшафтом місцевості, де розташований даний об'єкт.

Існує інший підхід розпізнавання образів - це наявність апріорної інформації про об'єкти моніторингу. Однак в умовах кризової ситуації (висока динаміка зміни обстановки) тут з'являються такі недоліки: апріорна недостатність інформації, наявність значної кількості предметів (об'єктів) які не уявляють інтерес. Результатом чого є підвищення витрат часу на обробку зображення і підвищення ймовірності помилки дешифровщіка при дешифруванні аерофотознімки.

Для вирішення другого проблемного аспекту пропонується створити методи і метрики семантичного аналізу аерофотознімків, спрямованих на виявлення і виділення значущих об'єктів, які представляють інтерес. При семантичному аналізі аерофотознімка здійснити перехід від ідентифікації об'єктів до підходу ідентифікації семантичної інформативності сегментів аерофотознімка. Пропонується розпізнавати не просто об'єкти, а ввести правило оцінки присутності цих об'єктів на аерофотознімку. Для здійснення такого підходу необхідно розробити теоретичну базу і методи

інтелектуальної ідентифікації сегментів аерофотознімки за ступенем інформативності в семантичному аспекті.

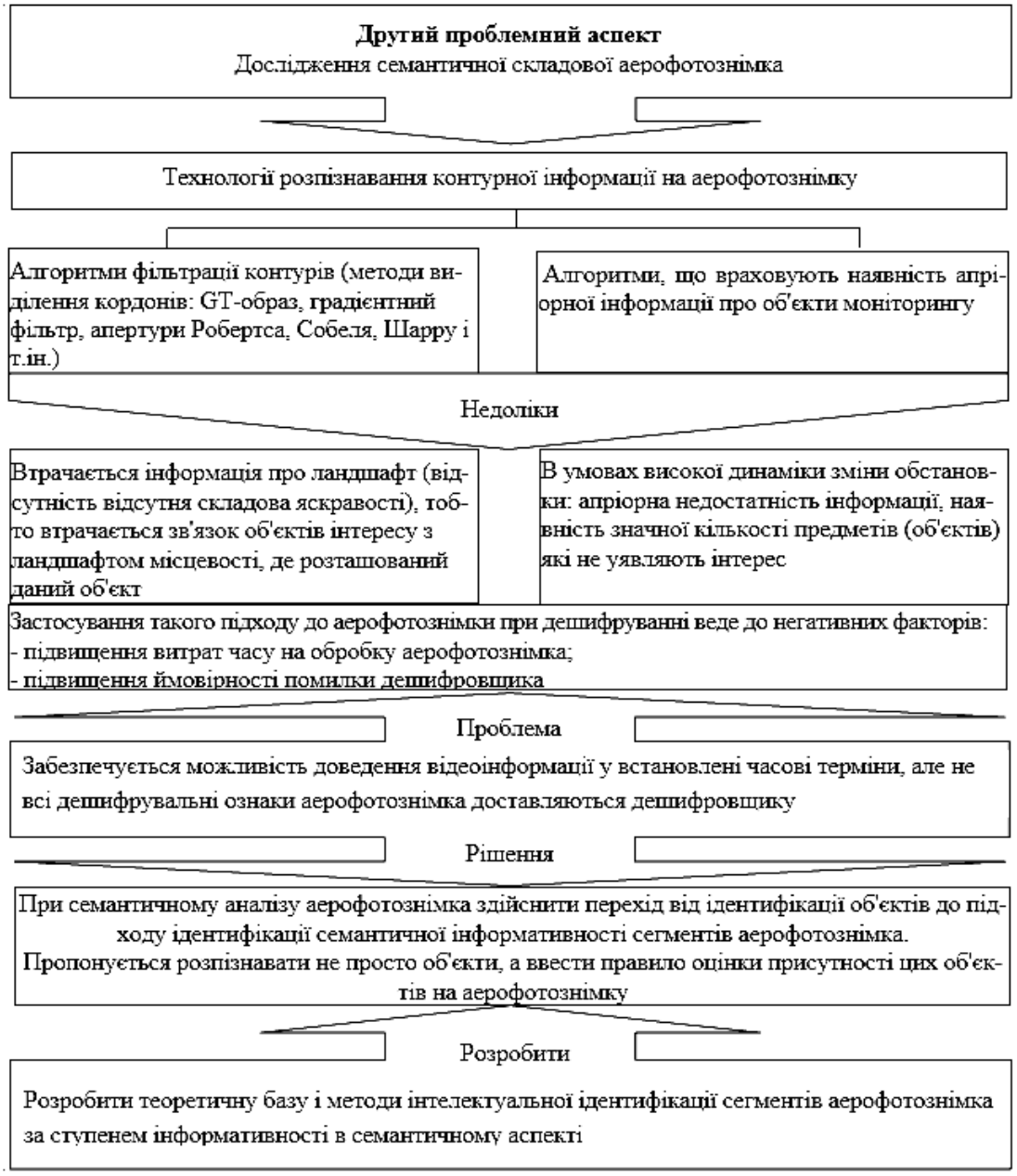


Рис. 7. Структурна схема другого проблемного аспекту (дослідження семантичної складової аерофотознімка)

Третій проблемний аспект відноситься до питання теоретичної платформи оцінювання семантичної інформативності ділянок (сегментів) аерофотознімка з урахуванням виявлення дешифрувальних ознак об'єктів інтересу на цьому аерофотознімки (рис. 8).

Існуючі методики визначення важливості інформації на певних ділянках (сегментах) аерофотознімки базуються на оцінці експертів. Однак такий підхід вносить суб'єктивний характер щодо прийняття рішення по виявленню об'єктів інтересу, оскільки залежить від кваліфікації

дешифровщіка і наявності апріорної інформації. Це веде до внесення помилкових рекомендацій і збільшення витрат часу на етапи дешифрування. У той же час є методики оцінювання вагових коефіцієнтів сегментів зображення. Але методів оцінювання семантично важливих сегментів зображення і кількості інформації в них не існує.

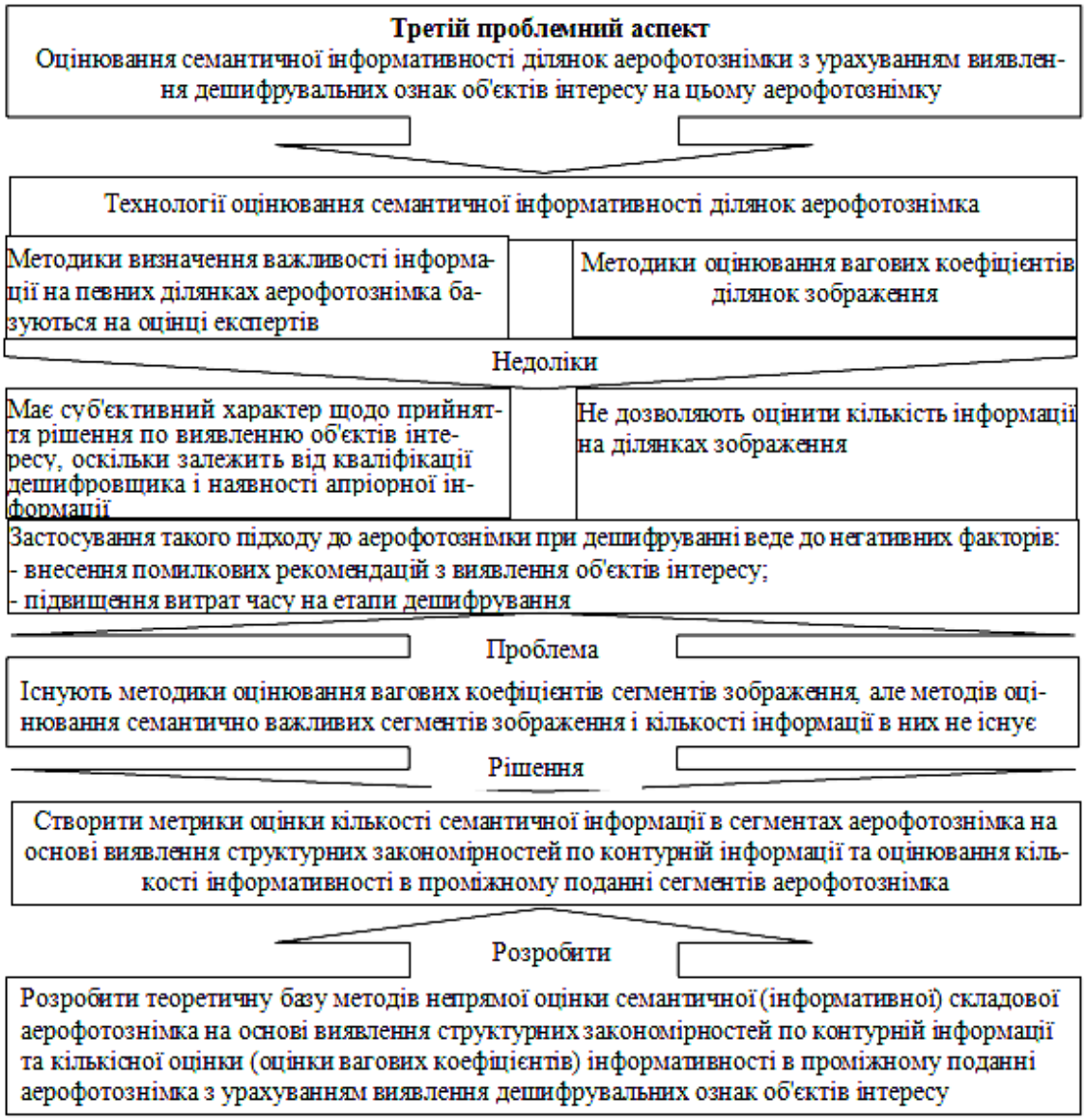


Рис.8. Структурна схема третього проблемного аспекту (оцінювання семантичної інформативності ділянок аерофотознімки з урахуванням виявлення дешифрувальних ознак об'єктів інтересу)

Для вирішення третього проблемного аспекту пропонується створити метрики оцінки кількості семантичної інформації в сегментах аерофотознімки на основі виявлення структурних закономірностей по контурній інформації та оцінювання кількості інформативності в проміжному поданні сегментів аерофотознімки. Для цього необхідно розробити теоретичну базу методів непрямої оцінки семантичної складової зображення на основі виявлення структурних закономірностей по

контурній інформації та кількісної оцінки (оцінки вагових коефіцієнтів) інформативності в проміжному поданні аерофотознімки з урахуванням виявлення дешифрувальних ознак об'єктів інтересу.

Четвертий проблемний аспект відноситься до питання ефективного синтаксичного опису семантичних (інформативних) складових аерофотознімку з урахуванням розкриття дешифрувальних ознак об'єктів інтересу (рис. 9).

Рис. 9. Структурна схема четвертого проблемного аспекту (ефективне синтаксичне опис семантичних складових аерофотознімки з урахуванням розкриття дешифрувальних ознак об'єктів інтересу)

Для виявлення об'єкта дешифровщик оперує не тільки контурної інформацією, а й складової яскравості об'єктів інтересу, а також загальним перепадом квітів ландшафту (визначення непрямих дешифрувальних ознак), однак існуючі методи, що формують інформативне виділення контурів об'єктів зображення і одночасний облік (опис) складової яскравості аерофотознімки (зображення) знаходяться на недостатньому рівні розвитку.

Для вирішення четвертого проблемного аспекту пропонується розробити методи обробки зображень, які зможуть одночасно виділити контурну інформацію і забезпечити облік складової яскравості об'єктів інтересу.

П'ятий проблемний аспект відноситься до питання помилок виділення інформативних відомостей в результаті семантичної обробки аерофотознімки.

Аналіз існуючих методів семантичної обробки зображень показує, що існує суперечність між обчислювальною складністю алгоритмів обробки і ймовірністю помилкового сприйняття дешифровщиком об'єктів інтересу відновленого (після застосування методу обробки зображень) аерофотознімка.

Методи обробки зображень без втрати якості і не високої обчислювальної складності (на фундаменті виявлення довжин серій, коди Хаффмана) не дозволяють на достатньому рівні підвищити інформативну щільність аерофотознімка. У той же час методи обробки зображень на платформі JPEG з одного боку дозволяють підвищити інформативну щільність аерофотознімка до заданого рівня, однак існує висока ймовірність втрати ключової складової аерофотознімка при дешифруванні, тим самим ставлячи під загрозу збереження семантично значимої інформації в моделі аерофотознімка.

Це спричиняє високу ймовірність не виконання вимоги щодо збереження її актуальності для роботи дешифровщика в режимі реального часу. Застосування методів на основі вейвлет-перетворень тягнуть за собою застосування алгоритмів високої обчислювальної складності, що, в свою чергу, позначиться на зростанні часових параметрів при обробці аерофотознімків. Це спричиняє високу ймовірність не виконання вимог щодо збереження актуальності отриманої інформації для роботи дешифровщика в режимі реального часу. Загальне для всіх розглянутих методів оброблення зображень – це не здатність виділення ключової інформації для дешифрування (рис. 10).

Для вирішення п'ятого проблемного аспекту пропонується розробити теоретичну базу і методи обробки аерофотознімків дозволяють, з одного боку, виділяти ключову інформацію про об'єкти інтересу, з іншого – мають обчислювальну платформу не високої складності (порівнянну з продуктивною потужністю бортового генератора).

Тому для вирішення цих п'яти проблемних аспектів необхідно створити теоретичні основи і методи ефективного синтаксичного опису утримання аерофотознімка з урахуванням інтелектуалізації процесу ідентифікації ключовий семантичної інформації про об'єкти інтересу за ступенем інформативності семантичного змісту сегментів аерофотознімка.

Пропонується створити технологію обробки відеоданих, які спрямовані на підвищення інформативності аерофотознімка і зменшення

сумарного часу дешифрування даних з метою прийняття рішення в інтересах аеромоніторингу. Тобто запровадити принципово новий підхід в обробці зображень, а саме пошук і виділення (підсвічування) потрібного (важливої) об'єкта на аерофотознімку.

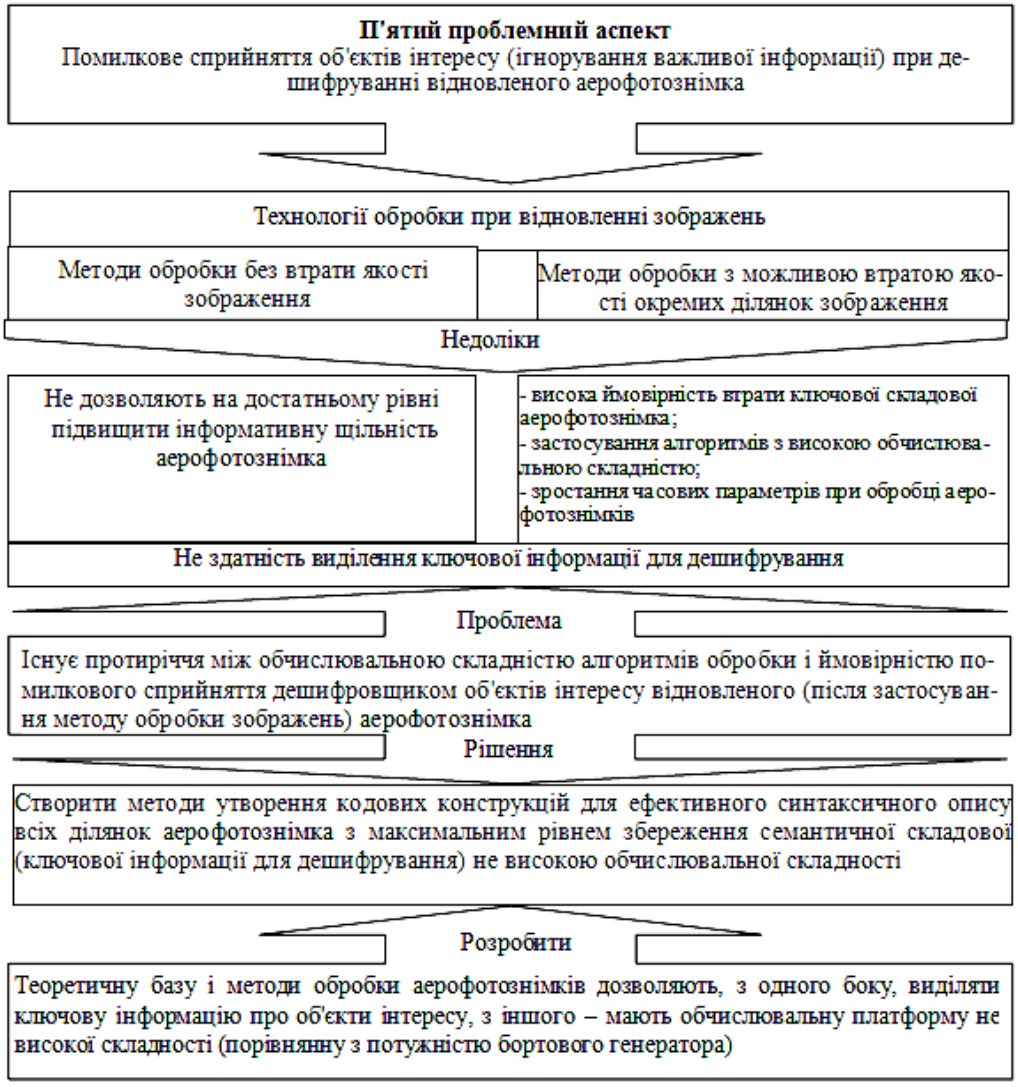


Рис. 10. Структурна схема п'ятого проблемного аспекту (помилки виділення інформативних відомостей у результаті семантичної оброблення аерофотознімка)

Побудова і реалізація такої концепції, щодо доставки відеоматеріалу дешифровщику представлена на рис. 11. Пропонується виконати початкову функцію дешифрування на борту літального апарату і вже корисну інформацію про об'єкт моніторингу передавати по каналу зв'язку на наземний пункт управління для остаточного прийняття рішення по дешифровці об'єкта. Це дозволить знизити інформаційну інтенсивність з урахуванням збереження семантично значущої для дешифрування інформації і «підсвітити» важливі блоки зображення.

Таким чином, пропонується створити таку технологію представлення відеоданих, спрямовану на збереження семантично значимої інформації моделі аерофотознімка і забезпечення її актуальності для роботи дешифровщика в режимі реального часу, для якого одночасно буде забезпечуватися:

– надання відеоінформації в зручному вигляді для дешифровщика, що забезпечить зниження часу роботи дешифровщика і підвищить ефективність дешифрування отриманої інформації;

– зниження інформаційної інтенсивності аерофотознімків, за умови збереження семантично важливою інформаційною складовою про об'єкти інтересу, що забезпечить необхідний рівень оперативності доставки інформації.

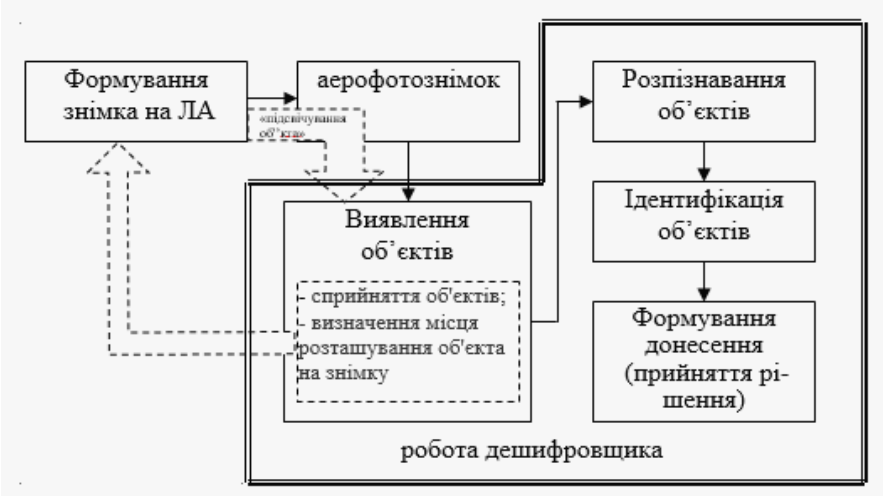


Рис. 11. Схема доставки відеоматеріалу дешифровщику об'єктів інтересу

Звідси випливає концепція створення теоретичних основ і методів де-шифрувального кодування аерофотознімків як послуги надання дистанційних відеоінформаційних сервісів з використанням бортових комплексів в умовах кризових ситуацій.

Висновки

1. Обґрунтовано, що розробку наукових основ підвищення оперативності доставки відеоінформації з борту літального апарату в системі аеромоніторингу, особливо в умовах кризових ситуацій необхідно проводити в правлінні вирішення суперечності, в основі якого лежить дисбаланс між необхідним рівнем оперативності доставки аерофотознімка, але з сумнівною вірогідністю інформації отриманого аерофотознімка, і забезпечення необхідного рівня надання інформації на доставленому аерофотознімку, але з втратою його оперативності доставки, що, у свою чергу відбивається на достовірності одержуваної відеомоделі аерофотознімка щодо реальних подій і зростанні часового циклу управління.

2. Доведено, що ядром такого дисбалансу є особливість цифрового аерофотознімка і особливість сучасних технологій обробки зображень. Отже, для підвищення оперативності доставки інформації з одночасним збереженням її інформативності та актуальності необхідно розробити новий підхід до бортової обробці аерофотознімків, з метою збереження важливої інформації для дешифрування.

3. Встановлено, що особливістю сучасних технологій обробки зображень заснованих на виявленні різних закономірностей з подальшим етапом скорочення надмірності не розкривають семантичну інформацію, а відповідно і не спрямовані на її збереження, що накладає певні обмеження на використання аеромоніторингу в умовах кризової ситуації.

4. Обґрунтовано, що на аерофотознімку є інформація, яка є надлишковою з позиції дешифрування. Тому введено нове поняття дешифрувальна надмірність аерофотознімка і виявлення, а надалі і усунення такої надлишковості, буде впливати як на оперативність доставки відеоматеріалу, так і на якість надання інформативного відеоматеріалу дешифровщику.

6. Проведено аналіз проблемних аспектів для методів обробки аерофотознімків, спрямованих на розробку концептуальних аспектів по вирішенню проблеми надання інформації на аерофотознімку дешифровщику в умовах кризової ситуації, виявив п'ять проблемних аспектів. Обґрунтовано, що системний підхід для вирішення сформульованих проблемних аспектів для методів обробки аерофотознімків полягає в розробці технології обробки відеоданих, які спрямовані на підвищення інформативності аерофотознімка і зменшення сумарного часу дешифрування даних з метою прийняття рішення в інтересах аеромоніторингу, яка дозволить здійснити пошук і виділення ключової інформації на аерофотознімку з наступним процесом афективного синтаксичного опису всього аерофотознімка з такої інформації з метою її максимального збереження.

Література

1. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов / Под ред. И.Б. Фоменко. - М.: Связь, 1980. - 248 с.

2. Бондарев В.Н., Трестер Г., Чернега В.С. Цифровая обработка сигналов: методы и средства. Учебное пособие для вузов. 2-е изд. – Х.: Конус, 2001. – 398с.

3. Баранник В.В. Метод повышения доступности видеоинформации аеромониторинга / В.В. Баранник, О.С. Кулица //Радиоэлектронные компьютерные системы. – 2013. – №3. – С. 17 – 20.

4. Власов А.В. Анализ методов обнаружения границ объектов на изображениях и их классификация / А.В. Власов, В.В. Баранник, А.В. Яковенко // Сучасна спеціальна техніка. – 2012. – вип. 3 (30). – С. 17 – 27.

5. V Barannik, A.Krasnorutsky, J.Gancarczyk. // VI International conference of students and doctoral students, "The engi-neer of the XXIst century", 2016, (Biel-sko-Biala, December 2, 2016) / Bielsko-Biala: 2016. – P. 185-190.

ШЛЯХИ МІНІМІЗАЦІЇ ІНФОРМАЦІЙНИХ ВТРАТ В ЦЕНТРАХ ОБРОБКИ ДАНИХ

Оксіюк О.Г.

Вступ

На ряду з розвитком інформаційних технологій збільшується не лише технічний прогрес, але і різноманітні згубні процеси, включаючи незаконне проникнення, руйнування, копіювання інформації, її ствотворення, знищення та. інш. На підставі цього актуалізується питання про захист від кібератаки, яка має бути своєчасним і достатнім.

Дана наукова робота присвячується дослідженню процесу мінімізації інформаційних втрат в Центрах Обробки Даних, як одного з пріоритетних напрямів захисту від кібератак.

Сьогодні розвиток інформаційних технологій (Information Technologies, IT) пов’язаний із реалізацією концепції центрів оброблення даних (ЦОД) – комплексних організаційно-технічних рішень для створення високопродуктивної, відмовостійкої IT-інфраструктури. До головних завдань ЦОД належать консолідоване зберігання і опрацювання даних користувачів, надання їм прикладних сервісів, підтримка функціонування застосувань. В Україні концепція ЦОД має перспективи реалізації в міністерствах і відомствах, корпораціях і великих організаціях, насиченість IT-інфраструктури яких породжує проблему пошуку способів збільшення ефективності її використання. Схожі проблеми виникнуть і у будь-якої компанії, яка вкладає кошти у створення системи ЦОД для обслуговування користувачів інших компаній, тобто набуває статусу хостингової компані.

У глобальному розумінні, комплексне організаційно-технічне рішення, призначене для створення високопродуктивної і відмовостійкої інформаційної інфраструктури, являє собою центр обробки даних (ЦОД). Проте, на сьогодні, наукова література дає визначення ЦОДу, як окремому приміщенню, яке призначене для розміщення обладнання для обробки і зберігання даних, що забезпечує підключення до швидких каналів зв'язку [1,3].

Сьогоднішній стан інформаційного захисту ЦОДу є не вирішеним, в силу того, що на ряду з розвитком інформаційних технологій збільшується не тільки технічний процес, а й різноманітні пагубні процеси, включаючи впливи на інформаційні масиви, що зберігаються, функціонують та розповсюджуються у межах користувачів ЦОД. До вирішення питань захисту, на протязі кількох років, підходило чимало вчених, як зарубіжних так і вітчизняних. Знайдено та описано головні принципи та механізми захисту ЦОДу, проте питання мінімізації інформаційних втрат залишається відкритим та потребує додаткового дослідження.

В умовах глобальної інформатизації суспільства і бізнесу зростають вимоги користувачів до рівня сервісу, керованості, надійності, доступності і масштабованості ІТ-інфраструктури [1]. Це ускладнює управління ІТ-інфраструктурою. Створення ЦОД вимагає значних коштів, ефективної підтримки ІТ-інфраструктури, наявності у штаті висококваліфікованих фахівців. Тому все більше фірм стають клієнтами хостингових компаній.

Хоча є певний досвід у створенні й експлуатації ЦОД, залишаються проблеми, які ще потребують розв'язання. Експлуатація ЦОД можлива лише за наявності організаційної структури, оснащеної сучасним інструментарієм та комплексом методик збору, аналізу інформації, прийняття рішень та управління їхнім відпрацюванням. Створення зазначених інструментарію та комплексу методик вимагає глибокого розуміння процесів, які відбуваються в хостингових організаціях, функціонування ІТ- інфраструктури, чіткого формулювання конкретних задач дослідження, розроблення математичних моделей і відповідних методів розв'язання задач та, зрештою, реалізації згаданих інструментарію та методик у структурі системи управління ІТ- інфраструктурою.

ЦОД представляє з себе об'єднання великої кількості програмних і апартаних платформ різного типу - серверів, СЗД, ОС, систем управління навантаженням і засобами резервування даних. При цьому проектування необхідно здійснювати таким чином, щоб забезпечити високу готовність системи (англ. - висока доступність). Вона досягається через резервування обчислювального і селевого комплексу з впровадженням нових технологій автоматичного оновлення при збоях. Коли ми говоримо про безвідказність системи такого рівня, то маємо на увазі готовність роботи центру обробки даних на 99,95% (або 4,5 годин простою на рік). При цьому можуть бути заплановані зупинки на проведення планових робіт, профілактики системи і т.п.

До головних об'єктів ЦОДу, які потребують якісного захисту, слід віднести: інформацію, яка зберігається і обробляється в системі; програмне забезпечення, встановлене в рамках ЦОД; елементи системи, обладнання центру [2].

Структура сучасного ЦОД включає в себе:

- серверний комплекс;
- систему зберігання даних;
- систему експлуатації;
- систему інформаційної безпеки.

Дві останні інтегровані між собою і об'єднані високопродуктивною локальною обчислювальною мережею [2].

Центри обробки даних, сьогодення, вирішують наступні завдання, які представлені на (рис. 1):

Головним аспектом дієвості будь-якого центру обробки даних є правильна організація структурної складової. Створення індивідуальних

центрів обробки даних, або заміна систем зберігання інформації вимагає професіонального підходу до вибору обладнання, установки систем та впровадження інноваційних рішень.

У випадку не правильної організації центру обробки даних або економії на одному з перерахованих вище пунктів може стати причиною серйозних наслідків. По-перше, це блокування доступу до інформації, а по-друге, це безповоротна втрата, що є у багатьох випадках неприпустимим фактором (банки, державні установи, таке інше).

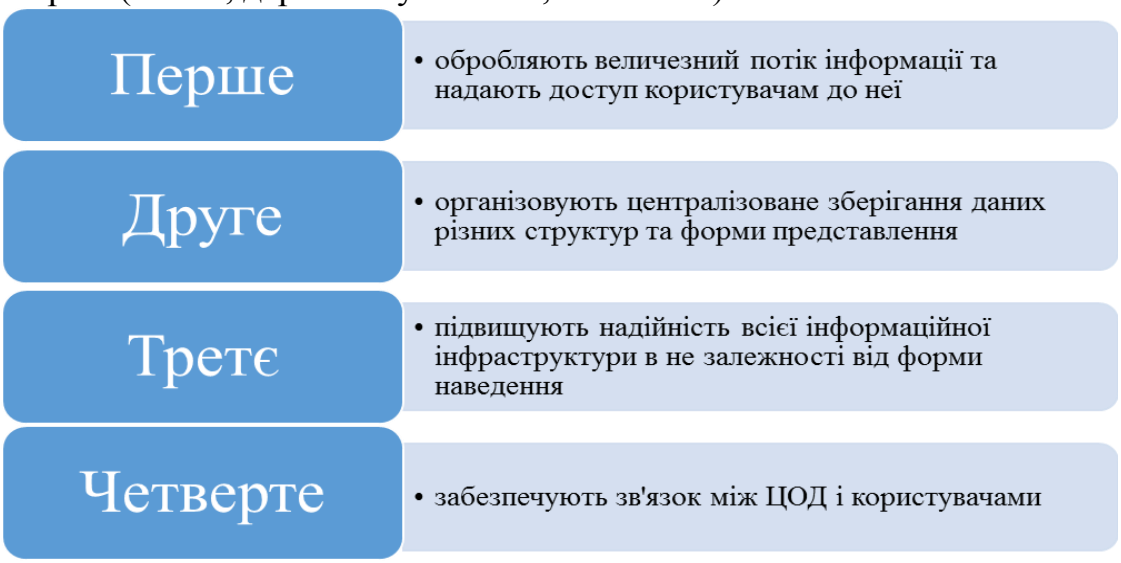


Рис. 1. Завдання центрів обробки даних

Головні стадії забезпечення безпеки ЦОД окреслені на (рис. 2).

Рис. 2. Головні стадії забезпечення безпеки ЦОД

Загалом сучасні ЦОД володіють потужною системою захисту інформації, до її видів варто віднести (на рис.3):

– програмно-апаратний захист, що передбачає використання програмного забезпечення ЦОД, комплексів програм, а також апаратних пристроїв, вбудованих у складі технічних засобів ЦОД;

– технічний захист, який ґрунтується на використанні технічних пристроїв, вузлів, блоків, елементів, систем, як у вигляді окремих засобів, так і вбудованих в процесі єдиного технологічного циклу створення засобів обробки інформації в ЦОД [3];

Організаційний захист, який ґрунтується на реалізації організаційних і організаційно-технічних заходів, що здійснюються для захисту інформації [4-5].

В якості окремого виду засобів захисту інформації виділяються крипто графічні засоби, що реалізуються у вигляді технічних, програмних і програмно апаратних засобів.

Проте, якщо не приділяти аспекту захисту особливу увагу, то не виключені розкрадання даних з архівів. Більш того, знаходження серверів в приміщенні, яке доступно для різних видів користувачів, може стати причиною витоку важливої інформації або виведення з ладу ЦОД. При розгалуженні IT-інфраструктури та розширення масштабів діяльності виникає необхідність у виділенні додаткових приміщень для серверів і збільшенні штату фахівців.

Рис. 3. Система захисту інформації ЦОД

Головними пріоритетними вимогами, вживаними до ЦОДам, нашого часу, виступають цілісність інформації, що зберігається, її доступність і конфіденційність [6-7]. Крім того, у рамках фінансових установ, порушення функціонування ЦОД може привести до безповоротних фінансових і політичних процесів, що, у багатьох випадках, є неприпустимим явищем.

Мета забезпечення кіберзахищенного інформаційного простору зводиться до мінімізації інформаційних втрат в ЦОД, за рахунок зменшення ризиків і зниження рівня можливих зовнішніх і внутрішніх дій.

По структурному складу ЦОД, до головних об'єктів, що вимагають якісного захисту, слід віднести: інформацію, яка зберігається і обробляється в системі; програмне забезпечення, встановлене у рамках ЦОД; елементи системи, устаткування центру.

Аналізуючи вищесказане, стає зрозумілою актуальність здійснення, що без побудови СУИБ (системи управління інформаційною безпекою) як усього підприємства, так і елементу (ЦОД) адекватного захисту не забезпечити, тобто без інформація, яка циркулює в системі, устаткування (елементи) та програмного забезпечення [8-9].

З математичної точки зору модель центру обробки даних є у вигляді на (рис. 4):

Рис. 4. Математична модель захисту ЦОД від кібератак

Визначення цілей проведеної роботи

Цілі проведеної роботи полягають у наведенні сукупності шляхів мінімізації втрат у центрах обробки даних, що, у разі їх застосування на реальному ЦОДі, призведе до підвищення надійності ЦОД в цілому, та окремих його частин зокрема, збереже надійність критично важливого обладнання, підвищить максимальний рівень можливого навантаження, мінімізує рівень як зовнішніх так і внутрішніх атак, скоротить витрати на утримання центру та у сукупності підніме авторитет ЦОДу у потенційних клієнтів.

Постановка задачі

У рамках дослідження теми мінімізації інформаційних втрат в центрах обробки даних (далі ЦОД) пропонується запропонувати основні шляхи мінімізації інформаційних втрат в ЦОД, для покращення надійності ЦОДу, мінімізації як внутрішніх так і зовнішніх атак, підвищення рівня можливого навантаження та ін.

Шляхи мінімізації втрат

Пріоритетним напрямком підвищення ефективності роботи центрів обробки даних, мінімізації інформаційних втрат, забезпечення надійності

систем зберігання інформації і зниження сукупних витрат є впровадження послуги колокації. Послуга колокації, полягає в тому, що провайдер розміщує обладнання клієнта на своїй території, підключає його до електрики, забезпечує обслуговування і підключення до каналів зв'язку з високою пропускною спроможністю [10-11].

Мета забезпечення кіберзахищеного інформаційного простору зводиться до мінімізації інформаційних втрат в ЦОД, за рахунок зменшення ризиків і зниження рівня можливих зовнішніх і внутрішніх дій.

По структурному складу ЦОД, до головних об'єктів, що вимагають якісного захисту, слід віднести: інформацію, яка зберігається і обробляється в системі; програмне забезпечення, встановлене у рамках ЦОД; елементи системи, устаткування центру [12-13]:

$$\text{ЦОД} = \left(\text{М}_{\text{бо}} \cup \text{М}_{\text{бд}} \cup \text{М}_{\text{бм}}\right), \quad (1)$$

де ЦОД – центр обробки даних;

$\text{М}_{\text{бо}}$ – безліч обчислювальних вузлів ЦОД;

$\text{М}_{\text{бд}}$ – безліч сховищ даних;

$\text{М}_{\text{бм}}$ – безліч комутаційних елементів мережі обміну і фізичних каналів передачі даних.

$$\text{РЗ} = \text{М}_{\text{вм}} \cup \text{М}_{\text{мэ}}, \text{М}_{\text{вк}}, \quad (2)$$

де РЗ – ресурсний запит;

$\text{М}_{\text{вм}}$ – безліч віртуальних машин;

$\text{М}_{\text{мэ}}$ – безліч елементів;

$\text{М}_{\text{вк}}$ – безліч віртуальних каналів передачі даних між віртуальними машинами і елементами запиту.

Призначення ресурсного запиту формується у вигляді:

$$\text{А: РЗ} \rightarrow \text{ЦОД} = \left\{\text{М}_{\text{бв}} \rightarrow \text{М}_{\text{бв}}, \text{М}_{\text{бе}} \rightarrow \text{М}_{\text{бд}}, \text{М}_{\text{бк}} \rightarrow \text{М}_{\text{бм}}\right\}. \quad (3)$$

При цьому, для актуальної роботи, необхідним є виконання наступних умов:

– кожен обчислювальний вузол зобов'язаний мати продуктивність і загальну сумарну пам'ять, яка відповідає сумарній складовій усіх віртуальних машин що відносяться до нього [14-15];

– кожен віртуальний канал може бути відображений на фізичний канал за умови, що загальна безліч віртуальних каналів відображених на фізичний канал, менше номінальної пропускної спроможності каналу передачі даних [16-17];

– кожен віртуальний канал може проходити через комутаційний елемент, за умови, що безліч віртуальних каналів, що проходять через комутаційний елемент, менше ніж загальна пропускна спроможність цього елементу(байт/с) [18-19];

– кожен елемент загального інформаційного простору може бути розміщений в сховищі даних, за умови, що кожен елемент, а також його

тип співпадають з типом сховища даних і загальна безліч усіх елементів, що зберігаються, не перевищує об'єму усієї пам'яті.

З метою мінімізації інформаційних втрат в ЦОД, за рахунок зменшення ризиків і зниження рівня можливих зовнішніх і внутрішніх дій, пропонується використати механізми оптимізації розміщення віртуальних машин на фізичних серверах, тобто застосувати принцип мінімального заповнення серверів [20-21]:

$$\min \sum_{i=1}^{n} y_i ; \tag{4}$$

$$y_i = \begin{cases} 1 - \text{на сервері є одна віртуальна машина} \\ 0 - \text{сервер не задіяний} \end{cases}.$$

Практична реалізація цього підходу дозволяє мінімізувати рівень як зовнішніх так і внутрішніх дій, а також понизити витрати на утримування загального парку серверів ЦОД, у разі ідентичних технічних характеристик [20-21].

Умови проведених в ході дослідження експериментів і їх результати

Інформаційна безпека ЦОД базується на сукупності спеціальних технологій побудови, до яких обов’язково входять програмні та програмно-апаратні засоби захисту інформації. За для реалізації дієвої системи захисту застосовуються технології аутентифікації, кластеризації, розмежування доступу, віртуалізації таке інше. Стосовно програмних засобів захисту, обов’язковими є [23-24]:

- операційна система;
- гіпервізор;
- гостьова операційна система;
- засоби захисту інформації від несанкціонованого доступу.

Висновки

Актуальним питанням сучасного інформаційного простору центрів обробки даних(далі ЦОД) виступає побудова ефективного захисту від кібератак.

ЦОД - це централізована комплексна система, яка є відмовостійкою і забезпечує якісне обслуговування бізнес-процесів що проходять в її рамках, з високим рівнем послуг, що надаються. Основною фундаментальною умовою надійної працездатності сучасного ЦОД є його безпека. Чим вище рівень безпеки усіх інформаційних ресурсів, що зберігаються в системі і знаходяться в обороті, що надається, тим вище міра гарантованого забезпечення потрібного якість сервісу.

Головними пріоритетними вимогами, вживаними до ЦОДам, нашого часу, виступають цілісність інформації, що зберігається, її доступність і

конфіденційність. У рамках фінансових установ, порушення функціонування ЦОД може привести до безповоротних фінансових і політичних процесів, що у багатьох випадках, є неприпустимим явищем.

У сукупності, використання запропонованих шляхів мінімізації втрат в центрах обробки даних призведе до підвищення надійності ЦОД в цілому, та окремих його частин зокрема, збереже надійність критично важливого обладнання, підвищить максимальний рівень можливого навантаження, мінімізує рівень як зовнішніх так і внутрішніх атак, скоротить витрати на утримання центру та у сукупності підніме авторитет ЦОДу у потенційних клієнтів.

Високонадійні і сучасні інженерні рішення дозволяють створювати ЦОД для будь-якої компанії з урахуванням специфіки її бізнес-процесів, адаптувати до конкретних умов і вибудовувати системи зберігання і обробки даних в розрахунку на перспективу і довгострокову експлуатацію, що позитивно впливає на сучасні тенденції розвитку та адаптації до інформаційної глобалізації суспільства.

Наукова література дає визначення ЦОДу, як окремому приміщенню, яке призначене для розміщення обладнання для обробки і зберігання даних, що забезпечує підключення до швидких каналів зв'язку.

До головних об'єктів ЦОДу, які потребують якісного захисту, слід віднести: інформацію, яка зберігається і обробляється в системі; програмне забезпечення, встановлене в рамках ЦОД; елементи системи, обладнання центру.

Практичне значення представленої моделі захисту ЦОД від кібератак важко переоцінити, в силу яскраво вираженої актуалізації кіберпростору сучасного інформаційного суспільства. Наслідуючи усі описані напрями(механізмам) захисту передбачається зниження рівня можливих кібератак, а також підвищення рівня захисту цента обробки даних.

Література

1. Теленик С.Ф. Генетичні алгоритми вирішення задач управління ресурсами і навантаженням центрів оброблення даних / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, С.А. Андросов // Автоматика. Автоматизація. Електротехнічні комплекси та системи. – 2010. – №1 (25). – С. 106 – 120.

2. Мельников Д. А. Информационная безопасность открытых систем. – Москва: ФЛИНТА, 2012. – С. 448.

3. Оценка защищенности информационных процессов в территориальных ОВД: модели исследования: монография / под ред. С.В. Скрыля. – Воронеж: Воронежский институт МВД России, 2010. – 217 с.

4. Литвинов Д.В., Скрыль С.В., Тямкин А.В. Исследование механизмов противодействия компьютерным преступлениям: организационно-правовые и криминалистические аспекты: монография – Воронеж: Воронежский институт МВД России, 2009. – 218 с.

5. Matusitz, Jonathan The Role of Intercultural Communication in Cyberterrorism // Journal of Human Behavior in the Social Environment. – 2014. – Vol. 24. – P. 775 – 790.

6. Вдовин П.М., Костенко В.А. Алгоритм распределения ресурсов в центрах обработки данных с раздельными планировщиками для различных типов ресурсов // Известия РАН. Теория и системы управления. – 2014. – № 6. – С. 56 – 68.
7. Шестак Я. В. Модель построения киберзащищенного информационного пространства ЦОД: математический аспект / Я. В. Шестак, Д. О. Огбу, А. Г. Оксиюк // Кібербезпека в Україні: правові та організаційні питання: Науково-практична конференція, Одеса, 21 жовтня 2016 р.: тези доповідей. – Одеса: ОДУВС, 2016. – С. 159 – 160
8. Герасимов Б.М., Оксіюк О.Г., Шворов С.А. Проектування та застосування експертно-навчальних систем: Монографія – Київ: Європейський институт, 2008. – 218 с.
9. T. A. Longstaff, Clyde Chittister, Rich Pethia, Yacov Y. Haimes, "Are We Forgetting the Risks of Information Technology", IEEE Computer, pp 43 – 51, December, 2000.
10. N. Ye, C. Hosmer, J. Giordano, J. Feldman, "Critical Information Infrastructure Protection through Process Modeling and Model-based Information Fusion", Proceedings of the Information Survivability Workshop, 1998.
11. F. Cohen "Simulating Cyber Attacks, Defenses, and Consequences", IEEE Symposium on Security and Privacy Special 20th Anniversary Program, Berkeley, CA, May, 1999.
12. Yebin Zhang, and Connie M. Borror" Robustness of the Markov-Chain Model for Cyber-Attack Detection" IEEE Trans. Reliability, vol. 53, no. 1, March 2004.
13. Rex B. Hughes, "NATO and Cyber Defence: Mission Accomplished?" Netherlands Atlantic Association, Amsterdam, Atlantisch Perspectief 8 (2008).
14. "The EU Internal Security Strategy in Action: Five steps toward a more secure Europe," European Commission, November 22, 2010.
15. T. A. Wadlow, The Process of Network Security, Addison-Wesley, Nov 30, 2005.
16. W. Kaufman, Running LINUX, Oreilly, 1999.
17. S. Northcutt, Network Intrusion Detection An Analyst's Handbook, New Riders, 1999.
18. Quoted in Jim Garamone, "Lynn: NATO Must Get Ahead of Cyber Threat," American Forces Press Service, January 25, 2011.
19. Павлов О.А. Інформаційні технології та алгоритмізація в управлінні / О.А.Павлов, С.Ф.Теленик. – К.: Техніка, 2002. – 344 с.
20. Теленик С.Ф. Определение распространения влияния неисправностей в сети доступа на качество предоставляемых сервисов / С.Ф. Теленик, А.И. Ролик, М.М. Букасов, М.В. Ясочка// Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – К.: Век+, – 2009. – №50. – С. 164 – 173.
21. Теленик С.Ф. Моделі і методи розподілу ресурсів в системах з серверною віртуалізацією / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, О.А. Косован, О.І. Кобець // Зб. наук. праць ВІТІ НТУУ «КПІ». – К., 2009. – №3. – С.100 – 109.
22. Теленик С.Ф. Моделі управління розподілом обмежених ресурсів в інформаційно-телекомунікаційній мережі АСУ / С.Ф. Теленик, О.І. Ролік, М.М. Букасов // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: «Экотех». – 2006. – №44. – С. 234 – 239.
23. Теленик С.Ф. Моделі управління віртуальними машинами при серверній віртуалізації / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, А.Ю. Лабунський // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: «ВЕК+», – 2009. – № 51. – С. 150 – 155.
24. K.S. Leung, "A New Model of Simulated Evolutionary Computation – Convergence Analysis and Specifications"/ K.S. Leung, Q.H. Duan, Z.B. Xu, and C. K. Wong// IEEE Transactions on Evolutionary Computation, Vol. 5, No. 1. – February, 2001 – P. 3 – 16.

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ОБНАРУЖЕНИЯ СИГНАЛОВ НА ФОНЕ НЕГАУССОВСКИХ ПОМЕХ ПО МОМЕНТНОМУ КРИТЕРИЮ ТИПА НЕЙМАНА-ПИРСОНА

Палагин В.В., Лелеко С.А., Зорин А.С.

Введение

Передовые системы диагностики, наблюдения, мониторинга и управления характеризуются высокими требованиями к качеству обработки данных, повышенной сложностью и функциональностью. Основываясь на вышеуказанных факторах, возникает необходимость в разработке эффективных систем обнаружения и распознавания сигналов. В общем случае этим системам приходится иметь дело со стохастическими процессами. Традиционно, проектирование таких систем основано на классических методах теории проверки статистических гипотез. Как правило, такие методы не имеют ограничений на использование типа функции плотности распределения случайных процессов [1,2]. Однако, на практике широко применяется гауссовская функция плотности распределения случайных процессов, которая во многих случаях она не описывает реальные процессы с нужной точностью и является удобной математической идеализацией реального стохастического процесса [3,4]. Классические методы теории статистической проверки гипотез характеризуются значительными ограничениями обработки негауссовских процессов и связаны как со сложностью их алгоритмической реализации, так и с увеличением необходимых вычислительных ресурсов [5,6].

Известно, что свойства решающих функций могут быть описаны с использованием других математических характеристик, таких как математическое ожидание и дисперсия решающих правил (РП). Например, критерий отклонения был развит в классе линейно-квадратичных (L-Q) систем [7] и дальнейшее развитие этого направления показано в [8]. Однако критерий отклонения и его модификации слабо связаны с классическими критериями качества и не описывают всех свойств решающих функций.

В работе представлен другой подход к описанию статистических свойств негауссовских процессов, основанных на использовании статистик высших порядков (Higher-Order Statistics - HOS) [9,10]. В качестве частичного описания случайных процессов используются моменты и кумулянты. Такие характеристики позволяют с достаточной точностью описать статистические свойства негауссовских процессов [11-14] и повысить точность обработки негауссовских сигналов по сравнению с традиционными методами, а также уменьшить сложность алгоритмов обнаружения и распознавания сигналов на фоне негауссовских помех [13-16].

Основной целью работы является синтез и анализ методов и алгоритмов обнаружения сигналов на фоне негауссовских помех на основе моментно-кумулянтного описания случайных величин, полиномиальных РП, оптимальных по моментному критерию качества типа Неймана-Пирсона. Такой подход дает возможность создать эффективные компьютерные средства для функционирования систем приема и обработки данных.

1. Адаптация моментного критерия принятия решений

Пусть случайные сигналы $\xi(t)$ наблюдаются в интервале времени $(0,T)$. Необходимо синтезировать алгоритмы обработки сигналов при анализе входного стохастического процесса $\xi(t)$ на основе принятия решения: принимается сигнал $s(t)$ (реализация гипотезы H_1) или сигнала нет (реализация гипотезы H_0), где, $\xi(t) = s(t) + \eta(t)$, $\eta(t)$ - негауссовский стационарный случайный процесс, описанный конечной последовательностью моментов и кумулянтов. Предположим, что такая последовательность, при выполнении гипотезы H_1, выглядит как - m_i и для гипотезы H_0 - u_i соответственно, где u_i, m_i - моменты порядка i при реализации гипотез H_0 и H_1 соответственно.

Пусть из стохастического процесса $\xi(t)$ производится выборка $\mathbf{X} = \{x_1, x_2, ... x_n\}$, тогда их дискретные значения для соответствующих гипотез H_i $(i = 0,1)$ будут выглядеть как:

$$H_1: \xi_v = s_v(\alpha_k) + \eta_v(\gamma_k),$$
$$H_0: \xi_v = \eta_v(\gamma_k),\ v = \overline{1,n},$$

где $s_v(\alpha_k)$ – сигнал с известными параметрами α_k, $\eta_v(\gamma_k)$ – негауссовская случайная величина с известными параметрами в виде дисперсии χ_2 и кумулянтов γ_k $k = \overline{1,\mu}$.

Согласно классическому подходу оптимальный байесовский алгоритм обнаружения сигнала определяется из условия минимума среднего риска [2,3]. Минимальная достаточная статистика для простой проверки гипотез определяется как отношение правдоподобия и может быть найдена в виде:

$$\Lambda(\mathbf{X}) = P(\mathbf{X} \mid H_1) / P(\mathbf{X} \mid H_0). \qquad (1)$$

Решение таких задач в основном выполняется в предположении гауссовской плотности распределения вероятности случайных величин. В общем случае для произвольных распределений проблематично определить плотность распределения вероятности и найти решения в форме (1). Поэтому можно использовать другой подход, когда отношение правдоподобия представляется в виде полиномиальной функции,

оптимальные коэффициенты которой определяются согласно выбранного критерия качества.

Пусть отношение правдоподобия является непрерывной функцией. Тогда согласно теореме Вейерштрасса о приближении непрерывных функций многочленами логарифм отношения правдоподобия можно представить в виде ряда:

$$\ln \Lambda(\mathbf{X}) = \sum_{i=1}^{\infty}\sum_{v=1}^{n} k_{iv} x_v^i + k_0 .$$

В этом случае отношение правдоподобия для независимых выборочных значений будет представлено как стохастическое полиномиальное РП степени s:

$$\ln \Lambda(\mathbf{X}) = \sum_{v=1}^{n}\sum_{i=1}^{s} k_{iv} x_v^i + k_0 \underset{H_0}{\overset{H_1}{\gtrless}} 0. \qquad (2)$$

Неизвестные коэффициенты k_{iv} и k_0 (2) могут быть найдены из минимума известного вероятностного критерия качества (Байеса, Неймана-Пирсона и т.п.), но в общем реализовать это невозможно. Поэтому предлагается новый моментный критерий качества проверки статистических гипотез [14], основанный на использовании HOS.

Предположим, что существует решающая функция

$$f(\mathbf{X}) = \gamma(\mathbf{X}) + k_0 \underset{H_0}{\overset{H_1}{\gtrless}} 0, \qquad (3)$$

где $\gamma(\mathbf{X})$ – функция от выборочных значений $\mathbf{X}$, k_0 выбранная таким образом что:

$$M_0 = E[f(\mathbf{X})/H_0] = \int_{-\infty}^{\infty} f(\mathbf{X})\, p(\mathbf{X}/H_0)\Pi dx < 0,$$

$$M_1 = E[f(\mathbf{X})/H_1] = \int_{-\infty}^{\infty} f(\mathbf{X})\, p(\mathbf{X}/H_1)\Pi dx \geq 0.$$

В соответствии с неравенством Чебышева, вероятности ошибок первого и второго рода (3) определяются как:

$$\alpha = P[f(\mathbf{X}) \geq 0/H_0] \leq G_0 / M_0^2 = \alpha_0,$$

$$\beta = P[f(\mathbf{X}) < 0/H_1] \leq G_1 / M_1^2 = \beta_0,$$

где $G_i(\gamma) = \int_{-\infty}^{\infty} [f(\mathbf{X}) - M_i]^2 p(\mathbf{X}/H_i) \Pi dx$ – дисперсия решающей функции $\gamma(\mathbf{X})$ при гипотезах H_i, $i = 0,1$.

Тогда критерий суммы вероятности ошибок (3) можно записать в виде следующего неравенства

$$F_1(\alpha, \beta) = \alpha + \beta \le \alpha_0 + \beta_0 = \frac{G_0}{M_0^2} + \frac{G_1}{M_1^2} = \Phi(G, M).$$

Предположим, что для M_0 и M_1 коэффициент k_0 определяется как

$$k_0 = -0.5(E_0 + E_1), \quad (4)$$

где $E_i(\gamma) = E[\gamma(\mathbf{X})|H_i]$ – математическое ожидание решающей функции $\gamma(\mathbf{X})$ при гипотезах H_i, $i = 0,1$.

Тогда, функция $\Phi(G, M)$ для коэффициента k_0 (4) определяется как $\Phi(G, M) = 4\,Ku\,1(G, E)$, где

$$Ku(E, G) = \frac{G_0[\gamma] + G_1[\gamma]}{(E_1[\gamma] - E_0[\gamma])^2}. \quad (5)$$

Функционал $Ku(E, G)$ является критерием качества принятия решений (3). Этот критерий называется “Моментным критерием качества верхних границ вероятностей ошибок” или кратко “Ku критерием” [14].

Адаптируем моментный критерий качества (5) под вероятностный критерий типа Неймана-Пирсона. Для того, чтобы по аналогии с вероятностным критерием Неймана-Пирсона иметь возможность зафиксировать одну из вероятностей ошибок, выберем порог РП (2) k_0 из условия

$$k_0 = -(E_0 C + (1 - C)E_1), \quad (6)$$

где C – нормирующий коэффициент, который принимает значение на интервале $C \in (0,1)$ и дает возможность изменять порог k_0.

Таким образом, видно, что при $C = 0.5$, k_0 принимает значение (4) для критерия (5). Легко показать, что при изменении порога k_0 РП происходит и изменение величины верхних границ вероятностей ошибок.

Согласно синтезу критерия Неймана-Пирсона [1,4], предлагается наложить ограничения на вероятность ошибки первого рода α, когда ее значение должно быть не более некоторого заданного числового значения ρ. Тогда, учитывая выражение (6), вероятность ошибки первого рода α запишется в виде:

$$\frac{\frac{G_0}{(1 - C)^2}}{[E_1 - E_0]^2} \le \rho, \quad (7)$$

а оптимальное РП будет определяться из условия минимума вероятности ошибки второго рода β с учетом выражения (6).

Тогда, оптимальное РП вида (3) находится из условия минимума функционала:

$$KuP(E,G) = \frac{\frac{G_0}{(1-C)^2} + \frac{G_1}{C^2}}{[E_1 - E_0]^2}, \qquad (8)$$

где E_i, G_i – математическое ожидание и дисперсия РП при гипотезах H_i, $i = 0,1$ соответственно.

В качестве РП будем использовать стохастические полиномы обобщенного вида (2), оптимальные коэффициенты которого будут находиться из минимума (8).

Определение 1. Возьмем функционал $KuP(E,G)$ за критерий качества выбора РП вида (2) и будем считать наилучшим то правило, которое при k_0, равном (6), минимизирует правую часть (8) при заданном значении вероятности ошибки α. Данный критерий будем называть моментным критерием качества проверки статистических гипотез типа Неймана-Пирсона или кратко критерием $KuP(E,G)$.

Коэффициенты k_i, минимизирующие правую часть (8), находятся из решения системы уравнений:

$$\sum_{j=1}^{s} k_j \left[\frac{F_{i,j}(H_0)}{(1-C)^2} + \frac{F_{i,j}(H_1)}{C^2} \right] = m_i - u, i = \overline{1,s}, \qquad (9)$$

где $F(H_1)_{i,j} = m_{(i+j)} - m_i m_j$, $F(H_0)_{i,j} = u_{(i+j)} - u_i u_j$ – корреляты размерностью (i,j) наблюдаемой случайной величины ξ при гипотезе H_1 и альтернативе H_0 соответственно.

Тогда, согласно (7), неизвестный нормирующий коэффициент C находится из условия заданной вероятности ошибки первого рода:

$$\alpha = \frac{G_0}{(1-C)^2 [E_1 - E_0]^2}, \qquad (10)$$

а минимизированная вероятность ошибки второго рода РП (2) примет вид

$$\beta = \frac{G_1}{C^2 [E_1 - E_0]^2}. \qquad (11)$$

При использовании полиномиального РП общего вида (2) со степенными преобразованиями выборочных значений математические ожидания E_i и дисперсии G_i РП при гипотезе H_i, $i = 0,1$ запишутся в виде

$$E_0 = n\sum_{i=1}^{s} k_i u_i, \; E_1 = n\sum_{i=1}^{s} k_i m_i,$$

$$G_0 = n\sum_{i=1}^{s}\sum_{j=1}^{s} k_i k_j F_{i,j}(H_0),\ G_1 = n\sum_{i=1}^{s}\sum_{j=1}^{s} k_i k_j F_{i,j}(H_1). \qquad (12)$$

Критерием качества в теории оценки параметров является дисперсия оценки параметров случайных величин. Показано, что минимальная дисперсия обратно пропорциональна информации Фишера [1,17]. Показано, что математическое ожидание и дисперсия полиномиального стохастического РП (2) могут быть представлены в виде информационного числа Кульбака-Лейблера с использованием функции плотности распределения для гипотезы и альтернативы. В этом случае для моментного критерия качества принятия решений минимальное значение также можно определить с помощью функции плотности распределения. Поэтому, как и в теории оценки параметров, целесообразно ввести понятие количества извлекаемой информации из выборки объема n о различии гипотез H_1 и H_0.

Назовем величину, обратную критерию $KuP(E,G)$, как количество извлекаемой информации о различии гипотез H_0, H_1 и обозначим в виде:

$$KuP(E,G) = I_{KuPsn}^{-1}. \qquad (13)$$

Показано, что I_{KuPsn} также определяется как:

$$I_{KuPsn} = \frac{1}{KuP(E,G)} = \frac{G_0}{(1-C)^2} + \frac{G_1}{C^2} = E_1 - E_0.$$

Для классического вероятностного критерия Неймана-Пирсона справедливо соотношение [18]:

$$k_0 = E_0 + \chi_2 x_\alpha / \sqrt{n},$$

где x_α - процентная точка гауссовкого распределения (квантиль), определяющая заданную величину вероятности первого рода. Величины k_0, E_0 имеют такой же физический смысл, как и для моментного критерия Неймана-Пирсона, а именно порога принятия решения и математического ожидания РП. Порог k_0 выбирается таким образом, что бы определить зависимость от вероятности ошибки первого рода α и количества выборочных значений.

При этом

$$x_{1-\beta} = x_\alpha - \frac{E_1 - E_0}{\chi_2}\sqrt{n}. \qquad (14)$$

Из соотношения (14) видно, что для вероятностного критерия типа Неймана-Пирсона существует соотношение между вероятностями ошибок α и β. Из (14) следует что при $n \to \infty$ значение $\beta \to 0$. Другими словами, выражение (14) при заданных α и β определяет минимально возможную величину $(E_1 - E_0)/\chi_2\sqrt{n}$. Тогда, при заданных α, β и отношении сигнал

шум $q = a^2/\chi_2$, существует определенный минимальный объем выборки для проверки статистической гипотезы. Минимальный объем выборки находиться из соотношения:

$$n \geq \frac{\chi_2 (x_{1-\beta} - x_\alpha)^2}{[E_1 - E_0]^2}.$$

Таким образом, разработанный новый метод обнаружения сигналов на фоне негауссовских помех основывается на новых моментно-кумулянтных моделях случайных процессов и адаптированного моментного критерия качества проверки статистических гипотез. Данный метод будет использован для синтеза и анализа нелинейных полиномиальных алгоритмов обнаружения сигналов на фоне негауссовских помех.

2. Синтез полиномиальных алгоритмов обнаружения сигналов на фоне негауссовских помех

Рассмотрим эффективность предлагаемого метода, используя пример обнаружения постоянного сигнала на фоне негауссовских помех. Пусть случайный сигнал $\xi(t)$ наблюдается на интервале времени $[0, T]$ и состоит из полностью известного сигнала a [19] и помехи $\eta(t)$:

$$\xi(t) = a + \eta(t) \ ,$$

где $\eta(t)$ - негауссовская помеха с нулевым математическим ожиданием, дисперсией χ_2 и описывается последовательностью моментов и кумулянтов.

Пусть из сигнала $\xi(t)$ производится выборка $\mathbf{X} = \{x_1, x_2, \ldots, x_n\}$, тогда их дискретные значения для гипотез H_i $(i = 0,1)$ примут вид:

$$H_1: x_v = a + \eta_v(\gamma_k), \ H_0: x_v = \eta_v(\gamma_k), \ v = \overline{1,n}.$$

В работе рассмотрено обнаружение постоянного сигнала с заданной амплитудой a на фоне асимметрично-эксцессной помехи 2-го типа 1-го вида, для которой характерно отличие от нуля кумулянтов третьего и четвертого порядка, а остальные кумулянтные коэффициенты высших порядков имеют произвольные значения. Обнаружение постоянных сигналов в виде прямоугольных видеоимпульсов характерно для таких случаев, как последетекторная обработка в радиолокационных станциях, обнаружение сигнала синхронизации в системах связи, лазерное измерения расстояния и т.д. [20, 21].

При реализации гипотезы H_0 начальные моменты для асимметрично-эксцессного случайного процесса 2-го типа 1-го вида [13], когда учитываются коэффициенты асимметрии и эксцесса, имеют вид:

$$u_1 = 0, \ u_2 = \chi_2, \ u_3 = \gamma_3 \chi_2^{3/2}, \ u_4 = (3 + \gamma_4)\chi_2^2,$$

где γ_3, γ_4 - кумулянтные коэффициенты асимметрии и эксцесса соответственно, которые учитывают негауссовское распределение случайной величины.

При реализации гипотезы H_1 начальные моменты для этого случайного процесса имеют вид:

$$m_1 = a,\ m_2 = a^2 + \chi_2,\ m_3 = a^3 + 3a\chi_2 + \gamma_3\chi_2^{3/2},$$
$$m_4 = a^4 + 6a^2\chi_2 + 4a\gamma_3\chi_2^{3/2} + (3+\gamma_4)\chi_2^2 .$$

Если выборочные значения одинаково распределены, то в общем виде при степени полинома $S = 1$ линейное РП запишется в виде:

$$\Lambda(\mathbf{X})_{1n} = k_0 + k_1 \sum_{v=1}^{n} x_v \underset{H_0}{\overset{H_1}{\gtrless}} 0, \qquad (15)$$

где из уравнения (9) легко получить значение коэффициента k_1:

$$k_1 = \frac{q^{0.5}}{\chi_2^{0.5}\left(\frac{1}{C^2} + \frac{1}{(1-C)^2}\right)},$$

где $q = \frac{a^2}{\chi_2}$ – отношение сигнал/шум по мощности.

Для получения порога k_0 в РП (14) воспользуемся выражением (6), которое примет вид:

$$k_0 = \frac{nq(1-C)}{\left(\frac{1}{C^2} + \frac{1}{(1-C)^2}\right)}.$$

Таким образом, линейное РП (15) при степени полинома $S = 1$ примет вид:

$$\Lambda(\mathbf{X})_{1n} = \frac{1}{n}\sum_{v=1}^{n} x_v - (1-C)a \underset{H_0}{\overset{H_1}{\gtrless}} 0. \qquad (16)$$

Отметим, что при С=0,5 получаем РП, которое полностью совпадает с линейным РП, полученным как по моментному критерию минимума верхней границы вероятностей ошибок [13, 16], так и по вероятностному критерию идеального наблюдателя в предположении гауссовской помехи [1, 18].

Нормирующий коэффициент С находится из условия заданной вероятности ошибки (10) и после несложных преобразований получим

$$C = 1 - \frac{1}{\sqrt{nq\alpha}} \ .$$

Минимизированная вероятность ошибки второго рода РП, согласно выражению (11) при $S = 1$, принимает вид

$$\beta_{1n} = \frac{1}{C^2 nq} \ .$$

Из выражений вероятностей ошибок первого и второго рода РП видно, что при возрастании одной вероятности ошибок другая уменьшается и наоборот, что совпадает с теоретической трактовкой их взаимодействия.

Для линейного РП количество извлекаемой информации о различии гипотез запишется:

$$I_{1n} = n \frac{q}{\left(\frac{1}{C^2} + \frac{1}{(1-C)^2} \right)}$$

и согласно (13), является обратной величиной критерию качества (8) для синтезированного линейного РП (16).

Отметим, что линейное РП (16) не учитывает негауссовского распределения помехи, так как в качестве априорного описания случайного процесса используются начальные моменты только первого и второго порядка, описывающие математические ожидание и дисперсию случайных величин. Поэтому увеличим степень полинома до $S = 2$ и используем для описания случайного процесса при гипотезе и альтернативе начальные моменты до 4 порядка.

Для построения нелинейного РП при степени $S = 2$ постановка задачи будет совпадать с постановкой задачи для случая $S = 1$. Тогда, используя уравнение (2), получим следующее РП

$$\Lambda(\mathbf{X})_{2n} = k_1 \sum_{v=1}^{n} x_v + k_2 \sum_{v=1}^{n} x_v^2 - k_0 \underset{H_0}{\overset{H_1}{\gtrless}} 0. \qquad (17)$$

Неизвестные коэффициенты РП (17) находятся из решения системы алгебраических уравнений (9), а порог РП определяется согласно (6) и имеет вид:

$$k_0 = n[Ck_2 + (1-C)\left(k_1 q^{0.5} + k_2 (q+1)\right)].$$

Количество извлекаемой информации о различии гипотез, в соответствии с (13), примет вид:

$$I_{2n} = n\left(k_1 q^{0.5} + k_2 q\right).$$

Неизвестная нормирующая величина C находится из условия заданной вероятности ошибки (10) и имеет вид:

$$\alpha_{2n} = \frac{k_1^2 + 2k_1k_2\gamma_3 + k_2^2(\gamma_4 + 2)}{n(1-C)^2\left[k_1q^{0.5} + k_2q\right]^2}.$$

Тогда минимизированная вероятность ошибки второго рода, согласно (11), примет вид:

$$\beta_{2n} = \frac{A + B}{nC^2\left[k_1q^{0.5} + k_2q\right]^2},$$

где:

$$A = k_1^2 + 2k_1k_2\left(\gamma_3 + 2q^{0.5}\right),$$

$$B = k_2^2\left(4q^{0.5}\gamma_3 + 4q^{0.5} + \gamma_4 + 2\right).$$

Используя моментно-кумулянтные модели случайных величин и новый метод синтеза полиномиальных РП, основанный на использовании нового адаптированного моментного критерия качества типа Неймана-Пирсона, построены и проанализированы нелинейные РП при различных степенях полинома s.

Проведено исследование свойств синтезированных полиномиальных РП. На рис. 1, 2 приведены графики зависимости вероятности правильного обнаружения $(1-\beta)$ нелинейных РП β_{sn} $(S=2,3)$ от соотношения сигнал/шум q при заданной вероятности ошибки первого рода $\alpha = 10^{-3}$, объеме выборки n=200 и разных значениях коэффициентов асимметрии γ_3 и эксцесса γ_4 негауссовского процесса.

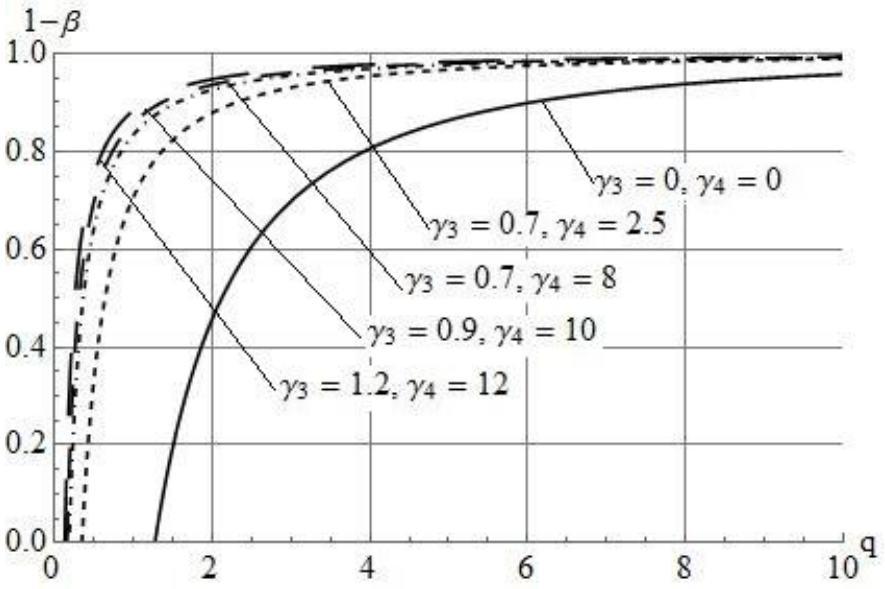


Рис. 1. График зависимости вероятности правильного обнаружения от соотношения сигнал/шум q для РП при степени полинома $S = 2$.

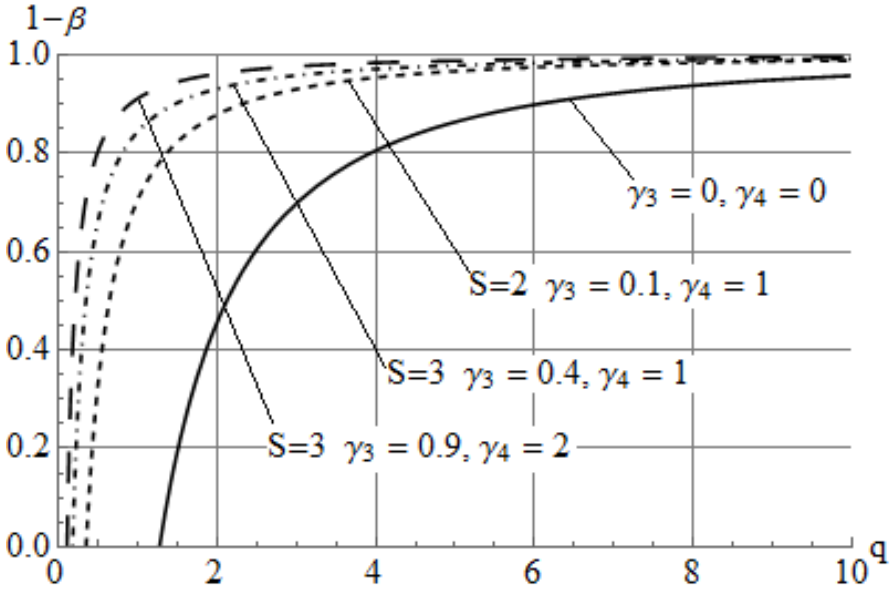


Рис. 2. График зависимости вероятности правильного обнаружения от соотношения сигнал/шум q для РП при степени полинома $S=3$.

Кривая для нулевых значений коэффициентов асимметрии γ_3 и эксцесса γ_4 (сплошная линия, рис. 1) соответствует характеристике правильного обнаружения линейного РП (16), когда $S=1$. На графиках видно, что учет коэффициентов асимметрии γ_3 и эксцесса γ_4 приводит к увеличению вероятности правильного обнаружения. Если сравнивать графики при q=2 то видно, что вероятность правильного обнаружение возрастает более чем на 40% при учете характеристик негауссовского распределения случайного процесса. Из графиков на рис. 2 видно, что при возрастании степени полинома РП до $S=3$ также происходит увеличение вероятности правильного обнаружения по сравнению с нелинейной обработкой РП при степени полинома $S=2$. Так, при степени полинома РП $S=3$ выигрыш в увеличении вероятности правильного обнаружения составляет 5% по сравнению с нелинейным РП при степени полинома $S=2$, значении отношения сигнал/шум $q=1$ и одинаковых параметрах γ_3 и γ_4.

На рис. 3 представлены ROC характеристики для синтезированных обнаружителей сигналов. Из рисунков видно, что учет негауссовских параметров случайного процесса и увеличение степени полинома РП ведут к улучшению ROC характеристик обнаружителей сигналов. Кривая classic получена для РП, оптимального по классическому вероятностному критерию Неймана-Пирсона при использовании гауссовской плотности распределения случайного процесса. Кривая $S=1$ соответствует линейному РП вида (15). Как видно, эти кривые имеют общую точку пересечения, которая соответствует значению коэффициента C=0,5. При

увеличении степени полинома (S=2,3) учитываются такие параметры негауссовского распределения, как начальные моменты третьего и выше порядков в виде коэффициентов асимметрии и эксцесса ($\gamma_3 = 0.9, \gamma_4 = 4$), что в целом увеличивает качество обнаружения сигналов при негауссовых помехах.

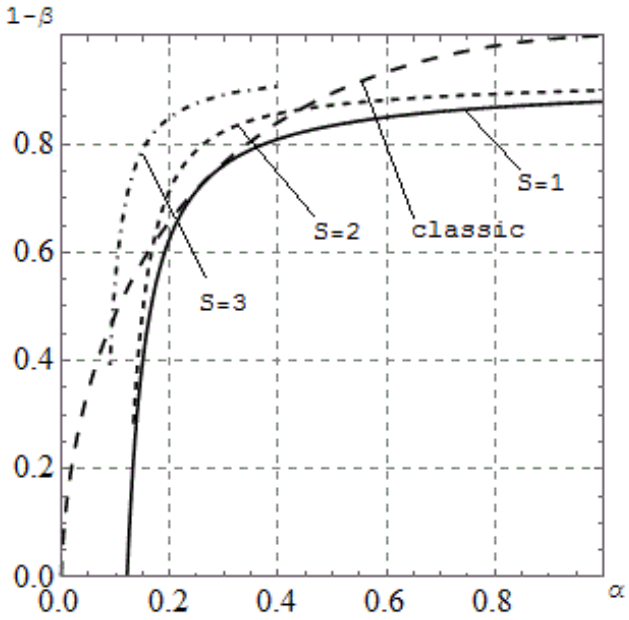


Рис. 3. Сравнение ROC характеристик для линейного и нелинейных РП при n=30, q=0.5.

На основе полученных РП построена обобщенная структурная схема полиномиальных обнаружителей постоянного сигнала на фоне негауссовских помех (рис. 4). Структурная схема состоит из 6 параллельных степенных блоков обработки сигнала. Количество степенных блоков определяется степенью полинома s РП. В степенных блоках обработки происходит суммирование и умножение выборочных значений x_v на соответствующие коэффициенты k_i, значение которых определяется из минимума моментного критерия качества типа Неймана-Пирсона (6). Сумма перемноженных выборочных значений x_v сравнивается со значением порогового коэффициента k_0 и система принимает решение о наличии на входе устройства только помехи (реализуется гипотеза H_0) или аддитивной смеси сигнала и помехи (реализуется гипотеза H_1).

При проведении имитационного моделирования были подтверждены полученные теоретические результаты. На рис. 5 сопоставлены теоретические (сплошная линия) и экспериментальные (точки) результаты

моделирования отношения моментных критериев качества KuP_1/KuP_s (6) при реализации РП степени полинома $S = 1,2$.

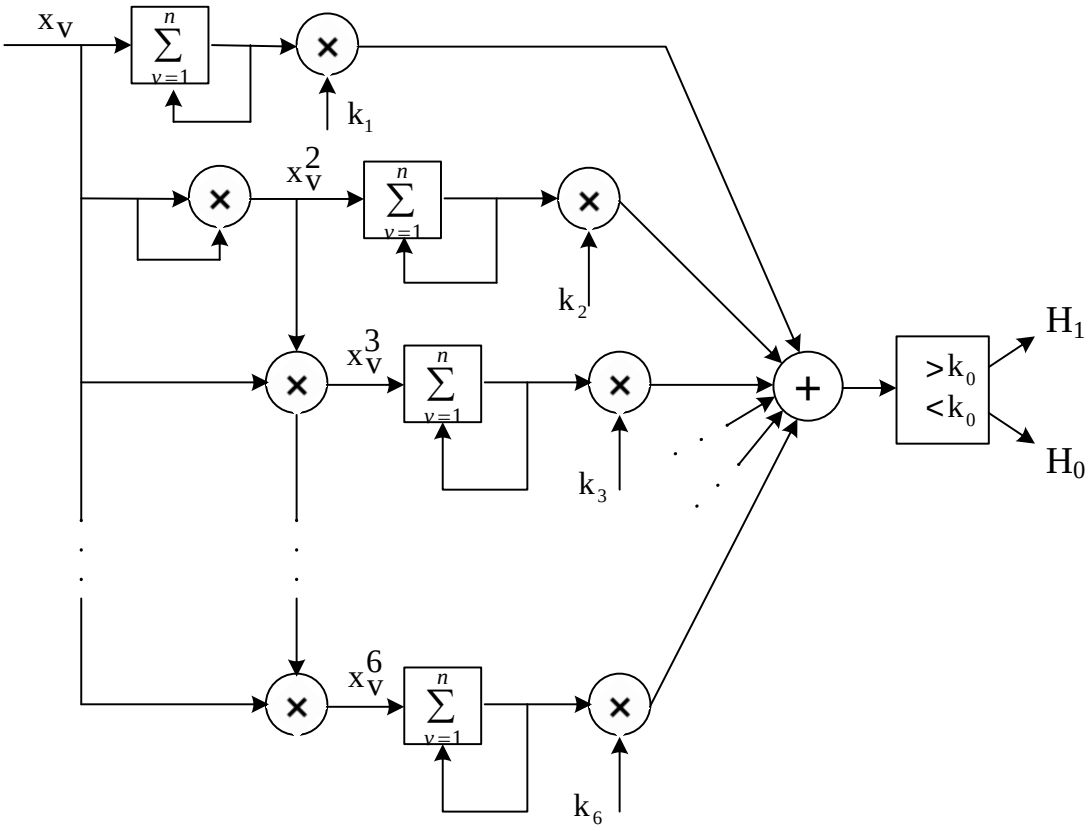


Рис. 4. Структурная схема полиномиального обнаружителя полностью известного сигнала на фоне помех

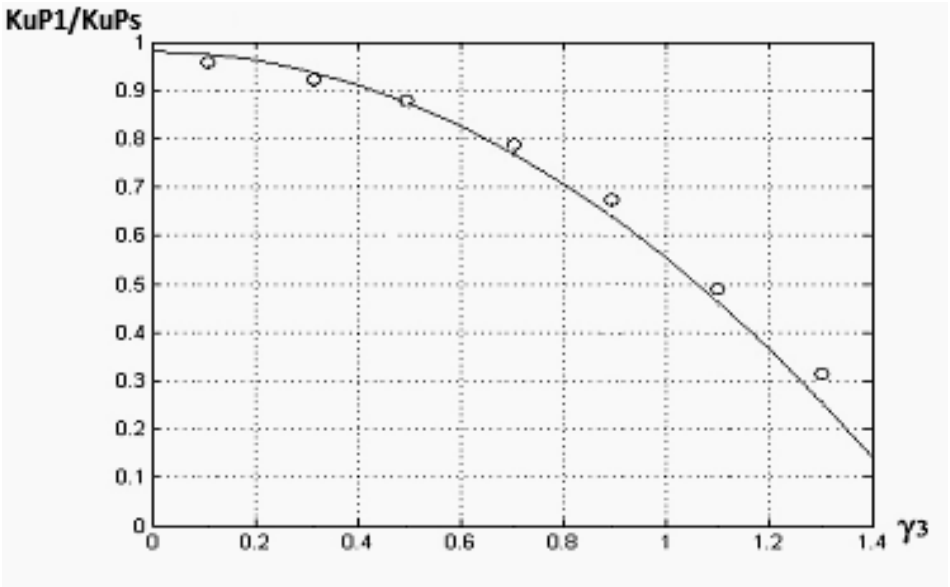


Рис. 5. Результаты сравнения теоретических (сплошная линия) и экспериментальных (точки) значений отношения моментных критериев качества для линейных и нелинейных РП от коэффициента асимметрии γ_3 при $q=1$ и $\gamma_4 = 2$

Из графиков видно, что сравнение теоретических и экспериментальных результатов совпадает. При рассмотрении гауссовских помех, когда $\gamma_3 = 0$, отношение $\mathrm{KuP}_1/\mathrm{KuP}_s$ принимает значение равное 1, что свидетельствует о равенстве линейного и нелинейного РП, а также их характеристик в виде отношения критериев качества. Из графиков видно, что учет параметров негауссовской помехи в виде коэффициента асимметрии γ_3 позволяет повысить точностные характеристики систем обнаружения сигналов. На пример, при значении $\gamma_3 = 0.8$ значение критерия качества KuP_2 уменьшается на 30%, что свидетельствует об уменьшении вероятностей ошибок нелинейного РП (17) при степени полинома $S = 2$ по сравнению с линейным РП (16).

Таким образом, полученные результаты анализа синтезированных полиномиальных РП, оптимальных по моментному критерию качества типа Неймана-Пирсона свидетельствуют о том, что нелинейная обработка случайного негауссовского процесса и учет его параметров позволяет повысить качество обнаружения сигналов.

Выводы

В данной работе представлен новый метод обнаружения сигналов на основе разработки и использования моментно-кумулянтных моделей негауссовских случайных величин и использования полиномиальных стохастических РП, оптимальных по адаптированному моментному критерию качества проверки статистических гипотез типа Неймана-Пирсона. Синтезированы нелинейные алгоритмы обнаружения сигналов на фоне негауссовских помех, приведены их качественные характеристики.

Представлено компьютерное моделирование эффективности синтезированных полиномиальных РП. Показано, что учёт негауссовского распределения помех в виде кумулянтных коэффициентов третьего и выше порядков, а также увеличение степени стохастического полинома РП позволяет увеличить вероятность правильного обнаружения сигналов и уменьшить вероятность ошибки второго рода, что в целом приводит к увеличению эффективности синтезированных полиномиальных алгоритмов обработки сигналов.

Результаты исследований показывают, что нелинейная обработка выборочных значений позволяет повысить эффективность синтезированных РП и дает возможность их использования при проектировании эффективных систем обработки негаусовских процессов в радиолокации, системах связи, диагностики и управления.

Литература

1. Van Trees, H.L. Bell, K. L. Tiany, Z. Detection Estimation and Modulation Theory, 2nd Edition, Part I, Detection, Estimation, and Filtering Theory. – John Wiley & Sons, 2013.

2. Kay S. Fundamentals of Statistical Signal Processing: Detection Theory. – Prentice-Hall, 1993.

3. Kassam S. Signal Detection in Non-Gaussian Noise. – Springer Verlag, 1988.

4. Tuzlukov V.P. Signal Processing Noise. – CRC Press LLC, 2002.

5. Duana F., Chapeau-Blondeaub F., Abbott D. Non-Gaussian noise benefits for coherent detection of narrow band weak signal // In Physics Letters A., 378, pp.1820 – 1824, 2014.

6. Guo G., Mandal M., Jing Y. A robust detector of known signal in non-Gaussian noise using threshold systems. Signal Processing. 92(11), 2676 – 2688, 2012.

7. Picinbono, B.: On deflection as a performance criterion in detection. IEEE Trans // Aerosp. Electron. Syst. 31(3), 1072 – 1081, 1995.

8. Biglieri E., Lops M. Linear–Quadratic Detectors for Spectrum Sensing // Journal of Communications and Networks. 16(5) 485-492, 2014.

9. Nandi A.K. Blind Estimation Using Higher-Order Statistics. – Springer-Verlag, 1999.

10. Primak S., Kontorovich V., Lyandres V. Stochastic Methods and Their Applications to Communications Stochastic Differential Equations Approach. – John Wiley & Sons, 2004.

11. Orosco E., Die, P., Laciar E., Mut V., Soria C., Sciascio F. On the use of high-order cumulant and bispectrum formuscular-activity detection // Biomedical Signal Processing and Control. 18, 325 – 333, 2015.

12. D. Denkovski, V. Atanasovski, L. Gavrilovska, "HOS Based Goodness-of-Fit Testing Signal Detection," // IEEE Communications Letters, 2012. – vol. 16(3). – pp.310-313.

13. Y. Kunchenko, Polynomial Parameter Estimations of Close to Gaussian Random Variables. – Aachen: Shaker Verlag, 2002.

14. Y. Kunchenko, "A moment Performance Criterions of a Deccision-making for Testing Simple Statistical Conjecture," // ISIT 1997, Ulm, Germany, June 29-July 4., pp.407.

15. V.Palahin, O.Palahina, V.Filipov, S.Leleko, A. Ivchenko, "Modeling of Joint Signal Detection and Parameter Estimation on Background of Non-Gaussian Noise," // Journal of Applied Mathematics and Computational Mechanics, issue 14 (3) , 2015. – pp. 87 – 94.

16. Palahina, E., Palahin,V.: Signal Detection in Additive-Multiplicative non-Gaussian Noise Using Higher Order Statistics. // Proceedings of the 26th International Conference Radioelectronika - 2016 (April 19-20, 2016, Košice, Slovakia), 2016. – pp. 262 – 267.

17. Hannelore Liero, Silvelyn Zwanzig. Introduction to the Theory of Statistical // CRC Press LLC, 2001.

18. Levin B. Teoreticheskie osnovy statisticheskoi radiotekhniki. – Izd. 3-e, pererab. i dop. – M.: Radio i svyaz', 1989. – 696 s.

19. James D., Taylor P.E. Ultra-wideband radar technology // CRC Press LLC, 2001.

20. W. A. Lintz and J. C. McEachen A Method for Emphasizing Signal Detection in Wireless Sensor Network Radio Frequency // Array Operation, System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on System Sciences, Big Island, HI, 2009, pp. 1 – 10.

21. K. Pierre, W. Moreno and C. S. Jeong, "System Testability threshold design effectiveness via signal detection theory". – SoutheastCon 2015, Fort Lauderdale, FL. – 2015. – pp. 1 – 9.

ФАКТОРИАЛЬНОЕ КОДИРОВАНИЕ С ИСПРАВЛЕНИЕМ ОШИБОК. ТЕОРЕТИЧЕСКОЕ ОБОСНОВАНИЕ И ПРИМЕРЫ РЕАЛИЗАЦИИ

Фауре Э.В.

Введение

Проблемы повышения эффективности систем передачи данных (СПД), включая повышение достоверности передачи и пропускной способности, всегда привлекали внимание специалистов информационных технологий и телекоммуникаций. В данном контексте перспективными являются методы факториального кодирования информации [1-9], позволяющие совместить операции помехоустойчивого кодирования, крипто- и имитозащиты и тем самым уменьшить вносимую передатчиком избыточность, повысить быстродействие и увеличить эффективную пропускную способность. Вместе с тем возможности факториального кодирования, изложенные в [1-9], далеко не исчерпаны, что и определяет круг решаемых в данной работе задач.

1. Выделение нерешенных задач

В работе [7] предложен метод факториального кодирования с восстановлением данных по перестановке (ФКВД, FCDR), основанный на биективном преобразовании множества информационных слов, составленных из k бит, (2^k векторов $A(x)$) в разрешенное множество из 2^k перестановок $R_{FCDR}(x)$ порядка M ($M! \geq 2^k$). Перестановка $R_{FCDR}(x)$ представляет собой последовательность закодированных равномерным двоичным кодом чисел $\{0;1;K;M-1\}$, очередность следования которых определяется информационной последовательностью и алгоритмом кодирования. Если порядок формирования перестановки по информационному слову источника держится в секрете, ФКВД, помимо обнаружения ошибок в канале связи, обеспечивает защиту данных от несанкционированного чтения. Кроме того, такой код является самосинхронизирующимся и не требует разделителя кодовых слов.

В работе [8] для ФКВД введен показатель избыточности (по мощности) α, определяемый следующим образом:

$$\alpha = M!/2^k. \quad (1)$$

В [8] также показано, что при $k>1$ справедливо $M! \geq 2^k$ и, соответственно, $\alpha > 1$, что приводит к избыточности кода. При этом введение дополнительных проверочных бит перед преобразованием информационного вектора в перестановку позволяет повысить обнаруживающую способность ФКВД.

С другой стороны, избыточность ФКВД обеспечивает возможность увеличения расстояния между перестановками - носителями информации и

создает предпосылки для создания факториального кода с исправлением ошибок.

Целью данной работы является разработка метода факториального кодирования информации, который реализует функцию защиты информации от несанкционированного доступа, а также функцию помехоустойчивого кодирования, сочетающего обнаружение и исправление ошибок, возникающих в канале связи.

2. Решение задачи

Как показано в [7], приемник содержит блок проверки корректности принятой из канала кодовой комбинации и декодер ФКВД. Проверка корректности сводится к проверке того факта, что в принятой кодовой комбинации каждый символ множества $\{0;1;\mathrm{K};\mathrm{M}-1\}$ применяется ровно по одному разу. В случае, если принятая последовательность является некорректной, она не допускается к декодированию, а на передающую станцию по обратному каналу связи передается запрос повторной передачи блока.

Корректная последовательность подлежит декодированию – обратному преобразованию $\mathrm{f}_{\mathrm{FCDR}}^{-1}:\mathrm{R}_{\mathrm{FCDR}}(\mathrm{x})\to\mathrm{A}(\mathrm{x})$. Согласно [7], поскольку $\mathrm{M}!\geq 2^{\mathrm{k}}$ при $\mathrm{k}>1$, множество перестановок на входе декодера состоит из двух подмножеств – разрешенного и запрещенного. К разрешенному подмножеству относятся 2^{k} перестановок (в простейшем случае их синдромы S_{F} соответствуют целым числам $[0;2^{\mathrm{k}}-1]$ числовой оси), а к запрещенному – подмножество из $(\mathrm{M}!-2^{\mathrm{k}})$ остальных перестановок (в простейшем случае их синдромы S_{F} соответствуют целым числам $[2^{\mathrm{k}};\mathrm{M}!-1]$ числовой оси). Таким образом, прием любой перестановки из неразрешенной части множества также инициирует команду переспроса.

Рассмотрим возможность исправления ошибок за счет вносимой ФКВД избыточности. Такой код будем называть факториальным кодом с восстановлением данных и исправлением ошибок – ФКВДио (FCDRec – FCDR with error correction).

Введем следующие определения.

Определение 1. Сигнальными векторами называются представленные в двоичном виде перестановки разрешенного множества.

Множество сигнальных векторов кода образует его сигнально-кодовую конструкцию (СКК).

Определение 2. Сигнальными точками называются точки на числовой оси $[0;\mathrm{M}!-1]$, которые соответствуют сигнальным векторам кода.

Множество сигнальных точек кода образует его сигнальное созвездие.

Рассмотрим два способа формирования СКК для ФКВДио:

1) СКК, основанная на минимальном расстоянии Эвклида между сигнальными точками. Такие СКК будем называть СКК первого типа и обозначать через СКК-1;

2) СКК, основанная на минимальном расстоянии Хэмминга между сигнальными векторами. Такие СКК будем называть СКК второго типа и обозначать через СКК-2.

ФКВДио с СКК-1

Поскольку мощность множества значений вектора $A(x)$ равняется 2^k, минимальное расстояние между сигнальными точками на оси $[0; M!-1]$

$$D_{min} \le \left[\frac{M!-1}{2^k-1}\right]. \qquad (2)$$

В простейшем случае 2^k сигнальных точек располагаются на числовой оси $[0;2^k-1]$ с шагом $D_{min}=1$. Такой факториальный код не предназначен для исправления ошибок. Он может быть применен для обнаружения ошибок, причем только тех, которые приводят к преобразованию переданной перестановки в «не перестановку» или в перестановку из запрещенного множества.

Напомним, что для ФКВД $k \le [\log_2 M!]$. Выполним оценку D_{min} при $k=[\log_2 M!]$, для которого достигается максимальная скорость кода. Поскольку $\log_2 M!-1<[\log_2 M!]\le \log_2 M!$, имеет место $M!/2<2^k\le M!$, а $1\le \frac{M!-1}{2^k-1}<\frac{M!-1}{M!/2-1}=2+\frac{2}{M!-2}$. Таким образом, при $k=[\log_2 M!]$ минимальное расстояние между сигнальными точками $D_{min}\le 2$. Такой ФКВД не способен исправлять все ошибки, приводящие даже к минимальному смещению сигнальных векторов по числовой оси, и поэтому его целесообразно применять для обнаружения ошибок. При этом, как и для $D_{min}=1$, обнаруживаются только те ошибки, которые приводят к преобразованию переданной перестановки в «не перестановку» или в перестановку из запрещенного множества.

Для обеспечения возможности исправления ошибок необходимо увеличить минимальное расстояние между сигнальными точками. При этом ошибки могут быть исправлены при $D_{min}\ge 3$.

Для увеличения расстояния между сигнальными точками необходимо увеличивать показатель избыточности (по мощности) α. Очевидно, что показатель $\alpha=M!/2^k$ является монотонно возрастающей функцией по M и монотонно убывающей по k. Поэтому увеличение значения α может быть достигнуто как увеличением M, так и уменьшением k. При этом, как показано в [8], уменьшение длины

информационного вектора на Δk бит при фиксированном M приводит к увеличению показателя избыточности (по мощности) в $2^{\Delta k}$ раз.

Графически расположение сигнальных точек на числовой оси представлено на рис. 1. При этом расстояние от нуля до сигнальной точки i будем обозначать через D_i, а между сигнальными точками i и j – через $D_{i,j} = D_j - D_i$.

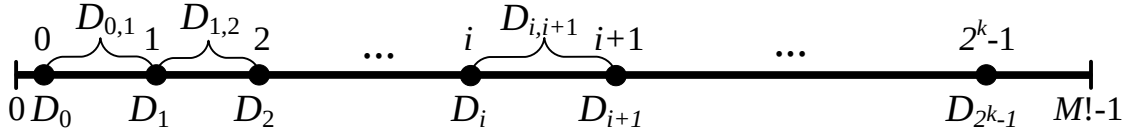


Рис. 1. Расположение сигнальных точек на числовой оси

Положение сигнальных точек на числовой оси определяется используемой СКК. В простейшем случае сигнальные точки располагаются равномерно с шагом D_{min}, при этом $D_{i,i+1} = D_{min} = const$ для $i \in [0; 2^k - 2]$. В более общем случае $D_{i,i+1} \neq const$, а $D_{i,j} \geq D_{min}$, $i, j \in [0; 2^k - 1]$, $i \neq j$.

При передаче сигнального вектора действующая в канале связи помеха может преобразовать переданную перестановку в любую другую перестановку, тем самым сместив сигнальную точку передатчика в любую другую точку отрезка $[0; M!-1]$. Данная точка может быть как сигнальной, так и не сигнальной, а принятая перестановка может относиться как к разрешенному, так и к запрещенному множеству.

Приемник принимает решение о переданном сигнальном векторе на основании критерия максимального правдоподобия путем нахождения сигнальной точки, ближайшей (в метрике Эвклида) к точке, соответствующей принятому вектору. Для этого декодер вычисляет расстояния от соответствующей принятому вектору точке числовой оси до соседних сигнальных точек. При равенстве этих расстояний формируется сигнал переспроса.

Таким образом, если помеха сместила сформированный передатчиком i-ый вектор не более чем на $-\left[\frac{D_{i-1,i}-1}{2}\right]$ и $+\left[\frac{D_{i,i+1}-1}{2}\right]$ точек числовой оси, то эта ошибка исправляется, а принятый вектор корректируется приемником в перестановку, соответствующую *i*-ой сигнальной точке. Если смещение равняется $-\left[\frac{D_{i-1,i}}{2}\right]$ (или $+\left[\frac{D_{i,i+1}}{2}\right]$) и при этом $\frac{D_{i-1,i}}{2} \in \breve{y}$ $\left(\frac{D_{i,i+1}}{2} \in \breve{y}\right)$, т.е. расстояние до соседних сигнальных

точек одинаково, ошибка обнаруживается кодом и исправляется путем переспроса. Если же смещение превышает $-\left[\frac{D_{i-1,i}}{2}\right]$ (или $+\left[\frac{D_{i,i+1}}{2}\right]$), такие ошибки код исправить не может. В этом случае, если расстояния от соответствующей принятому вектору точке числовой оси до соседних сигнальных точек одинаково, ошибка обнаруживается и исправляется путем переспроса, если же это расстояние различное, имеет место ошибка декодирования и, как следствие, не обнаруженная кодом ошибка.

Рассмотрим процесс декодирования, если принятый вектор соответствует точке из диапазона $[0; D_0 - 1]$ или $[D_{2^k-1} + 1; M!-1]$. В этом случае возможны два варианта правила принятия решения декодером:

1) все точки диапазона $[0; D_0 - 1]$ корректируются в нулевую сигнальную точку, а диапазона $[D_{2^k-1} + 1; M!-1]$ – в $(2^k - 1)$ сигнальную точку;

2) коррекция в нулевую сигнальную точку выполняется для диапазона $\left[D_0 - \left[\frac{D_{min} - 1}{2}\right]; D_0 - 1\right]$, в $(2^k - 1)$ сигнальную точку – для диапазона $\left[D_{2^k-1} + 1; D_{2^k-1} + \left[\frac{D_{min} - 1}{2}\right]\right]$, остальные точки крайних диапазонов являются запрещенными.

В любом случае все ошибки, приводящие к смещению сигнальной точки на расстояние $D \le \left[\frac{D_{min} - 1}{2}\right]$, исправляются.

Положим $D_{i,i+1} = D_{min}$ для $\forall i \in [0; 2^k - 2]$, $D_0 = \left[\frac{D_{min} - 1}{2}\right]$, а $M!-1 - D_{2^k-1} \ge \left[\frac{D_{min} - 1}{2}\right]$. В этом случае имеет место оценка

$$(2^k - 1)D_{min} + 2\left[\frac{D_{min} - 1}{2}\right] + 1 \le M!. \qquad (3)$$

При заданных k и M минимальное расстояние $D_{min} \le \max(D) : (2^k - 1)D_{min} + 2\left[\frac{D - 1}{2}\right] + 1 \le M!$. Например, если k=40, а M=16, то $D_{min} \le 19$. Таким образом, выбор параметров k и M однозначно определяют максимальную исправляющую способность кода.

Выражение (3) также может служить для выбора k или M при других известных параметрах кода. Например, если k=16, а $D_{min} = 3$, то $M! \ge 196608$, откуда $M \ge 9$. Если же M=8, а $D_{min} = 6$, то $k \le 12$.

Кроме того, выражение (3) для представленного выше второго правила принятия решения декодером показывает, что все точки, лежащие правее пороговой точки $(2^k - 1)D_{min} + 2\left[\frac{D_{min} - 1}{2}\right] + 1$, относятся к неиспользуемой (запрещенной) части числового множества (числовой оси). Поэтому все принятые из канала связи кодовые комбинации после проверки корректности проходят сравнение с пороговым значением. Если соответствующая кодовой комбинации точка числовой оси расположена выше пороговой точки, производится переспрос блока данных, в противном случае выполняется поиск ближайшей сигнальной точки и отождествление с ней принятой кодовой комбинации.

Структурная схема приемника ФКВДио представлена на рис. 2.

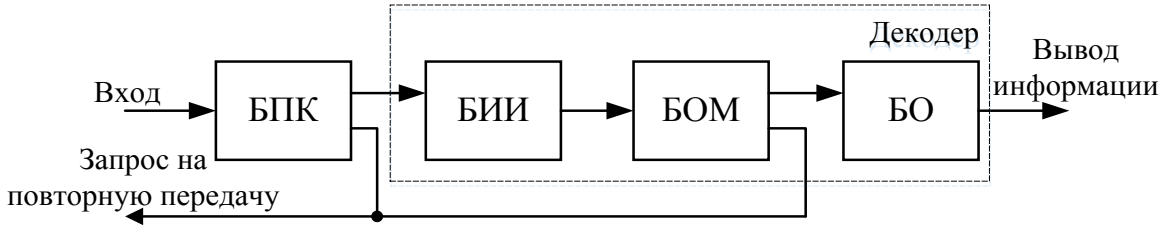


Рис. 2. Структурная схема приемника ФКВДио

На рис. 2 введены следующие обозначения: БПК – блок проверки корректности принятой комбинации; БИИ – блок извлечения информации из перестановки; БОМ – блок оценки принадлежности принятой перестановки к разрешенному множеству; БО – блок отождествления принятой перестановки с ближайшим разрешенным вектором данных.

Таким образом, предложенный факториальный код обеспечивает сочетание свойств кодов с прямым исправлением ошибок (ECC) и кодов с обнаружением ошибок и их исправления путем переспроса (EDC).

Вероятностные характеристики ФКВДио с СКК-1

Ошибки, не обнаруживаемые кодом, обусловлены только теми преобразованиями, при которых переданная перестановка трансформировалась в перестановку разрешенной части множества, соответствующую любой точке из множества точек числовой оси, для которых ближайшими являются сигнальные точки других перестановок.

Обозначим с помощью $f_{per}^{ud}(i,t)$ количество ошибок веса t, приводящих к ошибочному декодированию i-ого сигнального вектора. Примем, что канал связи – симметричный двоичный постоянный с переходной вероятностью p_0, а битовые ошибки возникают в нем независимо. Тогда вероятность не обнаруженной ФКВД или ФКВДио ошибки

$$P_{ud}(FCDR(ec), p_0) = \sum_{i=0}^{2^k-1}\left(P_w(i) \cdot \sum_{t=1}^{r} f_{per}^{ud}(i,t) p_0^t q_0^{r-t} \right), \qquad (4)$$

где $r = l_r \cdot M$ – длина кодового слова;

$l_r = \text{entier}(\log_2 M) + 1$ – количество бит для кодирования равномерным кодом одного символа перестановки;

$P_w(i)$ – вероятность применения источником i-ого слова, $i \in [0; 2^k - 1]$.

Определим долю φ ошибок, приводящих к ошибочному декодированию:

1) для ФКВД в режиме обнаружения ошибок: $\varphi = \varphi_{обн} = \frac{2^k - 1}{M!}$;

2) для ФКВДио в режиме исправления и обнаружения ошибок:

$$\varphi = \varphi_{испр} \geq \varphi_{обн} + \frac{2 \cdot (2^k - 1) \cdot \left[\frac{D_{min} - 1}{2} \right]}{M!} = \frac{(2^k - 1) \cdot \left(2 \left[\frac{D_{min} - 1}{2} \right] + 1 \right)}{M!}.$$

Поскольку множество ошибок, приводящих к ошибочному декодированию в режиме исправления и обнаружения ошибок, содержит множество ошибок, приводящих к ошибочному декодированию в режиме обнаружения ошибок, при $D_{min} \geq 3$ выполняется неравенство $P_{ud}(FCD\,Rec, p_0) > P_{ud}(FCDR, p_0)$.

Учтем, что для простейшей системы с РОС динамическая составляющая потери скорости вследствие переспросов $\nu_2 = Q + P_{ud}$ [10, с. 676], где Q – вероятность приема блока данных без ошибок, P_{ud} – вероятность необнаруженной ошибки. В случае использования кода с исправлением и обнаружением ошибок справедливо выражение

$$Q + P_{EC} + P_{det} + P_{ud} = 1, \qquad (5)$$

где P_{EC} – вероятность того, что ошибка будет исправлена, а кодовая комбинация принята верно;

P_{det} – вероятность обнаруженной ошибки.

Тогда динамическая составляющая потери скорости вследствие переспросов для кода с исправлением и обнаружением ошибок

$$\nu_2 = 1 - P_{det} = Q + P_{EC} + P_{ud}. \qquad (6)$$

Вероятность исправления ошибок для ФКВДио определяется следующим образом:

$$P_{EC}(FCDRec, p_0) = \sum_{i=0}^{2^k-1}\left(P_w(i) \cdot \sum_{t=1}^{r} f_{per}^{EC}(i,t) p_0^t q_0^{r-t} \right), \qquad (7)$$

где $f_{per}^{EC}(i,t)$ – количество ошибок веса t, исправляемых ФКВДио для i-ой сигнальной точки.

Сравнивая выражение $\nu_2 = Q + P_{ud}$ для ФКВД и выражение (6) для ФКВДио, можно видеть, что для режима обнаружения ошибок динамическая составляющая потери скорости ν_2 и, как следствие, относительная скорость передачи $\nu_0 = \nu_1 \cdot \nu_2$, где ν_1 – скорость кода, ниже, чем для режима исправления и обнаружения ошибок.

Таким образом, при одинаковых параметрах и $D_{min} \geq 3$ ФКВДио обеспечивает бóльшую относительную скорость передачи по сравнению с ФКВД, однако проигрывает в помехоустойчивости.

Определим энергетический выигрыш ΔP при применении ФКВДио для некогерентного приемника, характеризующегося вероятностью битовой ошибки $p = 0.5 \cdot e^{-0.5h^2}$ [11, с. 45], где h^2 – соотношение сигнал/шум. В этом случае

$$\Delta P = 10 \lg \frac{h_{eq}^2}{h_0^2} = 10 \lg \frac{\ln(2p_{0eq})}{\ln(2p_0)}, \qquad (8)$$

где $h_0^2 = 2\ln(2p_0)$ – соотношение сигнал/шум на входе некогерентного приемника, обеспечивающее вероятность битовой ошибки p_0 на его выходе;

h_{eq}^2 – соотношение сигнал/шум на входе некогерентного приемника, обеспечивающее эквивалентную вероятность битовой ошибки на его выходе p_{0eq}, определенную в [10, с. 676] как вероятность ошибки в гипотетическом симметричном постоянном двоичном канале, при которой вероятность безошибочного приема достаточно длинного сообщения такая же, как и в рассматриваемой системе. Согласно [10, с. 677],

$$(1 - p_{0eq})^N = (1 - P_{ud})^{\left\lceil \frac{N}{k} \right\rceil \frac{1}{1 - P_{det}}}. \qquad (9)$$

Решая уравнение (9) относительно p_{0eq} при $p_{0eq}, P_{ud} = 1$, можно видеть, что

$$p_{0eq} \approx \frac{P_{ud}}{k(1 - P_{det})}. \qquad (10)$$

Учтем, что в соответствии с (5) для рассматриваемой системы ФКВД с исправлением и обнаружением ошибок $1 - P_{det} = Q + P_{EC} + P_{ud}$. Тогда для ФКВДио выражение (10) принимает вид:

$$p_{0eq} \approx \frac{P_{ud}}{k(Q + P_{EC} + P_{ud})}. \qquad (11)$$

Остаточная вероятность ошибочного приема [10, с. 678], под которой понимается вероятность того, что комбинация, выданная получателю, содержит хотя бы одну ошибку, для ФКВДио равна

$$P_{res} = \frac{P_{ud}}{1 - P_{det}} = \frac{P_{ud}}{Q + P_{EC} + P_{ud}}. \qquad (12)$$

ФКВДио с СКК-2

Определенное для СКК-1 расстояние Эвклида $D_{i,j}$ между сигнальными точками i и j в общем случае не равняется расстоянию Хэмминга между кодовыми словами, соответствующими этим сигнальным точкам. Вместе с тем из теории корректирующих кодов [12] известно, что для исправления ошибки в двоичных разрядах кратности t минимальное расстояние Хэмминга d_{min} между кодовыми словами должно удовлетворять условию $d_{min} \geq 2t + 1$.

Расстояние Хэмминга между сигнальными точками i и j будем обозначать с помощью $d_{i,j}$. СКК-2 для ФКВДио предусматривает выполнение условия $d_{i,j} \geq d_{min}$, $i, j \in [0; 2^k - 1]$, $i \neq j$.

В простейшем случае для ФКВД сигнальные вектора соответствуют сигнальным точкам с шагом $D_{i,i+1} = D_{min} = 1$ и $D_0 = 1$. Тогда $d_{min} = 2$, а ФКВД только обнаруживает ошибки, приводящие к преобразованию переданной перестановки в «не перестановку» или в перестановку из запрещенного множества.

Определение связи между M, k и d_{min} является актуальной задачей, однако выходит за рамки данной работы.

Очевидно, что для обеспечения возможности исправления ошибок необходимо увеличить минимальное расстояние d_{min} между сигнальными точками. Для увеличения расстояния между сигнальными точками необходимо увеличивать показатель избыточности (по мощности) α. При этом ошибки могут быть исправлены при $d_{min} \geq 3$.

При передаче сигнального вектора по каналу связи на него воздействует помеха. Модифицированный помехой вектор поступает на вход приемника ФКВДио с СКК-2, структура которого показана на рис. 3.

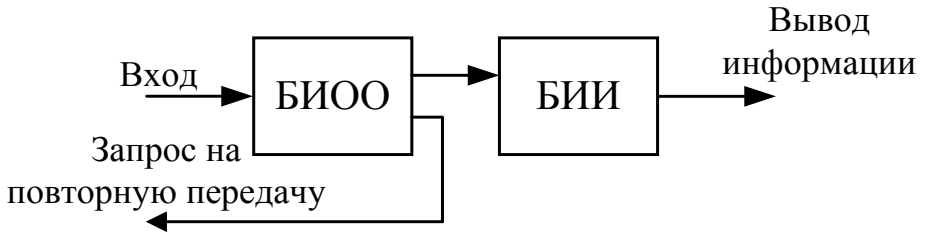


Рис. 3. Структурная схема приемника ФКВДио с СКК-2

Приемник ФКВДио с СКК-2 содержит блок исправления и обнаружения ошибок БИОО и блок извлечения информации из перестановки БИИ.

Блок исправления и обнаружения ошибок БИОО реализует следующие функции:

1) определяет расстояния Хэмминга r_i между принятым вектором и всеми сигнальными векторами, $i \in [0;2^k - 1]$;

2) находит минимальное расстояние $r_{min} = \min\{r_i\}$;

3) если минимальное расстояние r_{min} соответствует расстоянию только до одного i-го сигнального вектора, т.е. существует единственное $i \in [0;2^k - 1]$: $r_i = r_{min}$, принятая комбинация отождествляется с i-ым сигнальным вектором;

4) если минимальное расстояние r_{min} соответствует расстоянию до двух или более сигнальных векторов, т.е. существует как минимум два значения $i, j \in [0;2^k - 1]$: $r_i = r_j = r_{min}$, формируется сигнал переспроса.

Таким образом, правила декодирования, реализованные в блоке БИОО, основываются на критерии максимального правдоподобия.

В блоке извлечения информации из перестановки БИИ производится преобразование перестановки в k-битную информационную последовательность.

Учтем, что расстояние Хэмминга между сигнальными векторами четно и, следовательно, d_{min}, M. Поэтому при передаче i-го сигнального вектора ФКВДио с СКК-2 справедливы следующие утверждения:

1) ошибка с весом $t \le \left[\frac{d_{min} - 1}{2}\right] = \frac{d_{min} - 2}{2} = \frac{d_{min}}{2} - 1$ исправляется, а принятый вектор корректируется приемником в переданный сигнальный вектор;

2) ошибка с весом $t = \frac{d_{min}}{2}$ может быть как исправлена (если минимальное расстояние r_{min} соответствует расстоянию только до одного *i*-го сигнального вектора), так и обнаружена и исправлена путем переспроса (если минимальное расстояние r_{min} соответствует расстоянию до двух или более сигнальных векторов);

3) если вес ошибки $t > \frac{d_{min}}{2}$, может иметь место исправленная ошибка (если минимальное расстояние r_{min} соответствует расстоянию только до одного i-го сигнального вектора), обнаруженная ошибка (если минимальное расстояние r_{min} соответствует расстоянию до двух или более сигнальных векторов) или необнаруженная ошибка (если минимальное

расстояние r_{min} соответствует расстоянию только до одного сигнального вектора, отличного от i-го).

Вероятностные характеристики ФКВДио с СКК-2 определяются по тем же формулам, что и ФКВДио с СКК-1: (4), (6), (7), (8), (12).

Таким образом, данный код позволяет комбинировать исправление наиболее частых сочетаний ошибок и обнаружение с последующей повторной передачей для более редких сочетаний ошибок.

Актуальным, однако выходящим за рамки данной работы, вопросом при выборе СКК второго типа является вопрос о максимальном количестве $N_{sv}(d_{min}, M)$ сигнальных векторов, обеспечивающих заданное минимальное расстояние d_{min} при известном M (данный вопрос тесно связан с теорией решеток и задачей наилучших упаковкой шаров в пространствах различных размерностей [13]). Зная значение $N_{sv}(d_{min}, M)$, можно сконструировать эффективный код, передающий с помощью одной перестановки $k = \log_2 N_{sv}$ бит информации и исправляющий все ошибки кратности $t \le \frac{d_{min}}{2} - 1$. Вероятностные характеристики такого кода определяются по представленный выше выражениям, в которых вместо значения $2^k - 1$ принимается значение $N_{sv}(d_{min}, M) - 1$, а $k = \log_2 N_{sv}$.

3 Примеры ФКВДио

ФКВДио с СКК-1

Примем k=3, а M=4. В соответствии с (3) $D_{min} \le 3$. Тогда сигнальными точками являются точки 1, 4, 7, 10, 13, 16, 19, 22, а СКК для базовой перестановки $\pi(0) = \{0;1;2;3\}$ представлена в табл. 1. Из таблицы видно, что такая СКК обеспечивает $d_{min} = 2$.

Расположение сигнальных точек показано на рис. 4.

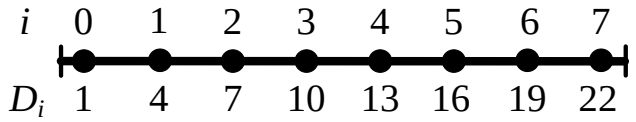


Рис. 4. Расположение сигнальных точек для СКК-1

Определим вероятность необнаруженной ошибки ФКВДио с СКК-1, представленной в табл. 1.

Ошибка не обнаруживается кодом, если кодовое слово, соответствующее i-ой сигнальной точке, будет преобразовано помехой в комбинацию, соответствующую любой точке числовой оси, не принадлежащей диапазону $[D_i - 1; D_i + 1]$.

Для СКК-1 построим матрицу расстояний Хэмминга между множеством сигнальных точек и множеством всех точек числовой оси (табл. 2).

Таблица 1

СКК-1 для ФКВДио при $k = 3,\ M = 4$

Сигнальные точки	СКК-1
1	00 01 11 10
4	00 11 01 10
7	01 00 11 10
10	01 11 00 10
13	10 00 11 01
16	10 11 00 01
19	11 00 10 01
22	11 10 00 01

Таблица 2

Матрица расстояний между сигнальными точками и точками числовой оси

Точки числовой оси		D_i							
		1	4	7	10	13	16	19	22
0	00 01 10 11	2	4	4	4	4	4	4	6
1	00 01 11 10	0	2	2	4	4	6	6	8
2	00 10 01 11	4	2	4	4	4	4	6	4
3	00 10 11 01	4	4	4	6	2	4	4	4
4	00 11 01 10	2	0	4	2	6	4	8	6
5	00 11 10 01	4	4	6	4	4	2	4	4
6	01 00 10 11	4	6	2	4	4	6	2	4
7	01 00 11 10	2	4	0	4	4	8	4	6
8	01 10 00 11	6	4	4	2	6	4	4	2
9	01 10 11 00	4	4	2	4	4	6	4	4
10	01 11 00 10	4	2	4	0	8	4	6	4
11	01 11 10 00	4	4	4	2	6	4	4	4
12	10 00 01 11	4	4	4	6	2	4	4	4
13	10 00 11 01	4	6	4	8	0	4	2	4
14	10 01 00 11	4	4	6	4	4	2	4	4
15	10 01 11 00	2	4	4	6	2	4	4	6
16	10 11 00 01	6	4	8	4	4	0	4	2
17	10 11 01 00	4	2	6	4	4	2	6	4
18	11 00 01 10	4	4	2	4	4	6	4	4
19	11 00 10 01	6	8	4	6	2	4	0	2
20	11 01 00 10	4	4	4	2	6	4	4	4
21	11 01 10 00	4	6	4	4	4	4	2	4
22	11 10 00 01	8	6	6	4	4	2	2	0
23	11 10 01 00	6	4	4	4	4	4	4	2

Серым цветом для сигнальной точки i в таблице выделены расстояния для точек диапазона $[D_i - 1; D_i + 1]$.

Представленные результаты полностью согласуются с приведенной в [7] теоретической оценкой количества ошибок веса t, преобразующих перестановку в перестановку: $f_{per}(0) = 1$, $f_{per}(2) = 4$, $f_{per}(4) = 14$, $f_{per}(6) = 4$, $f_{per}(8) = 1$.

Вероятность не обнаруженной ФКВДио ошибки определяется по (4). Примем, что все слова применяются источником с одинаковой вероятностью, равной $P_w(i) = P_w = \frac{1}{2^k}$. Для рассматриваемого примера $P_w = \frac{1}{8}$, $l_r = 2$, а r=8. Учтем также, что вес ошибок, приводящих к ошибочному декодированию, четен. Тогда $P_{ud}(FCDRec, p_0) = \frac{1}{8}\sum_{i=0}^{7}\sum_{t=1}^{4} f_{per}^{ud}(i,2t) p_0^{2t} q_0^{8-2t}$.

Значения $f_{per}^{ud}(i,2t)$ для данных из табл 2 приведены в табл 3.

Результаты вычисления $P_{ud}(FCD\,Rec, p_0)$ в соответствии с (4) для различных p_0 представлены в табл. 4.

Таблица 3

Значения $f_{per}^{ud}(i,2t)$ для ФКВДио с СКК-1

t	Сигнальная точка							
	0	1	2	3	4	5	6	7
1	3	4	3	3	3	3	4	3
2	13	12	13	13	13	13	12	13
3	4	4	4	4	4	4	4	4
4	1	1	1	1	1	1	1	1

Таблица 4

Вероятность необнаруженной ошибки для ФКВДио с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{ud}(FCD\,Rec, p_0)$	$1.81\cdot10^{-2}$	$3.06\cdot10^{-4}$	$3.23\cdot10^{-6}$	$3.25\cdot10^{-8}$	$3.25\cdot10^{-10}$	$3.25\cdot10^{-12}$

Динамическая составляющая потери скорости вследствие переспросов определяется по (6):

$$\nu_2(FCDRec, p_0) = 1 - P_{det}(FCDRec, p_0) = Q + P_{EC}(FCDRec, p_0) + P_{ud}(FCDRec, p_0), \quad (13)$$

где $Q = (1 - p_0)^r$ – вероятность приема кодового слова без ошибок;

$P_{det}(FCD\,Rec, p_0)$ – вероятность переспроса;

$P_{EC}(FCD\,Re\,c, p_0)$ – вероятность исправления ошибок, определяемая по (7).

С учетом четности ошибок, преобразующих перестановку в перестановку, для СКК-1

$$P_{EC}(FCD Re\,c, p_0) = \sum_{i=0}^{2^k-1}\left(P_w(i) \cdot \sum_{t=1}^{[r/2]} f_{per}^{EC}(i,2t) p_0^{2t} q_0^{r-2t} \right).$$ Значения $f_{per}^{EC}(i,2t)$ для данных из таблицы 2 приведены в табл. 5.

Результаты вычисления $P_{EC}(FCD\,Re\,c, p_0)$ для различных p_0 представлены в табл. 6.

Значения энергетического выигрыша в результате применения ФКВДио с СКК-1 представлены в табл. 7.

Таблица 5

Значения $f_{per}^{EC}(i,2t)$ *для ФКВДио с СКК-1*

t	Сигнальная точка							
	0	1	2	3	4	5	6	7
1	1	0	1	1	1	1	0	1
2	1	2	1	1	1	1	2	1

Таблица 6

Вероятность исправления ошибки для ФКВДио с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{EC}(FCD\,Re\,c, p_0)$	$4.07 \cdot 10^{-3}$	$7.06 \cdot 10^{-5}$	$7.46 \cdot 10^{-7}$	$7.50 \cdot 10^{-9}$	$7.50 \cdot 10^{-11}$	$7.50 \cdot 10^{-13}$

Таблица 7

Энергетический выигрыш ФКВДио с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\Delta P(FCD\,Re\,c, p_0)$, дБ	3.525	3.328	3.219	3.164	3.132	3.111

Значения остаточной вероятности ошибочного приема в результате применения ФКВДио с СКК-1, вычисленные по (12), представлены в табл. 8.

Таблица 8

Остаточная вероятность ошибочного приема ФКВДио с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{res}(FCD\,Re\,c, p_0)$	$4.00 \cdot 10^{-2}$	$3.32 \cdot 10^{-4}$	$3.26 \cdot 10^{-6}$	$3.25 \cdot 10^{-8}$	$3.25 \cdot 10^{-10}$	$3.25 \cdot 10^{-12}$

Значения динамической составляющей потери скорости для ФКВДио с СКК-1 представлены в табл. 9.

Таблица 9

Динамическая составляющая потери скорости ФКВДио с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\nu_2(\mathrm{FCD\,Re\,c}, p_0)$	0.453	0.923	0.992	0.9992	0.99992	0.999992

ФКВД с СКК-1.

Определим вероятность необнаруженной ошибки и значение динамической составляющей потери скорости для обнаруживающего ошибки ФКВД, который использует СКК-1.

Количество $f_{per}^{ud}(i,2t)$ ошибок веса 2t, приводящих к ошибочному декодированию, для каждой сигнальной точки i данного кода приведено в табл. 10.

Таблица 10

Значения $f_{per}^{ud}(i,2t)$ *для ФКВД с СКК-1*

t	Сигнальная точка							
	0	1	2	3	4	5	6	7
1	2	2	1	1	1	1	2	2
2	2	2	4	4	4	4	2	2
3	2	2	1	1	1	1	2	2
4	1	1	1	1	1	1	1	1

Результаты вычисления по (4) $P_{ud}(\mathrm{FCDR}, p_0)$ для ФКВД с СКК-1 при различных p_0 представлены в табл. 11.

Таблица 11

Вероятность необнаруженной ошибки для ФКВД с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{ud}(\mathrm{FCDR}, p_0)$	$9.47\cdot10^{-3}$	$1.65\cdot10^{-4}$	$1.74\cdot10^{-6}$	$1.75\cdot10^{-8}$	$1.75\cdot10^{-10}$	$1.75\cdot10^{-12}$

Значения энергетического выигрыша в результате применения ФКВД с СКК-1 представлены в табл. 12.

Таблица 12

Энергетический выигрыш ФКВД с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\Delta P(\mathrm{FCDR}, p_0)$, дБ	4.211	3.636	3.420	3.314	3.251	3.210

Значения остаточной вероятности ошибочного приема в результате применения ФКВД с СКК-1 представлены в табл. 13.

Таблица 13

Остаточная вероятность ошибочного приема ФКВД с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{res}(FCDR, p_0)$	$2.15 \cdot 10^{-2}$	$1.79 \cdot 10^{-4}$	$1.75 \cdot 10^{-6}$	$1.75 \cdot 10^{-8}$	$1.75 \cdot 10^{-10}$	$1.75 \cdot 10^{-12}$

Значения динамической составляющей потери скорости вследствие переспросов $\nu_2(FCDR, p_0) = 1 - P_{req}(FCDR, p_0) = Q + P_{ud}(FCDR, p_0)$ приведены в табл. 14.

Таблица 14

Динамическая составляющая потери скорости для ФКВД с СКК-1

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\nu_2(FCDR, p_0)$	0.44	0.923	0.992	0.9992	0.99992	0.999992

ФКВДио с СКК-2.

В табл. 15 представлена СКК-2 из 8 сигнальных векторов с $d_{min} = 4$, обеспечивающая исправление однократной битовой ошибки в кодовом слове.

Таблица 15

СКК-2 для ФКВДио при $k = 3$, $M = 4$

СКК	Сигнальные точки
00 01 10 11	0
01 00 11 10	7
10 11 00 01	16
11 10 01 00	23
11 01 10 00	21
01 11 00 10	10
10 00 11 01	13
00 10 01 11	2

Данная СКК построена следующим образом:

– в качестве первого кодового слова выбрана представленная в двоичном виде перестановка $\{0;1;2;3\}$;

– второе кодовое слово образовано путем перестановки 1 и 2, а также 3 и 4 символов первого кодового слова;

– третье кодовое слово образовано путем перестановки 1 и 3, а также 2 и 4 символов первого кодового слова;

– четвертое кодовое слово образовано путем перестановки 1 и 4, а также 2 и 3 символов первого кодового слова;

– кодовые слова 5-8 образованы путем записи справа налево кодовых слов 1-4.

Рассмотрим вероятностные характеристики ФКВДио для СКК-2 из табл. 15.

Экспериментально установлено, что данный код позволяет исправить только любые ошибки кратности t=1, а ошибка не обнаруживается тогда и только тогда, когда кодовое слово, соответствующее i-ому сигнальному вектору, преобразовывается помехой в комбинацию, для которой расстояние Хэмминга до любого другого сигнального вектора не превышает единицу.

Для данной СКК-2 построим матрицу расстояний Хэмминга между i-ым $(i \in [0,7])$ сигнальным вектором и множеством векторов, образующих в метрике Хэмминга сферы единичного радиуса с центрами в сигнальных точках. Результаты сведем в табл. 16.

Серым цветом для сигнальной точки i в таблице выделены расстояния до точек собственной единичной сферы.

Вероятность не обнаруженной кодом ошибки вычисляется по (4). Примем также, что все слова применяются источником с одинаковой вероятностью $P_w = \frac{1}{8}$. Значения $f_{per}^{ud}(i,t)$ для данных из табл. 16 приведены в табл. 17 ($f_{per}^{ud}(i,t) = 0$ при $t \le 2$).

Таблица 16

Матрица расстояний между сигнальными векторами и векторами, приводящими к ошибочному декодированию

Единичные сферы		D_i							
		0	2	7	10	13	16	21	23
Центр	00 01 10 11	0	4	4	8	4	4	4	4
	10 01 10 11	1	5	3	7	3	5	3	5
	01 01 10 11	1	3	5	7	3	3	5	5
	00 11 10 11	1	5	3	7	5	3	5	3
	00 00 10 11	1	3	5	7	5	5	3	3
	00 01 00 11	1	5	3	7	5	3	5	3
	00 01 11 11	1	3	5	7	5	5	3	3
	00 01 10 01	1	5	3	7	3	5	3	5
	00 01 10 10	1	3	5	7	3	3	5	5
Центр	01 00 11 10	4	0	8	4	4	4	4	4
	11 00 11 10	5	1	7	3	3	5	3	5
	00 00 11 10	3	1	7	5	5	5	3	3
	01 10 11 10	5	1	7	3	5	3	5	3
	01 01 11 10	3	1	7	5	3	3	5	5
	01 00 01 10	5	1	7	3	5	3	5	3
	01 00 10 10	3	1	7	5	3	3	5	5
	01 00 11 00	5	1	7	3	3	5	3	5
	01 00 11 11	3	1	7	5	5	5	3	3

Продолжение таблицы 16

Единичные сферы		D_i							
		0	2	7	10	13	16	21	23
Центр	10 11 00 01	4	8	0	4	4	4	4	4
	00 11 00 01	3	7	1	5	5	3	5	3
	11 11 00 01	5	7	1	3	3	3	5	5
	10 01 00 01	3	7	1	5	3	5	3	5
	10 10 00 01	5	7	1	3	5	5	3	3
	10 11 10 01	3	7	1	5	3	5	3	5
	10 11 01 01	5	7	1	3	5	5	3	3
	10 11 00 11	3	7	1	5	5	3	5	3
	10 11 00 00	5	7	1	3	3	3	5	5
Центр	11 10 01 00	8	4	4	0	4	4	4	4
	01 10 01 00	7	3	5	1	5	3	5	3
	10 10 01 00	7	5	3	1	5	5	3	3
	11 00 01 00	7	3	5	1	3	5	3	5
	11 11 01 00	7	5	3	1	3	3	5	5
	11 10 11 00	7	3	5	1	3	5	3	5
	11 10 00 00	7	5	3	1	3	3	5	5
	11 10 01 10	7	3	5	1	5	3	5	3
	11 10 01 01	7	5	3	1	5	5	3	3
Центр	11 01 10 00	4	4	4	4	0	4	4	8
	01 01 10 00	3	3	5	5	1	3	5	7
	10 01 10 00	3	5	3	5	1	5	3	7
	11 11 10 00	5	5	3	3	1	3	5	7
	11 00 10 00	5	3	5	3	1	5	3	7
	11 01 00 00	5	5	3	3	1	3	5	7
	11 01 11 00	5	3	5	3	1	5	3	7
	11 01 10 10	3	3	5	5	1	3	5	7
	11 01 10 01	3	5	3	5	1	5	3	7
Центр	01 11 00 10	4	4	4	4	4	0	8	4
	11 11 00 10	5	5	3	3	3	1	7	5
	00 11 00 10	3	5	3	5	5	1	7	3
	01 01 00 10	3	3	5	5	3	1	7	5
	01 10 00 10	5	3	5	3	5	1	7	3
	01 11 10 10	3	3	5	5	3	1	7	5
	01 11 01 10	5	3	5	3	5	1	7	3
	01 11 00 00	5	5	3	3	3	1	7	5
	01 11 00 11	3	5	3	5	5	1	7	3
Центр	10 00 11 01	4	4	4	4	4	8	0	4
	00 00 11 01	3	3	5	5	5	7	1	3
	11 00 11 01	5	3	5	3	3	7	1	5
	10 10 11 01	5	5	3	3	5	7	1	3
	10 01 11 01	3	5	3	5	3	7	1	5
	10 00 01 01	5	5	3	3	5	7	1	3

Продолжение таблицы 16

Единичные сферы		D_i							
		0	2	7	10	13	16	21	23
	10 00 10 01	3	5	3	5	3	7	1	5
	10 00 11 11	3	3	5	5	5	7	1	3
	10 00 11 00	5	3	5	3	3	7	1	5
Центр	00 10 01 11	4	4	4	4	8	4	4	0
	10 10 01 11	5	5	3	3	7	5	3	1
	01 10 01 11	5	3	5	3	7	3	5	1
	00 00 01 11	3	3	5	5	7	5	3	1
	00 11 01 11	3	5	3	5	7	3	5	1
	00 10 11 11	3	3	5	5	7	5	3	1
	00 10 00 11	3	5	3	5	7	3	5	1
	00 10 01 01	5	5	3	3	7	5	3	1
	00 10 01 10	5	3	5	3	7	3	5	1

Таблица 17

Значения $f_{per}^{ud}(i,t)$ *для ФКВДио с СКК-2*

t	Сигнальная точка							
	0	1	2	3	4	5	6	7
3	24	24	24	24	24	24	24	24
4	6	6	6	6	6	6	6	6
5	24	24	24	24	24	24	24	24
6	0	0	0	0	0	0	0	0
7	8	8	8	8	8	8	8	8
8	1	1	1	1	1	1	1	1

Как можно видеть распределение количества $f_{per}^{ud}(i,t)$ ошибок веса *t*, приводящих к ошибочному декодированию ФКВДио с СКК-2, не зависит от сигнальной точки.

Результаты вычисления $P_{ud}(FCD\,Rec, p_0)$ для различных p_0 представлены в табл. 18.

Таблица 18

Вероятность необнаруженной ошибки для ФКВДио с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{ud}(FCD\,Rec, p_0)$	$1.47\cdot10^{-2}$	$2.29\cdot10^{-5}$	$2.39\cdot10^{-8}$	$2.4\cdot10^{-11}$	$2.4\cdot10^{-14}$	$2.4\cdot10^{-17}$

Вероятность исправления ошибок $P_{EC}(FCD\,Rec, p_0)$ определяется по (7). С учетом того, что данный код ФКВДио с СКК-2 исправляет только

ошибки единичного веса, $f_{per}^{ud}(i,t)=\begin{cases}8, \text{если } t=1,\\ 0, \text{если } t\neq 1.\end{cases}$ Результаты вычисления $P_{EC}(FCD\,Rec, p_0)$ для различных p_0 представлены в табл. 19.

Таблица 19

Вероятность исправления ошибки для ФКВДио с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{EC}(FCD\,Rec, p_0)$	0.383	$7.46\cdot10^{-2}$	$7.94\cdot10^{-3}$	$7.99\cdot10^{-4}$	$8\cdot10^{-5}$	$8\cdot10^{-6}$

Значения энергетического выигрыша в результате применения ФКВДио с СКК-2 представлены в табл. 20.

Таблица 20

Энергетический выигрыш ФКВДио с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\Delta P(FCD\,Rec, p_0)$, дБ	4.401	4.524	4.608	4.652	4.677	4.694

Значения остаточной вероятности ошибочного приема в результате применения ФКВДио с СКК-2 представлены в табл. 21.

Таблица 21

Остаточная вероятность ошибочного приема ФКВДио с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{res}(FCD\,Rec, p_0)$	$1.78\cdot10^{-2}$	$2.29\cdot10^{-5}$	$2.39\cdot10^{-8}$	$2.40\cdot10^{-11}$	$2.40\cdot10^{-14}$	$2.40\cdot10^{-17}$

Определенные по (13) значения динамической составляющей потери скорости вследствие переспросов для ФКВДио с СКК-2 представлены в табл. 22.

Таблица 22

Динамическая составляющая потери скорости ФКВДио с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\nu_2(FCD\,Rec, p_0)$	0.828	0.997	0.99997	$1-2.8\cdot10^{-7}$	$1-2.8\cdot10^{-9}$	$1-2.8\cdot10^{-11}$

ФКВД с СКК-2.

Для обнаруживающего ошибки ФКВД, который использует СКК-2 из таблицы 15, количество $f_{per}^{ud}(i,2t)$ ошибок веса 2t, приводящих к ошибочному декодированию, для каждой сигнальной точки *i* приведено в табл. 23.

Как можно видеть распределение количества $f_{per}^{ud}(i,2t)$ ошибок веса 2t, приводящих к ошибочному декодированию обнаруживающего ошибки ФКВД с СКК-2, не зависит от сигнальной точки.

Таблица 23

Значения $f_{per}^{ud}(i,2t)$ *для ФКВД с СКК-2*

t	Сигнальная точка							
	0	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0	0
2	6	6	6	6	6	6	6	6
3	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1

Результаты вычисления $P_{ud}(FCDR, p_0)$ для ФКВД с СКК-2 при различных p_0 представлены в табл. 24.

Таблица 24

Вероятность необнаруженной ошибки для ФКВД с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{ud}(FCDR, p_0)$	$3.94 \cdot 10^{-4}$	$5.76 \cdot 10^{-8}$	$5.98 \cdot 10^{-12}$	$6 \cdot 10^{-16}$	$6 \cdot 10^{-20}$	$6 \cdot 10^{-24}$

Значения энергетического выигрыша в результате применения ФКВД с СКК-2 представлены в табл. 25.

Таблица 25

Энергетический выигрыш ФКВД с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\Delta P(FCDR, p_0)$, дБ	6.628	6.379	6.256	6.194	6.158	6.134

Значения остаточной вероятности ошибочного приема в результате применения ФКВД с СКК-2 представлены в табл. 26.

Таблица 26

Остаточная вероятность ошибочного приема ФКВД с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{res}(FCDR, p_0)$	$9.14 \cdot 10^{-4}$	$6.25 \cdot 10^{-8}$	$6.02 \cdot 10^{-12}$	$6.00 \cdot 10^{-16}$	$6.00 \cdot 10^{-20}$	$6.00 \cdot 10^{-24}$

Значения динамической составляющей потери скорости вследствие переспросов $\nu_2(FCDR, p_0) = 1 - P_{req}(FCDR, p_0) = Q + P_{ud}(FCDR, p_0)$ приведены в табл. 27.

Таблица 27

Динамическая составляющая потери скорости для ФКВД с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\nu_2(FCDR, p_0)$	0.43	0.923	0.992	0.9992	0.99992	0.999992

ФКВДио с расширенной СКК-2.

Примем, как и в предыдущем примере ФКВДио с СКК-2, M=4 и $d_{min} = 4$. Определим, существует ли СКК, обеспечивающая $d_{min} = 4$ для M=4 для сигнальных векторов, количество которых превышает 8.

В табл. 28 представлена СКК-2 из 12 сигнальных векторов с $d_{min} = 4$.

Данная СКК построена следующим образом:

– в качестве первого сигнального вектора выбрана представленная в двоичном виде перестановка $\{0;1;2;3\}$;

– сформированы $f_{per}(2) = 4$ вектора, удаленных от первого сигнального на расстояние d=2: 01 00 10 11, 10 01 00 11, 00 11 10 01, 00 01 11 10;

– для каждого из полученных векторов сформированы еще по 3 вектора, удаленных от них на расстояние d=2. Два из них совпадают между собой, оставшееся множество из 10 векторов образуют сигнальные вектора;

– последним сигнальным вектором является вектор, удаленный от первого сигнального вектора на расстояние d=8.

Таблица 28

СКК-2 для ФКВДио при $M = 4$

СКК	Сигнальные точки
00 01 10 11	0
11 00 10 01	19
01 10 00 11	8
01 00 11 10	7
11 01 00 10	20
10 11 00 01	16
10 00 01 11	12
01 11 10 00	11
00 10 11 01	3
10 01 11 00	15
00 11 01 10	4
11 10 01 00	23

Заметим, что сигнальные точки для СКК из таблицы 28 расположены на числовой оси парами: 0, 3-4, 7-8, 11-12, 15-16, 19-20, 23. Таким образом, соседние вектора с расстоянием Эвклида $D_{i,i+1} = 1$ могут иметь расстояние Хэмминга $d_{i,i+1} = 4$.

Отметим также, что оставшиеся вектора, не вошедшие в СКК из табл. 28, образуют также СКК с $d_{min} = 4$.

Рассмотрим вероятностные характеристики ФКВДио для СКК-2 из табл. 28.

Экспериментально установлено, что данный код позволяет исправить только любые ошибки кратности t=1, а ошибка не обнаруживается тогда и только тогда, когда кодовое слово, соответствующее i-ому сигнальному вектору, будет преобразовано помехой в комбинацию, для которой расстояние Хэмминга до любого другого сигнального вектора не превышает единицу. Распределение количества $f_{per}^{ud}(i,t)$ ошибок веса t, приводящих к ошибочному декодированию, инвариантно по отношению сигнальному вектору i для приведенной в таблице 28 СКК. Значения $f_{per}^{ud}(i,t)$ для нее приведены в табл. 29 ($f_{per}^{ud}(i,t)=0$ при $t\le 2$).

Таблица 29

Значения $f_{per}^{ud}(i,t)$ для ФКВДио с расширенной СКК-2

t	3	4	5	6	7	8
$f_{per}^{ud}(i,t)$	40	10	40	0	8	1

Результаты вычисления $P_{ud}(FCD\,Rec, p_0)$ для различных p_0 представлены в табл. 30.

Таблица 30

Вероятность необнаруженной ошибки для ФКВДио с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{ud}(FCD\,Rec, p_0)$	$2.46\cdot10^{-2}$	$3.81\cdot10^{-5}$	$3.98\cdot10^{-8}$	$4.00\cdot10^{-11}$	$4.00\cdot10^{-14}$	$4.00\cdot10^{-17}$

Поскольку код исправляет только ошибки единичного веса, $f_{per}^{EC}(i,t)=\begin{cases}8, \text{если } t=1,\\ 0, \text{если } t\ne 1.\end{cases}$ Значения вероятности исправления ошибок $P_{EC}(FCD\,Rec, p_0)$ по (7) для различных p_0 совпадают со значениями из табл. 19.

Значения энергетического выигрыша в результате применения ФКВДио с расширенной СКК-2 из таблицы 28 представлены в табл. 31.

Таблица 31

Энергетический выигрыш ФКВДио с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\Delta P(FCD\,Rec, p_0)$, дБ	4.075	4.392	4.527	4.593	4.632	4.657

Значения остаточной вероятности ошибочного приема в результате применения ФКВДио с расширенной СКК-2 представлены в табл. 32.

Таблица 32

Остаточная вероятность ошибочного приема ФКВДио с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{res}(FCD\,Re\,c, p_0)$	$2.93\cdot10^{-2}$	$3.82\cdot10^{-5}$	$3.98\cdot10^{-8}$	$4.00\cdot10^{-11}$	$4.00\cdot10^{-14}$	$4.00\cdot10^{-17}$

Определенные по (13) значения динамической составляющей потери скорости вследствие переспросов для ФКВДио с расширенной СКК-2 представлены в табл. 33.

Таблица 33

Динамическая составляющая потери скорости ФКВДио с СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\nu_2(FCD\,Re\,c, p_0)$	0.838	0.997	0.99997	$1-2.8\cdot10^{-7}$	$1-2.8\cdot10^{-9}$	$1-2.8\cdot10^{-11}$

ФКВД с расширенной СКК-2.

Для обнаруживающего ошибки ФКВД, который использует расширенную СКК-2 из таблицы 28, количество $f_{per}^{ud}(i,2t)$ ошибок веса 2t, приводящих к ошибочному декодированию, для каждой сигнальной точки i приведено в таблице 34. Распределение количества $f_{per}^{ud}(i,2t)$ ошибок веса 2t, приводящих к ошибочному декодированию обнаруживающего ошибки ФКВД с приведенной в табл. 28 СКК-2, инвариантно по отношению сигнальному вектору i.

Таблица 34

Значения $f_{per}^{ud}(i,2t)$ *для ФКВД с расширенной СКК-2*

t	1	2	3	4
$f_{per}^{ud}(i,2t)$	0	10	0	1

Результаты вычисления $P_{ud}(FCDR, p_0)$ для ФКВД с расширенной СКК-2 при различных p_0 представлены в табл. 35.

Таблица 35

Вероятность необнаруженной ошибки для ФКВД с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{ud}(FCDR, p_0)$	$6.56\cdot10^{-4}$	$9.61\cdot10^{-8}$	$9.96\cdot10^{-12}$	$1.00\cdot10^{-15}$	$1.00\cdot10^{-19}$	$1.00\cdot10^{-23}$

Значения энергетического выигрыша в результате применения ФКВД с расширенной СКК-2 представлены в табл. 36.

Значения остаточной вероятности ошибочного приема в результате применения ФКВД с расширенной СКК-2 представлены в табл. 37.

Таблица 36

Энергетический выигрыш ФКВД с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\Delta P(FCDR, p_0)$, дБ	6.318	6.246	6.170	6.131	6.108	6.092

Таблица 37

Остаточная вероятность ошибочного приема ФКВД с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$P_{res}(FCDR, p_0)$	$1.52 \cdot 10^{-3}$	$1.04 \cdot 10^{-7}$	$1.00 \cdot 10^{-11}$	$1.00 \cdot 10^{-15}$	$1.00 \cdot 10^{-19}$	$1.00 \cdot 10^{-23}$

Значения динамической составляющей потери скорости вследствие переспросов $\nu_2(FCDR, p_0) = 1 - P_{req}(FCDR, p_0) = Q + P_{ud}(FCDR, p_0)$ приведены в табл. 38.

Таблица 38

Динамическая составляющая потери скорости для ФКВД с расширенной СКК-2

p_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\nu_2(FCDR, p_0)$	0.431	0.923	0.992	0.9992	0.99992	0.999992

4. Сравнительная оценка кодов

Сравнение ФКВД(ио) с СКК-1 и СКК-2.

Сведем полученные выше результаты и представим их графически. На рис. 5 показаны графики зависимостей вероятностей необнаруженной ошибки от вероятности битовой ошибки p_0 для рассмотренных кодов: ФКВДио с СКК-1 (FCDRec-1) и СКК-2 (FCDRec-2), а также ФКВД с СКК-1 (FCDR-1) и СКК-2 (FCDR-2).

На рис. 6 для этих же кодов показаны графики зависимостей энергетических выигрышей от вероятности битовой ошибки p_0.

Графики остаточных вероятностей ошибочного приема в результате применения рассмотренных кодов ФКВДио и ФКВД с СКК-1 и СКК-2 в зависимости от вероятности битовой ошибки p_0 представлены на рис. 7.

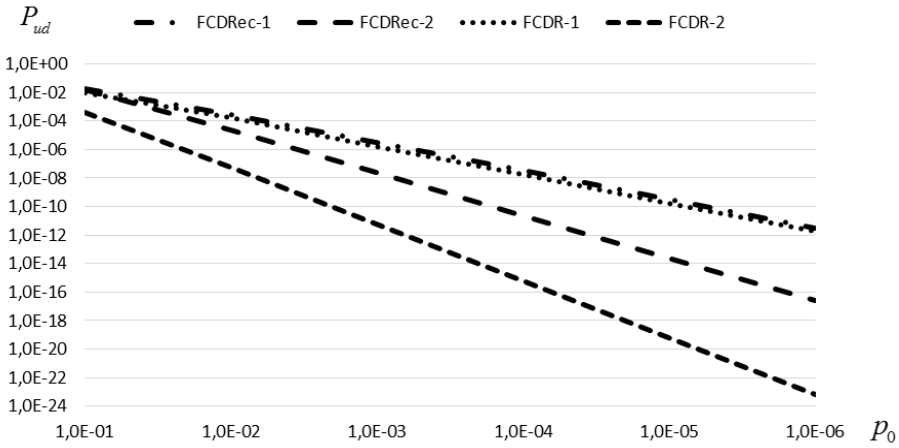


Рис. 5. Графики зависимостей вероятностей необнаруженной ошибки от вероятности битовой ошибки

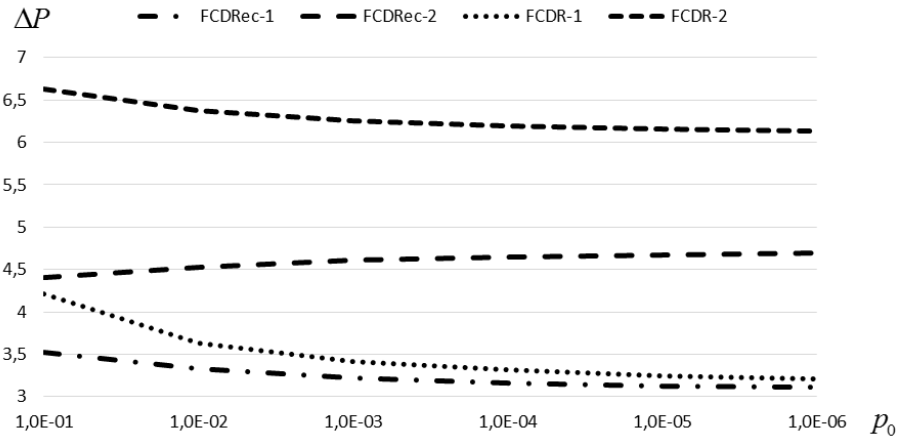


Рис. 6. Графики зависимостей энергетических выигрышей от вероятности битовой ошибки

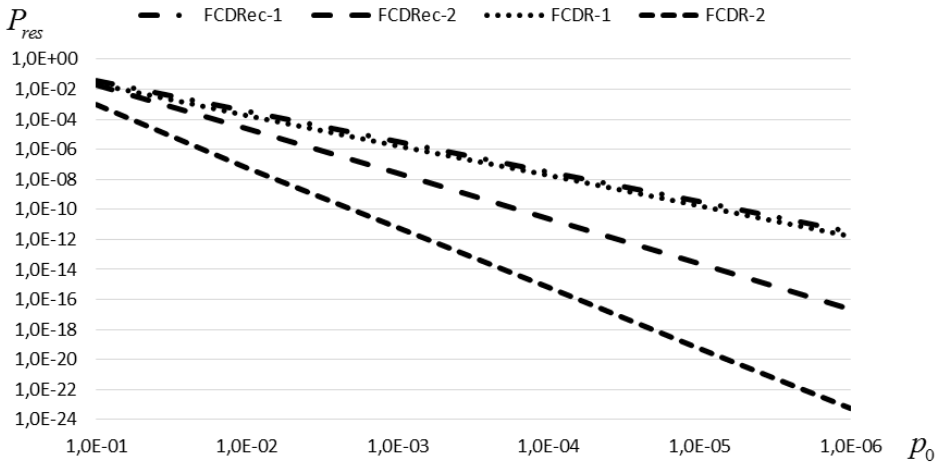


Рис. 7. Графики зависимостей остаточных вероятностей ошибочного приема от вероятности битовой ошибки

На рис. 8 для этих кодов показаны графики зависимостей от вероятности битовой ошибки p_0 величины $1-\nu_2$, показывающей,

насколько близко динамическая составляющая потери скорости приближается к своему максимальному значению.

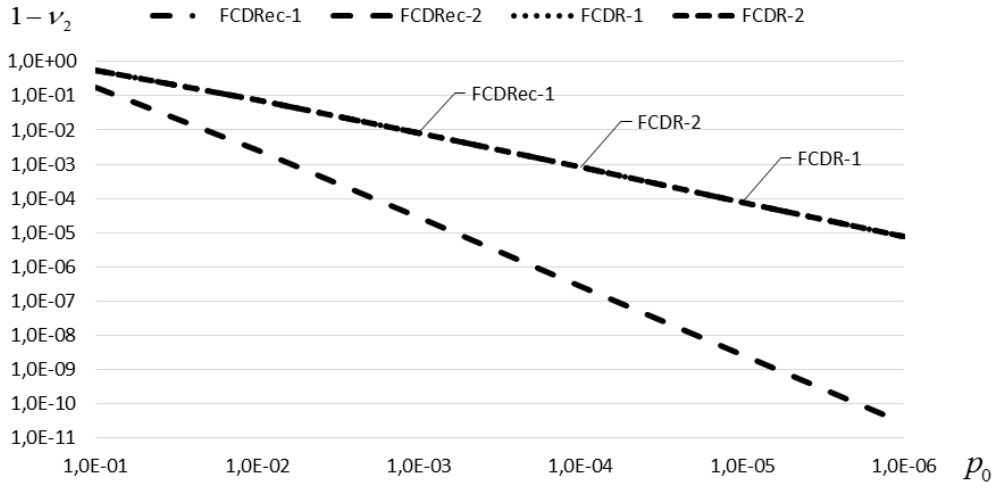


Рис. 8. Графики зависимостей величины $1-\nu_2$ *от вероятности битовой ошибки*

Из представленных графиков следует, что наибольшую вероятность необнаруженной ошибки и наименьший энергетический выигрыш из рассмотренных кодов имеет ФКВДио с СКК-1. Практически в два раза данная вероятность меньше для ФКВД с СКК-1. При этом динамическая составляющая потери скорости для ФКВДио и ФКВД с СКК-1 различаются незначительно (например, при вероятности битовой ошибки $p_0 = 0.1$ отличие составляет $1.27 \cdot 10^{-2}$ (менее 3%) и уменьшается с уменьшением p_0). Поэтому в данном случае при сравнении ФКВДио с СКК-1 и ФКВД с СКК-1 предпочтение следует отдать обнаруживающему ошибки ФКВД. Вместе с тем из этого пока не следует вывод в общем случае о меньшей эффективности ФКВДио по сравнению с ФКВД для СКК первого типа.

Наименьшую вероятность необнаруженной ошибки и наибольший энергетический выигрыш из рассмотренных кодов имеет ФКВД, обнаруживающий ошибки, с СКК-2. Отношение вероятностей необнаруженной ошибки для ФКВД с СКК-2 и СКК-1 пропорционально величине $0.1(p_0)^{-2}$ и при $p_0 = 0.1$ составляет $2.4 \cdot 10^1$, а при $p_0 = 0.001$ – $2.9 \cdot 10^5$ (разница в энергетическом выигрыше при $p_0 = 0.1$ составляет $\Delta P = 2.42$ дБ, а при $p_0 = 0.001$ – $\Delta P = 2.84$ дБ), указывая на бóльшую эффективность СКК-2 для ФКВД.

Использование же СКК-2 для ФКВДио увеличивает по сравнению с ФКВД вероятность необнаруженной ошибки пропорционально величине $(p_0)^{-1}$. Так, при $p_0 = 0.1$ отношение вероятностей необнаруженной ошибки для ФКВДио с СКК-2 и ФКВД с СКК-2 составляет $3.7 \cdot 10^1$, а при

$p_0 = 0.001$ – $4 \cdot 10^3$; разница в энергетическом выигрыше при $p_0 = 0.1$ составляет $\Delta P = 2.23$ дБ, а при $p_0 = 0.001$ – $\Delta P = 1.65$ дБ. Вместе с тем динамическая составляющая потери скорости для ФКВД с СКК-2 практически не отличается от динамической составляющей потери скорости для ФКВДио и ФКВД с СКК-1 (например, при вероятности битовой ошибки $p_0 = 0.1$ отличие от ФКВД с СКК-1 составляет $9.07 \cdot 10^{-3}$ (2,06%) и уменьшается с уменьшением p_0), в то время, как динамическая составляющая потери скорости для ФКВДио с СКК-2 может значительно превосходить динамическую составляющую потери скорости для ФКВД с СКК-2 (например, при вероятности битовой ошибки $p_0 = 0.1$ разность данных показателей для ФКВДио и ФКВД с СКК-2 составляет $0.828 - 0.431 \approx 0.4$ (более 92%) и увеличивается с увеличением p_0). Таким образом, для ФКВДио и ФКВД с СКК-2 становится больше заметен эффект обмена достоверности передачи на пропускную способность, объясненный выше.

Применение СКК-2 вместо СКК-1 для ФКВДио позволяет увеличить динамическую составляющую потери скорости до 82,9% при $p_0 = 0.1$ (на $8 \cdot 10^{-4}$ % при $p_0 = 10^{-6}$), при этом вероятность необнаруженной ошибки уменьшается в 1,23 раза (в $1.35 \cdot 10^5$ раз при $p_0 = 10^{-6}$).

Таким образом, приведенный анализ однозначно указывает, что для рассмотренных кодов ФКВД(ио) с СКК-1 и СКК-2 большей эффективностью обладают ФКВД(ио) с СКК-2. Вместе с тем из этого пока не следует вывод в общем случае о меньшей эффективности СКК первого типа по сравнению с СКК второго типа.

В любом случае, принцип построения СКК играет одну из наиболее важных ролей при проектировании ФКВД с обнаружением и (или) исправлением ошибок.

Сравнение ФКВД(ио) с СКК-2 и расширенной СКК-2.

На рис. 9 показаны графики зависимостей вероятностей необнаруженной ошибки от вероятности битовой ошибки p_0 для следующих кодов: ФКВДио с СКК-2 (FCDRec-2) и расширенной СКК-2 (FCDRec-2ext), а также ФКВД с СКК-2 (FCDR-2) и расширенной СКК-2 (FCDR-2ext).

На рис. 10 для этих же кодов показаны графики зависимостей энергетических выигрышей от вероятности битовой ошибки p_0.

Графики остаточных вероятностей ошибочного приема в результате применения кодов ФКВДио и ФКВД с СКК-2 и расширенной СКК-2 в зависимости от вероятности битовой ошибки p_0 представлены на рис. 11.

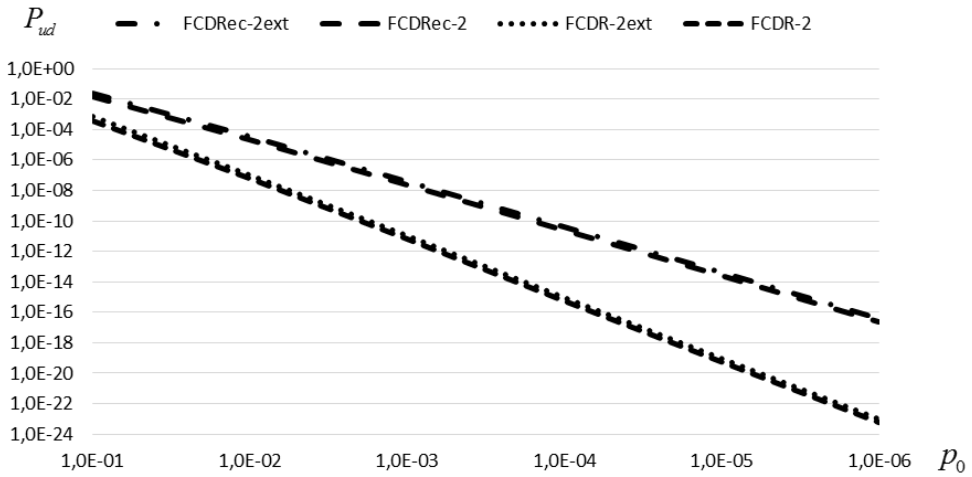


Рис. 9. Графики зависимостей вероятностей необнаруженной ошибки от вероятности битовой ошибки

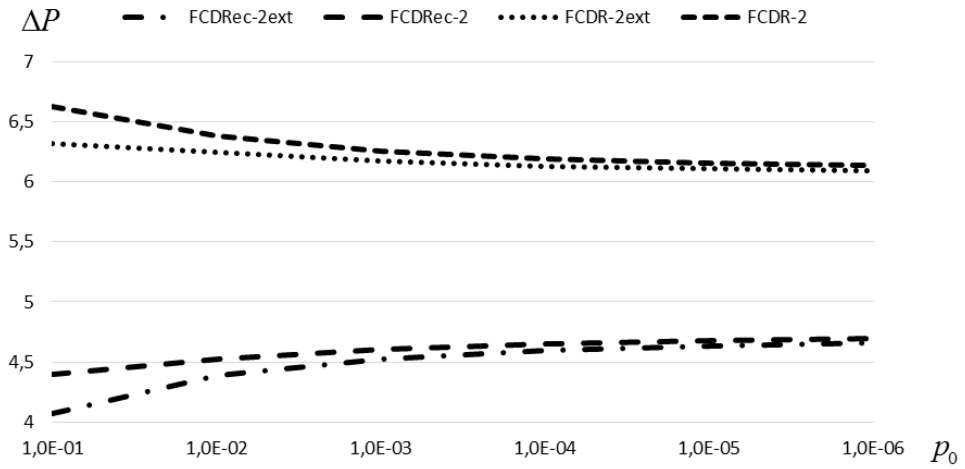


Рис. 10. Графики зависимостей энергетических выигрышей от вероятности битовой ошибки

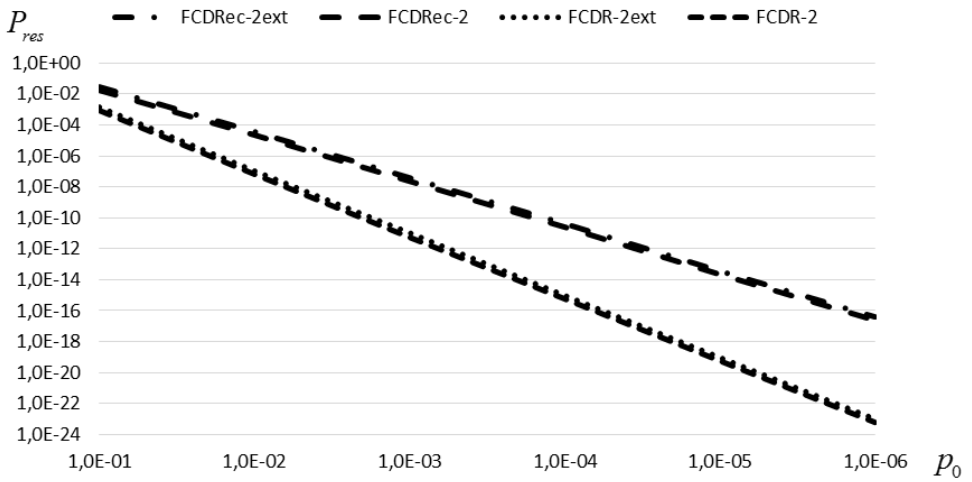


Рис. 11. Графики зависимостей остаточных вероятностей ошибочного приема от вероятности битовой ошибки

На рис. 12 для этих кодов показаны графики зависимостей величины $1-\nu_2$ от вероятности битовой ошибки p_0.

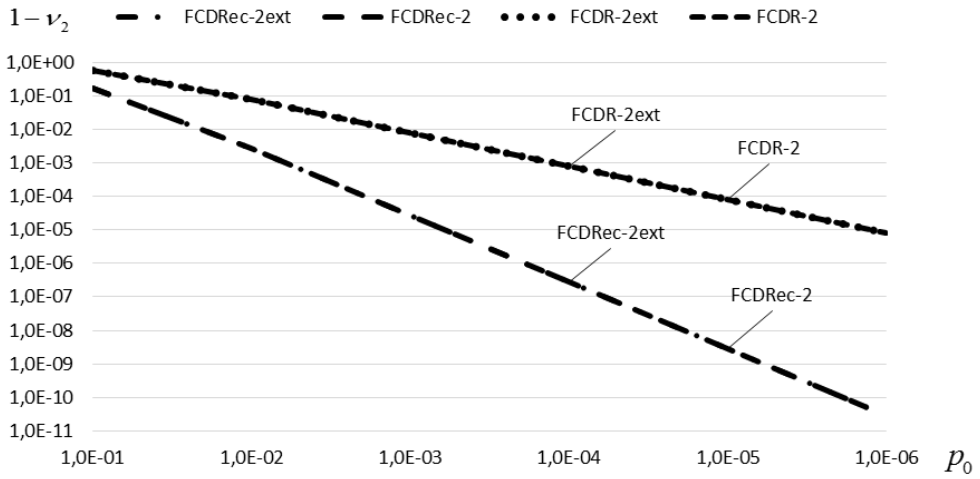


Рис. 12. Графики зависимостей величины $1-\nu_2$ от вероятности битовой ошибки

Из представленных графиков следует, что помехоустойчивость ФКВД(ио) с расширенной СКК-2 с 12 сигнальными векторами уступает помехоустойчивости ФКВД(ио) с СКК-2 с 8 сигнальными векторами.

Так, для ФКВДио:

– вероятность ошибочного декодирования, а также остаточная вероятность ошибочного приема увеличивается в 1,67 раза;

– энергетический выигрыш уменьшается от 0,326 дБ для $p_0 = 0.1$ до 0,038 дБ для $p_0 = 10^{-6}$;

для ФКВД:

– вероятность ошибочного декодирования, а также остаточная вероятность ошибочного приема увеличивается в 1,67 раза;

– энергетический выигрыш уменьшается от 0,310 дБ для $p_0 = 0.1$ до 0,041 дБ для $p_0 = 10^{-6}$.

При этом динамическая составляющая потери скорости для ФКВДио с расширенной СКК-2 практически не отличается от динамической составляющей потери скорости для ФКВДио с СКК-2 (например, при вероятности битовой ошибки $p_0 = 0.1$ относительная величина разности этих величин составляет 1,17% (абсолютная – $9.83 \cdot 10^{-3}$) и уменьшается с уменьшением p_0, достигая при $p_0 = 10^{-3}$ значения $1.59 \cdot 10^{-6}$% (абсолютное значение – $1.59 \cdot 10^{-8}$)). Аналогично динамическая составляющая потери скорости для ФКВД с расширенной СКК-2 практически не отличается от динамической составляющей потери скорости для ФКВД с СКК-2 (например, при вероятности битовой ошибки $p_0 = 0.1$ относительная

величина разности этих величин составляет $6.09 \cdot 10^{-2}\%$ (абсолютная – $2.62 \cdot 10^{-4}$) и уменьшается с уменьшением p_0, достигая при $p_0 = 10^{-3}$ значения $4.02 \cdot 10^{-10}\%$ (абсолютное значение – $3.98 \cdot 10^{-12}$)).

Таким образом, для ФКВД и ФКВДио расширенная СКК-2 из 12 сигнальных векторов проигрывает в помехоустойчивости по сравнению с СКК-2 из 8 сигнальных векторов при сохранении динамической составляющей потери скорости.

Рассмотрим дополнительно для каждого из кодов зависимость относительной скорости передачи $\nu_0 = \nu_1 \cdot \nu_2$ от вероятности битовой ошибки p_0. Графики этих зависимостей приведены на рис. 13. При этом для СКК-2 из 8 сигнальных векторов скорость кода $\nu_1 = \frac{3}{8}$, а для расширенной СКК-2 из 12 сигнальных векторов скорость кода $\nu_1 = \frac{\log_2 12}{8}$.

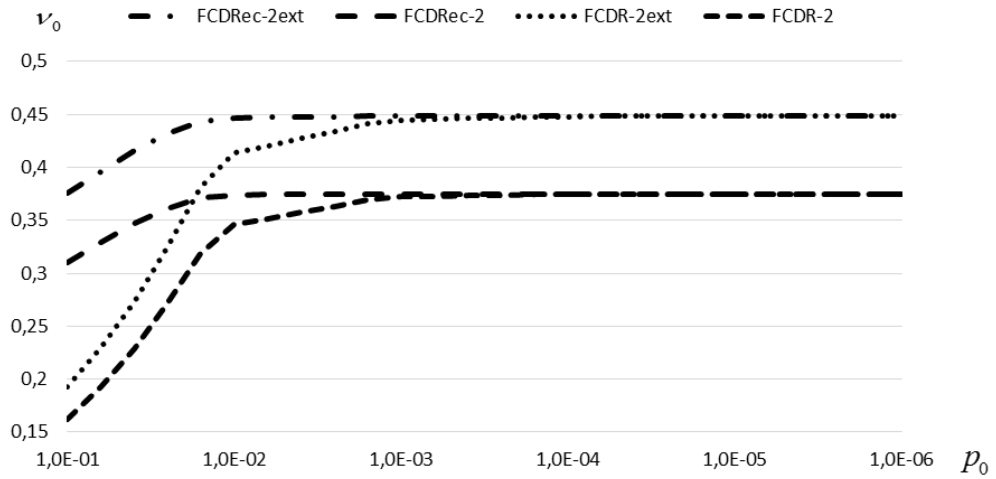


Рис. 13. Графики зависимостей относительной скорости передачи от вероятности битовой ошибки

Анализ графиков рис. 13 подтверждает, что увеличение числа сигнальных векторов увеличивает относительную скорость передачи кода. Рис. 13 также показывает, что:

– относительная скорость передачи для ФКВДио не уступает, а в некоторых случаях существенно превосходит относительную скорость передачи для ФКВД с такой же СКК-2 (например, при вероятности битовой ошибки $p_0 = 0.1$ разность данных показателей для ФКВДио и ФКВД с СКК-2 составляет $0.310 - 0.162 \approx 0.149$, что соответствует более 92% (для ФКВДио и ФКВД с расширенной СКК-2 эта разность равна $0.375 - 0.193 \approx 0.182$, что соответствует более 94%) и уменьшается с увеличением p_0);

– в большинстве случаев относительная скорость передачи для представленных кодов с расширенной СКК-2 с 12 сигнальными векторами превышает относительную скорость передачи кодов с СКК-2 с 8 сигнальными векторами. Вместе с тем существует диапазон значений p_0, для которых относительная скорость передачи ФКВДио с СКК-2 с 8 сигнальными векторами превышает относительную скорость передачи ФКВД с СКК-2 с 12 сигнальными векторами: $\nu_0(\text{FCD Rec} - 2, p_0) - \nu_0(\text{FCDR} - 2\text{ext}, p_0) > 0$ при $p_0 \geq 2.38 \cdot 10^{-2}$;

– увеличение числа сигнальных векторов с 8 до 12 для ФКВД и ФКВДио увеличивает относительную скорость передачи кода на величину от 19,5% при $p_0 = 10^{-6}$ до 20,9% (для ФКВДио) и 19,6% (для ФКВД) при $p_0 = 0.1$, при этом вероятность необнаруженной ошибки увеличивается в 1,67 раза.

Выводы

В процессе проведенного исследования показана возможность факториального кодирования информации, совмещающего функции исправления и обнаружения ошибок, возникающих в канале связи при передаче сообщения. Такое совмещение позволяет повысить динамическую составляющую потери скорости и, как следствие, относительную скорость передачи, по сравнению с обнаруживающим ошибки факториальным кодированием за счет снижения помехоустойчивости кода.

Установлено также, что показатели помехоустойчивости факториального кодирования с восстановлением данных, а также с восстановлением данных и исправлением ошибок не являются инвариантными по отношению к выбору сигнально-кодовой конструкции, если в качестве сигнальных векторов используется некоторое собственное подмножество множества векторов всех возможных перестановок порядка M.

Литература

1. Фауре Э.В. Метод формирования имитовставки на основе перестановок / Э.В. Фауре, В.В. Швыдкий, В.А. Щерба // Захист інформації. – 2014. – №4. – Т. 16. – С. 334 - 340. – Режим доступу: http://jrnl.nau.edu.ua/index.php/ZI/article/view/334/8755.

2. Пат. 106669 Україна, МПК G06F 21/64 (2013.01). Спосіб формування імітовставки / Фауре Е.В., Швидкий В.В., Щерба А.І.; заявник та патентовласник ЧДТУ. – № a201505934; заявл. 16.06.2015; опубл. 10.05.2016, Бюл. № 9.

3. Фауре Э.В., Швыдкий В.В., Щерба А.И. Контроль целостности информации на основе факториальной системы счисления / Э.В. Фауре, В.В. Швыдкий, А.И. Щерба // Journal of Qafqaz University. Mathematics and computer science. – 2016. – №2.

4. Пат. 107655 Україна, МПК G06F 21/64 (2013.01), H04L 1/16 (2006.01). Спосіб контролю цілісності інформації / Рудницький В.М., Фауре Е.В., Швидкий В.В., Щерба А.І.; заявник та патентовласник ЧДТУ. – № a201505937; заявл. 16.06.2015;

опубл. 24.06.2016, Бюл. № 12.

5. Фауре Э.В. Комбинированное факториальное кодирование и его свойства / Э.В. Фауре, В.В. Швыдкий, В.А. Щерба // Радіоелектроніка, інформатика, управління. – 2016. – №3. – С. 80-86. – Режим доступу: http://www.csit.narod.ru/ric/riu_2016_3.pdf.

6. Пат. 107657 Україна, МПК H03M 13/09 (2006.01), H04K 1/06 (2006.01), G09C 1/06 (2006.01). Спосіб комбінованого кодування інформації / Рудницький В.М., Фауре Е.В., Швидкий В.В., Щерба А.І.; заявник та патентовласник ЧДТУ. – № a201508148; заявл. 17.08.2015; опубл. 24.06.2016, Бюл. № 12.

7. Фауре Э.В. Факториальное кодирование с восстановлением данных / Э.В. Фауре // Вісник Черкаського державного технологічного університету. – 2016. – №2. – С. 33-39. – Режим доступу: http://vtn.chdtu.edu.ua/article/view/82932/78400.

8. Фауре Э.В. Метод повышения эффективности факториального кодирования с восстановлением данных / Э.В. Фауре // Вісник Черкаського державного технологічного університету. – 2016. – №4. – С. 57-61.

9. Фауре Э.В. Факториальное кодирование с несколькими контрольными суммами / Э.В. Фауре // Вісник Житомирського державного технологічного університету. – 2016. – №3. – С. 104-113. – Режим доступу: http://vtn.ztu.edu.ua/article/view/86481/82932.

10. Финк Л.М. Теория передачи дискретных сообщений. – М.: Советское радио, 1970. – 728 с.

11. Теплов Н. Л. Помехоустойчивость систем передачи дискретной информации / Н.Л. Теплов. – М.: Связь, 1964. – 360 с.

12. Питерсон У. Коды, исправляющие ошибки / У. Питерсон, Э. Уэлдон; [пер. с англ. под ред. Р.Л. Добрушина, С.И. Самойленко]. – М.: Мир, 1976. – 590 с.

13. Конвей Дж. Упаковки шаров, решетки и группы: в 2 т. / Дж. Конвей, Н. Слоэн; при участии Э. Баннаи и др.; перевод с англ. С. Н. Лицына и др. 2 т. – М. : Мир, 1990.

МЕТОД СЖАТИЯ ВИДОВЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ДВОХКОМПОНЕНТНОГО ПРЕДСТАВЛЕНИЯ АПЕРТУРНЫХ СОСТАВЛЯЮЩИХ

Хименко В.В., Додух А.Н., Супрун О.В., Окладной Д.Е.

Введение

Современное общество характеризуется высоким уровнем использования новейших достижений в сфере информационно-телекоммуникационных технологий. В соответствии с Национальной программой информатизации государства главная задача информационно-телекоммуникационных систем состоит в обеспечении своевременной доставки достоверной информации в независимости от расстояния между источником и получателем.

В свою очередь, средства телекоммуникаций имеют особенности, которые зависят от реализации их на базе проводных или беспроводных сетей передачи данных. Проведенный анализ показал, что наибольшие сложности на пути организации видеоинформационного обеспечения возникают для сектора дистанционного формирования и доставки видеоданных с использованием беспроводных инфокоммуникационных технологий [1].

Основная доля видео-приложений сектора дистанционной передачи видеоданных характеризуется требованием относительно обеспечения высокого качества изображений. Качество видеокадров определяется не только наличием искажений (шумов), но и разрешающей способностью. Поэтому требование относительно достижения высокого качества изображений связано с ростом объемов цифрового представления кадров.

На основе проведенного анализа результатов оценки времени на обработку и передачу видеоданных по каналам связи, можно заключить следующее:

1) в режиме отсутствия предварительного сжатия изображений, время доставки видеоданных достигает десятков минут для скорости передачи 9,6 Кбит/с и десятков секунд для скорости передачи порядка 4 Мбит/с;

2) в режиме предварительной компрессии с использованием существующих технологий, достигается снижение времени доставки видеоданных. Однако для пикового отношения сигнал/шум (ПОСШ) на уровне 40 дБ, что соответствует хорошему качеству визуального восприятия изображений, получены следующие результаты:

- для низких скоростей передачи и в случае обработки изображений размером, превышающем 3504x2336 элементов, время доставки данных находится на уровне 15 - 150 с.
- для наиболее распространенных технологий беспроводной передачи данных, обеспечивающих скорость передачи на уровне

256 Кбит/с, и наиболее широко используемых форматов кадров на уровне 2048x1536 элементов, время доставки изменяется в пределах 40 - 50 с.

- время передачи сжатых видеоданных резко возрастает в случае снижения скорости передачи по каналам связи и увеличения объемов передаваемых данных.

Отсюда можно утверждать, что существующие технологии доставки видеоданных, формируемых дистанционно и передаваемых с использованием беспроводных каналов связи, не соответствуют требованиям служб предоставляющих услуги видео-сервисов.

По результатам исследований можно подытожить следующее:

а) наиболее критичным с позиции обеспечения своевременной доставки информации является сектор дистанционной видеосъемки;

б) предварительное сжатие изображений позволяет снизить время доставки видеоданных. Однако для дистанционных технологий обработки и передачи изображений снижение задержки на доставку видеоданных не превышает 50%. Это не обеспечивает соответствие требованиям видеоинформационных приложений.

Поэтому возникает ***противоречие*** между характеристиками доставки (задержка на доставку, качество визуального восприятия) видеоданных с использованием существующих средств телекоммуникаций на базе беспроводных технологий и требованиями видео-приложений.

Значит, снижение временных задержек на дистанционную обработку и передачу видеоданных с использованием беспроводных средств телекоммуникаций является ***актуальной научно-прикладной задачей.***

Сформулированную задачу в условиях заданных характеристик беспроводных технологий обработки и передачи данных ***предлагается*** решать на базе развития технологий компрессии.

Основной материал

С одной стороны, использование технологий сжатия видеоданных (ТСВ) позволяет сократить объем передаваемых данных, а с другой стороны, использование технологий компрессии приводит к возникновению задержки на обработку видеоданных, обусловленную выполнением процессов сжатия и восстановления изображений. Такого рода задержки зависят от количества выполняемых операций, отводимых на сжатие и восстановление изображений, содержащих некоторое количество элементов.

Сжатие изображений достигается за счет устранения избыточности. В пространстве обработки кадра различают статистическую, психовизуальную и структурную виды избыточности. Наибольшая степень сжатия достигается в результате устранения психовизуальной избыточности. Процесс исключения психовизуальной избыточности сопровождается внесением искажений в изображениях. Поэтому чаше

всего методы обработки изображений разделяют на два класса в зависимости от наличия искажений. Первый класс методов не содержит в себе механизмов, связанных с устранением психовизуальной избыточности. Для методов такого класса достигается нулевой уровень искажений. Методы второго класса, наоборот, базируются на механизмах, обеспечивающих сокращение психовизуальной избыточности. В этом случае уровень искажений будет отличен от нулевого значения. Для оценки уровня искажений используются различные метрики как объективного (количественного) типа, так и субъективного (экспертные оценки). Количественные метрики базируются, в основном, на среднеквадратическом показателе погрешности.

Предлагается проводить совершенствование технологий компрессии с контролируемым уровнем внесения искажений.

Для этого исследуем причины возникновения следующих недостатков:

1) ухудшения качества реконструируемых изображений с ростом степени сжатия;

2) повышения количества операций на обработку видеоданных в условиях сохранения требуемого качества изображений после декомпресии.

Рассмотрим первую причину. Степень сжатия для технологий JPEG-платформы достигается на основе учета структурно-психовизуальных и статистических свойств трансформант. При этом для методов JPEG и JPEG2000 такие свойства проявляются по-разному. Это главным образом определяется этапом предварительной обработки.

Для рекомендаций JPEG предварительный этап состоит в выполнении дискретного косинусного преобразования (ДКП). Это объясняет то, что для технологий на базе JPEG, с одной стороны, достигается повышение степени сжатия для оптических изображений и текстурных участков изображений. В то время как резко снижается эффективность при сжатии изображений, насыщенных резкими перепадами и мелкими деталями. В противном случае рост степени сжатия для такого класса изображений сопровождается появлением блочного эффекта на разжатых изображениях, вплоть до разрушения их отдельных фрагментов. Значит, механизм выполнения ДКП для локально-равномерных сегментов в случае отсутствия обратной связи оказывается не чувствительным к степени насыщенности обрабатываемых фрагментов. Поэтому для фиксированной стратегии квантования для насыщенных изображений, искажения будут проявляться в большей степени как с позиции визуальной оценки, так и с позиции ПОСШ. Использование же обратной связи в процессе кодирования приводит к повышению количества операций на обработку и, как следствие, росту задержки на сжатие.

Технологии JPEG 2000 отличаются тем, что на предварительном этапе выполняется дискретное вейвлет-преобразование (ДВП). Для таких технологий незначительно повышается степень сжатия насыщенных изображений. С другой стороны, наоборот, при сжатии текстурных участков эффективность снижается.

Общим недостатком существующих технологий на JPEG платформе является то, что для насыщенных изображений не обеспечивается сжатие в режиме сохранения высокого качества декодированных изображений

Поэтому предлагается использовать альтернативное направление предварительной обработки, состоящее в построении апертурных структур [2, 3]. Апертура - участок изображения, значения элементов которого находятся в пределах ограниченного динамического диапазона (рис. 1).

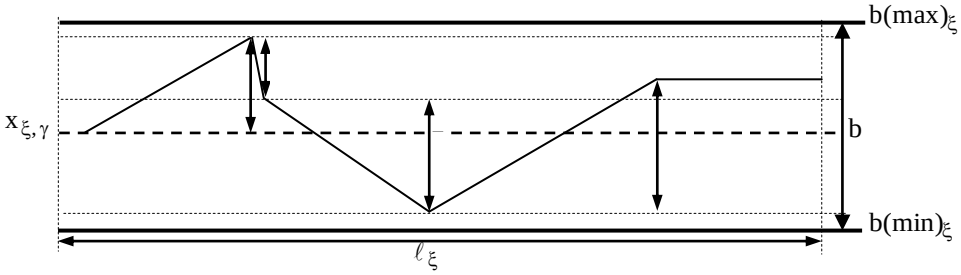


Рис. 1. Структура апертуры изображения

Преимуществом предложенного подхода является:

а) относительно снижения задержки на обработку то, что:

- на выявление апертур затрачивается относительно небольшое количество операций сравнения, а операции умножения и сложения не используются;

- сокращается время на дальнейшее компрессионное кодирование в результате снижения количества обрабатываемых данных (перехода от описания последовательностей элементов к их длинам);

б) относительно повышения степени сжатия и снижения задержки на передачу данных по каналу связи то, что:

- выявляются структурные и статистические закономерности, обусловленные наличием когерентных областей, содержащих элементы, значения которых либо одинаковые, либо отличаются незначительно. Это создает потенциальные возможности для устранения структурной и статистической избыточности изображений;

- апертурное представление изображений является более гибким к изменению структурных свойств фрагментов изображений по сравнению с ортогональными преобразованиями. Апертурное представление на изменение структурного содержания проявляется в изменении длины аппретуры, описывающей цепочку одинаковых элементов изображения. Это позволяет сохранить информацию о резких перепадах, и обеспечить

более высокое качество визуальных оценок изображений, насыщенных мелкими деталями и перепадами.

Рассмотрим особенности построения апертурных структур. Апертура характеризуется двухкомпонентной системой координат.

Первая компонента определяет позицию апертуры $X^{(\xi)}$ в кадре изображения. Для этого используется:

- координата первого элемента $x_{\xi,\gamma}$ апертуры, относительно которого осуществляется определение величин $b(min)_\xi$ и $b(max)_\xi$;

- длина ℓ_ξ апертуры, равная количеству подряд расположенных элементов, для которых выполняется условие $x_{\xi,\gamma+r} \in [b(min)_\xi; b(max)_\xi]$, $r = \overline{0, \ell_\xi - 1}$, где:

$$b(max)_\xi = x_{\xi,\gamma} + b/2;\ b(min)_\xi = x_{\xi,\gamma} - b/2,$$

где $b(min)_\xi$, $b(max)_\xi$ - значения соответственно нижней и верхней границ (ξ)-й апертуры; b - высота апертуры.

Вторая координата позиционирует апертуру по шкале яркости (цветности). Здесь используется вектор P_a параметров аппроксимирующей функции.

Выявление апертур позволяет локализовать свойства изображений в пространственно-временной области. Такие свойства могут быть статистическими, структурными и психовизуальными, на основе чего уменьшается избыточность изображений.

Для устранения избыточности в режиме контролируемых потерь качества восстанавливаемых изображений и ограниченности времени на их обработку ***предлагается*** использовать среднее значение $\bar{x}_\xi$ по всем элементам апертуры:

$$\bar{x}_\xi = \Big(\sum_{r=0}^{\ell_\xi - 1} x_{\xi,\gamma+r}\Big) / \ell_\xi .$$

Такой подход дает возможность сократить сложности процесса обработки изображений. Управление степенью вносимых искажений обеспечивается на основе изменения высоты b апертуры. Степень компрессии $k_{сж}$ будет также зависеть от высоты апертуры b и определяться по формуле:

$$k_{сж} = d\, Z_{стр}\, Z_{стб} / D = d\, Z_{стр}\, Z_{стб} / (D_x + D_\ell), \qquad (1)$$

где D_x, D_ℓ - суммарные объемы цифрового представления соответственно для координат начального элемента апертур и для их длин:

$$D_x = d\nu_a;\ \ D_\ell = \sum_{\xi=1}^{\nu_a} \ell og_2 \ell_\xi . \qquad (2)$$

Здесь ν_a - количество апертур.

В тоже время предложенный подход для предварительной обработки в технологиях компрессии имеет следующие недостатки:

1) обусловленные условиями формирования апертур:

- в случае обработки фрагментов изображений, насыщенных мелкими деталями, и в условиях заранее заданной высоты апертуры происходит уменьшение длин апертур, и как следствие, происходит резкое снижение степени сжатия, вплоть до увеличения первоначального объема. Тогда $D_x \approx d Z_{стр}\ Z_{стб}$, а $k_{сж} = 1$;

- для варианта выбора заранее высокого значения высоты апертуры, проявляется недостаток, связанный с увеличением ошибки аппроксимации и как следствие к ухудшению качества реконструируемых изображений;

2) вызванные тем, что существующие подходы относительно сжатия апертур базируются в основном на формировании статистических кодовых конструкций. Это является причиной таких недостатков [4]:

- в результате выполнения условия префиксности неравномерных кодовых слов, построенных для длин структурных составляющих изображений, происходит увеличение длины кодовых комбинаций и дополнительное снижение помехоустойчивости кодограмм к ошибкам в канале связи;

- формируются неравномерные кодовые комбинации с неконтролируемой длиной, что обуславливает необходимость использования разделителей, и как результат повышение объема сжатых данных;

- не учитываются закономерности, которые можно выявить в случае блочной обработки.

Следовательно, можно заключить, что:

- с одной стороны предварительная обработка изображений на основе построения апертурных структур позволяет снизить влияние недостатков, относительно степени сжатия и качества реконструируемых изображений, и создать возможность для выявления структурных закономерностей. Это обеспечивает снижение задержки на обработку и повышение степени сжатия в условиях требуемого уровня качества декодированных изображений;

- с другой стороны выявление апертур в условиях заданной ее высоты и последующего использования технологий статистического кодирования характеризуется рядом недостатков. Это ограничивает возможности предложенной технологии на основе выявления апертур.

Основная причина снижения степени сжатия изображений на основе апертурной аппроксимации фрагмента изображения связана с увеличением

объема сжатого представления построчно-масштабирующей составляющей [5]. В соответствии с этим, необходимо повышать эффективность обработки построчно-масштабирующей составляющей.

Представим последовательность аппроксимирующих величин h_ξ апертурного описания в виде n-мерного вектора H_n, т.е.

$$H_n = \{h_1, ..., h_j, ..., h_n\}.$$

Полученная совокупность в условиях апертурного описания является ***координатно-яркостной*** составляющей архитектуры изображения. На основе координатно-яркостной составляющей формируется ***построчно масштабирующая компонента*** (ПМК). Данная компонента несет наибольшее количество информации о яркостных характеристиках фрагмента изображения. Физический смысл масштабирования заключается в том, что последовательность элементов изображений заменяется одной аппроксимирующей величиной. Для апертурного представления масштабирование применяется для строк изображения и описывает неравномерное количество элементов.

Дальнейшую обработку построчно-масштабирующей составляющей предлагается проводить по интегрированному принципу. Такой подход относительно скалярной обработки обеспечивает потенциал для учета большего количества информации. Это повышает в конечном итоге возможность для выявления новых закономерностей, свойственных совместному анализу компонент вектора H_n. В силу нестационарности насыщенных реалистических изображений увеличивается сложность для построения адекватных статистических моделей, описывающих построчно-масштабирующие составляющие. При этом закон распределения аппроксимирующих величин стремится к равномерному. С другой стороны, в силу описания апертурами когерентных областей происходит снижение корреляции между элементами ПМК. По этим причинам подход, основанный на выявлении статических закономерностей, не приведет к созданию потенциала для дополнительного повышения степени сжатия. Поэтому ***предлагается*** использовать новые структурные характеристики, свойственные для построчно-масштабирующей составляющей.

Для ***предложенного*** подхода методология сокращения избыточности базируется на формировании фрагменту изображения двух составляющих, а именно:

– неравномерной координатно-структурной составляющей. Такая составляющая формирует локально-структурную архитектуру фрагмента изображений. Компонентами такой составляющей являются длины апертур, выявляемых вдоль строк изображения;

– построчно-масштабирующей составляющей, которая определяет яркостную и цветовую насыщенность архитектурной формы фрагмента

изображения. Компонентами такой формы являются аппроксимирующие яркостные (цветовые) величины апертур.

В процессе построения метода кодирования ***предлагается*** организовывать следующие подходы, а именно [6]:

1. Формировать кодовое описание заданной длины. Например, для хранения в компактном виде видеоинформации в системах резервного копирования, хранилищах данных, на внешних носителях информационно-вычислительных систем. Здесь кодовым словом $D_{nec} = D_{proc}$ будет машинное слово равномерной длины, принимающая значения от 16 до 64 бит в зависимости от системы.

2. Формировать двухкомпонентное кодовое представление на базе совместного использования элементов координатно-структурного и построчно-масштабного представления фрагмента изображения. Это обеспечит обработку целостной информации о фрагменте изображения. Формирование кодовой комбинации ***предлагается*** осуществлять на основе ***двухкомпонентного интегрированного принципа***. В этом случае в отличие от бит-ориентированного принципа добавочная группа разрядов формируется на основе взвешенного добавления компоненты апертурно-яркостного описания фрагмента изображения.

Чтобы устранять интегрированную структурную избыточность, учитывая особенности массивов аппроксимирующих величин, ***предлагается*** осуществлять их ***построчную*** обработку.

Построчная обработка массивов аппроксимирующих величин:

– создает возможность для выявления дополнительных структурных закономерностей, обусловленных неравномерностью соседних элементов, т.е. наличием яркостных (цветовых) перепадов между соседними компонентами построчно-масштабирующей формы;

– сократить количество операций для определения количества элементов, для которых формируется первая часть ДК в процессе отбора элементов для формирования составляющих обобщенного кода на основе соответственно элементов массива аппроксимирующих величин и массивов длин апертур.

Построчная обработка координатно-структурной составляющей обеспечивает возможность учитывать ограниченность динамических диапазонов.

Важным условием также является обеспечение кодирования с ограниченной вычислительной сложностью. Здесь использование построчной обработки позволяет сократить сложность вычислений за счет исключения необходимости учета не стационарности переходов между строками.

Требуется учитывать, что по условию формирования апертурного описания изображения допускается, что апертура содержит элементы видеоданных, отличающиеся друг от друга в некотором диапазоне (в

диапазоне равном высоте апертуры). В этом случае в процесс обработки вносятся погрешности. Значит, внесение погрешностей в процессе обработки форм апертурного описания приводит к размножению ошибок, в результате чего, ухудшается качество восстановленных изображений. Поэтому для обеспечения заданного уровня достоверности необходимо осуществлять обработку массивов апертурного описания изображения без внесения погрешности.

Предлагается использовать новые структурные характеристики, свойственные для построчно-масштабирующей составляющей.

Проведем выявление таких свойств. В условиях формирования апертур для элементов ПМК, характерны следующие закономерности:

1. Первая закономерность определяется семантическим содержанием изображений и неравномерным распределением яркости в кадре. Закономерность такого типа состоит в том, что динамический диапазон вектора H_n будет иметь ограниченные значения. Тогда, если выделить в ПМК нижний $h_{i,min}$ и верхний $h_{i,max}$ уровни его динамического диапазона, получим следующее неравенство:

$$h_{i,min} \le h_1,\dots,h_j,\dots,h_n \le h_{i,max}. \qquad (3)$$

2. Вторая закономерность вытекает из особенностей формирования ПМК в процессе выявления апертур. Такое свойство проявляется в том, что для аппроксимирующих величин смежных апертур $X^{(\xi)}$ и $X^{(\xi+1)}$ (где $\xi = \overline{1,n}$) выполняется условие

$$h_\xi \ne h_{\xi+1},\ \xi = \overline{1,n}. \qquad (4)$$

Проведем теперь группировку последовательности аппроксимирующих величин апертур в локальные области. Для чего построим теперь из отдельных векторов H_n двумерные массивы $H^{(\nu)}_{m,n} = \{h_{ij}\}$ размером m строк и длиной n элементов (где ν - количество массивов, которое можно сформировать на основе выявления апертур для всего изображения). В этом случае построчно-масштабирующая составляющая будет сегментироваться по двумерным массивам.

Выявленные свойства построчно-масштабирующей составляющей позволяют использовать новое представление, основанное на таких принципах.

Тогда первый принцип заключается в существовании ограничений на нижний и верхний уровни массива ПМК. Это задается следующими неравенствами:

1) относительно ограничений на верхний уровень

$$h_{ij} \le h_{i,max},\ \ h_{i,max} = \max_{1\le j\le n}\{h_{ij}\},\ \ i=\overline{1,m}; \qquad (5)$$

2) относительно ограничений на нижний уровень

$$h_{i,min} > 0,\ h_{i,min} = \min_{1 \le j \le n} \{h_{ij}\},\ i = \overline{1,m}. \quad (6)$$

С учетом выражения (6) можно утверждать, что будет выполняться неравенство:

$$h'_{ij} = h_{ij} - h_{i,min} < h_{ij}. \quad (7)$$

Отсюда выгоднее перейти к рассмотрению массивов ПМК в дифференциальном пространстве, используя для этого левую часть неравенства (7). Тогда введем дифференциально-представленную последовательность ΔH_n:

$$\Delta H_n = \{h_{i1} - h_{i,min}, \ldots, h_{ij} - h_{i,min}, \ldots, h_{in} - h_{i,min}\}. \quad (8)$$

Второй принцип относится к особенностям выявленных ограничений на динамические диапазоны массивов ПМК. Он состоит в том, что динамические диапазоны в строках могут быть неравными друг другу, т.е.:

$$h_{i,max} \neq h_{u,max}, \text{ где } i \neq u \text{ и } i,u = \overline{1,m}.$$

Третий принцип состоит в рассмотрении значений аппроксимирующих величин h'_{ij} как элементов, имеющих следующие динамические диапазоны:

- динамический диапазон элемента ПМК массива $\Delta H^{(v)}_{m,n}$ с координатами (1;1) будет равен $w(h)_{11} = h_{1,max} - h_{1,min} + 1$, т.к. $h'_{11} \in [0;\ h_{1,max} - h_{1,min}]$;

- для всех остальных элементов ПМК будет выполняться условие (4). Значит, область значений каждого последующего элемента будет исключать из своего содержания значение предыдущего элемента. Соответственно динамический диапазон уменьшается на единицу и равен $w(h)_{ij} = h_{i,max} - h_{i,min}$, где $i = \overline{2,m}$ для $j = 1$ и $i = \overline{1,m}$ для $j \ge 2$. Значения элементов будут находиться в диапазоне $h'_{ij} \in [0;\ h_{i,max} - h_{i,min} - 1]$.

В результате обобщения предложенных принципов можно заключить, что для вектора аппроксимирующих величин апертур формируются массивы $\Delta H^{(v)}_{m,n}$, элементы которых удовлетворяют следующим условиям:

$$h'_{11} \le w(h)_{11} = h_{1,max} - h_{1,min} + 1;\ h'_{ij} \le w(h)_{ij} = h_{i,max} - h_{i,min}, \quad (9)$$

$$i = \overline{2,m} \text{ для } j = 1 \text{ и } i = \overline{1,m} \text{ для } j \ge 2.$$

В этом случае для массива $\Delta H^{(v)}_{m,n}$ можно сформулировать следующую интерпретацию с позиции структурно-интегрированного описания.

Определение. Строки массива $\Delta H^{(v)}_{m,n}$, для элементов которых выполняются условия (9), так, что в общем случае $w(h)_{ij} \neq w(h)_{uv}$, $i \neq u$,

$j \neq v$ и $i, u = \overline{1, m}$, $j, v = \overline{1, n}$, называются одномерными адаптивными позиционными числами (АПЧ) с неравными соседними элементами и с системой оснований, задающуюся вектором $W(h) = \{w(h)_{ij}\}$.

В соответствии с определением, ***адаптивное позиционное число с неравными соседними элементами*** (АПЧ) образуется на основе одномерных позиционных чисел с неравными соседними элементами, расположенными в строках обрабатываемых массивов. Это позволяет рассматривать процесс кодирования адаптивного позиционного числа как формирование кодов для отдельных строк массива аппроксимирующих величин, которые являются ***одномерными адаптивными позиционными числами* (ОАПЧ) *с неравными соседними элементами.***

Поэтому решение задачи, а именно формирование кодового описания предлагается осуществлять в рамках структурного подхода на базе кодовых конструкций для позиционных чисел.

Построение выражения для кодирования ОАПЧ осуществляется в два этапа. Первый этап заключается в определении позиции строки массива $\Delta H_{m,n}^{(\nu)}$ во множестве допустимых позиционных чисел с адаптивным основанием. На втором этапе организуется получение индекса текущего ОАПЧ во множестве позиционных чисел с дополнительным запретом на равенство соседних элементов. Для этого требуется исключить количество запрещенных позиционных чисел, которые содержат равные соседние элементы. Принцип назначения индексов позиций ОАПЧ в допустимом множестве организуется по лексикографическому правилу.

Избыточное количество разрядов, т.е. $\Delta D > 0$, в процессе формирования двухкомпонентного кода (ДК) для элементов массивов аппроксимирующих величин, обусловлено не кратностью величины динамических диапазонов $w(h)_i - 1$ элементов массива аппроксимирующих величин степени двойки, в случае представления кодового описания машинным словом.

Начальное значение ДК $E(h)_{i,j}^{(i,\gamma)}$ формируется на базе элементов строки массива $\Delta H_{m,n}^{(\nu)}$, рассматриваемых как адаптивное одномерное позиционное число с неравными соседними элементами. В общем случае значение $E(h)_{i,j}^{(i,\gamma)}$ определяется по формуле:

$$E(h)_{i,j}^{(i,\gamma)} = \sum_{\phi=\gamma}^{j} (h_{i,\phi} - \text{sign}\,(1 - \text{sign}\,(h_{i,\phi-1} - h_{i,\phi}))) (w(h)_i - 1)^{j-\gamma+1-\phi}, \quad (10)$$

где $(i;\gamma)$, $(i;j)$ - координаты соответственно начального и конечного элементов i-й строки, на базе которых формируется ДК.

Отсюда получаем, количество $\nu(h,i)_\xi$ элементов i-й строки, для которых формируется ξ-й код, где $\nu(h,i) = j - \gamma + 1$. Величина динамического диапазона для элементов строки фиксирована, и равна $w(h)_i - 1$. Тогда на основе соотношения (2.21) количество элементов $\nu(h)_\xi$ определяется из условия

$$\ell og_2 E(h)_{i,j}^{(i,\gamma)} \le D(h)_{i,j}^{(i,\gamma)} = \nu(h,i)_\xi \ \ell og_2 (w(h)_i - 1) \le D_{nec}, \quad (11)$$

что позволяет исключить возможность переполнения кодового слова, имеющего заранее заданную длину D_{nec}.

Откуда величина $\nu(h,i)_\xi$ равна

$$\nu(h,i)_\xi = [D_{nec} / \ell og_2 (w(h)_i - 1)]. \quad (12)$$

Строчная обработка массивов аппроксимирующих величин апертур позволяет сократить вычислительные затраты для определения количества $\nu(h,i)_\xi$ элементов, и проводить процесс кодирования за один проход.

В случае наличия избыточных разрядов, т.е. $\Delta D \ne 0$, формирование второй компоненты ДК проводится на базе элементов текущей α-й строки массива $\Delta L_{m,n}^{(\nu)}$ длин апертур.

Рассмотрим основные этапы построения обобщенного двухкомпонентного кода.

Первый этап заключается в формировании координатно-структурной и построчно-масштабных составляющих фрагмента изображении. Для этого осуществляется выявление апертур и построение массивов $\Delta H_{m,n}^{(\nu)}$ аппроксимирующих величин и $\Delta L_{m,n}^{(\nu)}$ длин апертур.

Выявление апертур проводится по строкам кадра в направлении строчной развертки. Используется условие $x_{\xi,\gamma+r} \in [b(min)_\xi; b(max)_\xi]$, $r = \overline{0, \ell_\xi - 1}$, где ℓ_ξ - длина текущей апертуры, $b(min)_\xi$, $b(max)_\xi$ - значения соответственно нижней и верхней границ (ξ)-й апертуры, которые зависят от высоты b апертуры. В противном случае, когда $x_{\xi,\ell_\xi} \notin [b(min)_\xi; b(max)_\xi]$, то начинает строиться следующая апертура. Выявление аппретур заканчивается тогда, когда обработан последний элемент $x_{z_{lin}, z_{col}}$ кадра изображения.

Образование массивов $\Delta H_{m,n}^{(\nu)}$ и $\Delta L_{m,n}^{(\nu)}$ проводится в направлении строк, что позволяет выявить дополнительные структурные закономерности, и обеспечить потенциальные возможности для устранения избыточности.

Целостность реконструкции фрагмента изображения на основе структурной и масштабирующих составляющих достигается равенством

размеров массивов $\Delta H_{m,n}^{(\nu)}$ и $\Delta L_{m,n}^{(\nu)}$ и однозначным порядком их образования. Это позволит исключить необходимость использования дополнительных служебных данных и временной задержки для позиционирования апертур и фрагментов изображений. Формирование массивов величинами $\ell'_{\phi,\xi}$ и $h_{\phi,\xi}$ (ϕ - номер строки кадра, $\phi=\overline{1,Z_{lin}}$) на $(i;j)$-м шаге реализуется на основе следующего правила:

1) если $j \le n$ и выполняется неравенство $(i-1)n+j \le \nu_\phi$, где $((i-1)n+j)$ - количество апертур ϕ-й строки, на базе компонент которых сформировано текущее количество элементов массивов $\Delta H_{m,n}^{(\nu)}$ и $\Delta L_{m,n}^{(\nu)}$; ν_ϕ - количество апертур в строке кадра изображения, то $\ell_{i,j} = \ell'_{\phi,(in+j)}$ и $h_{i,j} = h_{\phi,(in+j)}$;

2) если $j \le n$, но $(i-1)n+j > \nu_\phi$, то не отобранные апертуры ϕ-й строке отсутствуют, и отбор компонент апертур проводится для $(\phi+1)$-й строки кадра, т.е. $\ell_{i,j} = \ell'_{\phi+1,1}$ и $h_{i,j} = h_{\phi+1,1}$;

3) если $j > n$, $(i+1) \le m$, то для $(i-1)n+j \le \nu_\phi$ получим $\ell_{i+1,1} = \ell'_{\phi,(in+1)}$ и $h_{i+1,1} = h_{\phi,(in+1)}$, и наоборот для $(i-1)n+j > \nu_\phi$ - $\ell_{i+1,1} = \ell'_{\phi+1,1}$ и $h_{i+1,1} = h_{\phi+1,1}$;

4) если $(i+1) > m$, то построение массивов $\Delta H_{m,n}^{(\nu)}$ и $\Delta L_{m,n}^{(\nu)}$ завершено.

Второй этап. Определение оснований элементов массивов $\Delta H_{m,n}^{(\nu)}$ и $\Delta L_{m,n}^{(\nu)}$, рассматриваемых соответственно как адаптивное позиционное число с неравными соседними элементами и позиционное число. Выполняются следующие действия:

1) для формирования системы оснований $W(h)$, $W(h)=\{w'(h)_i\}$, $i=\overline{1,m}$ элементов АПЧ с неравными соседними элементами:

$$w'(h)_i = w(h)_i - \mathrm{sign}(j-1) = h_{i,max} - h_{i,min} + 1 - \mathrm{sign}(j-1),$$
$$h_{i,max} = \max_{1\le j\le n}\{h_{i,j}\}+1;\ h_{i,min} = \min_{1\le j\le n}\{h_{i,j}\};$$

2) для системы оснований $W(\ell)$, $W(\ell)=\{w(\ell)_i\}$, $i=\overline{1,m}$ элементов ПЧДП:

$$w(\ell)_{ij} = \ell_{max} - \ell_{min} + 1 = w(\ell),$$
$$\ell_{max} = \max_{\substack{1\le j\le n\\ 1\le i\le m}}\{\ell_{i,j}\}+1;\ \ell_{min} = \min_{\substack{1\le j\le n\\ 1\le i\le m}}\{\ell_{i,j}\}.$$

Третий этап. Организуется оценка количества элементов $\nu(h,i)_\xi$ и $\nu(\ell)_\xi$ двухкомпонентных составляющих для построения обобщенного

кода (ДК). Длина D_{nec} кодового слова для построения текущего двухкомпонентного кода считается заданной. По условию формирования ДК выбор первой составляющей на основе построения кода проводится для элементов одной строки массива $\Delta H_{m,n}^{(\nu)}$. Отсюда $\nu(h,i)_\xi = [D_{nec} / \ell og_2(w(h)_i - 1)]$.

Вторая составляющая формируется на основе кодового описания элементов массива $\Delta L_{m,n}^{(\nu)}$, расположенных в общем случае на разных строках. Поэтому величина $\nu(\ell)_\xi$ определяется по следующей технологии:

1. Находится общее количество элементов массива апертурно-координатной составляющей. Для этого используется следующая формула:

$$\nu(\ell)_\xi = [\Delta D / ([\ell og_2 w(\ell)] + 1)].$$

2. Найденное количество элементов $\nu(\ell)_\xi$ распределяется по строкам массива длин апертур. При этом учитывается, что длина строки равна n, а позиция первого элемента - $(\alpha;\gamma)$. Тогда такое распределение заключается в получении количества β полных строк и количества j элементов в последней включаемой строки массива длин апертур, на которые распространяется количество $\nu(\ell)_\xi$. Для этого выполняются следующие этапы:

1) если $\nu(\ell)_\xi > n - \gamma + 1$, то величина $\nu(\ell)_\xi$ превышает количество свободных элементов в текущей α-й строке, и требуется оценить количество полных строк. В противном случае количество необходимых добавляемых элементов будет принадлежать текущей α-й строки, и последняя позиция будет определяться как $(\alpha;\ \gamma + \nu(\ell)_\xi + 1)$;

2) определяем количество β полных строк по формуле

$$\beta = [\frac{\nu(\ell)_\xi - n + \gamma - 1}{n}];$$

3) если $\beta n < \nu(\ell)_\xi - n + \gamma - 1$, то вычисляем количество j элементов в $(\beta+1)$-й строке, т.е. $j = \beta n - [\frac{\nu(\ell)_\xi - n + \gamma - 1}{n}] n$.

В результате получаем распределение общего количества $\nu(\ell)_\xi$ добавляемых элементов массива длин апертур по строкам, а именно:

$$\nu(\ell)_\xi = (n - \gamma + 1) + \beta n + j.$$

В итоге получаем количество $\nu(h,i)_\xi$ элементов массива аппроксимирующих величин апертур и количество $\nu(\ell)_\xi$ элементов массивов длин апертур, участвующих в образовании двухкомпонентного кода. Причем выполняется обобщенное неравенство:

$$[\ell og_2\left((w(h)_i - 1)^{\nu(h,i)\xi}\ w(\ell)^{(n-\gamma+1)+\beta n+j}\right)]+1 \le D_{nec}.$$

Четвертый этап. Осуществляется построение ДК. Первая кодовая составляющая $E(h)^{(i,\gamma)}_{i,\gamma+\nu(h,i)\xi-1}$, формируемая на основе $\nu(h,i)_\xi$ элементов строки массива аппроксимирующих величин, будет равна

$$E(h)^{(i,\gamma)}_{i,\gamma+\nu(h,i)\xi-1} = \sum_{j=\gamma}^{\gamma+\nu(h,i)\xi-1} (h_{i,j} - sign\,(1 - sign\,(h_{i,j-1} - h_{i,j})))\,(w(h)_i - 1)^{\nu(h,i)\xi+\gamma-1-j}, \quad (13)$$

Рекуррентное выражение для формирования $E(h)^{(i,\gamma)}_{i,\gamma+\nu(h,i)\xi}$ принимает вид

$$E(h)^{(i,\gamma)}_{i,\gamma} = h_{i,\gamma};\ E(h)^{(i,\gamma)}_{i,\gamma+j} = E(h)^{(i,\gamma)}_{i,\gamma+j-1}(w(h)_i - 1) + h_{i,\gamma+j},$$
$$\overline{j=1,\ \nu(h,i)_\xi - 1}, \quad (14)$$

где $(i;\gamma)$, $(i;\ \gamma+\nu(h,i)_\xi - 1)$ - координаты соответственно начального и конечного элементов первой составляющей ДК на основе i-й строки массива аппроксимирующих величин апертур;

$E(h)^{(i,\gamma)}_{i,\gamma+j}$, $E(h)^{(i,\gamma)}_{i,\gamma+j-1}$ - значение кода первой составляющей соответственно на $(\gamma + j)$-м и на $(\gamma + j - 1)$-м шагах обработки.

Структура кода для формирования ДК на основе первой компоненты задается таким выражением:

$$E(h;\ell)_\xi = E(h)^{(i,\gamma)}_{i,\gamma+\nu(h,i)\xi-1}\ w(\ell)^{(n-\gamma+1)+\beta n+j} + \Delta E(\ell)^{(\alpha,\gamma)}_{\alpha+\beta+1,\tau}, \quad (15)$$

где $V(\ell)^{(\alpha,\gamma)}_{\alpha+\beta+1,\tau} = w(\ell)^{(n-\gamma+1)+\beta n+\tau}$ - весовой коэффициент первой компоненты $E(h)^{(i,\gamma)}_{i,\gamma+\nu(h,i)\xi-1}$ двухкомпонентного кода.

Здесь величина $V(\ell)^{(\alpha,\gamma)}_{\alpha+\beta+1,\tau}$ - определяется как накопленное произведение оснований элементов массива длин апертур, начиная с основания элемента на позиции $(\alpha;\gamma)$ и заканчивая основанием элементом на позиции $(\alpha+\beta+1;\tau)$.

При этом обеспечивается выполнение следующих неравенств:

$$\Delta E(\ell)^{(\alpha,\gamma)}_{\alpha+\beta+1,\tau} < V(\ell)^{(\alpha,\gamma)}_{\alpha+\beta+1,\tau};$$

$$[\ell og_2(\ (w(h)_i - 1)^{\nu(h,i)\xi}\ w(\ell)^{(n-\gamma+1)+\beta n+j}\ )\ ]+1 \le D_{nec}.$$

С учетом сказанного выше можно утверждать, что на основе выражений (14) – (15) осуществляется формирование двухкомпонентного кода на базе неравнозначного вклада элементов массива аппроксимирующих величин апертур и элементов массива длин апертур.

Это позволяет:

1) дополнительно повысить степень сжатия за счет сокращения количества незначимых старших разрядов в кодовых комбинациях. Это

достигается в результате добавления элементов массивов длин апертур, имеющих меньшие значения динамических диапазонов, в процессе формирования обобщенного двухкомпонентного кода (ДК).

2) достичь наибольшей степени сжатия в результате устранения избыточных разрядов, обусловленных не кратностью степени двойки значений весовых составляющих кодовых компонент ДК.

3) повысить оперативность обработки фрагментов изображений в результате: проведения восстановления фрагмента изображения сразу при реконструкции кодового представления (информация для этого будет содержаться не в разных кодах, а полностью в целостном виде в одном коде); отсутствия операций обращения к отдельным разрядам кодового слова, как это делается для бит-ориентированного принципа слияния кодов.

Строки массивы построчно-масштабирующей составляющей фрагмента изображения представляются в виде ***адаптивных позиционных чисел с неравными соседними элементами*** (АПЧ). Это позволяет рассматривать процесс кодирования АПЧ как формирование последовательности кодов для отдельных строк массива аппроксимирующих величин.

Система выражений обеспечивает:

а) формирование кода для строки массива аппроксимирующих величин, рассматриваемой как адаптивное одномерное позиционное число с ограниченным динамическим диапазоном;

б) исключение избыточного количества позиционных чисел, которые содержат равные соседние элементы;

в) устранение количества запрещенных последовательностей на произвольном шаге обработки включая количество избыточных последовательностей:

– содержащих равные соседние элементы на позициях не старше позиции обрабатываемого элемента;

– содержащих равные элементы для предшествующих элементов на текущей позиции и элемента обрабатываемой последовательности на предыдущей позиции относительно обрабатываемой.

Научная новизна проведенных исследований заключается в следующем:

1. Впервые разработан метод кодирования на основе позиционного представления построчно-масштабирующей составляющей апертурного описания. Отличие заключается в том, что в процессе кодирования учитываются структурные особенности, состоящие в неравенстве соседних элементов построчно-масштабирующей составляющей. Это позволяет повысить степень сжатия апертурного описания без внесения потери информации.

2. Впервые разработан метод сжатия на основе двухкомпонентного представления апертурных составляющих. Базовое отличие от известных методов состоят в том, что равномерное кодовое представление формируется на основе наращивания кода-номера адаптивного позиционного числа на основе элементов неравномерно-структурной составляющей. Это позволяет повысить степень сжатия видеоданных и снизить задержку на их передачу в телекоммуникационных системах.

Практическое значение полученных результатов исследований в результате интегрирования реализации метода двухкомпонентного кодирования апертурных составляющих для систем дистанционного формирования и передачи изображений с использованием беспроводных средств телекоммуникаций состоят в следующем:

1. Для разработанного метода относительно известных методов данного класса обеспечивается сокращение задержки на обработку в среднем в 12 раз, что достигается за счет снижения количества обрабатываемых данных в результате выявления апертур, уменьшения количества операций умножения (от 2 до 4 раз) и сложения (в среднем в 5 раз). Это позволяет относительно известных методов компрессии с контролируемой потерей качества обеспечивать возможность обработки и передачи в реальном времени статических изображений, размером 2048x1536 и 3504x2336 элементов с использованием вычислительной аппаратуры технологий дистанционного формирования изображений (тактовая частота МП соответственно равна 369 МГц и 800 МГц).

2. В зависимости от уровня пикового отношения сигнал/шум и характеристик вычислительных систем дистанционного формирования видеоданных для разработанного метода относительно существующих обеспечивается снижение задержки на доставку сжатых видеоданных от 1,19 до 2,3 раз.

Литература

1. Кашкин В.Б. Цифровая обработка аэрокосмических изображений: Конспект лекций.- Красноярск : ИПК СФУ, 2008. – 121 с.

2. Баранник В.В. Методологический анализ системы аэрокосмического видеомониторинга чрезвычайных ситуаций / В.В. Баранник, А.В. Яковенко, А.Ю. Школьник // Сучасна спеціальна техніка. – К.: 2011. – №4(27). – С. 12 – 22.

3. *Баранник В.В.* Структурно-комбинаторное представление данных в АСУ / В.В. Баранник, Ю.В. Стасев, Н.А. Королева - Х.: ХУПС, 2009. – 252 с.

4. Баранник В.В., Додух А.Н. Технология сжатия цифровых изображений на основе двухкомпонентного кодирования // Автоматизированные системы управления и приборы автоматики. – №157. – 2012. – С. 14 – 24.

5. Баранник В.В., Додух А.Н. Метод двухкомпонентного кодирования апертурных составляющих изображений в средствах телекоммуникаций // Радиоэлектроника и информатика. - №2. – 2012. – С. 11 – 18.

6. Баранник В.В., Додух А.Н. Адаптивное позиционное кодирование с неравными соседними элементами // Информационно-управляющие системы на железнодорожном транспорте. - №4. – 2012. – С. 23 – 28.

МЕТОДИ СЕГМЕНТУВАННЯ ЗОБРАЖЕНЬ, ЩО ОТРИМАНІ З БОРТОВИХ СИСТЕМ ОПТИКО-ЕЛЕКТРОННОГО СПОСТЕРЕЖЕННЯ

Худов В.Г., Худов Г.В.

Вступ

Відомо [1-3], що сегментування зображення, що отримано з бортових систем оптико-електронного спостереження, є розділення зображення на області, що мають приблизно однаковий рівень яскравості (для полутонових зображень) або однакові кольорові характеристики (для кольорових зображень). Для сегментування зображень (визначення границь або контурів об'єктів розроблено багато методів [4, 5]. Але на зображеннях, що отримані з бортових систем оптико-електронного спостереження, в більшості присутні контури об'єктів з різними швидкостями зміни яскравості (для полутонових зображень) або кольору (для кольорових зображень) [1, 2]. У зв'язку з цим, неможливо найкращим чином визначити усі присутні на зображенні границі з використанням якогось конкретного методу сегментування. Тому для підвищення якості сегментування будемо використовувати методи, які дозволяють побудувати картину контурів об'єктів на зображенні на основі інформації, що отримується в результаті обробки зображень різних масштабів.

Мета роботи – розглянути методи сегментування зображень, що отримані з бортових систем оптико-електронного спостереження.

Аналіз останніх досягнень і публікацій. У теперішній час не існує загальної теорії сегментування зображень бортових систем оптико-електронного спостереження, яка дозволяє отримати вичерпні рекомендації щодо оптимальному вибору методу сегментування та набору вхідних даних [1]. Виділяють наступні ознаки якісного сегментування [1]:

– однорідність області по характеристикам (в першу чергу, по кольору та текстурі);

– відмінність значень обраних характеристик для суміжних областей зображення;

– гладкість границь кожного сегменту зображення;

– незначна кількість «дірок» у сегменті.

Враховуючи перераховане вище, витікають три основні види можливих недоліків сегментування зображень бортових оптико-електронних систем спостереження [1]:

– неправильне сегментування, коли контури розподілу не співпадають з границями об'єктів на зображенні;

– пересегментування, коли має місце збільшений розподіл зображення на області;

– недосегментування, коли має місце недостатній розподіл зображення на області.

Зазвичай методи сегментування використовують декілька параметрів, підбираючи які можна уникнути останніх двох недоліків. Однак, перший недолік можна уникнути лише вибором методу сегментування. Відомо, що найбільш ефективні методи сегментування розроблені для конкретних завдань з урахуванням специфіки зображення [1, 6-8].

Загальна класифікація класичних методів сегментування наведена на рис. 1 [4].

Найбільш відомі методи сегментації – порогова сегментація, центроїдне зв'язування, метод водорозділу. Всі ці методи використовують один і той же принцип: групування в області пікселів, що розташовані поруч та мають рівень яскравості, що відрізняється не більш ніж на визначене число. Це число є порогом сегментування. В залежності від порога сегментування можна отримати різні результати сегментування зображення: різну кількість сегментів, різні параметри сегментів і т.і.

Пороговий метод сегментування зображення. Для сегментування пікселя (i, j) зображення з яскравістю f_{ij} пороговий метод сегментування має вигляд (1):

$$\begin{cases} 1, \text{ у випадку } f_{ij} \le T, \\ 0, \text{ в іншому випадку,} \end{cases} \quad (1)$$

де T - величина порогу сегментування.

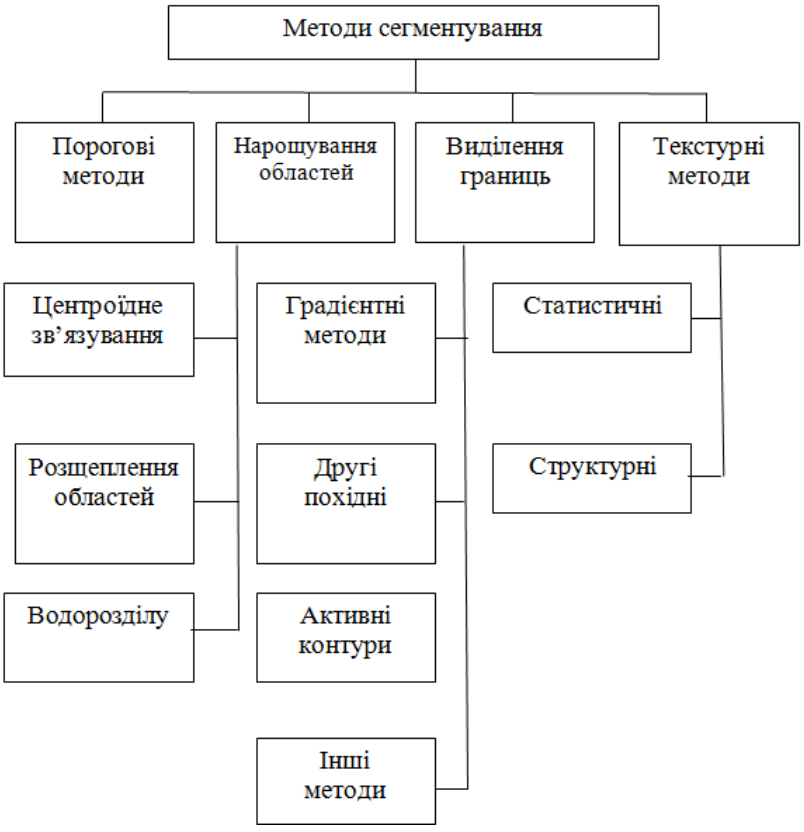


Рис. 1. Загальна класифікація класичних методів сегментування зображення [4]

При сегментуванні кольорових зображень, наприклад у просторі RGB, пороговий метод сегментування має вигляд (2):

$$\begin{cases} 1, \text{ у випадку } D(z,a) \le T, \\ 0, \text{ в іншому випадку,} \end{cases} \quad (2)$$

де $D(z,a) = \left[(z_R - a_R)^2 + (z_G - a_G)^2 + (z_B - a_D)^2\right]^{1/2}$;

a - центр кластеру, що відповідає області кольору зображення визначеного класу об'єктів у кольоровому просторі RGB;

z - колір пікселю зображення.

У випадку, коли на вхідному зображенні є декілька об'єктів, використовується метод січеній, коли задаються два пороги t та T, і сегментуються пікселі зображення об'єктів, яскравості яких знаходяться в межах заданих порогових значень за виразом (3):

$$\begin{cases} 1, \text{ у випадку } 1 \le f_{ij} \le T, \\ 0, \text{ в іншому випадку.} \end{cases} \quad (3)$$

Порогові методи дозволяють проводити сегментування на простих зображеннях, але, як правило, не дають необхідного результату на зображеннях з наявністю нерівного освітлення, тіней та різного роду завад.

Для зменшення впливу указаних недоліків розроблені методи, які реалізують аналіз вагових значень екстремумів (інтенсивність та градієнт).

В загальному випадку для коректного використання порогових методів:

– необхідно уникати «зміщення» при виборі порогового значення шляхом жорсткого контролю однаковості розподілу в темних та світлих областях гістограми яскравості;

– необхідно розбивати зображення на можливо малі елементи, таким чином, щоб гістограма яскравості мала ярко виражені екстремуми;

– такі малі елементи, з іншого боку, повинні бути достатньо великими, щоб об'єм статистичної вибірки дозволяв задовільно оцінювати місцеположення екстремумів та описувати околицю.

Методи нарощування областей. Якщо на зображенні є стійка зв'язність окремих сегментів, то використовують методи нарощування областей – проводиться групування сусідніх елементів з однаковими або близькими рівнями яскравості, які потім об'єднуються в однорідні області. Найбільш відомими методами нарощування областей є центроїдне зв'язування, розщеплення областей та водорозділу (рис. 1).

При центроїдному зв'язуванні з використанням інформації щодо об'єкту обираються стартові точки, яким присвоюється різні мітки. Точки з однаковими мітками утворюють окремі множини. Такий метод може використовуватися лише для сегментування простих зображень.

Для більш складних зображень вибір точок проводиться по ітераціям, на кожній з яких розглядається набір точок на предмет

належності їх сусідів даній множині. Точки, що включені у множину на попередніх ітераціях, не розглядаються. Так проводиться аналіз всіх множин по черзі. Точки, що додані до множині на даній ітерації, називаються фронтом, а об'єднання фронтів – хвилею, тому такий метод отримав назву – хвильовий.

Метод розщеплення областей розділяє точки зображення шляхом розбивання визначеним чином зображення на квадрати, які потім аналізуються для їх перевірки на однорідність (частіше за все це однорідність за яскравістю). Якщо квадрат не задовольняє умовам однорідності, він замінюється чотирма «підквадратами», а чотири квадрати, що підходять за умови однорідності, можуть бути об'єднані в одну область.

Суть методу водорозділу полягає в тому, що після побудови поля контрасту зображення необхідно побудувати водорозділи (області високої контрастності).

На рис. 2 наведено вихідне зображення, а на рис. 3, 4 – сегментоване зображення з використанням методу водорозділу з різним значенням порогового рівня.

Рис. 2. Вихідне зображення [8]

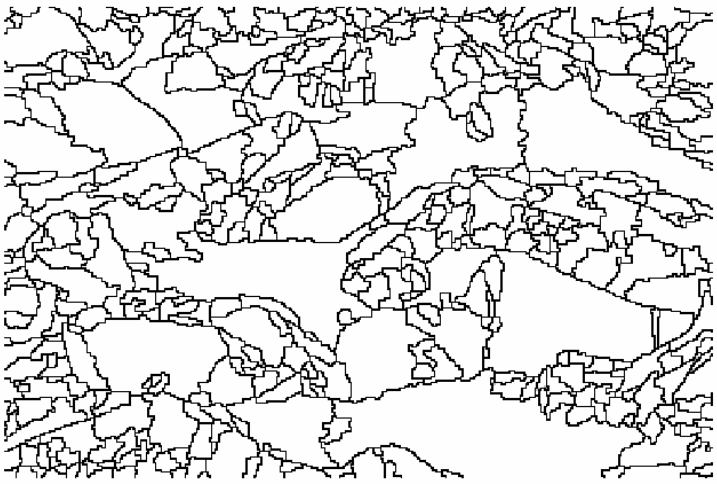


Рис. 3. Сегментоване зображення методом водорозділу (пороговий рівень дорівнює 5 одиницям) [8]

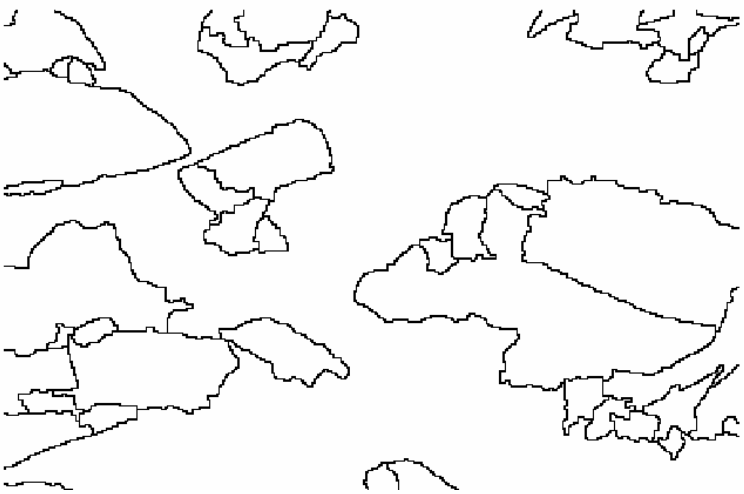


Рис. 4. Сегментоване зображення методом водорозділу (пороговий рівень дорівнює 15 одиницям) [8]

До переваги методів нарощування областей відноситься їх ефективність при роботі з зображеннями, на яких присутні шуми.

До недоліків методів нарощування областей відносять в першу чергу те, що вони виділяють загальні фрагменти, в багатьох випадках не показуючи інформації щодо змін яскравості всередині області та можливих внутрішніх границях.

В роботі [8] запропоновано метод структурного сегментування, який дозволяє вирішати завдання сегментування без вибору порога. При цьому задача структурного сегментування поділяється на наступні етапи [8]:

- виділення початкових сегментів і контурів за допомогою модифікованого методу водорозділу;
- представлення контурного зображення у вигляді графа;
- знаходження структури сегментів.

Результати роботи модифікованого методу водорозділу наведено на рис. 5, 6.

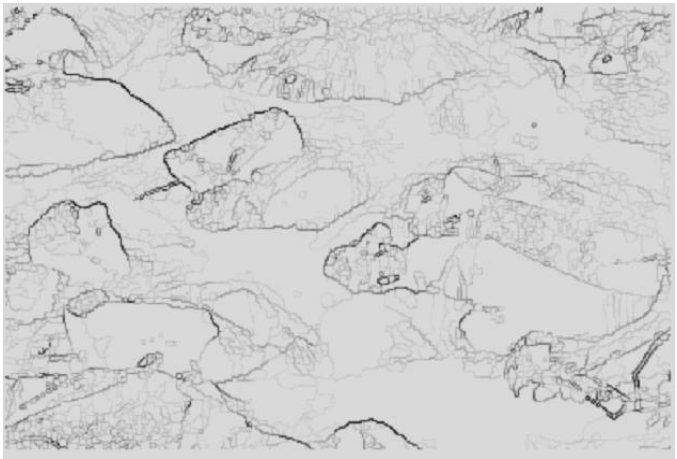


Рис. 5. Сегментоване зображення модифікованим методом водорозділу [8]

Основою методів сегментування залишаються контурні методи (методи сегментування границь зображення) в силу їх стійкості до незначних варіацій рівня яскравості та контрастності зображення. При цьому важливою є та особливість контурних методів, що по результатах їх застосування при необхідності може бути отримано не тільки границя, але і область зображення [4].

Методи виділення границь складаються з фільтрації (покращення виділення границь при наявності шумів), підсилення (акцент на точках, де існує перепад яскравості), виділення (з використанням порогового значення приймається рішення щодо включення точок в границю) і локалізації (визначення місця розташування та напрямку). На даний час існує багата кількість методів виділення границь, наприклад, з використанням операторів Робертса, Собела, Превітта, Канні та інші [4, 9].

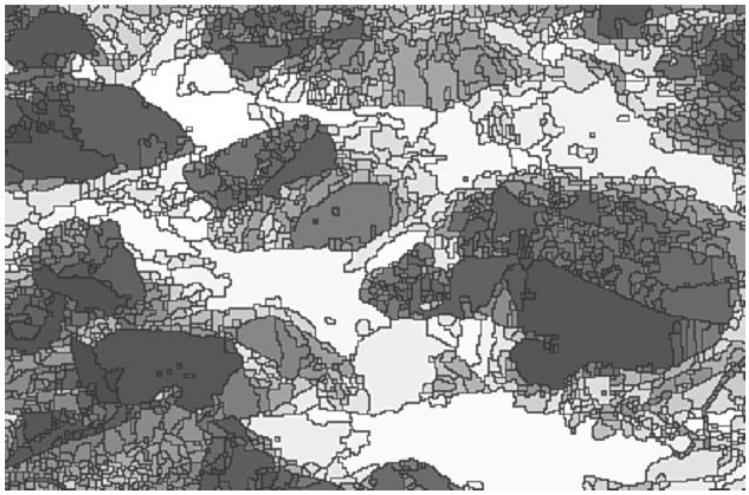


Рис. 6. Сегментоване зображення модифікованим методом водорозділу з виділенням середньої яскравості точок сегментів [8]

Існуючі методи виділення границь поділяються на методи порівняння з еталоном та диференціально-градієнтні методи [4, 9, 10]. Обидва методи визначають, коли коливання градієнта яскравості становиться достатньо великим, щоб стверджувати, що на цьому місці знаходиться границя об’єкта. Принципова різниця методів полягає в способі локальної оцінки градієнтного значення та визначенні локальної направленості границь.

В цілому методи виділення границь дають непогані результати для інтерпретації зображення. Карти границь можуть бути побудовані в різних масштабах, що дозволяє отримувати корельовано результати. Також методи виділення границь потребують менше ресурсів для проведення обчислень, а результуюча інформація займає суттєво менше місця для зберігання. Методи виділення границь рекомендовано використовувати тоді, коли границі мають достатню чіткість та стабільність.

До недоліків методів виділення границь відносяться велика обчислювальна складність, використання різних масок, проблеми при роботі на зображеннях з шумами.

Окремою групою виділяються методи сегментування, що засновані на кластеризації. Їх переваги – автоматичні та можуть бути використані для любої кількості ознак та класів. Існуючі методи кластеризації, такі як К-середніх, медоїдний, CURE, ROCK, DBSCAN, створені для знаходження кластерів, які відповідають будь якій статичній моделі. Такі методи можуть дати збій, якщо параметри моделі обрані некоректно, по відношенню до класифікованих даних, або якщо модель не враховує у повній мірі характеристики кластерів. Також деякі методи допускають помилки, якщо дані складаються з кластерів різної форми, щільності та розмірів.

В теперішній час існує декілька підходів до аналізу багатомасштабної інформації, тобто до *побудови картини контурів об'єктів градієнтних зображень різного масштабу* [11]. Існують підходи, в яких аналіз градієнтних зображень проводиться від грубих масштабів до точних [12, 13] та від точних до грубих [14, 15]. Методи розрізнюються по принципах побудови градієнтного зображення одного масштабу, але при цьому відкритим є питання, яким чином необхідно комбінувати багатомасштабну інформацію для побудови кінцевої картини границь. В роботі Бергольма [12] запропоновано метод, який полягає у послідовному аналізу багатомасштабної інформації від грубих масштабів до точних. Такий підхід дозволяє значно зменшити вплив шуму і, таким чином, уникнути хибного визначення контурів під впливом шумів. Недоліком методу [12] є можливе розділення контурів, що визначаються на грубих масштабах, на декілька окремих при переході до більш точного масштабу. Стратегія розгляду градієнтних масштабів від грубих до точних також відмічається в роботі [13]. Однак в тих випадках, коли на зображенні присутні невеликі об'єкти з різкими границями, точне визначення границь цих об'єктів при переході від грубих масштабів до точних є ускладненим, так як на градієнтних зображеннях грубого масштабу виникає значне зміщення положення різких контурів.

В роботах [14, 15] кінцева картина границь складається на основі аналізу градієнтних зображень від точних масштабів до грубих. При цьому основними задачами є зменшення впливу шуму, до якого чутливі оператори градієнту малого розміру, та комбінування границь, що отримані на точних масштабах, з плавними границями, які визначаються на грубих масштабах. При успішному рішенні таких проблем підхід до аналізу градієнтних зображень від точних масштабів до грубих є найбільш ефективним для багатьох практичних випадків, в яких необхідно достатньо точно визначити контури об'єктів. Однак, методи, що наведені в [12-15] можуть бути застосовані для сегментування сканованих зображень

сторінок книг, газет, журналів з великою кількістю об'єктів невеликого розміру, наприклад, букв та символів.

Методи обробки багатомасштабної послідовності цифрових зображень в промислових системах контролю якості наведені в роботах [11-13]. Однак, розроблені в роботах [16-18] методи обробки багатомасштабної послідовності цифрових зображень можуть бути використані при:

– зменшенні часу на розшифровку рентгенографічних знімків зварних з'єднань;

– подавлені шуму на рентгенограмах без внесення додаткових спотворень;

– виділяти дефекти зварних з'єднань;

– виявляти групові дефекти зварних швів;

– проводити якісний аналіз мікроструктури металів;

– відновлювати томографічні зображення по неповним даним.

Методи, що запропоновані в [16-18], неможна напряму використовувати для обробки багатомасштабної послідовності зображень, отриманих з бортових систем оптико-електронного спостереження.

1. Постановка задачі та викладення матеріалів дослідження

В теперішній час для вирішення різних завдань, що виникають при обробці зображень, використовуються *генетичні алгоритми*, наприклад [19-21]. Генетичні алгоритми – самостійний розділ теорії штучного інтелекту – еволюційних обчислень, які засновані на математичному моделюванні процесів біологічної еволюції. Генетичні алгоритми застосовуються для вирішення оптимізаційних задач, їх предметна область включає проблеми комбінаторики, біоінформатики, теорії ігор, а також – обробка і розпізнавання образів, зокрема зображень.

При використанні генетичних алгоритмів пошук рішення проблеми проходить на підмножині точок простору пошуку, що досягається створенням множини потенційних рішень, яке формує популяцію. Популяція удосконалюється за допомогою генетичних операторів, які відповідають за змінність та фітнес-функції, які моделюють природний відбір. Спадщина забезпечується тім, що нові хромосоми формуються з хромосом попереднього покоління і, відповідно, мають загальні з ними гени. Якщо генетичний алгоритм реалізований коректно, то з кожним новим поколінням середнє значення фітнес-функції популяції та найкраще значення фітнес-функції зростають в сторону глобального оптимуму.

Для правильної роботи генетичного алгоритму необхідно обрати кодування даних і фітнес-функцію [19]. Кодування даних – спосіб представлення потенційного рішення. Передбачається, що потенційне рішення можна представити у вигляді параметрів (генів), які можна з'єднати в прості структури даних (хромосоми). Традиційно гени

кодуються двоїчними числами, і хромосоми представляють собою бінарні строки. Крім того, гени можуть бути представлені за допомогою алфавіту з більшою розмірністю або числами з плаваючою точкою [19], а хромосоми можуть бути представлені, наприклад, як дерева або матриці [19]. Фітнес-функція є цільовою функцією, яка для вирішення кожної задачі обирається індивідуально. Роботу простого генетичного алгоритму можна представити наступним чином [19].

1. Створюється початкова популяція (набір хромосом), звичайно випадковим чином. Обчислюється фітнес-функція кожної хромосоми популяції та середня адаптивність популяції. Встановлюється рахунок епох.

2. Нарощується рахунок епох, за допомогою оператора репродукції формується проміжна популяція – популяція батьків з урахуванням їх адаптації.

3. Формується наступне покоління. Випадковим чином з проміжної популяції обирається пара батьків, з заданою імовірністю проводиться над генотипами обраних хромосом кросинговер, обирається один з потомків. До нього послідовно застосовується оператор інверсії, а потім – мутації з заданими ймовірностями. Отриманий генотип потомка зберігається в новій популяції.

4. Якщо в проміжному поколінні ще є батьки, то здійснюється повернення до пункту 3, в противному випадку – пункт 5.

5. Якщо рахунок поколінь досяг заданого значення, то здійснюється перехід до пункту 6, якщо ні, то здійснюється перехід до пункту 2.

6. Вибір найкращих рішень, кінець роботи.

З наведено вище генетичного алгоритму видно, що основними генетичними операторами є репродукція, кросинговер, мутація та інверсія.

Репродукція – процес формування проміжного покоління [19]. Біологічний зміст кросинговеру – передача ознак батьків потомкам [19]. Простий оператор кросинговеру виконується наступним чином [19]. Обираються дві хромосоми:

$$A = a_1, a_2, a_3, ..., a_L, \tag{4}$$

$$B = a_1', a_2', a_3', ..., a_L', \tag{5}$$

де L - довжина хромосоми, обирається точка кросинговеру - k.

Дві нові хромосоми формуються з А і В (вирази (4), (5)) наступним чином: частина хромосоми А до точки кросинговеру сполучається з частиною хромосоми В після точки кросинговеру та формує першу хромосому-потомок, і, аналогічно, частину хромосоми В до точки кросинговеру сполучається з частиною хромосоми А після точки кросинговеру і формує другу хромосому-потомок:

$$A' = a_1, a_2, a_3, ..., a_k, a_{k+1}', a_{k+2}', a_{k+3}', ..., a_L', \tag{6}$$

$$B' = a_1', a_2', a_3', ..., a_k', a_{k+1}, a_{k+2}, a_{k+3}, ..., a_L. \tag{7}$$

Оператор мутації призначений для того, щоб підтримувати різнорідність складу популяції, який реалізується наступним чином: в кожній строчці мутації довільний біт з імовірністю P_m змінюється на протилежний. При виконанні оператора інверсії хромосома розбивається на дві частини, які потім змінюються місцями. Як правило, імовірність використання операторів мутації та інверсії досить мала (приблизно 0,0001) [19].

Використання генетичних алгоритмів для обробки та розпізнавання зображень розглянуто в [19]. Так, в [19] розглянуто підхід автоматичної розмітки сцени за допомогою генетичних алгоритмів. Комбінація використання семантичних мереж для представлення обмежень області і нечіткої логіки для досягнення відповідності міток цим обмеженням породили нову стратегію обчислення фітнес-функцій для роботи генетичних алгоритмів. В [19] показано можливість використання даного підходу для ідентифікації знімків хмар на мультиспектральних супутникових знімках.

Можливість використання адаптивного генетичного алгоритмі для вирішення задачі сегментування кольорового зображення, ускладненого необхідністю прийняття рішення щодо оптимальної кількості сегментів і точного визначення текстурних областей, розглянуто в [19]. Так як в багатьох випадках при сегментуванні топологічним областям можуть бути поставлені у відповідність області ознак, дану задачу можна вирішити як оптимізаційну і використати генетичні алгоритми для кластерізації невеликих районів простору ознак [19].

В [19] також проаналізовані роботи, що присвячені використанню генетичних алгоритмів для вирішення наступних задач обробки зображень:

– квантування зображення з використанням комбінованого генетичного алгоритму, який об’єднує традиційний генетичний алгоритм і метод оптимального квантування зображення;

– визначення різних класів текстури на зображенні по їх кореляції зі спектром Фур’є. При цьому генетичний алгоритм використовується для вибору оптимальної маски із множини можливих, що використовується для сегментування магнітно-резонансних зображень мозку [19];

– поєднання етапів сегментування і розпізнавання зображення за допомогою генетичного алгоритму. Пошук ведеться на просторі можливих сегментів зображення, які порівнюються з шаблонними сегментами;

– оптимального визначення набору ознак для класифікації з метою розпізнавання зображення;

– розпізнавання образів на зображенні з використанням класифікатора Байєса.

Багатомасштабне перетворення вихідного зображення $f(x,y)$, де (x,y) - просторові координати зображення будемо представляти у вигляді (8):

$$L(x,y,t) = g(x,y,t) * f(x,y), \quad (8)$$

де $L(x,y,t)$ - багатомасштабне перетворення вихідного зображення $f(x,y)$;

$g(x,y,t)$ - ядро перетворення;

t - масштабний коефіцієнт;

$*$ - оператор згортки.

Необхідно зауважити, що у виразі (8) згортка виконується по просторових координатах (x,y), а масштабний коефіцієнт t лише указує, для якого масштабу проводиться операція згортки.

Ядро перетворення $g(x,y,t)$ будемо обирати у вигляді гаусіана (9):

$$g(x,y,t) = \frac{1}{2\pi t} e^{-\frac{\left(x^2+y^2\right)}{2t}}. \quad (9)$$

Вибір саме такого виду ядра перетворення обумовлений унікальністю гаусова ядра, яке включає лінійність, інваріантність до здвигу, не підсилення локальних екстремумів на вихідному зображенні, інваріантністю до масштабним спотворенням та інваріантністю до обертання зображення [22].

Масштабний коефіцієнт t відіграє роль дисперсії у виразі для гаусіана (9). При $t=0$ ядро перетворення $g(x,y,0)$ становиться імпульсною функцією, такою, що $L(x,y,0) = f(x,y)$, тобто масштабне перетворення вихідного зображення є саме вихідне зображення $f(x,y)$.

При збільшенні масштабного коефіцієнта t багато масштабне перетворення $L(x,y,t)$ є результат згладжування вихідного зображення $f(x,y)$.

Після отримання багатомасштабного перетворення вихідного зображення $L(x,y,t)$ проведемо сегментування кожного з зображень при різних значеннях масштабного коефіцієнта t. Для проведення сегментування будемо використовувати еволюційний метод, запропонований в роботі [3].

Отже, в найпростішому випадку сегментування зображення можна представити як сукупність наступних ділянок руху агентів (рис. 7): вихідна точка маршруту (ВТМ), прямі ділянки, кінцева точка маршруту (КТМ). Прямі ділянки проходять через поворотні точки маршруту (ПТМ), в яких відбувається зміна напрямку руху агенту. У подальшому вважаємо, що положення ВТМ, КТМ та поворотних точок маршруту повністю визначає маршрут руху агенту.

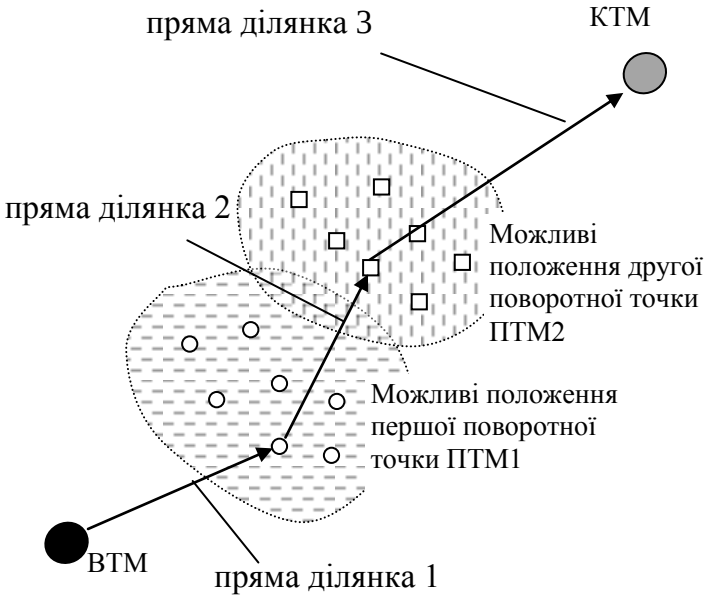


Рис. 7. Приклад представлення маршруту руху агенту на зображенні при сегментуванні зображення [3]

Рух по кожній з ділянок маршруту, а також здійснення повороту в вибраних ПТМ, має певні небезпеки та вимагає певних витрат ресурсів, що призводить до наявності переваги одного маршруту руху перед іншим. Оскільки варіантів розташування ПТМ може бути дуже багато, кількість можливих маршрутів руху буде надзвичайно великою, що ускладнює вибір маршруту руху методом перебору. Продемонструємо, як прокласти маршрут руху з використанням простішого еволюційного методу (ЕМ).

ЕМ, використаний в [3], оснований на імітації природного механізму пошуку найкоротшого шляху до джерела їжі колонією мурах (агентів). Самоорганізація системи забезпечується низькорівневою взаємодією агентів, при цьому агенти обмінюються тільки локальною інформацією, для передачі якої вони використовують спеціальний секрет, феромон, що відкладається агентом на своєму маршруті. Наступний агент, який буде знаходитись поблизу маршруту руху першого, сприймає феромон та з високою ймовірністю продовжить рух по шляху першого агента, в свою чергу відкладаючи феромон (підвищуючи його концентрацію на маршруті). Чим вище концентрація феромону на маршруті, тим вища привабливість цього маршруту для наступних агентів. Розподіл феромону в навколишньому середовищі являється немовби динамічною пам’яттю системи. Кожний агент в певний момент часу сприймає та змінює одну гратку цієї пам'яті – рівень феромону в околиці точки, в якій агент знаходиться.

Концентрація феромону, відкладеного на маршруті, пропорційна привабливості (якості, ефективності) маршруту. Чим привабливіший буде маршрут, тим більшою буде концентрація феромону на ньому, в результаті

кращі маршрути зберігаються в глобальній пам'яті колонії агентів і з вищою ймовірністю будуть обрані наступними агентами.

З часом феромон випаровується, що забезпечує зворотній зв'язок. Оскільки, як зазначено вище, концентрація феромону буде поступово збільшуватись на привабливих маршрутах, а швидкість випаровування феромону є постійною, через деякий час невдалі маршрути зникнуть, і все більше агентів будуть здійснювати рух лише по вдалих маршрутах. Використання зворотного зв'язку (випаровування) попереджує завчасну сходимість рішень – вибір агентами одного і того ж субоптимального маршруту.

В простішому ММ в кожній ітерації ітераційного процесу m агентами здійснюється пошук рішення та оновлення феромонів на знайденому маршруті. Кожний m-й агент при сегментуванні зображення починає шлях з ВТМ, послідовно проходить вибрані методом ПТМ і завершує шлях в одній з КТМ. Вибір ПТМ з J можливих здійснюється на основі ймовірнісного правила, що визначає ймовірність $P_i^m(t)$ переходу m-го агента в i-у ПТМ з врахуванням привабливості i-ї ділянки маршруту L_i та концентрації феромонів на цій ділянці F_i в момент часу t наступним чином:

$$P_i^m(t) = \frac{F_i(t)^\alpha \cdot L_i^\beta}{\sum_{j=1}^{J} F_j(t)^\alpha \cdot L_j^\beta}, \qquad (10)$$

де α і β – параметри, що задають вагу феромона і привабливості ділянки, відповідно.

Вважаємо, що привабливість ділянки маршруту L_i в ЕМ обернено пропорційна затратам на подолання ділянки, тобто

$$L_i = \frac{1}{D_i}, \qquad (11)$$

де D_i – довжина i-ї ділянки маршруту.

На початку ітераційного процесу кількість феромону на ділянках маршруту приймається однаковою і рівною деякому невеликому числу F_0. Після кожної ітерації концентрація феромонів на вибраних агентами ділянках оновлюється за правилом:

$$F_i(t+1) = (1-\rho)F_i(t) + \sum_{m=1}^{M} \Delta F_i^m, \qquad (12)$$

де $\rho \in [0,1]$– швидкість випаровування феромону;

ΔF_i^m - концентрація феромону на i-й ділянці маршруту, що створюється проходженням m-го агента.

В результаті проведення певної кількості ітерацій визначаються найпривабливіші за вибраним критерієм маршрути, концентрація феромону на яких максимальна. Феромон на непривабливих маршрутах поступово "висихає" і непривабливі маршрути зникають.

На рис. 8 наведено результати використання ЕМ для сегментування зображення [3]: рис. 9а - після проведення 100 ітерацій, рис. 9б - після проведення 300 ітерацій. Більш кращі маршрути позначені більш жирними лініями. Колами відмічені ділянки зображення, де необхідно змінити маршрут руху агентів (їх фізичний смисл необхідно визначити у подальших дослідженнях).

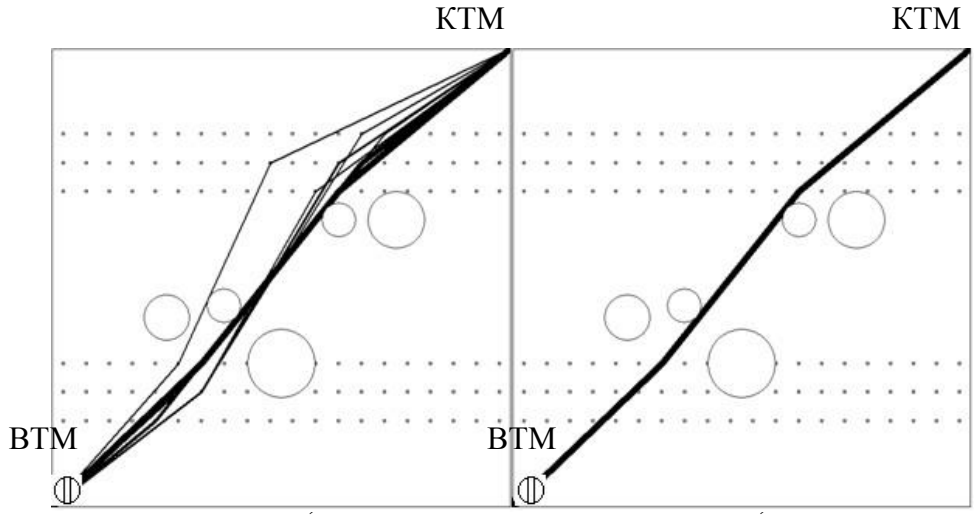


Рис. 9. Результати використання ЕМ для сегментування зображення [3]: а) після 100 ітерацій; б) після 300 ітерацій

З аналізу рис. 9 видно, що після 100 ітерацій рівень феромону на всіх маршрутах значно зменшується внаслідок випаровування (тонші лінії). Після 300 ітерацій кращий маршрут, який і є оптимальним для проведення сегментування, виділяється явно (рис. 9б).

Висновки

Таким чином, в роботі виділені основні ознаки якісного сегментування та основні види можливих недоліків сегментування зображень, що отримані з бортових оптико-електронних систем спостереження. Проведено аналіз відомих методів сегментування зображень, що отримані з бортових систем оптико-електронного спостереження. Встановлені основні недоліки, що притаманні класичним методам сегментування зображень. Для підвищення якості сегментування запропоновано використання генетичних алгоритмів та багатомасштабне перетворення зображень, при якому у якості методу сегментування запропоновано використання еволюційного методу.

У подальших дослідженнях необхідно провести оцінку ефективності запропонованого методу сегментування та його порівняльну оцінку з іншими методами сегментування оптико-електронних зображень.

Література

1. Малогабаритные беспилотные авиационные комплексы (Mini UVS) / Башинский В.Г., Бзот В.Б., Жилин Е.И. и др. / Монография. – Запорожье: изд. АО «Мотор-Сич». – 2014. – 261 с.
2. Барталев С.А. Анализ возможностей применения методов сегментации спутниковых изображений для выявления изменений в лесах / С.А.Барталев, Т.С.Ховратович // Современные проблемы дистанционного зондирования Земли из космоса. – 2011. – Т. 8. - № 1. – С. 44 - 62.
3. Худов В.Г. Мультиагентний метод сегментування зображень, що отримані з бортових систем оптико-електронного спостереження / В.Г.Худов // Системи озброєння і військова техніка. – 2016. – № 3 (47). – С. 116 - 119.
4. Смеляков К.С. Методы сегментации изображений объектов нерегулярного вида, особенности их применения и перспективы развития / К.С.Смеляков, И.А.Романенко, И.В.Рубан, Н.И.Кириллова, О.В.Шитова // Збірник наукових праць ХУПС. – 2010. – Вип. 2 (24). – С. 92 - 97.
5. Худов В.Г. Аналіз відомих методів сегментування зображень, що отримані з бортових систем оптико-електронного спостереження / В.Г.Худов, Г.А.Кучук, О.М.Маковейчук, А.В.Крижний // Системи обробки інформації. – 2016. – Вип. 9 (146). – С. 77 - 80.
6. Барталев С.А. Исследование возможностей оценки состояния поврежденных пожарами лесов по данным многоспектральных спутниковых измерений / С.А.Барталев, В.А.Егоров, А.М.Крылов, Ф.В.Стыценко, Т.С.Ховратович // Современные проблемы дистанционного зондирования Земли из космоса. – 2010. – Т. 7. - № 3. – С. 215 - 225.
7. Златопольский А.А. Выделение на изображении однородных участков с неполными границами / А.А.Златопольский // Исследование Земли из космоса. – 1985. – № 1. – С. 94 - 102.
8. Левашкина А.О. Исследование супервизорных критериев оценки качества сегментации изображений А.О.Левашкина, С.В.Поршнев // Известия Томского политехнического университета. – 2008. – Т. 313. – № 5. – С. 28 - 33.
9. Смеляков К.С. Модели и методы сегментации границ изображений нерегулярного вида на основе адаптивных масок: дис. канд. техн. наук: 09.03.05 / Смеляков Кирилл Сергеевич – Харьков. – 2005. – 162 с.
10. Самойленко Д.Е. Структурная сегментация изображений / Д.Е.Самойленко // Штучний інтелект. – 2004. - № 4. – С. 521 - 528.
11. Ziou D. Edge Detection Techniques / D.Ziou, S.Tabbone // An Overview technical report: Dept Math & Informatique. Universit de Sherbrooke, 1997. - № 195. – PP. 567-578.
12. Bergholm F. Edge Focusing / F.Bergholm // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987. - № 9. – PP. 726 - 741.
13. Williams D.J. Edge Contours Using Multiple Scales / D.J.Williams, M.Shas // Computer Vision, Graphics and Image Processing, 1990. - № 51. – PP. 256 - 274.
14. Lacroix V. The Primary Raster: A Multiresolution Image Description / V.Lacroix // In Proceedings of the 10th International Conference on Pattern Recognition, 1990. - PP. 903 - 907.
15. Canny J.F. A Computational Approach to Edge Detection / J.F.Canny // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986. - № 8. – PP. 679 - 698.

16. Жизняков А.Л. Формализация некоторых понятий теории обработки многомасштабных последовательностей цифровых изображений / А.Л.Жизняков // Системы управления и информационные технологии. – 2007. - № 3.3 (29). – С. 354-358.

17. Жизняков А.Л. Теоретические основы обрабоки многомасштабных последовательностей цифровых изображений / А.Л.Жизняков, С.С.Садыков. – Владим. гос. ун-т. – Владимир: Изд-во Владим. гос. ун-та. – 2008. – 121 с.

18. Жизняков А.Л. Теория и методыобработки многомасштабных последовательностей цифровых изображений в промышленных системах контроля качества: автореферат дис. докт. техн. наук: 05.13.01 / Жизняков Аркадий Львович – Владимир. – 2008. – 35 с.

19. Сергеева О.П. Применение генетических алгоритмов для распознавания изображений / О.П.Сергеева // Искусственный интеллект. – 2002. - № 4. – С. 516 – 520.

20. Худов В.Г. Генетичні алгоритми для сегментування зображень систем оптико-електронного спостереження / В.Г.Худов, О.М.Маковейчук // Наука і техніка Повітряних Сил Збройних Сил України. – 2016. – № 2 (23). – С. 142 - 145.

21. Махно Т.А. Автоматизированная система обработки ультразвуковых изображений сонных артерий на основе эволюционных алгоритмов / Т.А.Махно // Электротехнические и компьютерные системы. – 2015. - № 18 (94). – С. 92 - 99.

22. Babaud J. Uniqueness of the Gaussian kernel for scale-space filtering / J.Babaud, A.P.Witkin, M.Baudin, R.O.Duda // IEEE Trans. Pattern Anal. Machine Intell., 1986. – № 8. – PP. 26 – 33.

МЕТОДИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ КЕРУВАННЯ ЗАХИЩЕНИМИ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИМИ СИСТЕМАМИ НА ОСНОВІ ІДЕНТИФІКАЦІЇ УПРАВЛЯЮЧИХ СИГНАЛІВ

Юдін О.К., Ільєнко А.В., Зюбіна Р.В.

Вступ

Сучасні інформаційно-комунікаційні технологій об'єднали корпоративні мережі в глобальне інформаційне середовище. Це призвело до появи такого унікального явища, як глобальні інформаційні системи та мережі передачі даних. Впровадження та інтеграція інформаційно-комунікаційних систем та мереж (ІКСМ) потребує високого рівня технічних і соціальних вимог до якості інформаційних ресурсів та безпосередньо до систем передачі, обробки, відображення, збереження та захисту даних (рис. 1).

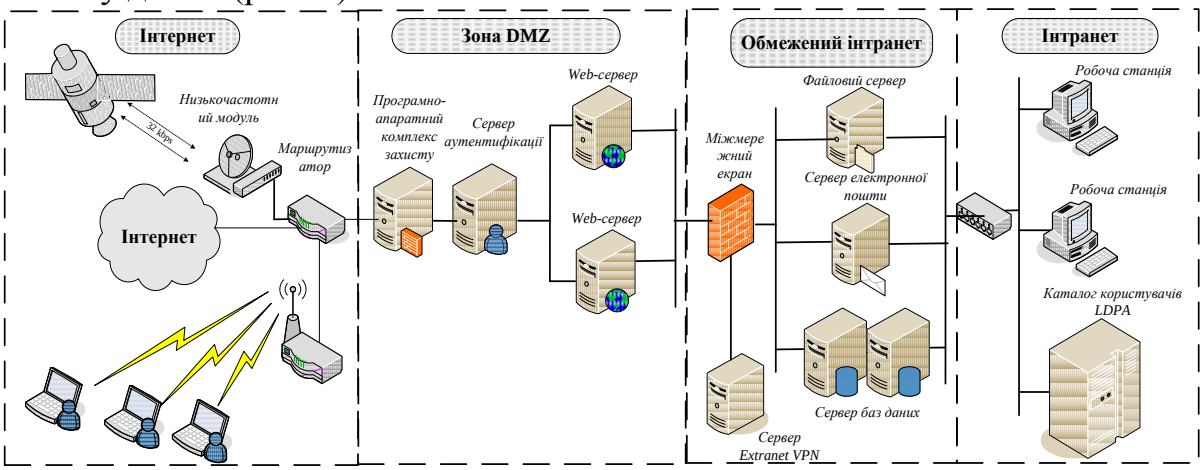


Рис. 1 Організаційна структура сучасної ІКСМ

Базовими властивостями інформаційних ресурсів з точки зору інформаційної безпеки та побудови захищених ІКСМ, є їх цілісність, конфіденційність і доступність. Випадкові, а також штучні завади спотворюють інформаційні потоки, які надходять від джерела повідомлення безпосередньо до споживача. Спотворення властивостей інформації при її передачі чи обробці веде до порушення системи доступу, відмов в обслуговуванні, призводить до неякісних процедур прийняття рішень або зовсім унеможливлює цей процес у всіх сферах діяльності сучасного розвиненого суспільства. Захищені інформаційні системи передачі даних є базовою платформою сучасного процесу інформатизації суспільства, що активно впливає на стан національної безпеки держави. Концепція розвитку захищених ІКСМ є логічним результатом розвитку нових інформаційних технологій та їх упровадження в усі сфери діяльності сучасного суспільства.

Базовою функцією захищених інформаційних систем передачі даних є здійснення процесів оперативного та надійного обміну інформацією між джерелом повідомлення та його користувачем, а також у забезпеченні ефективності функціонування всієї інформаційної системи – мінімізації часу передачі інформації, з умов зростання впливу завад (навмисне або не навмисне спотворення), при фіксованій вірогідності інформаційного потоку даних. Тому завдання забезпечення базових властивостей інформаційних ресурсів є однією із найактуальніших науково-технічних задач при розробці, впровадженні та експлуатації захищених інформаційних систем та їх елементів. Виконання цих вимог забезпечується традиційним удосконалюванням технічних характеристик виявляється економічно не вигідним або просто неможливим. До найбільш ефективних методів рішення даного роду задач варто, перш за все, віднести застосування завадостійкого канального кодування як технології забезпечення достовірності і цілісності інформаційних повідомлень. Основним завданням при побудові процедур прийняття рішення захищених ІКСМ є оцінка точності ідентифікації керуючих повідомлень з урахуванням забезпечення максимальної ймовірності правильного та надійного відтворення та розпізнавання як цифрової кодової конструкції так і голосу людини у якості управляючих сигналів.

Більшість задач ідентифікації інформаційних сигналів вирішуваних у захищених ІКСМ основані на структурних або статистичних методах канального кодування. Впровадження даних методів забезпечує вирішення важливих технічних та соціальних задач. Таким чином, актуальною є задача підвищення ефективності та надійності управління сучасними захищеними ІКСМ на основі розробки новітніх методів відновлення та ідентифікації кодових конструкцій, з урахуванням скорочення часу на обробку та достовірну неспотворену передачу даних, а також ідентифікації користувача за базовими характеристиками особливостей голосу людини.

Мета роботи. Метою роботи є підвищення ефективності та надійності керування захищеними інформаційними системами на основі розробки новітніх методів відновлення та ідентифікації кодових конструкцій, а також розробка методів прийняття рішення в процесі ідентифікації сигналу з використанням багатоальтернативних процедур прийняття рішень.

1. Захист інформаційних ресурсів на базі завадостійкого кодування

При проектуванні сучасних захищених ІКСМ одним з найважливіших є завдання забезпечення високої вірогідності та достовірності передачі даних. До найбільш ефективних методів рішення даного завдання варто віднести застосування завадостійкого канального кодування, як технології забезпечення достовірності і цілісності

інформаційних повідомлень.

У зв'язку з розширенням можливостей обміну інформацією між віддаленими абонентами та розподіленістю сучасних ІКСМ, різко зросла роль кодування, як базового засобу для забезпечення захисту інформації від спроб несанкціонованого впливу та модифікації інформації під час її передачі, оброблення та зберігання. Тому в процесі створення і експлуатації захищених ІКСМ важливе місце відводиться процесам завадостійкого кодування інформації та інформаційних ресурсів.

Для забезпечення цілісності та достовірності інформаційних потоків даних використовують різні види завадостійкого канального кодування, а саме блокові і згорткові коди. Блокові та згорткові коди мають різні характеристики. Вибір типу коду визначається числом факторів: характеристикою каналів, швидкість передачі, вид модуляції та ін. **Блокові коди** – для швидкого виявлення помилок (NMT-450, DECT) та **згорткові коди** – для виправлення одиночних помилок (GSM,CDMA). Найбільш поширеними серед блокових кодів є коди з перевіркою на парність, матричні коди, лінійні коди (коди Хемінга та Голея), поліномні коди (коди Ріда-Соломона, Ріда-Малера, булеві поліноми), циклічні коди (коди Боуза-Чоудхурі-Хоквінема, коди з контролем надлишковості, укорочені циклічні) [4-7]. Існуючі види завадостійкого кодування/декодування відрізняються один від одного наступними характеристиками: швидкістю, надмірністю, корегуючою здатністю, структурою, функціональним призначенням, енергетичною ефективністю, тощо

В результаті проведеного аналізу сучасних методів кодування/декодування встановлено систему критеріїв та вимог щодо формування сучасних методів й алгоритмів канального декодування з умови підвищення ефективності та надійності функціонування захищених ІКСМ:

- забезпечення мінімізації часу на обробку вибіркового простору кодових слів з урахуванням збільшення швидкодії та ефективності функціонування захищеної ІКСМ;
- мінімізація загального часу на прийняття остаточного твердого рішення, щодо відновлення повної кодової послідовності $t_{\sum}(n) \to \min$;
- зменшення складності реалізації процедури декодування та ідентифікації з урахуванням зменшення кількості математичних операцій $W(n) \to \min$ процедури відновлення кодової конструкції;
- зменшення ймовірності помилкової ідентифікації $P_{пом}(n,t) \to \min$

та повне усунення спотворень в інформаційному повідомленні, що дозволяє підвищити ефективність та надійність функціонування захищеної ІКСМ з умов збільшення вірогідності інформаційного потоку даних і забезпечення цілісності та достовірності;

− зменшення надлишковості завадостійкого коду за рахунок підвищення вірогідності процедур декодування та ідентифікації;

− використання послідовних правил прийняття рішень за рахунок накопичення достатньої кількості інформації на базі більш інформативних параметрів керуючих кодових конструкцій;

− можливість зупинки процесу декодування у разі відповідності функції правдоподібності встановленим порогам з урахуванням достатньої кількості інформації.

Отже, одним з ефективних шляхів підвищення надійності та ефективності функціонування сучасних захищених ІКСМ є розробка сучасних новітніх методів ідентифікації керуючих сигналів з урахуванням скорочення часу на обробку та неспотворену передачу інформаційних ресурсів [11-12].

2. Теоретичні основи побудови методів ідентифікації керуючих повідомлень захищених ІКСМ

Вирішуючи більшість практичних завдань декодування та ідентифікації інформаційних сигналів у захищених ІКСМ доводиться мати справу з керуючими кодовими конструкціями, що мають складну статистичну природу. Тому виникає необхідність вирішувати проблему декодування статистичними методами, використовуючи теорію статистичних вирішуючих правил, з урахуванням випадкової природи сигналів. Тобто вирішуючи більшість задач декодування інформаційних сигналів потрібно спиратися на існуючі математичні моделі інформаційних сигналів захищених ІКСМ. Результати рішення задачі ідентифікації в будь-якій області завжди носять допоміжний характер або такий, що забезпечує функціонування системи та прийняття правильного рішення згідно встановлених задач. Якщо на підставі цих результатів (або зроблених по них висновків) не приймаються конкретні практичні дії, саме не виконання чи некоректне виконання завдання ідентифікації виражає зміст .

Ідентифікація відкриває широкі можливості для оптимізації якісних наслідків практичних дій, тобто впливу на кінцевий результат, шляхом оптимізації самого процесу вирішення завдання ідентифікації, який можна представити наступними основними етапами:

1) визначення алфавіту класів;
2) визначення словника ознак;
3) опис алфавіту класів мовою словника ознак;
4) вибір критерію прийняття рішення;
5) моделювання та оцінка методів.

До числа переваг запропонованої структури досліджень відноситься можливість паралельної розробки й випробувань декількох реалізацій п.4 для кожного з варіантів узгодження по п.1, п.2, п.3. Разом з тим,

запропонована етапна загальна схема досліджень, будучи спрямованою на пошук оптимального варіанта функціональної структури системи декодування та ідентифікації кодових комбінацій, припускає відмову від методів динамічного програмування в силу складності цільової функції й обмежень, що носять суперечливий характер. Крім того, жоден з існуючих методів оптимізації, не гарантує визначення глобального екстремуму цільової функції.

Критерії ідентифікації повинні реалізовуватися залежно від умов функціонування системи й визначених вимог.

Існують два основних напрямки синтезу критерію:

– забезпечення $R \min\_ R$ при $v = \text{const}$ (Байєсовська концепція);

– забезпечення $\min\_ v$ при $R = \text{const}$ (послідовний підхід);

При порівнянні цих напрямків виявляється стійка закономірність: або $\_v < v$ при $\_R = R$, або $R < \_R$ при $\_v = v$. Остання обставина в значній мірі визначає вибір критерію на користь послідовного аналізу для рішення завдань декодування та ідентифікації. Інше питання полягає в тому, що навіть квазіоптимальність послідовних багатоальтернативних критеріїв, існування в них моменту зупинки, однозначність прийнятих рішень - найчастіше буває важко довести. Наступний важливий недолік, що різко обмежує область практичного застосування послідовних процедур - можливість затягування процесу ухвалення рішення на час, близький до нескінченності. Вихід запропонований у реалізації скорочених послідовних критеріїв, що забезпечують ухвалення рішення за час $v < v\_max$, де v_max - максимально припустимий час ухвалення рішення по окремій реалізації вимірюваних вибіркових значень. Однак, у послідовних скорочених процедурах важко виконати ті ж вимоги до ймовірності помилки, що й у послідовних процедурах без скорочення.

Крім того доведено, що найбільш перспективним є напрямок, пов'язаний із синтезом критеріїв, що дозволяє забезпечити якийсь оптимальний рівень між середнім ризиком $\_R$ і середнім числом спостережень $\_v$. У подальшому розглянуто процедури прийняття рішення з використанням мінімально достатньої кількості інформації. Припустимо, що є набір класів сигналів $A = \{A_k\}$, $k = 1,..., N$, який будемо характеризувати деяким розподілом

$$P_N = \{p_k, k = 1,..., N\}, \sum_{k=1}^{N} p_k = 1,$$

де p_k – апріорна ймовірність появи сигналу k - го класу.

Визначено, що процес добування інформації полягає в трансформації апріорного розподілу P_N в апостеріорне:

$$Q(x) = \{q_k(\bar{x}), k = 1,...N\},$$

де $$q_k(\bar{x}) = p_k \rho_k(\bar{x}) / \sum_{j=1}^{N} p_j \rho_j(\bar{x}),$$

$\bar{x} = \{x_1, ..., x_n\}$ - вибіркове значення випадкового векторного параметра $\bar{\xi}$.

Доведено, що процес одержання інформації повинен приводити до зменшення або повного зняття невизначеності про той або інший клас сигналів. Тому зазначено кількість інформації як невизначеність, що вдається зняти при трансформації апріорного розподілу P_N. Тоді якщо узагальнену міру невизначеності k-го класу по апріорних ймовірностях визначити як

$$\Psi(p_k) = \varphi(f(p_k)/(p_k)), k = 1, ..., N, \quad (1)$$

а по апостеріорних ймовірностях як

$$\Psi(q_k(\bar{x})) = \varphi(f(q_k(\bar{x}))/q_k(\bar{x})), k = 1, ..., N, \quad (2)$$

то як міра кількості інформації використовується різниця

$$I_k(\bar{x}) = \Psi(p_k) - \Psi(q_k(\bar{x})), k = 1, ..., N. \quad (3)$$

З виразу (3) доведено, що завдання вибору тієї або іншої міри кількості інформації зводиться до завдання вибору міри невизначеності $\Psi(\bullet)$, що, у свою чергу, визначається вибором функції невизначеності $f(\bullet)$. Тоді, використовуючи представлення про функції невизначеності й породжувані ними міри невизначеності, розглянуто завдання розробки методу ідентифікації інформаційного сигналу на базі побудови мір кількості інформації для декодування потоку даних. Показано, що в сучасній теорії прийняття рішення існує квадратична та експоненціальна функції невизначеності, які при відповідних накладених умовах формують міри кількості інформації Котельникова, Шеннона, Кульбака, Байєса та Фішера. Узагальнені міри кількості інформації показують, що кожна з них характеризує множину мір, кожна з якої може бути використана для оцінення інформаційних можливостей сигналів або їхніх параметрів та побудови процедури прийняття рішення.

Завдання теорії ідентифікації сигналів носять статистичний характер. В основу таких завдань закладений вибір тієї або іншої процедури прийняття рішень. Розглянемо процедуру прийняття рішень із позиції одержуваної кількості інформації. Введено послідовність чисел $\{h_k^{(N)}\}, \{s_{(k)}^{(N)}\}$ таких, що $s_k^{(N)} \le h_k^{(N)}, k = 1, ..., N$. На підставі даної послідовності чисел $\{h_k^{(N)}\}, \{s_{(k)}^{(N)}\}$ розглянуто наступні множини інформативності векторного параметра $\bar{\xi} = \{\xi_1, ...\xi_n\}$:

$W^{(+)} = \{\bar{x} : I_k(\bar{x}) \ge h_k^{(N)}, k = 1, ..., N\}$ – множина прийняття гіпотез;

$W^{(-)} = \{\bar{x} : I_k(\bar{x}) \le s_k^{(N)}, k = 1, ..., N\}$ – множина виключення гіпотез;

$W^{(*)} = \left\{\bar{x} : s_k^{(N)} < I_k\left(\bar{x}\right) < h_k^{(N)}, k = 1,..., N\right\}$ – множина невизначеності.

Відповідно до розбиття вибіркового простору X на відповідні множини інформативності, визначено наступний метод ухвалення рішення при завадостійкому декодуванні :

– $\{H_k\}, k = 1,....N$ – послідовність гіпотез щодо можливих кодових комбінацій на вході декодера згорточного декодування з умов, що гіпотеза H_k буде відповідати визначеному класу A_k, тобто k- й кодовій комбінації;

– для гіпотези H_k обчислюється значення кількості інформації $I_k(\bar{x})$, яка міститься у послідовності вибіркового простору X, що ідентифікуються;

– якщо порівняння розрахованої кількості інформації з порогом прийняття рішення мають співвідношення $I_k\left(\bar{x}\right) \geq h_k^{(N)}$, то гіпотеза H_k про k-й кодовий набір приймається як достовірна;

– якщо вирішальне співвідношення має вид: $I_k\left(\bar{x}\right) \leq s_h^{(N)}$, то гіпотеза H_k відкидається;

– якщо $s_k^{(N)} < I_k\left(\bar{x}\right) < h_k^{(N)}$, то рішення про клас A_k не виноситься.

Визначено, що залежно від того, яка з вказаних мір кількості інформації використовується в даній процедурі ухвалення рішення про наявність зазначеного кодового слова, можна визначити метод декодування на базі критеріїв Котельникова, Шеннона, Кульбака, Байєса і Фішера для прийняття «м'якого» рішення на відповідність гіпотези кодовій комбінації. Представлені підходи сформовані на базі інформативності параметрів кодових конструкцій з використанням традиційних, статистичних правил прийняття вірогідного рішення відносно тієї, чи іншої гіпотези. Данні критерії, характеризуються розрахунковою мірою мінімальної кількості інформації та на її основі сформованими порогами прийняття рішення, відносно представлених гіпотез. Апробація та впровадження зазначених методів, зазвичай використовувалась для вирішення задач радіотехніки, радіолокації і навігації з кількістю гіпотез, що не перевищувала $N = 3\text{-}10$ для інформаційних об'єктів різних класів [1, 2].

У випадку вирішення задач згорткового декодування на основі багатоальтернативних правил, процедури прийняття рішень повинні обробити інформаційні складові кодових конструкцій, можлива комбінація котрих складає 512 і більше та є можливість появи слабовідмінних гіпотез. Поріг прийняття рішення $V_k^{(1)}$ для кількості гіпотез від 3 до 10, змінюється в межах 0.67 – 0.2. Однак, данні правила розроблені для ідентифікації сигналів сформованих від різних класів інформаційних об'єктів з різко відмінними інформаційними параметрами (боїнг чи

спортивний літак) та мають конкретні недоліки:

— по-перше, у визначених правилах не беруться до уваги сцени де можлива ситуація появи двох або взагалі десяти однакових гіпотез одночасно, а також інформаційних об'єктів з схожими параметрами (слабовідмінні);

— по-друге, визначені статистичні пороги, що сформовані на базі мінімально-достатньої кількості інформації, при збільшенні кількості альтернативних гіпотез втрачають фізичний зміст та практично дорівнюють нулю.

Класичні багатоальтернативні правила прийняття рішення не адекватні у разі застосування для задач, що вирішуються з метою ідентифікації (відновлення) повної кодової конструкції при вирішення багато альтернативної задачі (з кількістю гіпотез більше 10). Визначенні пороги прийняття рішення статистично занижені і недосяжні в ситуаціях коли сцени ідентифікації інформаційних сигналів мають однакові чи схожі параметри (це стосується не тільки задач декодування кодових комбінацій, а взагалі всіх радіотехнічних задач).

Визначимо рівні мінімально достатньої кількості інформації різних мір кількості інформації, які приймають участь при визначені порогу прийняття рішення [1,2]. (табл. 1)

Якщо, нам невідомі апріорні дані приймемо, що поява всіх кодових в просторі рівно ймовірна.

Тому виникає необхідність побудови рівнів мінімально-достатньої кількості інформації, коли в якості апріорних ймовірностей $p_k = 1/N$, $k = 1...N$. На основі цього рівні мінімально достатньої кількості інформації приймуть наступний вид (табл.2)

Таблиця 1

Мінімально достатня кількість інформації

Міра кількості інформації	Мінімально достатня кількість інформації
Котельникова, Байеса	$I_k^{(1)} = \max\{(1-p_1),...(1-p_N)\} - \max\{p_1...p_N\}$ $I_b^{(2)} = \max\{(1-p_1),...(1-p_N)\} - \max\{p_1...p_N\}$
Шеннона	$I_{sh}^{(3)} = \ln\left(\frac{1-p_k}{p_k}\right),\ k = 1...N$
Фішера	$I_f^{(4)} = \ln p_k (1-p_k) \ln\frac{p_k}{1-p_k},\ k = 1...N$
Кульбака	$I_{kl}^{(5)} = 2\ln\frac{p_k}{1-p_k},\ k = 1...N$

Таблиця 2

Мінімально достатня кількість інформації розрахована по апріорним ймовірностям

Міра кількості інформації	Мінімально достатня кількість інформації
Котельникова, Байеса	$I_k^{(1)} = \frac{N-2}{N}$; $I_b^{(2)} = \frac{N-2}{N}$
Шеннона	$I_{sh}^{(3)} = \ln(N-1)$
Фішера	$I_f^{(4)} = \ln\left(\frac{1}{N-1}\right) \cdot \ln\left(\frac{N-1}{N^2}\right)$
Кульбака	$I_{kl}^{(5)} = 2\ln(N-1)$

3. Розробка методів ідентифікації керуючих кодових конструкцій та побудова моделей процедур статистичних правил

Вирішення встановленого протиріччя, між кількістю слабовідмінних гіпотез та величиною статистичного порогу прийняття рішення, можливо з урахуванням того, що поява всіх однакових гіпотез N одночасно: ймовірна, тобто одночасно можлива поява однієї й тієї ж кодової конструкції (граничний випадок $N = 10$). В даному випадку при присутності під множини n, співпадаючих або слабовідмінних кодових слів, значення статистичного порога прийняття рішення не може перевищувати $V_{min} = 1/n$. Приведений поріг є дійсно мінімальним, що характеризує мінімально-достатню кількість інформації про кодове слово.

Таким чином, дослідження показали, що для вирішення багато альтернативної задачі на базі мінімально – достатньої кількості інформації при відновленні повного кодового слова, існують два статистичних порога прийняття рішень:

- нижній поріг прийняття рішення, що дорівнює $V_{min} = 1/n$ при появі в множині N кодових слів n, що мають слабовідмінні параметри, тобто кодових конструкцій, які мають мінімальну Хемінгову відстань до 4 біт в ближніх позиціях або повністю співпадають;
- верхній поріг прийняття рішення дорівнює V_{max} та встановлений згідно стандартних статистичних правил для інформаційних сигналів з різко відмінними параметрами.

Отже, надалі ми будемо говорити про двох порогові процедури прийняття рішення для слабо та сильно відмінних кодових конструкцій захищених ІКСМ.

Вирішення даного питання, дозволяє нам перейти до наступного та найбільш важливого протиріччя, між зростаючою кількістю альтернативних гіпотез та величиною статистичного порогу прийняття

рішення. Вирішення даної задачі можливе на основі досягнення мінімально достатньої міри кількості інформації, що сформована з урахуванням найбільш інформативних параметрів інформаційних сигналів.

Під інформативними параметрами сигналу будемо розуміти спектральне представлення послідовності кодових слів. Даний вид представлення параметрів сигналу є найбільш інформативним для формування порогів прийняття рішення на базі мінімально-достатньої кількості інформації. Отже, за найбільш інформативний параметр інформаційного сигналу вибираємо не дискретні часові відліки сигналу, а дискретні значення енергетичних складових спектрів відповідних кодових конструкцій.

Спектральне подання сигналу буде використовуватися при розрахунках умовних ймовірностей і виборі найбільш ймовірної гіпотези, що свідчить про кількість інформації, яка відповідає кожній гіпотезі та позначена як: $I_k(x)$ відповідно до $\{H_k\}$, де $k = 1,..., N$. Для кожної гіпотези $\{H_k\}$ обчислюється кількість інформації $I_k(x)$, що міститься у спектрі. Позначимо $S_i(\omega_j)$ ($j = 1,..., N$; $i = 1,..., N$) як спектральне представлення прийнятого інформаційного сигналу при наявності в каналі зв'язку білого гаусівого шуму. Визначено, що міра кількості інформації $I_k(x)$ розрахована для кожної гіпотези $\{H_k\}$, має вид:

$$I_k(x) = 1 - \Psi(q_k(x)), \tag{4}$$

де $q_k(x)$ – апостеріорна ймовірність появи інформаційного сигналу; $\Psi(q_k(x))$ – узагальнена міра невизначеності для k-го інформаційного сигналу, розрахована по апостеріорних ймовірностях.

Визначено, що, враховуючи те, що за інформативний параметр інформаційного сигналу використовується спектральне представлення, (4) прийме вид:

$$I_k\left[P(H_k/S_i(\omega_j))\right] = 1 - \Psi\left[P(H_k/S_i(\omega_j))\right], \tag{5}$$

де $P(H_k/S_i(\omega_j))$ – апостеріорна ймовірність появи інформаційного сигналу; $\Psi\left[P(H_k/S_i(\omega_j))\right]$ – узагальнена міра невизначеності, розрахована по апостеріорній ймовірності:

$$\Psi\left[P(H_k/S_i(\omega_j))\right] = \varphi\left(\frac{f\left[P(H_k/S_i(\omega_j))\right]}{P(H_k/S_i(\omega_j))}\right).$$

У результаті того, що використано, як інформативний параметр спектральне представлення сигналу, розрахунок загального Байєсівського виразу знаходження умовної ймовірності появи кодової конструкції, тобто відповідної гіпотези $\{H_k\}$ за умови прийнятого інформаційного сигналу $S_i(\omega_j)$, визначено як:

$$P\left[H_k / S_i(\omega_j)\right] = \left. \frac{\prod_{j=1}^{N} \frac{1}{\sigma_k \sqrt{2\pi}} e^{-\frac{(S_k(\omega_j) - m_{ij})^2}{2\sigma_k^2}}}{\sum_{i=1}^{N} \prod_{j=1}^{N} \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(S_i(\omega_j) - m_{ij})^2}{2\sigma_i^2}}} \right|_{max}, \quad (6)$$

де $S_i(\omega_j)$ – прийнятий сигнал, за інформативний параметр якого взято спектральне представлення; j – поточний номер спектральних складових кожного $S_i(\omega)$ прийнятого сигналу ($j = 1...N$); i – поточне значення номеру інформаційного сигналу для множини гіпотез $i = 1,...,N$ ($k = 1,...,N$); m_{ij} – двопараметричне математичне очікування [13-14].

За двопараметричне математичне очікування m_{ij} взяте двомірне математичне очікування m_{ij} корисного інформаційного сигналу $S_i(\omega_j)$, де i – поточне значення номера сигналу для множини гіпотез, j – поточний номер спектральних складових кожного інформаційного сигналу). Використання спектрального представлення кодового слова на основі використання двопараметричного математичного очікування $m_{ij} \Leftrightarrow S_i(\omega_j)$, буде найкращим для побудови математичної процедури прийняття рішення з умов збільшення узагальненої міри кількості інформації. Введене математичне очікування використовується при побудові математичної процедури прийняття рішення, а саме: розрахунку умовної щільності розподілу сигналу, апостеріорної ймовірності правильної ідентифікації кодової конструкції, міри невизначеності та міри кількості інформації захищених ІКСМ.

Для прийняття рішення про ідентифікацію прийнятої послідовності $S_i(\omega_j)$ було сформовано наступну процедуру прийняття рішення:

$$I_k \left[P(H_k / S_i(\omega_j))\right] > V_{max}. \quad (7)$$

Доведено, що використання спектрального представлення кодового слова як функції невизначеності $f(x) \Leftrightarrow S_i(\omega_j)$ буде найкращим для побудови математичної процедури прийняття рішення з умов збільшення узагальненої міри кількості інформації та забезпечення ідентифікації кодових конструкцій. Подальше, введено функцію невизначеності, яка покликана зменшити інформаційну невизначеність та забезпечити достатню міру кількості інформації. Дана функція надалі буде використовуватися при побудові математичної процедури прийняття рішення, а саме: розрахунку міри невизначеності, міри кількості інформації та апостеріорної ймовірності правильної ідентифікації кодової конструкції. Визначено функцію невизначеності $f(x)$ як функцію

відповідного амплітудно-частотного спектру $S_i(\omega_j)$ пачки відеоімпульсів (8):

$$f(x) \Leftrightarrow S_i(\omega_j) = A_m \tau \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{nT}\right) n \cdot \frac{T}{2}\right]}{\sin\left[\left(\omega_j \pm \frac{2}{nT}\right) \cdot \frac{T}{2}\right]} \right| \cdot \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{nT}\right) \cdot \frac{\tau}{2}\right]}{\left[\left(\omega_j \pm \frac{2}{nT}\right) \cdot \frac{\tau}{2}\right]} \right| = S(n, \omega_j) B(n, \omega_j), \quad (8)$$

де T – період; τ – тривалість імпульсів; ω_j – основна частота дискретного спектру пачки; $j = 1,..., N$ – поточний номер спектральних складових; n – кількість імпульсів в пачці; $S(n, \omega_j)$ – функція амплітудно-частотного спектру одиночного імпульсу в пачці; $B(n, \omega_j)$ – функція частоти, яка не залежить від форми імпульсу та визначається лише їх числом та періодом слідування.

Для побудови послідовної процедури прийняття рішення, введено поняття ширини «частотних вікон» $\Delta\omega_i$ в задачах ідентифікації слабовідмінних кодових слів. Показано, що під *«частотним вікном»* варто розуміти ефективну ширину спектру сигналу з умов енергетичного або інформаційного вкладу його гармонійних складових в розрахунок апостеріорної ймовірності з умов ідентифікації. Тоді сумарне значення *«частотних вікон»*, які приймуть участь при ідентифікації, дорівнює $\Delta\Omega = \sum_{i=1}^{n} \Delta\omega_i$. Проведено графічний та аналітичний аналіз процедури формування *«частотного вікна»* та визначено інформативні складові спектру частот для 32-бітної кодової конструкції зі слабовідмінними параметрами (рис. 1, 2). Тобто визначено, що інформативними частотами є спектральні складові, що формуються біля основних частот дискретного спектру кодових конструкцій $\omega_j = \frac{j}{T}$. Дані частоти є центрами *«частотних вікон»* : $\omega_j = \frac{j}{T}$, $j = 0,2,4....$ Доведено, що необхідно розглядати ефективну ширину *«частотного вікна»* як співвідношення: $\Delta\omega_i = \omega_j \pm \frac{2}{n \cdot T} = \frac{j}{T} \pm \frac{2}{n \cdot T} = \frac{j \cdot n \pm 2}{n \cdot T}$. Визначено, що ефективна ширина «вікна», де знаходиться основна частота дискретизації, може бути графічно відображена згідно з рис.1,2.

Використовуючи введену функцію невизначеності (8), побудовано узагальнену міру невизначеності, на основі якої розраховано мінімально достатню кількість інформації $I_k(x)$ використовувану при формуванні статистичного правила прийняття рішення та розраховувану для кожної гіпотези $\{H_k\}$:

$$I_k\left[P\left(H_k/S_i\left(\omega_j\right)\right)\right]=1-\Psi\left[P\left(H_k/S_i\left(\omega_j\right)\right)\right]=1-\varphi\left(\frac{f\left[P\left(H_k/S\left(n,\omega_j\right)B\left(n,\omega_j\right)\right)\right]}{P\left(H_k/S\left(n,\omega_j\right)B\left(n,\omega_j\right)\right)}\right). \quad (9)$$

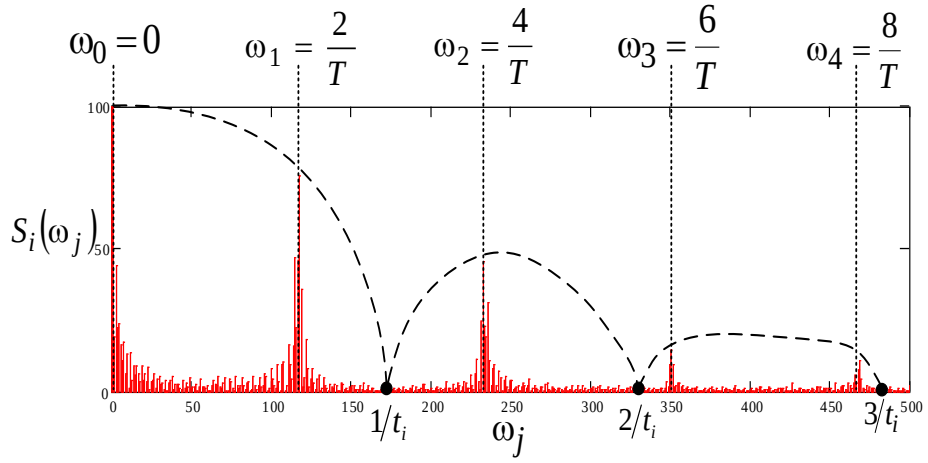


Рис.1. Амплітудно-частотний спектр 32-бітної кодової конструкції

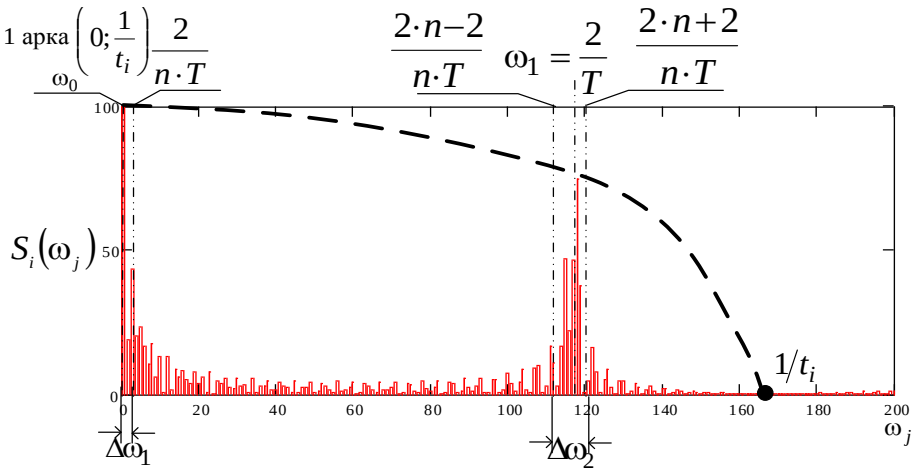


Рис.2. Ефективна ширина спектру $\Delta\omega_i$ *(«частотного вікна»), з центрами на частотах дискретизації* ω_j

Показано, що сформована процедура відповідає визначеній задачі – ідентифікації слабо відмінних кодових конструкцій на основі інформативних частот спектру з умов введення зазначеної функції невизначеності у розрахунку мінімально-достатньої кількості інформації (з урахуванням присутності в кодових послідовностях N «слабовідмінних» кодових слів прийме вид 10).

На основі проведених досліджень побудовано графік зростання ймовірності правильної ідентифікації залежно від кількості використаних інформативних спектральних складових. Доведено, що кожна спектральна складова вносить певну міру кількості інформації $I_k\left[P\left(H_k/S_i\left(\omega_j\right)\right)\right]$, таким чином відповідно збільшує апостеріорну ймовірність правильної ідентифікації $P\left(H_k/S_i\left(\omega_j\right)\right)$ кодової конструкції.

$$I_k\left\{\frac{\prod_{j=1}^{N}\frac{1}{\sigma_k\sqrt{2\pi}}\exp\frac{-\left(A_m\tau\left|\frac{\sin\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)n\cdot\frac{T}{2}\right]}{\sin\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)\cdot\frac{T}{2}\right]}\right|\cdot\left|\frac{\sin\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)\cdot\frac{\tau}{2}\right]}{\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)\cdot\frac{\tau}{2}\right]}\right|-m_{ij}\right)^2}{2\sigma_k^2}}{\sum_{i=1}^{N}\prod_{j=1}^{N}\frac{1}{\sigma_i\sqrt{2\pi}}\exp\frac{-\left(A_m\tau\left|\frac{\sin\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)n\cdot\frac{T}{2}\right]}{\sin\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)\cdot\frac{T}{2}\right]}\right|\cdot\left|\frac{\sin\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)\cdot\frac{\tau}{2}\right]}{\left[\left(\omega_j\pm\frac{2}{n\cdot T}\right)\cdot\frac{\tau}{2}\right]}\right|-m_{ij}\right)^2}{2\sigma_i^2}}\right\}>V_{max.} \quad (10)$$

З отриманих результатів зроблено висновки, що використання 20 гармонічних складових вносять таку ж саму міру кількості інформації, як і використання всієї вибірки з 2049 гармонічних складових. Апостеріорна ймовірність правильної ідентифікації при використанні всієї вибірки приймає значення $P^{(2049)}\left(H_k/S_i\left(\omega_j\right)\right)=0.39$, а при використанні інформативних складових трьох спектральних «вікон»: $P^{(20)}\left(H_k/S_i\left(\omega_j\right)\right)=0.37$ (рис. 3).

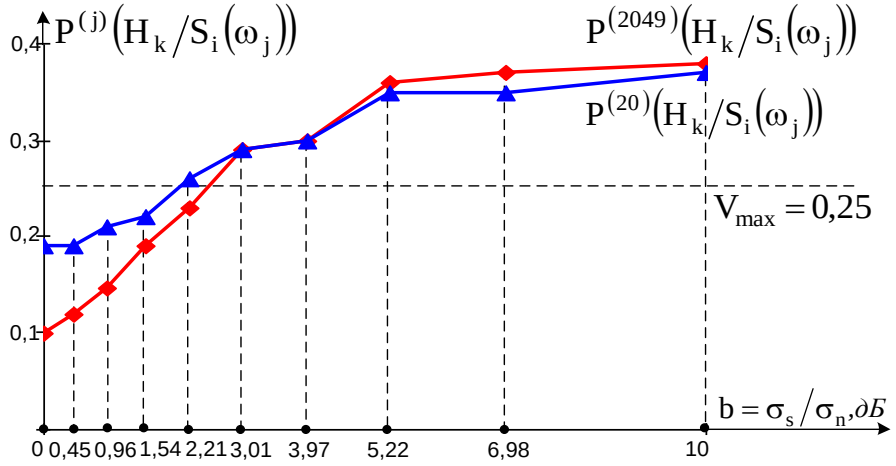


Рис.3. Ймовірність правильної ідентифікації залежно від співвідношення сигнал/шум (30 гіпотез)

Показано, що кількісний рівень ймовірності порогів правильної ідентифікації достатній і відповідає встановленим умовам відновлення та ідентифікації слабовідмінних керуючих кодових конструкцій захищених ІКСМ. Для оцінення якості нововведеного методу побудуємо графічні залежності ймовірності правильної ідентифікації від співвідношення сигнал/шум $b=\sigma_s/\sigma_n$. В роботі проведено порівняння ймовірності правильної ідентифікації при використанні в процедурі всіх складових спектру та тільки трьох «частотних вікон» з 20 інформативними частотами.

Графічні залежності побудовані з умов ідентифікації *30 гіпотез*

(кодових слів) при використанні 20 найбільш інформативних складових частотного спектру. Ґрунтуючись на проведених дослідженнях, можна дійти висновку: трьох «частотних вікон» достатньо для побудови ефективної процедури прийняття рішення на базі накопичення достатньої кількості інформації. На основі наведеного визначимо можливість зупинки послідовної процедури декодування на основі забезпечення достатньої кількості інформації [15-16].

На базі розробленого методу відновлення та ідентифікації кодових конструкцій розроблено модель процедури статистичних правил прийняття рішення у задачах відновлення повної кодової конструкції на базі мінімально-достатньої кількості інформації, що дало можливість виконання послідовної процедури зупинення процесу ідентифікації у разі відповідності отриманих результатів встановленим критеріям (рис. 4).

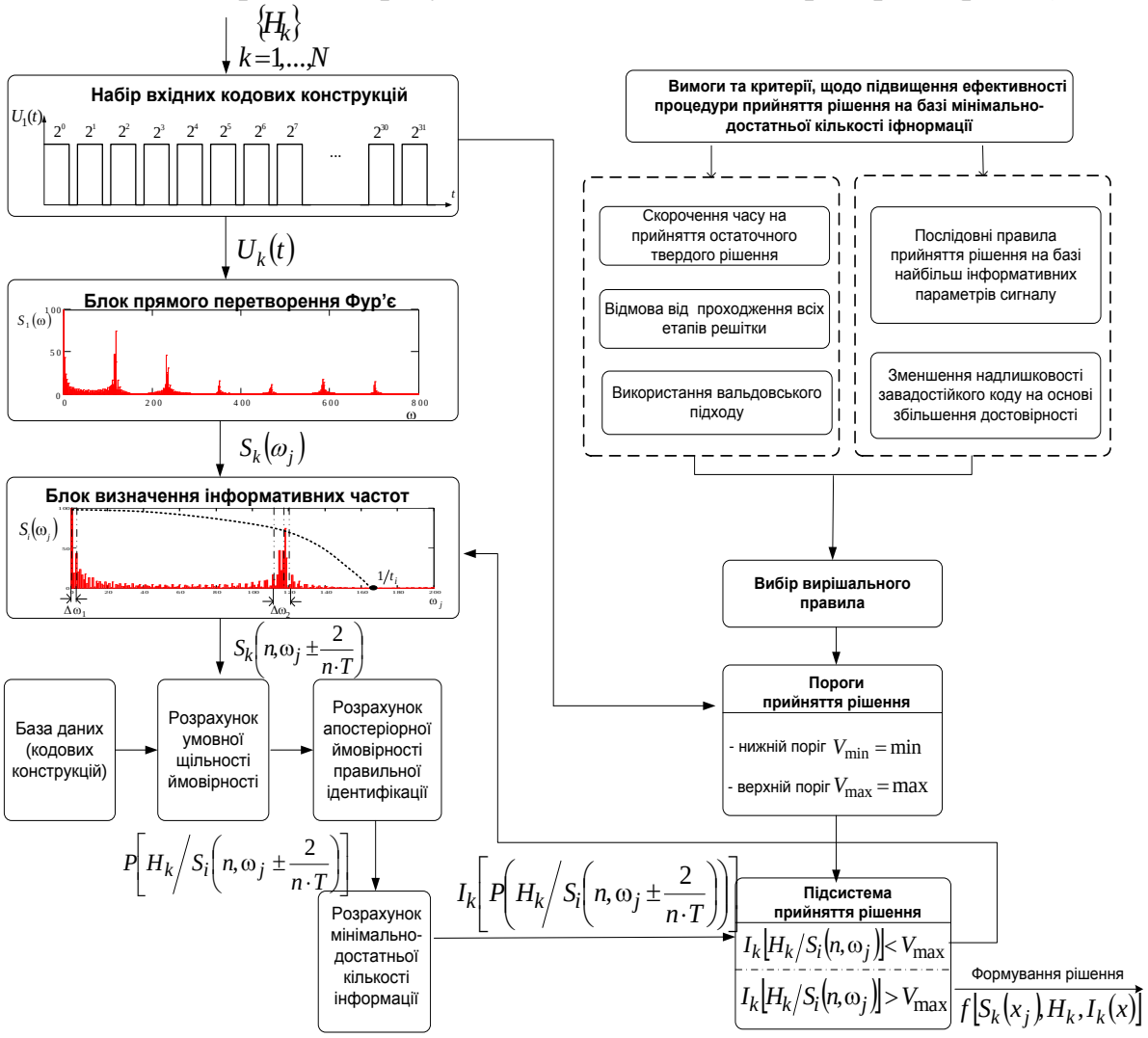


Рис. 4. Математична модель реалізації методу ідентифікації повної кодової конструкції з слабо відмінними параметрами для захищеної ІКСМ

В результаті використання розробленого методу зменшується час, затрачений на аналізування і оброблення всього спектру, підвищується

ефективність та результативність розробленої послідовної процедури прийняття рішення за рахунок використання тільки інформативних складових частотного спектру, тобто підвищується ефективність та надійність системи захисту інформації.

Розглянемо на основі розробленого методу приклад реалізації процедури ідентифікації та відновлення кодових конструкцій для організації системи управління доступом захищених ІКСМ (рис. 5) [8-9].

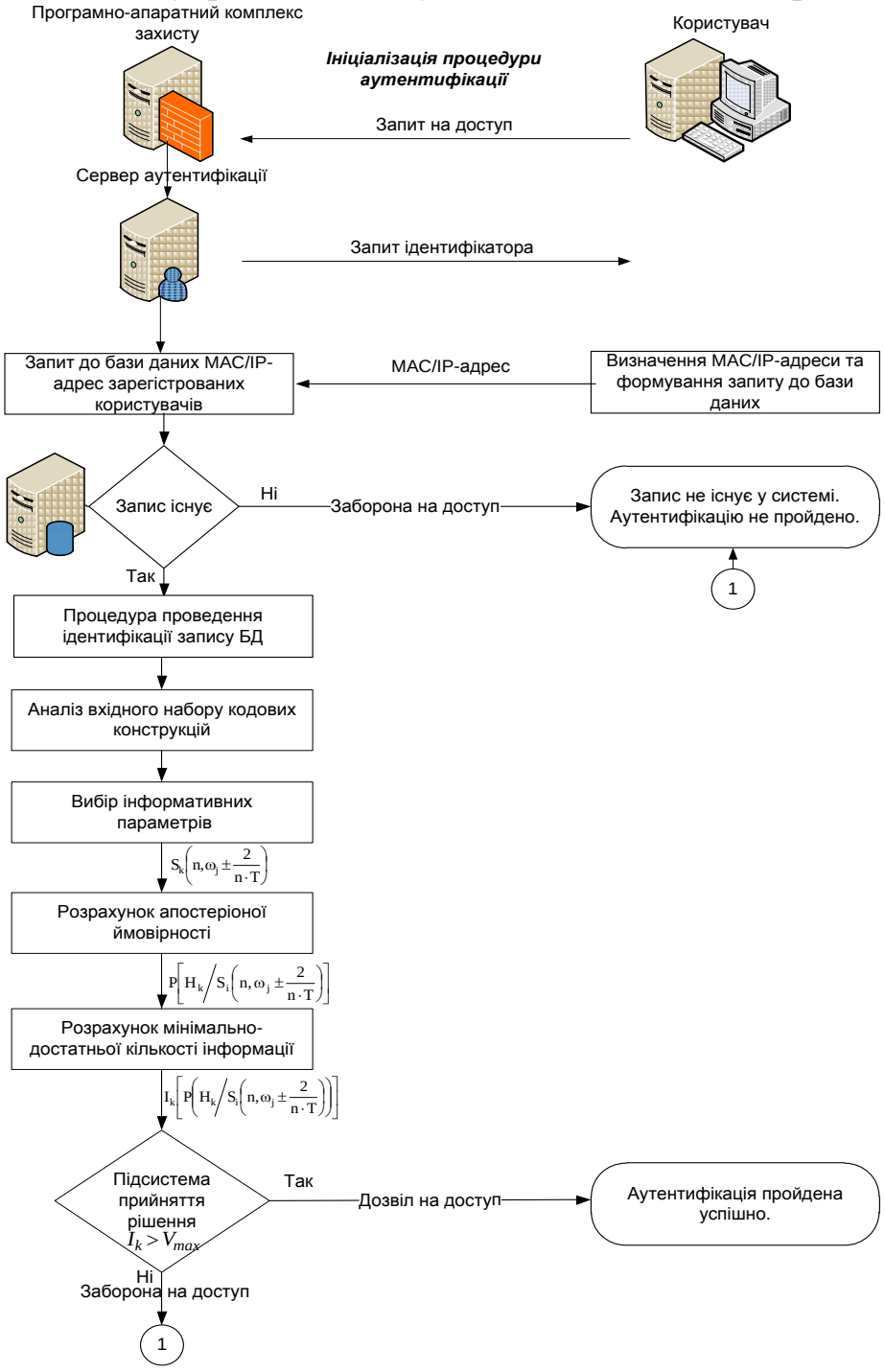


Рис.5. Реалізації методів ідентифікації та відновлення кодових конструкцій для організації системи управління доступом захищених ІКСМ

Висновки

1. Вперше розроблено метод відновлення кодових конструкцій та побудована модель процедури статистичних правил прийняття рішення на базі мінімально-достатньої кількості інформації, що дало можливість усунення надлишковості завадостійкого коду та звільнило від 6,25 до 12,5% обсягу кодового слова для передачі додаткової корисної інформації.

2. Вперше розроблено новий метод визначення інформативних складових сигналу при введенні заданої функції невизначеності з умов побудови послідовного правила прийняття рішення, що дозволило зменшити ймовірність не коригованих помилок для розробленого методу: при 1 бітній помилці ймовірність появи не коригованих помилок знизилася від 1,04 до 8,11 разу; при 2 бітній помилці – від 1,03 до 12,6 разу; при 3 бітній помилці – від 1,06 до 9,63 разу; при 4 бітній помилці – від 1,03 до 10,0 разу; при 5 бітній помилці – від 1,07 до 11,85 разу залежно від використаного методу завадостійкого кодування на основі збільшення достовірності та цілісності інформаційного потоку даних.

3. Вперше побудовано структурну та математичну модель процедури статистичних правил прийняття рішення у задачах декодування повної кодової конструкції, що дозволило мінімізувати час на процедури оброки сигналів без втрат якості інформаційного потоку даних від 1,5 до 5,9 разів.

4. Вдосконалено метод оцінення ефективності розроблених методів та побудованої моделі послідовних процедур прийняття рішення з урахуванням відновлення стандартних та слабовідмінних керуючих кодових конструкцій сучасних захищених ІКСМ. Проведене оцінювання часових характеристик та складності реалізації показало підвищення ефективності функціонування захищених ІКСМ з умови забезпечення скороченням часу на процедуру ідентифікації для 32-бітної кодової конструкції від 1,3 до 5,5 разу та для 64-бітної – від 1,25 до 10,9 разу при підвищенні достовірності прийняття рішення.

5. Розроблено програмно-апаратний комплекс реалізації розроблених методів, моделі, методик з урахуванням забезпечення достовірності та цілісності керуючих кодових конструкцій захищених ІКСМ.

Література

1. Косенко Г.Г. Критерии информативности при различении сигналов/ Г.Г. Косенко. – М.: Радио и связь, 1982. – 216 с.

2. Косенко Г.Г. Метод последовательного расширения областей принятия решений в задачах распознавания // Радиоелектроника. – М. – 1980. – №27 – С. 72–75.

3. Сидельников В. М. Теория кодирования/ В. М. Сидельников. – М.: ФИЗМАТЛИТ, 2008 – 324 с.

4. Скляр Б. Цифровая связь. Теоретические основы и практическое применение/ Б. Скляр. – М.: Издательский дом “Вильямс”, 2003. – 1104 с.

5. Никитин Г. И. Эффективные коды / Г. И. Никитин. – ЛИАП. Л., 1987. – 98 с.
6. Никитин Г. И. Корректирующие коды / Г. И. Никитин, И. Б. Антипова, А. В. Красновидов. – ЛИАП. Л., 1989. – 94 с.
7. Никитин Г. И. Помехоустойчивые циклические коды / Г.И. Никитин,С. С. Поддубный. – СПбГУАП. СПб., 1998. – 102 с.
8. Пат. 55214 Україна, МПК H03M 13/00 Спосіб відновлення кодових конструкцій на базі інформативних складових / Юдін О.К., Курінь К.О., Чунарьова А.В.; заявник та патентовласник Нац. авіац. ун-т. – u201006045; заявл. 19.05.2010; опубл. 10.12.2010, Бюл. №23. – 8 с.
9. Пат. 59534 Україна, МПК H03M 13/00 Система процедури статистичних правил прийняття рішення у задачах ідентифікації повної кодової конструкції з слабо відмінними параметрами / Юдін О.К., Курінь К.О., Чунарьова А.В.; заявник та патентовласник Нац. авіац. ун-т. – № u201010898; заявл. 10.09.2010; опубл. 25.05.2011, Бюл. №10. – 10 с.
10. Пат. 60394 Україна, МПК H03M 13/00 Спосіб спектрального визначення інформативних складових в процедурах усунення інформаційної невизначеності / Юдін О.К., Курінь К.О., Чунарьова А.В.; заявник та патентовласник Нац. авіац. ун-т. – u201006817; заявл. 02.06.2010; опубл. 25.06.2011, Бюл. № 12. – 6 с.
11. Юдін О.К. Кодування в інформаційно-комунікаційних мережах: монографія/ О.К. Юдін. – К.: НАУ, 2007. – 308с.
12. Юдін О.К. Оптимізація методів декодування інформаційних сигналів / О.К. Юдін, А.В. Чунарьова // Зб. наук. праць Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова. – К. : ІПМЕ НАН України. – 2009. – Вип. 51. – С. 84–90.
13. Юдін О.К. Математичні аспекти використання багатоальтернативних правил в задачах канального кодування інформаційних потоків / О.К. Юдін, А.В. Чунарьова // Проблеми інформатизації та управління : зб. наук. праць. – К. : Вид-во Нац. авіац. ун-ту «НАУ-друк». – 2008 – Вип.1. – С. 172 – 178.
14. Юдін О.К. Спектральні методи визначення інформативних складових в процедурах усунення інформаційної невизначеності / О.К. Юдін, А.В. Чунарьова // Радиотехника: всеукраинский межведомственный науч.-техн. сб. – Х.: ХНУРЭ. – 2009. – Вып. 159. – С. 209 – 214.
15. Юдін О.К. Оцінка ефективності методів ідентифікації кодових конструкцій в задачах канального декодування / О.К. Юдін, А.В. Чунарьова, Ю.Б. Чеботаренко // Вісник Інженерної академії України. – 2009. – Вип. 3–4. – С. 157 – 162.
16. Юдін О.К. Підвищення достовірності відновлення інформаційних потоків / О.К. Юдін, А.В. Чунарьова // Наукоємні технології. – 2010. – Вип. 4. – С. 84 – 88.

ЧАСТЬ 3
КИБЕРБЕЗОПАСНОСТЬ, ЗАЩИТА ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ РАЗВЕДКА В ИНФОКОММУНИКАЦИОННОМ ПРОСТРАНСТВЕ

АНАЛИЗ ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ МАСКИРОВАНИЯ ПРИ ВЫЯВЛЕНИИ ОБЛАСТЕЙ ДЛЯ СТЕГАНОГРАФИЧЕСКОГО ВСТРАИВАНИЯ

Баранник В.В., Бекиров А.Е., Баранник Д.В.

Введение

Из анализа существующих подходов стеганографического преобразования на основе использования изображения в качестве контейнера можно сформулировать общие недостатки. Такие недостатки характерны как для методов непосредственного, так и косвенного встраивания, а именно [1, 2, 3]:

- визуальные искажения, которые вносятся в изображение в результате стеганографического встраивания;
- низкая устойчивость встроенных данных к атакующим воздействиям, а именно компрессионные атаки и ошибки в канале передачи данных.

Для устранения выявленных недостатков при проектировании стеганографического метода предлагается разработать подход, для выявления областей изображения, которые устойчивы к атакующим воздействиям (в качестве модели атакующего воздействия предлагается использовать дискретное косинусное преобразование (ДКП) с различными коэффициентами квантования). Здесь под устойчивостью областей понимается следующее: если в результате атакующего воздействия (компрессионная атака ДКП с квантованием) значения элементов пространственного представления областей изображения остаются не измененными, то в этом случае область изображения, которая содержит такие элементы, называется ***устойчивой к атакующим воздействиям***.

Стеганографическое преобразование на основе использования устойчивых областей изображения с одной стороны позволит повысить стойкость встроенных данных к атакам, а соответственно и увеличить значение вероятности правильного изъятия встроенных данных в условиях наличия атак. С другой стороны, неподверженность элементов таких областей к воздействиям позволит осуществлять стеганографическое встраивание путем минимальной модификации, что в свою очередь отразиться на количестве вносимых искажений в результате встраивания.

Значит, ***цель исследования*** заключается в разработке подхода для выявления областей изображения, потенциально устойчивых к компрессионным атакам на основе ДКП.

1. Механизм выявления областей изображения, устойчивых к атакующим воздействиям

Для цифровых изображений наиболее полезной семантической нагрузкой обладают контуры объектов [4]. Контуры представляют собой линии, которые проходят на границах однородных областей. Элементы $\{z_{i,j}\}$ пространственно-временного представления изображения, значения которых не превышают определенного порога, формируют однородные области. Это задается следующим условием:

$$\left| z_{max} - z_{min} \right| \le l,$$

где z_{max} - элемент области изображения, который обладает наибольшим значением, определяется на основе следующего выражения:

$$z_{max} = \max_{1 \le i \le x} \{z_{i,j}\}, \quad j = \overline{1, y};$$

z_{min} - элемент области изображения, который обладает наименьшим значением, определяется на основе формулы:

$$z_{min} = \min_{1 \le i \le x} \{z_{i,j}\}, \quad j = \overline{1, y};$$

l - порог определения однородных областей.

Существующие компрессионные алгоритмы предназначены для сокращения избыточности изображений (психовизуальной, статистической, структурной, комбинаторной). При этом методы направлены на устранение избыточности одновременно с внесением наименьшего количества искажений, т.е. сохранения высокого качества изображения [4]. Для обеспечения невосприимчивости человеческого зрения к искажениям, компрессионное преобразование не вносит искажения в семантически значимые для зрения области [5]. Значит, для выявления областей, устойчивых к компрессионным воздействиям необходимо использовать методы выделения контуров изображения, для последующего их использования для стеганограифческого встраивания.

Контуры изображения формируются на границах однородных областей изображения. Для того что бы определить принадлежность элементов пространственного представления изображения к однородной области одновременно с проверкой наличия контура необходимо выполнение следующего условий:

- принадлежность элемента изображения $z_{i,j}$ к однородной области задается условием

$$\left| z_{i,j} - z_{i \pm 1, j \pm 1} \right| \le l \text{, где } i = \overline{1, x}; \quad j = \overline{1, y};$$

- принадлежность элемента изображения $z_{i,j}$ к соседней однородной области (формирование контура) определяется выполнением условия

$$\left| z_{i,j} - z_{i\pm1,j\pm1} \right| > l \text{, где } i=\overline{1,x};\ j=\overline{1,y}.$$

Наиболее распространенные и применяемые на практике подходы выявления контуров являются градиентные методы. Градиентные методы основаны на определении в каждой точке массива значения увеличения яркости (градиента) и направления их наибольшего изменения с последующим определением максимальных значений градиента яркости, их статистической обработки и деления на пороги (уровни).

Наиболее распространенным способом поиска контуров является обработка изображения Z скользящей маской K. Маска K представляет собой квадратную матрицу с коэффициентами $\{k\}$. Процесс обработки изображения Z на основе матрицы K называется фильтрацией или маскированием и задается следующим функционалом $f(\bullet)$ (рис. 1):

$$M = f(Z, K),$$

где M - изображение, полученное результате обработки изображения Z на основе маски K.

Процесс фильтрации основан на постепенном пространственном перемещении маски фильтра от элемента к элементу изображения. Из анализа рис. 1 видно, что значение элемента $m_{i,j}$ (отклика фильтрации) вычисляется с использованием значений предыдущих и последующих элементов в двумерной плоскости.

В этом случае значение элемента $m_{i,j}$ изображения M, полученного врезультате маскирования определяется по формуле:

$$m_{i,j} = \sum_{\xi=i-1}^{i+1} \sum_{\tau=i-1}^{\tau+1} z_{\xi,\tau} \cdot k,$$

или

$$m_{i,j} = z_{i-1,j-1} \cdot k_{-1,-1} + z_{i,j-1} \cdot k_{0,-1} + z_{i+1,j-1} \cdot k_{1,-1} + z_{i-1,j} \cdot k_{-1,0} +$$
$$+ z_{i,j} \cdot k_{0,0} + z_{i+1,j} \cdot k_{1,0} + z_{i-1,j+1} \cdot k_{-1,1} + z_{i,j+1} \cdot k_{0,1} + z_{i+1,j+1} \cdot k_{1,1},$$

или

$$m_{i,j} = z_{i-1,j-1} \cdot k_{-1,-1} + z_{i,j-1} \cdot k_{0,-1} + z_{i+1,j-1} \cdot k_{1,-1} + z_{i-1,j} \cdot k_{-1,0} +$$
$$+ z_{i,j} \cdot k_{0,0} + z_{i+1,j} \cdot k_{1,0} + z_{i-1,j+1} \cdot k_{-1,1} + z_{i,j+1} \cdot k_{0,1} + z_{i+1,j+1} \cdot k_{1,1}.$$

В качестве метода выделения контуров изображения предлагается использовать оператор Собеля. Данный оператор наиболее часто используется на практике, и имеет следующий вид:

$$m_{i,j} = \sqrt{G_i^2 + G_j^2};$$

$$G_i = K_i \cdot m_{i,j} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \cdot m_{i,j}; \quad G_j = K_j \cdot m_{i,j} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \cdot m_{i,j},$$

где K_i и K_j - операторы для определения приращения значения элемента изображения по горизонтали и вертикали соответственно.

G_i и G_j - блоки изображения, каждый элемент которого содержит приближенные значения производных по горизонтали и вертикали соответственно.

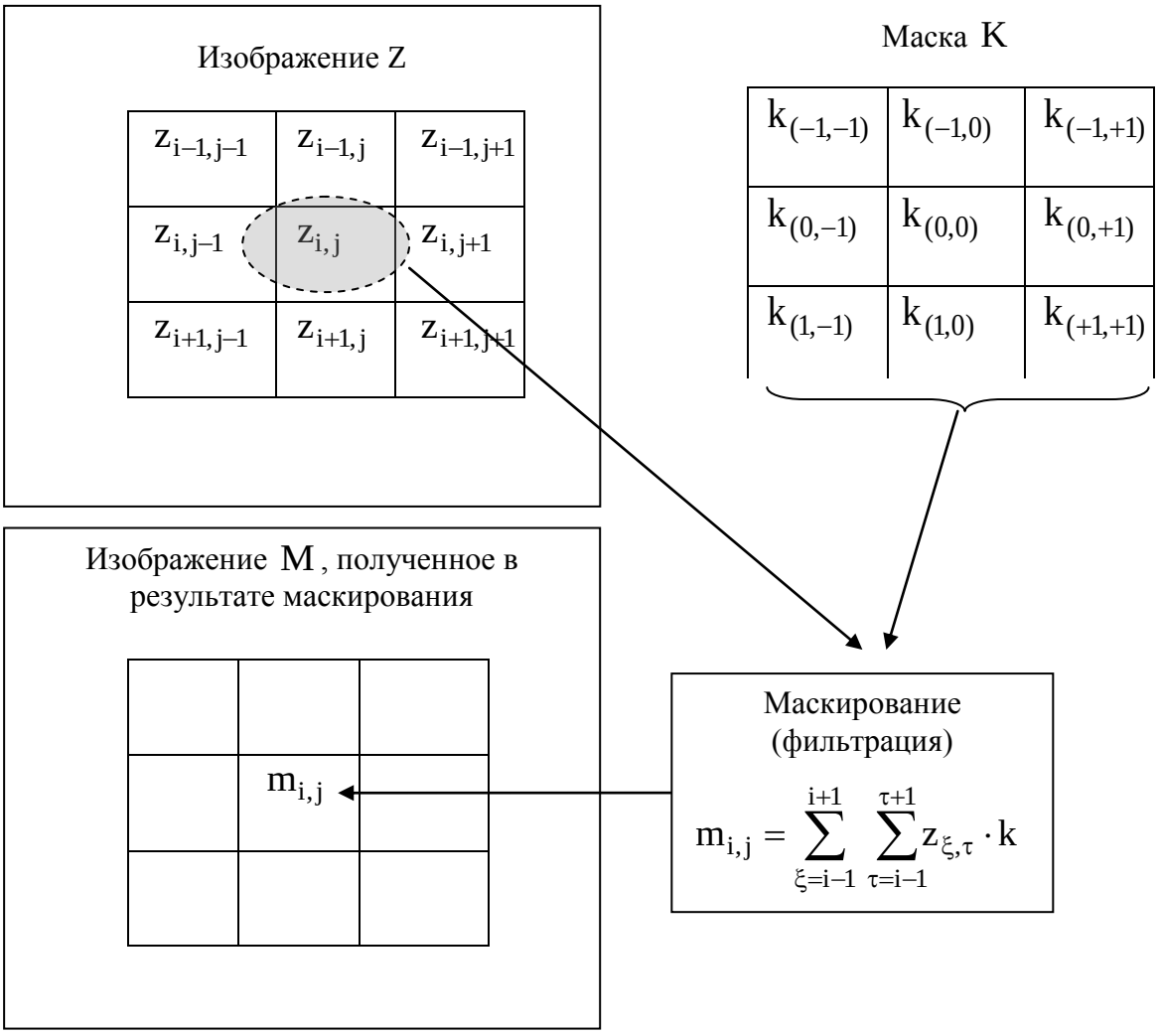


Рис. 1. Схема реализации фильтрации изображения на основе скользящей маски

Для маски Собеля K размерностью 3x3 элемента, которая представлена на рис. 1, значение элемента $m_{i,j}$ изображения после маскирования можно определить на основе формулы в линейном виде, а именно:

$$m_{i,j} = \sqrt{G_i^2 + G_j^2}\ , \text{ где}$$

$$G_i = (z_{i+1,j-1} + 2 \cdot z_{i+1,j} + z_{i+1,j+1}) - (z_{i-1,j-1} + 2 \cdot z_{i-1,j} + z_{i-1,j+1});$$

$$G_j = (z_{i-1,j+1} + 2 \cdot z_{i,j+1} + z_{i+1,j+1}) - (z_{i-1,j-1} + 2 \cdot z_{i,j-1} + z_{i+1,j-1}).$$

Рассмотрим роботу оператора Собеля K на примере маскирования (фильтрации) изображения Z размерность 6×6 элементов (рис. 2).

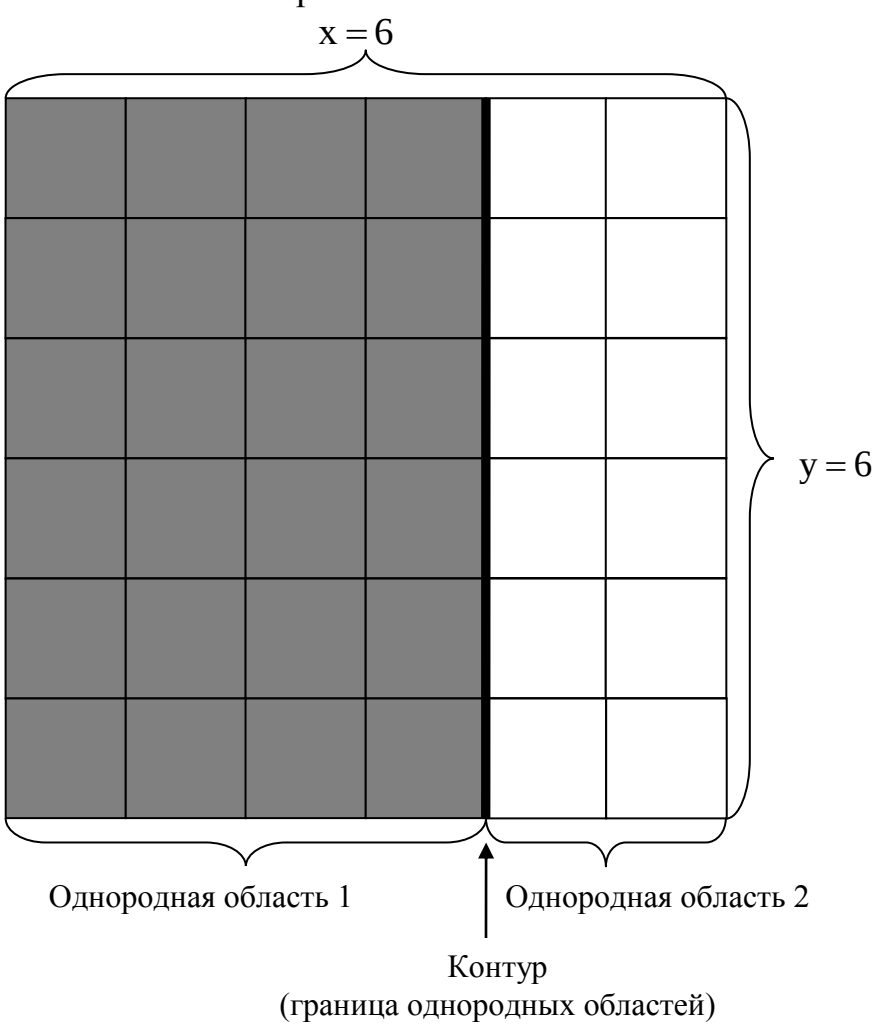


Рис. 2. Исходное изображение с двумя однородными областями

В качестве примера для изображения Z

$$Z = \begin{vmatrix} 200 & 201 & 201 & 204 & 120 & 121 \\ 201 & 203 & 200 & 200 & 121 & 122 \\ 203 & 203 & 201 & 201 & 123 & 121 \\ 204 & 204 & 201 & 204 & 121 & 120 \\ 202 & 204 & 201 & 202 & 120 & 124 \\ 201 & 202 & 204 & 200 & 120 & 121 \end{vmatrix}.$$

Вычислим значения элементов изображения после маскирования (значение откликов маскирования) с следующими координатами:

1. $i=1$ и $j=1$.

В этом случае есть особенность вычисления значения элемента $m_{1,1}$ изображения после маскирования. Учитывая, что элементы левее и выше $z_{1,1}$ отсутствуют, в этом случае при вычислении маски они заменяются нулями либо заменяются значениями соседних элементов. В первом случае края изображения будут определяться как контуры. Поэтому при вычислении элементов на краях изображения, недостающие элементы будут дополняться значениями соседних пикселей. Тогда элемент $m_{1,1}$ будет принимать значение:

$$m_{1,1}=\sqrt{G_i^2+G_j^2}=7;$$

$$G_i=(201+2\cdot 201+203)-(200+2\cdot 200+201)=5;$$

$$G_j=(201+2\cdot 201+203)-(200+2\cdot 200+201)=5.$$

2. $i=4$ и $j=4$.

В этом случае элемент $z_{4,4}$ находится на границе двух областей, а элемент $m_{4,4}$ примет следующее значение:

$$m_{1,1}=\sqrt{G_i^2+G_j^2}=\sqrt{706^2+(-319)^2}=774;$$

$$G_i=(201+2\cdot 202+120)-(201+2\cdot 201+123)=706;$$

$$G_j=(123+2\cdot 121+120)-(201+2\cdot 201+201)=-319.$$

В результате применения оператора Собеля K к изображению Z получена маски M, элементы $m_{i,j}$ которой принимают следующие значения:

$$Z=\begin{vmatrix} 7 & 4 & 7 & 322 & 327 & 5 \\ 12 & 7 & 7 & 318 & 317 & 5 \\ 10 & 11 & 11 & 317 & 318 & 6 \\ 2 & 10 & 7 & 320 & 324 & 5 \\ 12 & 4 & 6 & 326 & 319 & 8 \\ 7 & 8 & 8 & 333 & 315 & 11 \end{vmatrix}.$$

На рис. 3 представлена маска M исходного изображения Z.

Из анализа значений маски M можно сделать следующие выводы:

1. Элементы $m_{i,j}$ маски M, которые соответствуют элементам $z_{i,j}$ однородных областей изображения Z, принимают минимальные значения.

2. Значения элементов $m_{i,j}$, расположенные на границе однородных областей, принимают максимальное значение.

Таким образом, применение плавающей маски Собеля позволяет выделить контуры объектов на границах однородных областей. Предлагается использовать рассмотренную технологию для выявления областей, которые будут использоваться при стеганографическом встраивании.

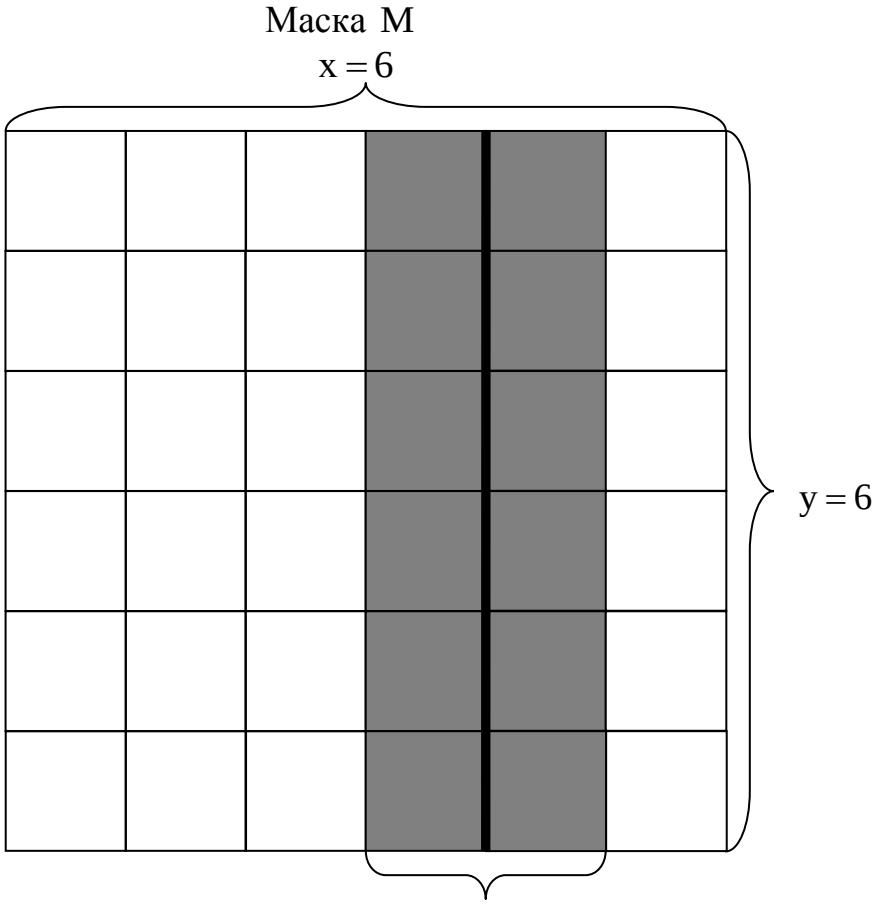


Рис. 3. Маска изображения, полученная в результате применения оператора Собеля

2. Сравнительная оценка устойчивости контуров изображений различной насыщенности к атакующим воздействиям

Проверку устойчивости элементов контуров изображения к дискретному косинусному преобразованию предлагается проводить на основе эксперимента. Суть эксперимента заключается в оценки визуальных искажений масок, полученных на основе изображений, которые будут подвержены дискретному косинусному преобразованию с различными коэффициентами квантования. Схематично этапы эксперимента представлены на рис. 4.

В качестве исследуемых изображений используются:

1. Сильнонасыщенное изображение рис. 5.

2.Средненасыщенное изображение рис. 6.
3. Слабонасыщенное изображение рис. 7.

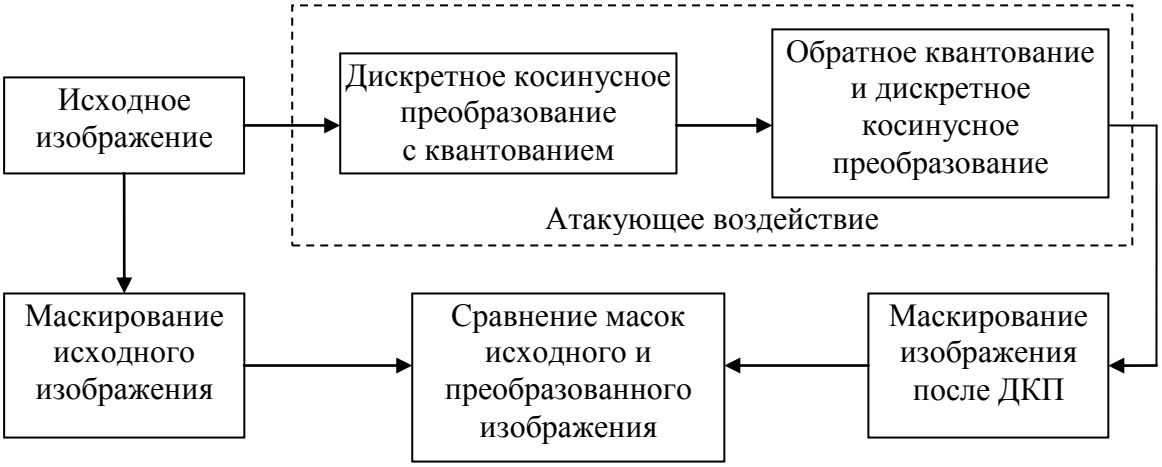


Рис 4. Этапы эксперимента по выявлению устойчивых областей изображения

Рис. 5. Сильнонасыщенное изображение

Рис. 6. Средненасыщенное изображение

Рис. 7. Слабонасыщенное изображение

Рассмотрим результаты эксперимента для слабонасыщенного изображения рис. 7. Для данного изображения проводится дискретное косинусное преобразование с различными коэффициентами квантования $q = 1, 2, 5, 10, 20$. На рис. 8 представлены изображения, полученные в результате применения к исходному слабонасыщенному изображению дискретного косинусного преобразования с различными коэффициентами квантования.

а) $q = 1$

б) $q = 2$

в) $q = 5$

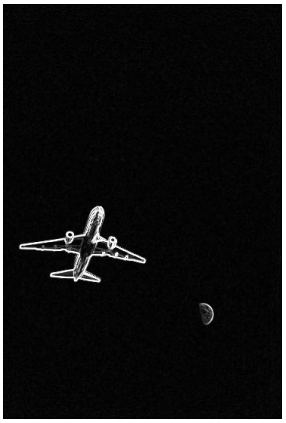
г) $q = 10$

д) $q = 20$

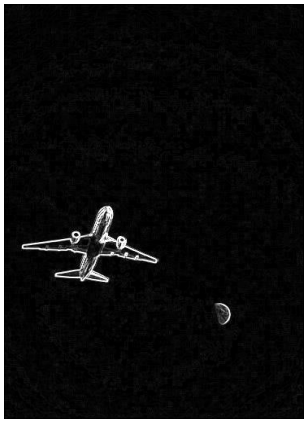
Рис. 8. Слабонасыщенные изображения, полученные в результате ДКП с различными коэффициентами квантования

Из анализа изображения на рис. 8 можно сделать вывод, что атакующее воздействие (ДКП и квантования) вносит визуальные искажения в исходное изображение. Сильнее всего визуальные искажения проявляются на изображении, которое получено после ДКП с коэффициентом квантования $q = 20$.

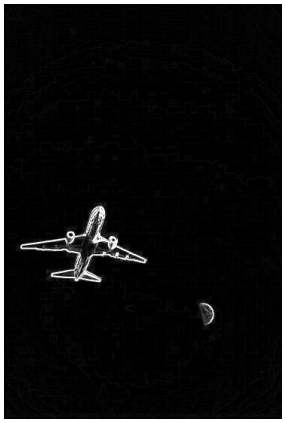
Теперь необходимо провести маскирование изображений после атакующего воздействия. Данный этап предусматривает применение плавающего оператора Собеля к каждому элементу пространственного представления слабонасыщенного изображения после дискретного косинусного преобразования с квантованием. Изображения, полученные после маскирования (выделения контуров изображений) представлены на рис. 9.



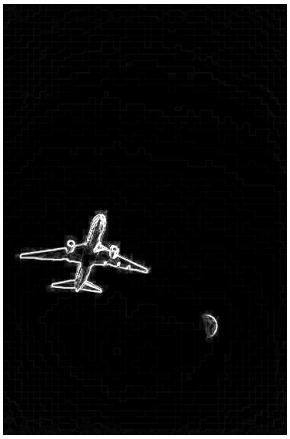
а) маска исходного изображения



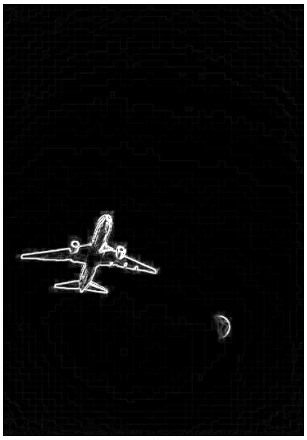
б) маска изображения после ДКП с квантованием $q = 1$



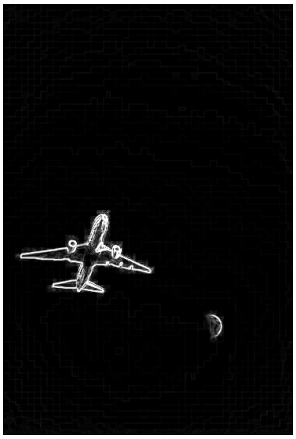
в) маска изображения после ДКП с квантованием $q = 2$



в) маска изображения после ДКП с квантованием $q = 5$



б) маска изображения после ДКП с квантованием $q = 10$



б) маска изображения после ДКП с квантованием $q = 20$

Рис. 9. Маски, полученные в результате применения оператора Собеля к изображениям после атакующего воздействия

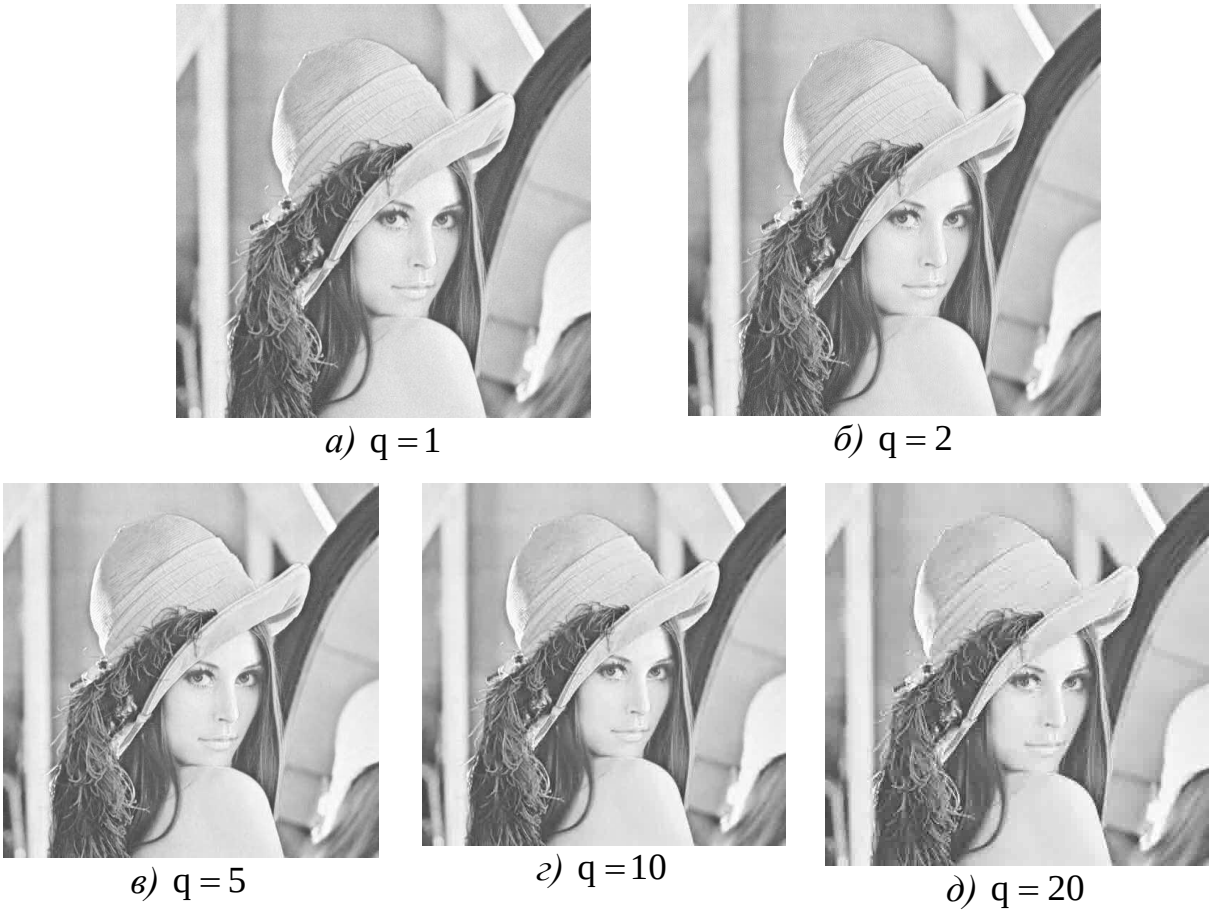
Из анализа масок, получены в результате применения оператора Собеля к изображениям на рис. 8, можно сделать следующие выводы:

1) в результате маскирования выделяются контуры на границах однородных областей;

2) визуальная оценка масок изображений после атакующего воздействия позволяет заключить, что контуры объектов не имеют значительных визуальных искажений;

3) контуры, полученные в результате маскирования изображений после атакующего воздействия, обладают устойчивостью к дискретному косинусному преобразованию с квантованием.

Теперь рассмотрим результаты эксперимента для средненасыщенного изображения. Изображения, полученные в результате применения дискретного косинусного преобразования к исходному изображению (рис. 6) с коэффициентами квантования $q = 1, 2, 5, 10, 20$ представлены на рис. 10.



а) $q = 1$ *б)* $q = 2$ *в)* $q = 5$ *г)* $q = 10$ *д)* $q = 20$

Рис. 10. Средненасыщенные изображения, полученные в результате ДКП с различными коэффициентами квантования

Из анализа изображений на рис. 10 следует, что ДКП с различными коэффициентами квантования вносит визуальные искажения в изображение. Наименее заметны искажения для изображения а), полученного в следствии ДКП с коэффициентом квантования $q = 1$, и наоборот наиболее визуально заметны искажения на изображении д), которое получено в результате ДКП с $q = 20$.

Для визуальной оценки устойчивости контуров изображений после ДКП, сравним результаты маскирования на основе оператора Собеля для средненасыщенного изображения после атакующего воздействия. Для этого необходимо применить маскирование на основе оператора Собеля

последовательно к каждому из изображений, представленных на рис. 10. В этом случае, если атакующее воздействие будет искажать значения элементов на границах однородных областей, то сравнительная оценка масок позволит выявить визуальные искажения.

Результаты маскирования средненасыщенного изображения после атакующего воздействия с различными коэффициентами квантования представлены на рис. 11.

а) маска исходного изображения

б) маска изображения после ДКП с квантованием $q = 1$

в) маска изображения после ДКП с квантованием $q = 2$

в) маска изображения после ДКП с квантованием $q = 5$

б) маска изображения после ДКП с квантованием $q = 10$

б) маска изображения после ДКП с квантованием $q = 20$

Рис. 11. Маски, полученные в результате применения оператора Собеля к изображениям после атакующего воздействия

Из анализа изображений на рис. 11 можно сделать следующие выводы:

1) маскирование средненасыщенного изображения выявило больше контуров в сравнении с слабонасыщенным изображением, что обусловлено большим количеством переходов однородных областей;

2) с увеличением коэффициента квантования наблюдается увеличение визуальных искажений, в том числе и для выявленных контуров;

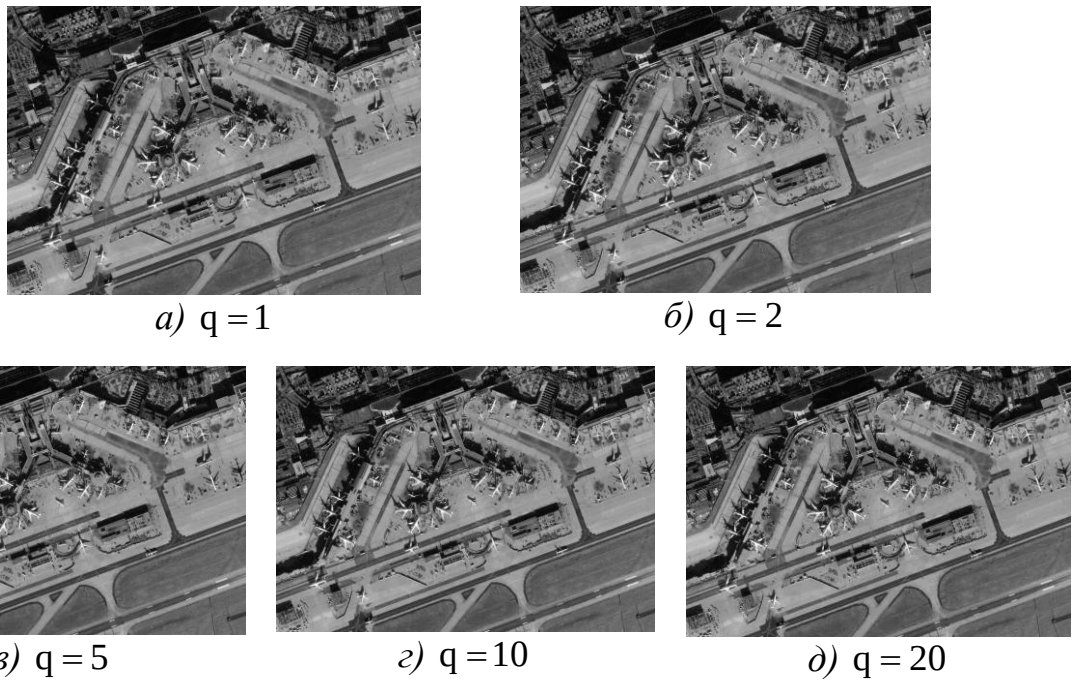
3) не все контуры являются устойчивыми к визуальным искажениям, при этом в первую очередь разрушаются контуры с меньшим уровнем перехода между однородными областями;

4) контуры, вяленые на границах областей с наибольшими перепадами, являются устойчивыми к атакующему воздействию ДКП с квантованием;

5) степень устойчивости контуров зависит от разности значений элементов пространственного представления изображения на границах однородных областей;

6) для выявления устойчивых контуров изображения для стеганографического встраивания возможно использование бинаризации элементов с различным порогом.

Рассмотрим теперь результаты эксперимента для сильнонасыщенного изображения (рис. 5). Сильнонасыщенное изображение характеризуется большим количеством мелких деталей и переходов яркостей на границах однородных областей. В случае использования сильнонасыщенного изображения для стеганографического преобразования возможно встраивать большее количество информации в сравнении с слабонасыщенным и средненасыщенным изображением. Это обусловлено большим количеством контуров объектов изображения.



а) $q = 1$ *б)* $q = 2$ *в)* $q = 5$ *г)* $q = 10$ *д)* $q = 20$

Рис. 12. Сильнонасыщенные изображения, полученные в результате ДКП с различными коэффициентами квантования

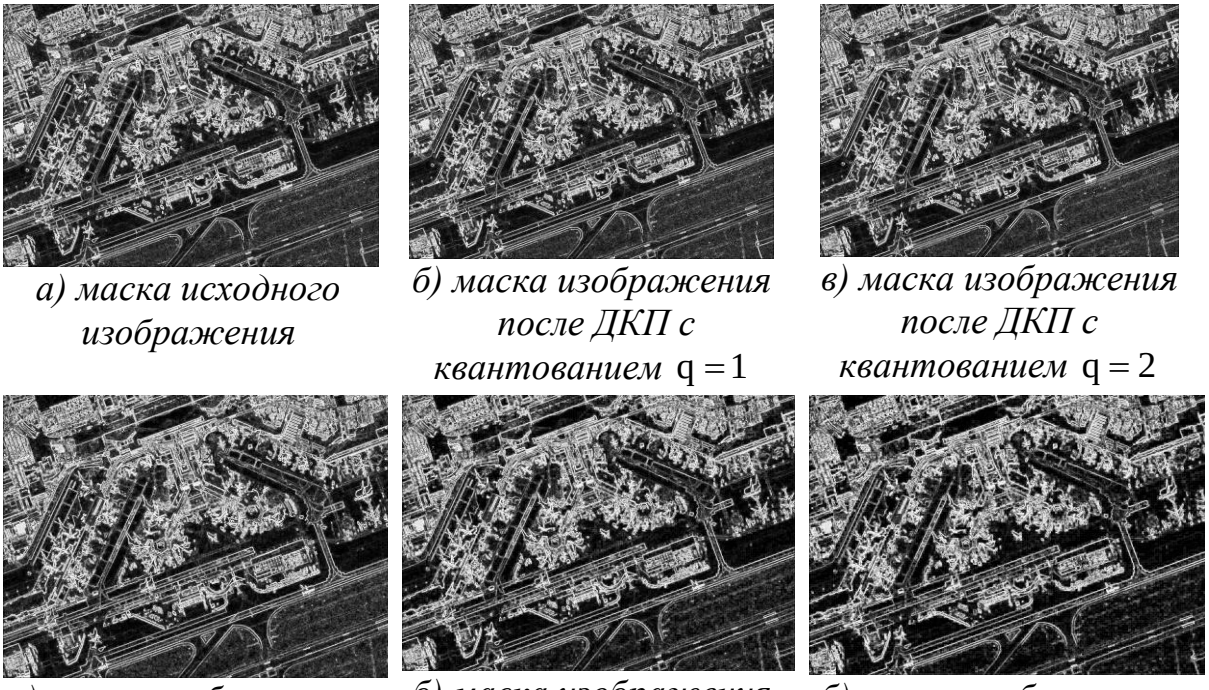
Результаты применения оператора Собеля для сильнонасыщенного изображения после атакующего воздействия представлены на рис. 13.

Анализ масок сильнонасыщенного изображения на рис. 13 позволяет заключить что:

1) для сильнонасыщенного изображение менее заметы визуальные искажения контуров однородных областей и перепадов яркости в сравнении с слабо и средненасыщенными изображениями;

2) маска, полученная в результате применения оператора Собеля к сильнонасыщенному изображению после атакующего воздействия, характеризуется большим количеством контуров объектов.

Таким образом, на основе проведенных экспериментов можно заключить, что для выявления областей изображения, устойчивых к атакующим воздействием, возможно использование маскирования. В этом случае реализуется выявление контуров однородных областей. Устойчивость контуров к атакующим воздействиям зависит от степени различия элементов на границе однородных областей.



а) маска исходного изображения

б) маска изображения после ДКП с квантованием $q = 1$

в) маска изображения после ДКП с квантованием $q = 2$

в) маска изображения после ДКП с квантованием $q = 5$

б) маска изображения после ДКП с квантованием $q = 10$

б) маска изображения после ДКП с квантованием $q = 20$

Рис. 13. Маски, полученные в результате применения оператора Собеля к сильнонасыщенным изображениям после атакующего воздействия

Дальнейшее исследование предлагается проводить в направлении оценки величины пикового отношения сигнал-шум масок изображений. Данная величина характеризует искажения, которые вносятся в контуры изображения в результате атакующего воздействия. При этом для выявления наиболее устойчивых областей возможно использование изменяемого порога бинаризации.

Выводы

1. Предложен подход для устранения выявленных недостатков существующих стеганографических подходов, который заключается в разработке механизма выявления устойчивых областей изображения. Стеганографическое встраивания информации в такие области изображения позволит обеспечить устойчивость встроенных данных к активным стеганографическим атакам.

2. Рассмотрен метод определения устойчивых областей, который заключается в использовании механизма выявления контуров изображения. Для определения контуров однородных областей предлагается использовать оператор Собеля.

3. Проведен эксперимент, который заключается в сравнительной оценке масок различных изображений, которые подвергались атакующим воздействиям с разными коэффициентами ухудшения качества. На основе анализа результатов эксперимента можно заключить, что наибольшей устойчивостью обладают контуры границ областей с наибольшими значениями переходов.

Научная новизна. Впервые предложен механизм выявления устойчивых областей изображения для стеганографического встраивания информации. В отличии от других методов, определение областей осуществляется на основе использования маскирования для выявления границ однородных областей изображения.

Литература

1. Грибунин В.Г., Оков И.Н., Туринцев И.В., Цифровая стеганография. – М.: Солон-Пресс, 2002. – 272 с.

2. Конахович Г.Ф., Пузыренко А.Ю., Компьютерная стеганография. Теория и практика.- К.: « МК-Пресс». – 2006. – 288с.

3. Тарасов Д.О., Мельник А.С., Голобородько М.М. Класифікація та аналіз безкоштовних програмних засобів стеганографії // Інформаційні системи та мережі. Вісник НУ “Львівська політехніка” № 673.– Львів 2010. – С. 365 – 374.

4. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1073 с.

5. Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон. – М: Техносфера, 2004. – 368 с.

МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ЗАЩИТЫ ВИДЕОИНФОРМАЦИОННОГО РЕСУРСА В ИНФОКОММУНИКАЦИОННОЙ СОСТАВЛЯЮЩЕЙ КРИТИЧЕСКОЙ ИНФРАСТРУКТУРЫ

Баранник В.В., Власов А.В.

Введение

Современные процессы внедрения новейших информационных технологий, формирование информационного сообщества усиливают важность обеспечения всех составляющих национальной безопасности, в том числе и информационной безопасности. В настоящее время все основные процессы управления для объектов критической структуры не обходятся без применения инфокоммуникационных систем и систем сбора, обработки и передачи информации.

Комплексное применение данных информационных систем, позволяет существенно повысить динамичность, гибкость и эффективность управления. При этом проблема информационной безопасности и защиты обрабатываемого информационного ресурса не может быть решена без усовершенствования существующих технологий обработки и защиты информационного ресурса, а также внедрения ноу-хау, новых знаний, новых технологий.

При информационном обеспечении процессов управления объектов критической структуры в настоящее время усиленное внимание уделяется внедрению и повышению эффективности применения систем дистанционного сбора и обработки информации (системы видеонаблюдения, комплексы видеоконференционной связи, бортовые комплексы авиационных летательных аппаратов, беспилотные летательные аппараты и др.).

Актуальными продолжают оставаться требования: обеспечения качественного предоставления видеоданных, оперативности доставки видеоинформации, обеспечения необходимого уровня безопасности и защиты видеоинформационного ресурса. В то же время организационно-технические проблемы, связанные с использованием в системах управления на объектах критической структуры низкоскоростных каналов связи, видеооборудования низкого качества, устаревших систем технической защиты информации, обуславливают поиск решения задачи обеспечения безопасности ресурса и его защиты в области разработки новых программно-аппаратных решений при минимизации экономических затрат на разработку и внедрение. Данные решения должны гарантировать простую их практическую адаптацию при внедрении на объектах критической структуры аппаратных средств более высокого технологического уровня.

Поэтому важными являются исследования, связанные с

обеспечением информационной безопасности и защиты видеоинформационного ресурса, формируемого и обрабатываемого в интересах управления на объектах критической структуры.

При этом, как показывают исследования [1, 4, 5, 7] наиболее значимыми угрозами безопасности видеоинформационному ресурсу являются угрозы доступности и целостности, а с точки зрения обеспечения защиты – перехват, вскрытие и распознавание ресурса.

Следовательно, комплексная задача защиты динамического видеоинформационного ресурса в инфоркоммуникационной составляющей систем управления объектов критической структуры с одновременным обеспечением информационной безопасности данного ресурса, является ***актуальной научно-прикладной проблемой*.**

Для решения данной проблемы необходимо провести исследования и разработать технологию обработки видеоинформационного ресурса, которая будет содержать следующие методы:

- автоматического маскирования семантически значимой информации видеокадров (на основе каскадной схемы обработки видеопотока);
- оценки информационной интенсивности видеопотока с учетом структуры видеопотока и классификации семантически значимых элементов видеокадра (фрагментов или макроблоков);
- классификации (определения) степени семантической насыщенности видеокадров или структурной единицы видеопотока на основе их двухиерархической кластеризации (как во временной, так и в спектрально-пространственной областях);
- обработки видеокадров на основе структурной обработки его макроблоков с применением селективных подходов (обеспечения селекции структурных единиц базового кадра) для снижения интенсивности передаваемых видеоданных с использованием каскадных решающих правил в спектральном пространстве (с учетом выявления и закрытия значимых структурных единиц, а также согласовании будующих кодовых конструкций (согласно классификации семантически значимых фрагментов));
- обработки видеоизображений на основе дифференциальной обработки трансформированного представления кадров согласно классификации семантически значимых фрагментов (кодирование с адаптацией параметров в зависимости от значения класса семантической насыщености);
- защиты динамического видеоинформационного ресурса с формированием нескольких каналов обработки и передачи видеопотока согласно классификации семантического содержания элементов макроблоков (к примеру: канала скрытой передачи для передачи закрытой информации на основе методов блочно-симметричного кодирования;

открытого канала передачи кодового представления видеоизображений с встроенным информационным контейнером (формируемые стегановкладки));

- реконструкции видеопотока (видеоизображений) на основе декодирования различных каналов с учетом класса семантической насыщенности элементов видеоизображений (фрагментов, макроблоков), внутрикадровой селекции структурных единиц базового кадра, сформированного стеганоканала и внедренных методов шифрования.

Комплексная разработка и применение данных методов позволит разработать технологию обработки видеоинформационного ресурса и улучшить защиту видеоинформационного ресурса и служебной информации в инфокоммуникационной составляющей критической инфраструктуры с обеспечением необходимого (не ниже заданного) уровня конфиденциальности и доступности ресурса, а также сохранением (контролем) ключевой (семантически значимой) информации.

Основная часть

На рис. 1 представлена структурная схема предлагаемой технологии обработки видеоинформационного ресурса в инфокоммуникационной составляющей критической инфраструктуры.

Рассмотрим кратко идеологию, которая будет положена в разработку данных методов.

Для обеспечения информационной безопасности и улучшения защищенности видеоинформационного ресурса при выполнении требований к оперативности, достоверности и конфиденциальности в системе управления критической структуры необходимо разработать предложенные методы на основе усовершенствования стандартизированных MPEG технологий и ГОСТированных алгоритмов криптографической защиты, а также методов формирования стеганоканала.

Необходимо предварительно оценить информационную интенсивность видеопотока в различных его вариантах передачи и обработки для оценки возможностей по улучшению его обработки и защиты. Здесь обобщающим показателем, учитывающим соответствие требований к методам обработки и аппаратной части, является пропускная способность канала обработки информационного ресурса с точки зрения исходного потока и технологии обработки (кодирование, шифрование) и передачи видеопотока.

С позиции исходного потока под пропускной способностью $H_{зк}$ рассматриваемого канала для инфокоммуникационной составляющей системы критической структуры будем понимать суммарную интенсивность $V_{гк_з}$ закрытого видеопотока, который формируется,

обрабатывается и передается за требуемое время $T_{тр,д}$ при выполнении условий обеспечения его безопасности (по конфиденциальности и доступности $PSNR_{нсд} \le PSNR_{тр,нсд}$, целостности $PSNR_{с} \ge PSNR_{тр,с}$).

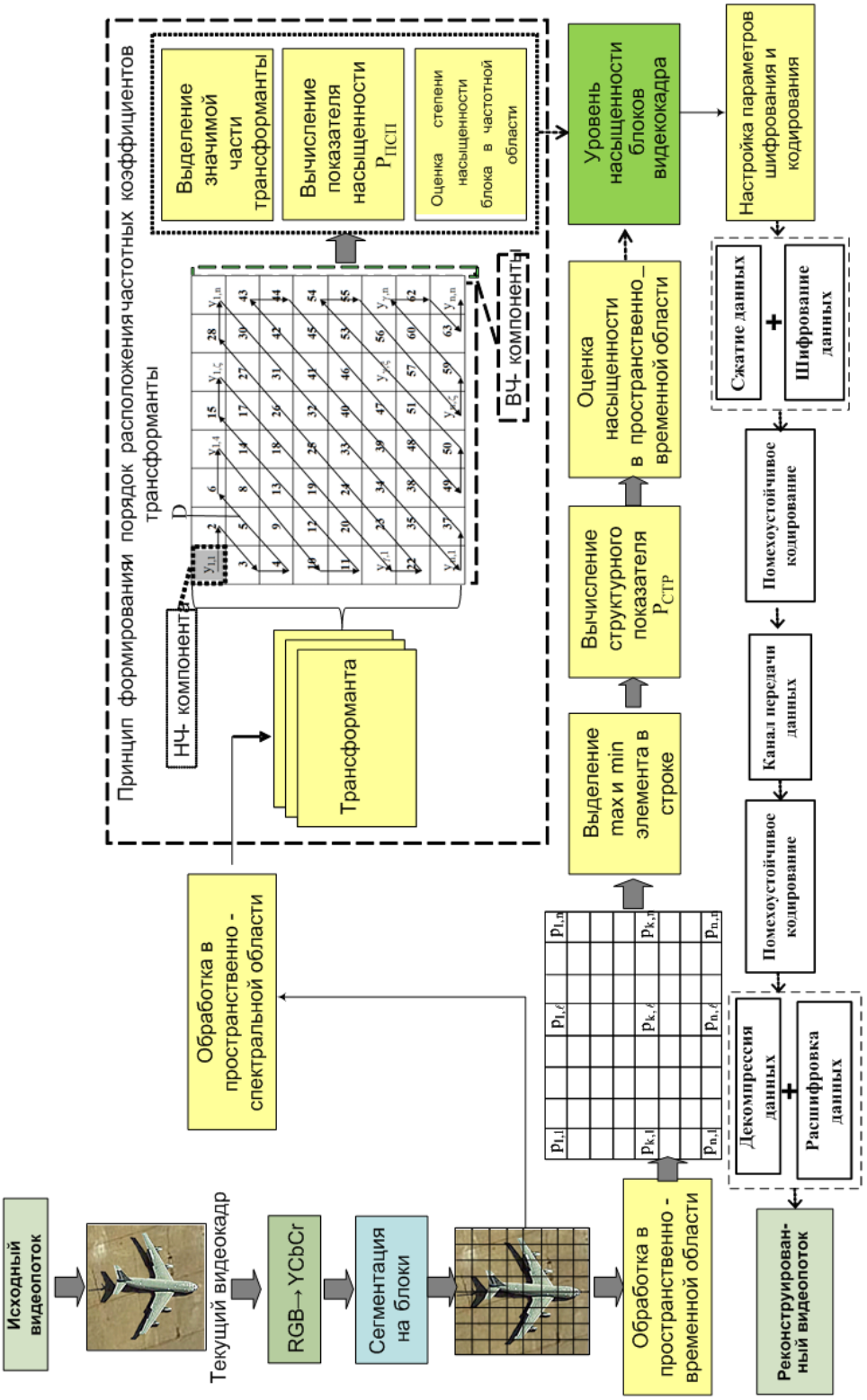


Рис. 1. Структурная схема предлагаемой технологии обработки видеоинформационного ресурса

Другими словами, под пропускной способностью $H_{зк}$ закрытого видеоканала понимается количество N_K исходных кадров $V_{K_{исх}}$, для которых система обеспечивает безопасность и доставку (обработку $T_{обр}$ и передачу $T_{п}$) за требуемое время $T_{тр,д}$ с необходимой достоверностью $PSNR_{тр,с}$.

С позиции технологии обработки (кодирование, шифрование) и передачи под пропускной способностью $H_{зк}$ закрытого видеоканала для инфокоммуникационной составляющей системы критической структуры подразумевается интенсивность $V_{гк_{код}}$ скрытых кодированных видеоданных, соответствующая такому количеству N_K кадров, которые необходимо обработать $T_{тр,обр}$ (с формирование скрытого контейнера) и передать $T_{тр,п}$ за требуемое время $T_{тр,д} = T_{тр,обр} + T_{тр,п}$ при обеспечении требований по безопасности и с учетом пропускной способности $L_{сети}$ сети.

Пропускная способность такого видеоканала зависит от:

1) времени обработки $T_{р,обр}$ (кодирование и шифрование) и передачи $T_{р,п}$ видеопотока. Чем меньше время доставки $T_{р,д}$ кодированного потока, тем больше количество N_K кадров будет обработано и сформировано за требуемое время $T_{тр,д}$;

2) целостности $PSNR_c$. Чем меньше значение пикового отношения сигнал/шум $PSNR_c$ при санкционированном доступе, тем меньше интенсивность $V_{гк_{код}}$ кодированного потока и меньше время $T_{р,д}$ доставки. Но при этом нарушается требование $PSNR_c < PSNR_{тр,с}$ по обеспечению заданного уровня $PSNR_{тр,с}$;

3) степени $PSNR_{нсд}$ закрытия. Чем меньше значение пикового отношения сигнал/шум $PSNR_{нсд}$ при неавторизированном доступе, тем выше степень закрытия $PSNR_{нсд} \to PSNR_{тр,нсд}$. Чем больше степень скрытия $PSNR_{нсд}$, тем больше разрушаются закономерности при кодировании и поэтому меньше избыточности устраняется, соответственно возрастает интенсивность $V_{гк_{код}}$. Это приводит к снижению пропускной способности $H_{зк}$ закрытого видеоканала.

4) степени насыщенности видеокадров. Чем выше степень насыщенности, тем больше количество структурных единиц $S_{зн}^{(\xi,\gamma)}$, которые обрабатываются и закрываются (кодируются по заданной схеме и

алгоритму согласно оценке степени насыщенности), и меньше количество структурных единиц $S_{незн}^{(\xi,\gamma)}$, которые кодируются по стандартному алгоритму. Это приводит к повышению суммарной интенсивности $V_{гк_{код}}$, а следовательно к снижению пропускной способности $H_{зк}$ данного видеоканала.

Отсюда, для выполнения требований к пропускной способности $H_{зк}$ закрытого видеоканала необходимо снизить интенсивность $N_K \cdot V_{K_{исх}}$ видеоданных в контексте снижения времени передачи, обеспечить их безопасность, защиту и доставку за требуемое время $T_{р,д} \le T_{тр,д}$ при необходимых условиях по целостности $PSNR_с \ge PSNR_{тр,с}$ и конфиденциальности $PSNR_{нсд} \le PSNR_{тр,нсд}$.

Таким образом, выполнить оценку пропускной способности $H_{зк}$ закрытого видеоканала в пересчете на исходный поток возможно следующим образом:

$H_{зк} = N_K \cdot V_{K_{исх}} : T_{р,д} \le T_{тр,д}; PSNR_с \ge PSNR_{тр,с}; PSNR_{нсд} \le PSNR_{тр,нсд}$,

где $V_{K_{исх}}$ – интенсивность исходного видеокадра;

N_K – количество видеокадров, которые формируются и передаются при заданных условиях;

$T_{тр,д}$ – требуемое время доставки, приходящееся на группу кадров;

$T_{р,д}$ – реальное время доставки, которое включает в себя время на кодирование, скрытие и передачу видеоданных;

$PSNR_с$ – значение пикового отношения сигнал/шум для доставленного видеокадра при санкционированном доступе;

$PSNR_{нсд}$ – значение пикового отношения сигнал/шум для доставленного видеокадра при несанкционированном доступе;

$PSNR_{тр,с}$ – требуемое значение пикового отношения сигнал/шум для доставленного видеокадра при санкционированном доступе;

$PSNR_{тр,нсд}$ – требуемое значение пикового отношения сигнал/шум для видеокадра при несанкционированном доступе, $L_{сети}$ – пропускная способность сети.

Для устранения максимального количества избыточности при обработке видеопотока предлагается использовать базовый I-кадр, так как в нем содержится максимальное количество информации, а кадры других типов содержат до 70% ссылок на него [1, 2] При этом разрабатываемые методы обработки и скрытия видеоданных (для канала скрытой передачи), будут базироваться на скрытии только I-кадров. Таким образом, обеспечивается полное скрытие всей видеопоследовательности при

минимальной его избыточности. Структура видеопотока представлена на рис. 2 [8].

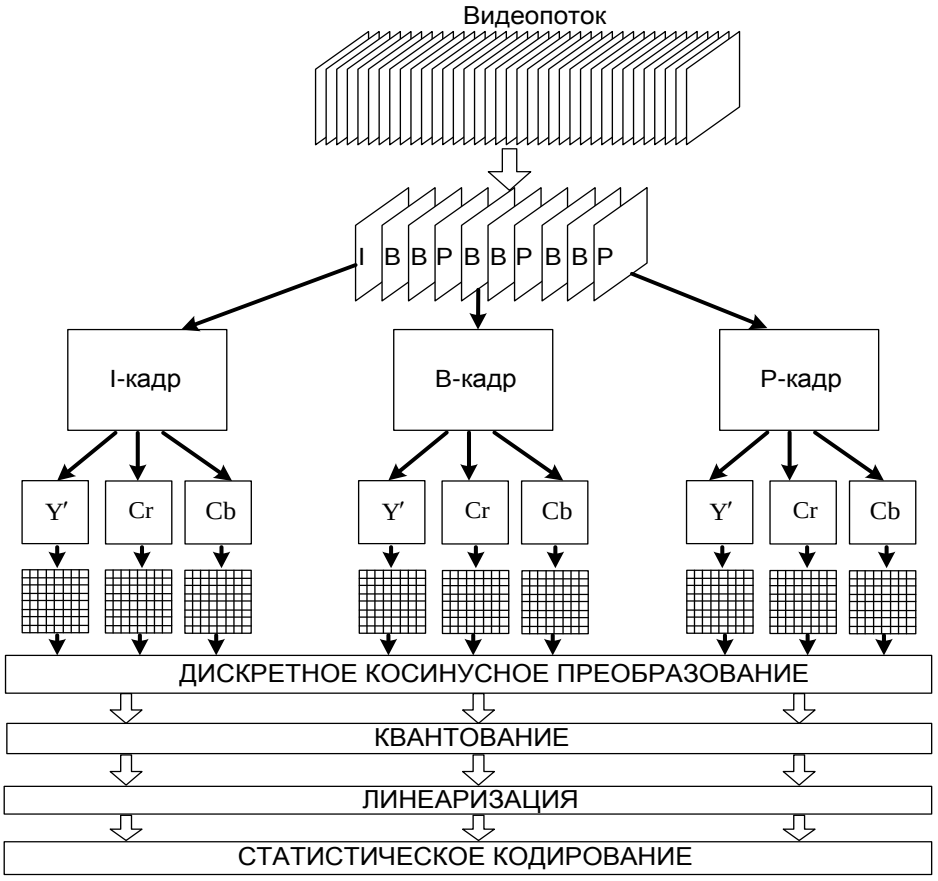


Рис. 2. Структурная схема видеопотока по технологии MPEG

Как видно из структуры (рис. 2), процесс формирования видеопотока основан на последовательном построении цепочки видеокадров разного типа. Под типом кадров видеопотока подразумевается способ кодирования и хранения информации об очередном кадре, отличающемся друг от друга наличием или отсутствием зависимостей этого кадра от предыдущего и последующего. При этом видеопоток состоит из последовательно сформированных групп видеокадров. Каждая группа видеокадров состоит из трех типов видеокадров (в соответствии с принятым MPEG-стандартом):

- I-кадры (intra) называются ключевыми (keyframes) или "базовыми" и содержат только независимо сжатые макроблоки.

- P-кадры (predicted) называются "разностными" и могут содержать как независимо сжатые макроблоки, так и макроблоки со ссылкой на другой I- или P-кадр.

- B-кадры (bi-predicted) – "двунаправленные", "обратные" кадры могут содержать следующие макроблоки: независимые, со ссылкой на один кадр или со ссылкой на 2 кадра. B-кадры ссылаются на ближайшие I-, P- или B- кадры.

Пример представления базового кадра представлен на рис. 3.

Предлагается выполнять анализ и обработку структурной единицы кадра для яркостной компоненты, а процедуры закрытия (шифрования) и кодирования выполнять для всех составляющих (макроблоков) кадра (рис. 3).

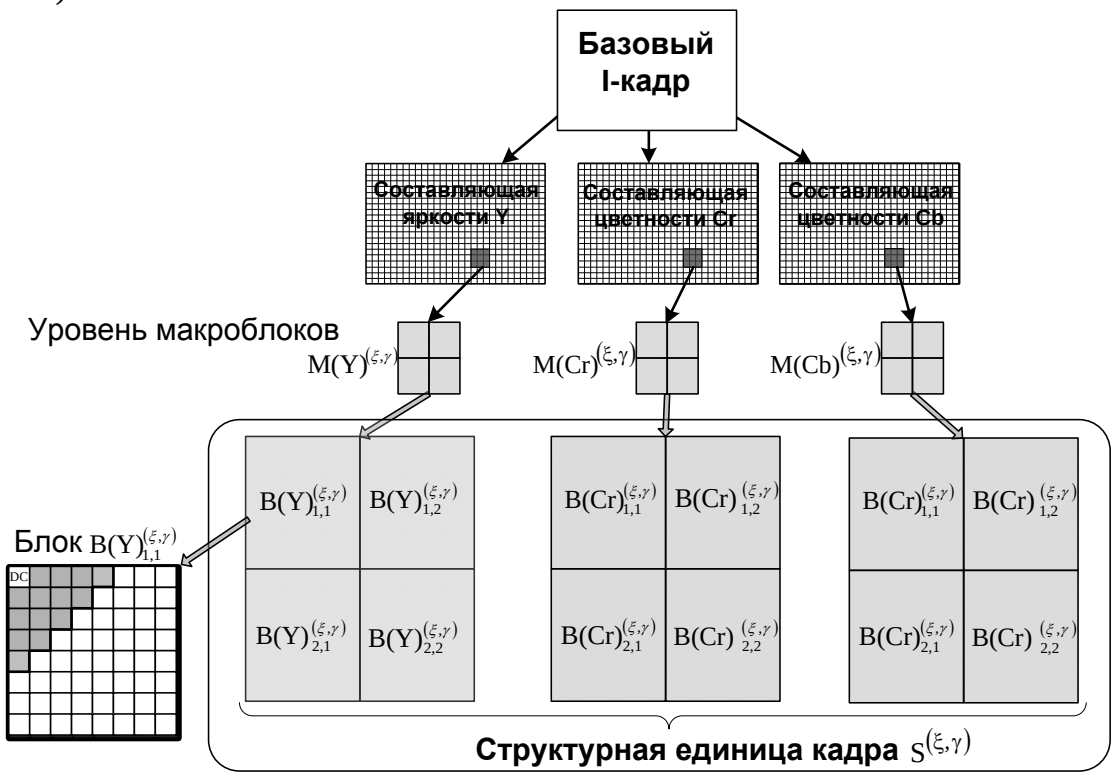


Рис. 3. Схема представления базового кадра

Процесс обработки может быть реализован на разных этапах формирования, обработки и передачи видеоданных, а именно:

1) до кодирования видеопотока (алгоритмы шифрования применяются к вновь созданным (не кодированным) исходным видеоданным; все операции по снижению интенсивности потока и помехоустойчивому кодированию выполняются с уже скрытыми видеоданными);

2) после того, как сформировано компрессионное представление видеоданных (перед тем, как кодированный видеопоток попадает в канал связи);

3) в процессе кодирования (алгоритмы шифрования интегрируются в стандартизированный процесс по обработке исходных видеоданных для снижения их интенсивности (на различных стадиях компрессии)).

На рис. 4 представлена структурно-функциональная схема обработки видеоданных в инфокоммуникационных сетях с возможными вариантами применения алгоритмов шифрования.

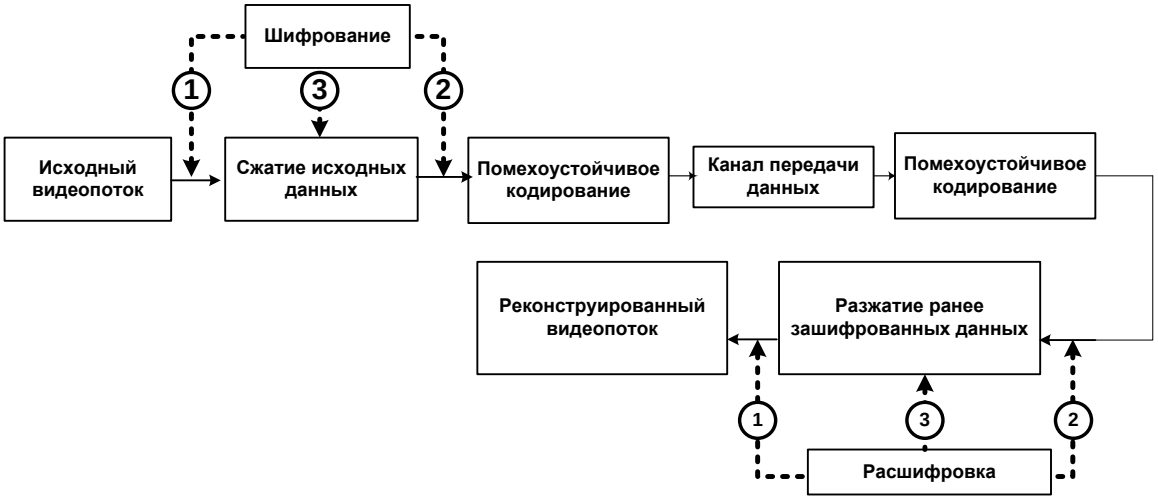


Рис. 4. Структурно-функциональная схема обработки видеоинформационного ресурса (существующие варианты)

При этом время $T^{(1)}_{\text{шид}}$ шифрования исходного видеопотока рассчитывается по формуле:

$$T^{(1)}_{\text{шид}} = \frac{\partial(N; R_k; G_k)^{(r)}_{\text{ш}}}{S_{\text{вк}}},$$

где $S_{\text{вк}}$ – производительность вычислительного комплекса, оцениваемая как количество операций в секунду; $\partial(N; R_k; G_k)^{(r)}_{\text{ш}}$ – количество операций на шифрование, которое зависит от используемого алгоритма r шифрования.

Время на кодирование $T^{(1)}_{\text{сшд}}$ шифрованного видеопотока рассчитывается по формуле:

$$T^{(1)}_{\text{сшд}} = \frac{\partial(V_{\text{шд}})^{(\alpha)}_{\text{сж}}}{S_{\text{вк}}},$$

где $\partial(V_{\text{шд}})^{(\alpha)}_{\text{сж}}$ – количество операций кодирования, которое зависит от используемого алгоритма α кодирования.

Вариант с шифрованием исходных данных до кодирования обладает следующими недостатками:

– не учитывается сокращение избыточности в исходных видеоданных;

– после кодирования происходит увеличение первоначальной интенсивности видеопотока в результате разрушения его структуры за счет предварительного шифрования;

– увеличение интенсивности кодированных шифрованных видеоданных влечет за собой увеличение времени на передачу этих данных в канале связи.

Вариант скрытия видеопотока после его компрессии, позволяет

сократить предварительную избыточность исходного видеопотока и снизить время на обработку (в том числе и на шифрование). Он обеспечивает высокий уровень закрытия информации, но при этом обладают существенными недостатками:

– при ошибках в канале связи происходит размножение ошибок;

– криптографической обработке подлежит весь видеоинформационный поток, из-за чего увеличивается суммарное время обработки формируемых видеоданных на передающей стороне и время обработки видеоданных на принимающей стороне.

Рассмотренные варианты обладают также общими недостатками:

– закрытие видеопотока происходит не в режиме реального времени;

– интенсивность закрытого видеопотока зачастую значительно превышает интенсивность исходного.

Поэтому для устранения данных недостатков предлагается использовать селективный подход (вариант, в котором данные закрываются в процессе их кодирования). Такая реализация представлена на рис.5 и реальна к разработке и внедрению в системы обработки данных, где возможна реализация дополнений (программных) и их интеграция в видеокодек. Для такого варианта кодирование и шифрование выполняются для исходных данных по мере поступления их на обработку.

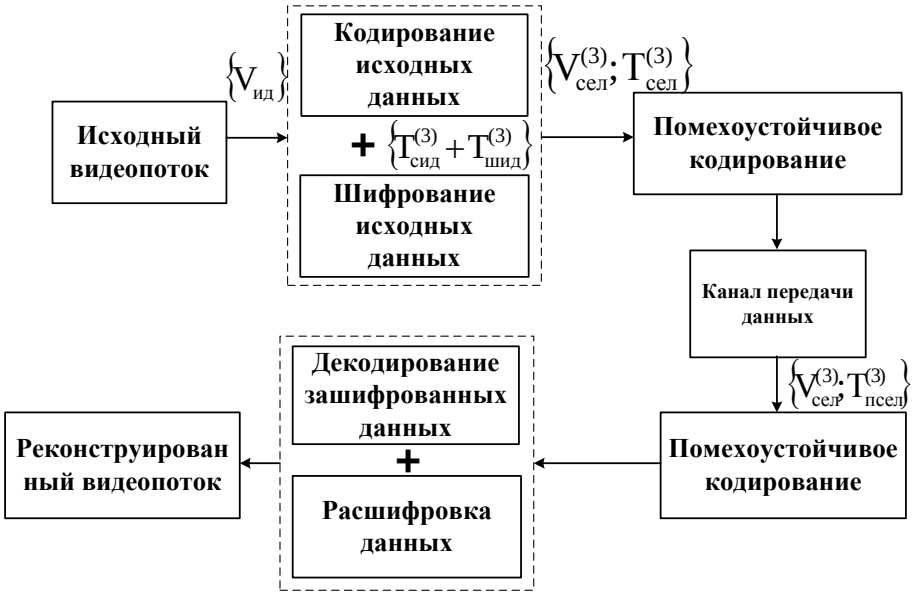


Рис. 5. Предлагаемая структурно-функциональная схема обработки видеоинформационного ресурса в системах критической структуры

В процессе формирования видеоинформационного потока при реализации селективного подхода достигается:

– повышение информативности передаваемых конструкций и сокращение первоначальной интенсивности;

– устраняется избыточность (снижается количество информации), которая может использоваться при криптоанализе;

– уменьшается время шифрования за счет уменьшения длины обрабатываемых сообщений.

Под механизмом селекции подразумевается закрытие не всего видеокадра, а только значимых $S_{зн}$ его составляющих [12].

Под значимой $S_{зн}$ составляющей понимается такая составляющая видеокадра K_I, которая несет в себе наибольшую семантическую и структурную информативность. В процессе автоматической селекции значимых $S_{зн}$ составляющих предлагается учитывать структурные особенности формирования видеопотока.

Для селекции значимых структурных единиц $S_{зн}$ предлагается выявлять наиболее информативные, в плане структурного и семантического содержания, составляющие базового кадра [13]. Поскольку наиболее полную информацию несет яркостная составляющая видеокадра K_I, то значимые структурные единицы предлагается выявлять на базе яркостных компонент. Поэтому принятие решения по закрытию структурной единицы предлагается осуществлять по результатам анализа информационной составляющей совокупности блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей [1, 6, 14].

Для определения энергетической насыщенности блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ предлагается ввести понятии блоков трех типов:

– слабонасыщенные блоки (блоки, в которых присутствуют равномерные участки изображения);

– средней насыщенности (блоки, в которых имеются незначительные отличия между пикселями, соответственно присутствуют плавные переходы контрастности);

– сильнонасыщенные блоки (блоки, в которых присутствуют резкие переходы яркости и контрастности изображения) [6].

Предлагается оценивать структурную и семантическую информативность структурной единицы (макроблоков) с позиции спектральных характеристик [1, 2] или на основе метрик [3, 6, 7, 15].

Следовательно, необходимо разработать систему показателей для выявления наиболее значимых блоков яркостной составляющей видеокадра по степени семантической и структурной насыщенности на основе оценки показателей: в пространственно-временной области и в пространственно-спектральной области [6, 7, 15].

Для повышения эффективности классификации фрагментов видеокадров (макроблоков) предлагается использовать двухбазисный принцип, который охватывает пространственно-временное и пространственно-спектральное представление видеокадра. Тогда для оценки уровня насыщенности в блоках видеокадров предлагается

использовать следующие показатели:

1) в пространственно-временной области - структурный показатель блока $P_{СТР}$, который определяет наиболее значимые элементы сроки, представлено выражением:

$$P_{СТР} = [\log_2(\prod_{i=1}^{n}(p_{k,max} - p_{k,min}))],$$

где $p_{k,max}$ - максимальное значение яркости элемента строки макроблока;

$p_{k,min}$ - минимальное значение яркости элемента строки макроблока;

i – порядковый номер строки.

2) в пространственно- спектральной области - показатель насыщенности $P_{ПСП}$ блока, который определяет наиболее значимые коэффициенты трансформанты в каждой означенной зоне, представлено выражением:

$$P_{ПСП} = [\log_2(\prod_{\gamma=1}^{D_d}\prod_{\xi=1}^{N_y}(y_{\gamma,\xi}))],$$

где $y_{\gamma,\xi}$ - частотный коэффициент трансформанты на позиции с координатами (γ, ξ) относительно диагонали D;

D_d - количество диагоналей в оцениваемой зоне частотных коэффициентов;

N_y - количество частотных коэффициентов, которые отвечают диагоналям в оцениваемой зоне.

Определение энергетической насыщенности блоков предлагается осуществлять после пространственно-спектрального представления (ПСП), к примеру, дискретного косинусного преобразования (ДКП). С помощью ПСП осуществляется переход от пространственно-временного представления видеокадра в пространственно-спектральное. Компоненты трансформанты ПСП (ДКП) являются интегральными характеристиками структурного содержания фрагмента изображения. Причем интегральные свойства компонент зависят от их положения в трансформанте.

На рис. 6 представлено расположение низкочастотных компонент трансформанты ДКП в блоках $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей макроблока. [1, 9] Из рис. 6 видно, что низкочастотные компоненты находятся в области первых пяти диагоналей.

Интегральная зависимость компонент трансформанты ПСП выглядит следующим образом:

1. Значение компоненты в верхнем левом углу трансформанты ПСП пропорциональны средней яркости изображения (фрагментов). Они характеризуют степень насыщенности блока изображения

низкочастотными перепадами (ступенчатые изменения уровня яркости или координаты цвета).

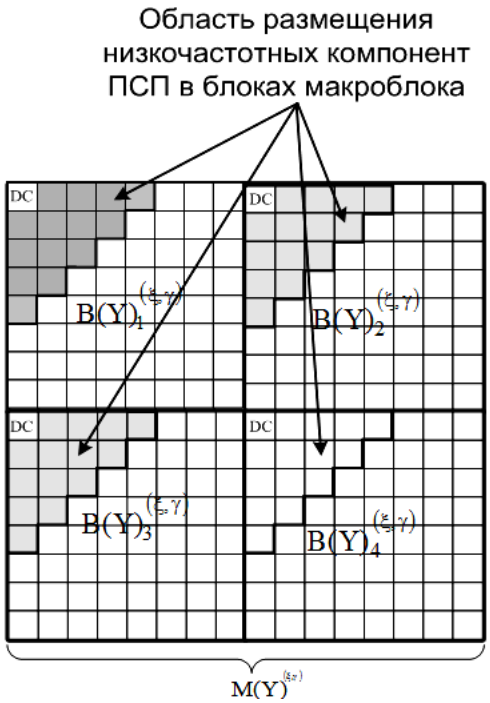


Рис. 6. Схема расположения низкочастотных трансформанты ПСП в блоках $B(Y)_{\varphi}^{(\xi,\gamma)}$ *яркостной составляющей макроблока*

2. Компоненты в средней части трансформанты определяют степень насыщенности блока изображения линейными, равномерными изменениями уровня яркости.

3. Значения компонент в нижней правой области трансформанты ПСП характеризуют степень насыщенности высокочастотными перепадами блока изображения. К высокочастотным перепадам относят импульсные изменения значений элементов изображений.

Значения компонент изменяются по мере преобладания в изображении различных структурных особенностей. Широкий класс изображений содержит в основном линейные, монотонные и ступенчатые структурные изменения уровня яркости. Импульсные изменения занимают меньшую площадь изображения. [1, 3, 6] Кроме того, они могут быть вызваны шумами дискретизации. Поэтому наибольшие значения имеют компоненты, расположенные в верхней левой части трансформанты. Компоненты в нижней части трансформанты соответствуют высокочастотным изменениям и поэтому имеют меньшие значения.

Селективные методы шифрования имеют простую реализацию, не требуют значительных вычислительных ресурсов, при этом повышают помехоустойчивость всего видеопотока. При несанкционированном перехвате такого видеопотока с ошибками, в процессе расшифровке количество этих ошибок будет только увеличиваться. [10]

Анализ различных вариантов селективного шифрования показал, что наиболее эффективным является шифрование после этапа дискретного косинусного преобразования (ДКП).

К рассмотрению предлагается разработка селективного метода скрытия видеоданных на основе шифрования базового видеокадра. Структурная схема кодирования видеопотока для селективного подхода с закрытием базового I-кадра представлена на рис. 7.

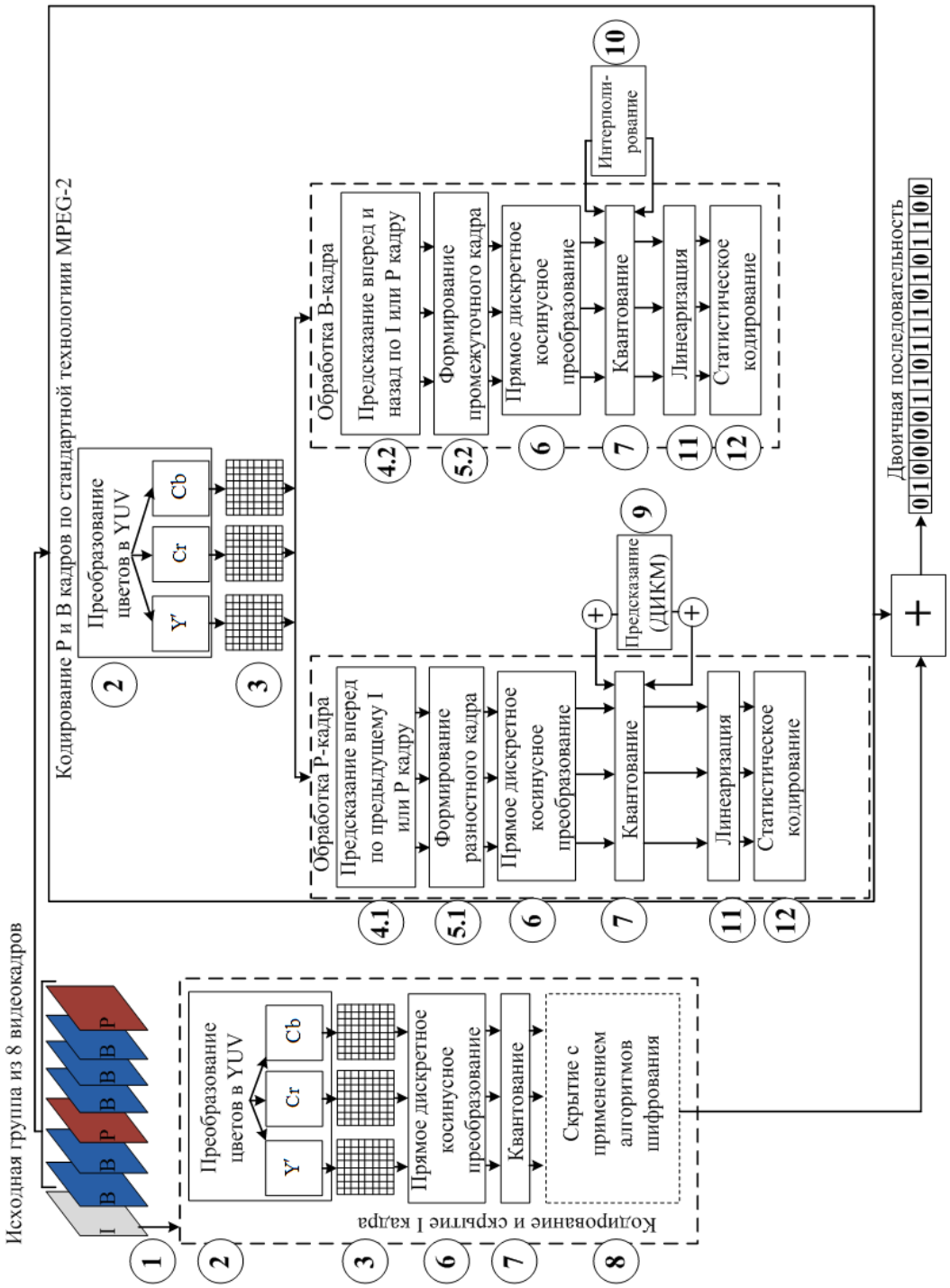


Рис. 7. Схема кодирования видеопотока в селективном подходе со скрытием I-кадра

Таким образом, выделены следующие этапы кодирования исходного видеопотока в селективном подходе:

1. Покадровое распределение исходного видеопотака (выделение кадров I, P и B типов из группы кадров для дальнейшей обработки).

2. Преобразование исходного видеокадра в цветовое пространство YUV (цветовая модель RGB преобразуется в YC_bC_r).

3. Субдискретизация компонентов яркости и цветности.

Составляющие цветности (C_b и C_r) содержат высокочастотную цветовую информацию, к которой глаз человека менее чувствителен. Поэтому определенная ее часть может быть отброшена и, тем самым, можно уменьшить количество учитываемых пикселей для каналов цветности, т.е обрабатывать варианты 4:4:4, 4:2:2 или 4:1:1 (рис.8);

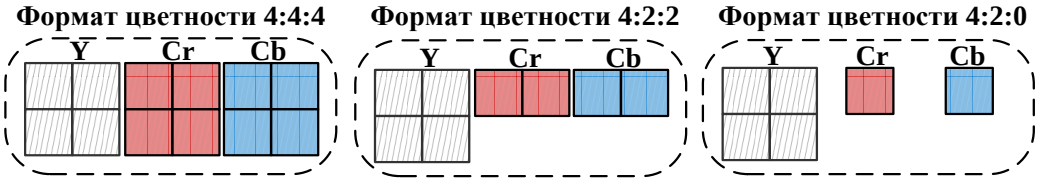


Рис. 8. Схема субдискретизации компонентов яркости и цветности

4.1. Предсказание вперед по предыдущему I или P кадру. P-кадры сжимаются с использованием предшествующих I- или P-кадров с помощью предсказывающего кодирования и компенсации движения (так называемое предсказание вперед, устраняющее временную избыточность), что обеспечивает увеличение степени сжатия.

4.2. Предсказание вперед или назад по I или P кадру. B-кадры сжимаются с использованием двунаправленного предсказания, т.е. с привлечением предшествующих и последующих I- и P-кадров.

5.1. Формирование разностного кадра с применением алгоритма дифференциальной импульсно-кодовой модуляции.

5.2. Формирование промежуточного кадра с использованием метода интерполирования.

6.1. Кодирование P-кадров с использованием алгоритмов компенсации движения и предсказания вперед по предшествующим I или P кадрам. Для такого кодирования применяется дифференциальная импульсно-кодовая модуляция (ДИКМ) – метод кодирования, который основывается на предположении наличия корреляционной связи между соседними отсчетами изображения. [5]

6.2. При кодирования B-кадров применяется компенсация движения и предсказание вперед по ближайшим предшествующим опорным I или P кадрам. При интерполяционном (двунаправленном) предсказании оценка выполняется по известным значениям предшествующих и последующих отсчетов с применением алгоритмов интерполяции.

7. Применение дискретных косинусных преобразований для

уменьшения избыточности изображения.

8. Для скрытия I-кадра к нему применяется гарантированное шифрование.

9. Линеаризация матриц квантовая.

$$[M * N] \Rightarrow (M_0, N_0), (M_0, N_1), (M_1, N_0), (M_2, N_0) \ldots (M_7, N_7),$$

где $[M * N]$ – размер матрицы квантования.

10. Статистическое кодирование результирующих коэффициентов с применением алгоритмов группового кодирования и алгоритма Хаффмана для удаления избыточности информации.

На основании количественной оценки показателей с использованием методов кластеризации (метод К-средних) выполняется классификация (выявление принадлежности) структурных элементов (макроблоков) на 3 класса: слабонасыщенные блоки, блоки средней насыщенности и сильнонасыщенные блоки [1, 2, 3, 6, 7, 15].

Построение кластеризации для оценки степени семантической и структурной насыщенности предлагается проводить из четырех этапов, представленных на рис. 9.

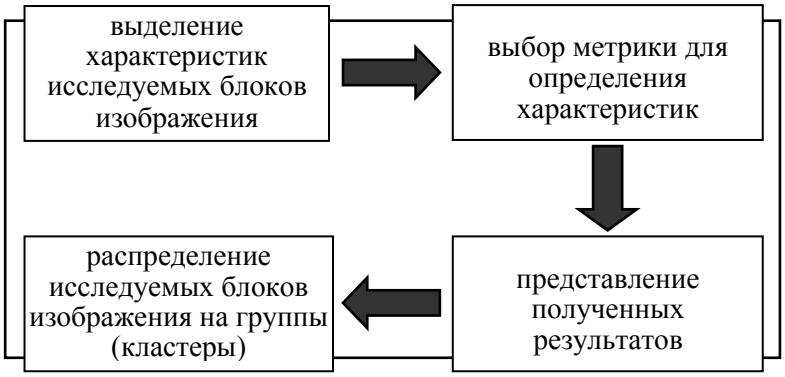


Рис. 9. Этапы построения кластеризации

Первым этапом необходимо осуществить определение свойств, наиболее точно характеризующих объекты: количественные значения (размеры, координаты, интервалы, измерение яркостных показателей и т.д.), а также характеристики, измеряемые качественными значениями (как правило, значения, выбираемые из списка: вид изображения, номер области изображения и т.д.).

На втором этапе, при построении кластеризации, необходимо выбрать метрику в зависимости от пространства, где располагаются объекты исследуемых областей, и от неявных (скрытых, неучтенных) характеристик кластеров.

Третий и четвертый этапы представляются в виде результатов разбиения, которые должно быть организованы в наглядном виде, при котором удобно осуществлять оценку результатов. Чаще всего кластеры представляют центроидами, набором характерных точек или их ограничениями.

Процесс кластеризации блоков видеокадров заканчивается тогда, когда распределены по кластерам все блоки или состав блоков в кластерах не меняется.

Пример кластеризации видеокадра приведен на рис. 10.

В табл. 1.1 представлены пример результата кластеризации блоков конкретного видеокадра (рис. 10) только по 2-м показателям: показателю насыщенности исходных блоков видеокадра и структурному показателю исходных блоков видеокадра.

Из табл. 1 видно, что в ходе применения метода кластеризации (метода *K*-средних) было получено три кластера (столбец 4 "Кластер"), по которым распределены блоки видеоизображения.

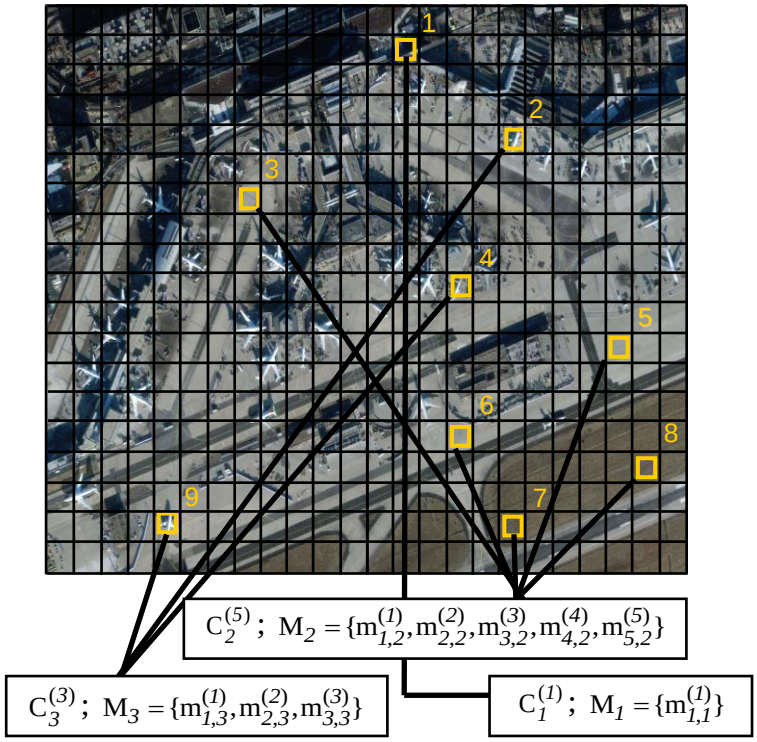


Рис. 10. Сформированные кластеры на первой итерации процесса кластеризации блоков видеокадра

Таблица 1

Пример кластеризации блоков видеокадра (к рис. 10)

Тип блока изображения	Обобщенные результаты				
	1 1-8	2 1-16	3 1-32	4 КЛАСТЕР	5 РАССТОЯНИЕ
Текстура_1	**279**	**188**	**116**	**2**	**0,00**
Контур_1	**115**	**189**	**277**	**3**	**12,12**
Однородность_1	**59**	**97**	**117**	**1**	**20,37**
Контур_2	**108**	**176**	**217**	**3**	**23,59**
Однородность_2	**67**	**113**	**143**	**1**	**38,50**
Однородность_3	**57**	**84**	**101**	**1**	**8,59**
Однородность_4	**32**	**37**	**37**	**1**	**39,50**
Однородность_5	**44**	**45**	**53**	**1**	**28,03**
Контур_3	**117**	**192**	**275**	**3**	**11,58**

Таким образом, на основе применения методов кластерного анализа будет усовершенствован метод классификация семантической насыщенности блоков изображения, который позволит выполнить классификацию блоков для формирования параметров кодирования и шифрования.

Для разработки метода кодирования значимой структурной единицы (макроблока) необходимо разработать основные этапы формирования двоичного кода зашифрованной значимой структурной единицы, которые базируются на трех технологических составляющих.

- первая составляющая заключается в формировании двоичного кода значения компоненты трансформанты для блока изображения;
- вторая составляющая заключается в формировании кодовой конструкции структурной единицы базового видеокадра, подлежащей шифрованию;
- третья составляющая заключается в формировании матриц двоичного кода значимой структурной единицы такого же размера, что и ключ шифрования.

Структура кодовой конструкции представления скрытой группы видеокадров представлена на рис. 11.

Предлагается разработать метод декодирования закрытого видеопотока на основе селекции (выявлении закрытых значимых структурных единиц) базового видеокадра. Для этого предлагается декодировать закрытый базовый видеокадр K_I с учетом определения значимых $S_{зн}^{(\xi,\gamma)}$ структурных единиц.

Структура метода декодирования закрытого видеопотока представлена на рис. 12.

Метод декодирования закрытого видеопотока должен включать в себя следующие этапы:

1. Выделение кодовой конструкции группы кадров из двоичной последовательности потока видеоданных.
2. Определение типа видеокадров в группе кадров.
3. Выделение цифрового представления закрытого базового видеокадра K_I из цифрового представления группы кадров.
4. Определение закрытых $S_{зн}^{(\xi,\gamma)}$ и не закрытых $S_{незн}^{(\xi,\gamma)}$ структурных единиц. Это происходит в результате анализа метки М, значение которой хранится в дополнительных данных цифровом описании структурной единицы.
5. Дешифровка закрытых значимых $S_{зн}^{(\xi,\gamma)}$ структурных единиц, при которой расшифровываются значения компонент трансформант ДКП блоков $B(Y)_{m,n}^{(\xi,\gamma)}$, $B(Cr)_{m,n}^{(\xi,\gamma)}$ и $B(Cb)_{m,n}^{(\xi,\gamma)}$.

Рис. 11. Структура кодовой конструкции представления скрытой группы видеокадров

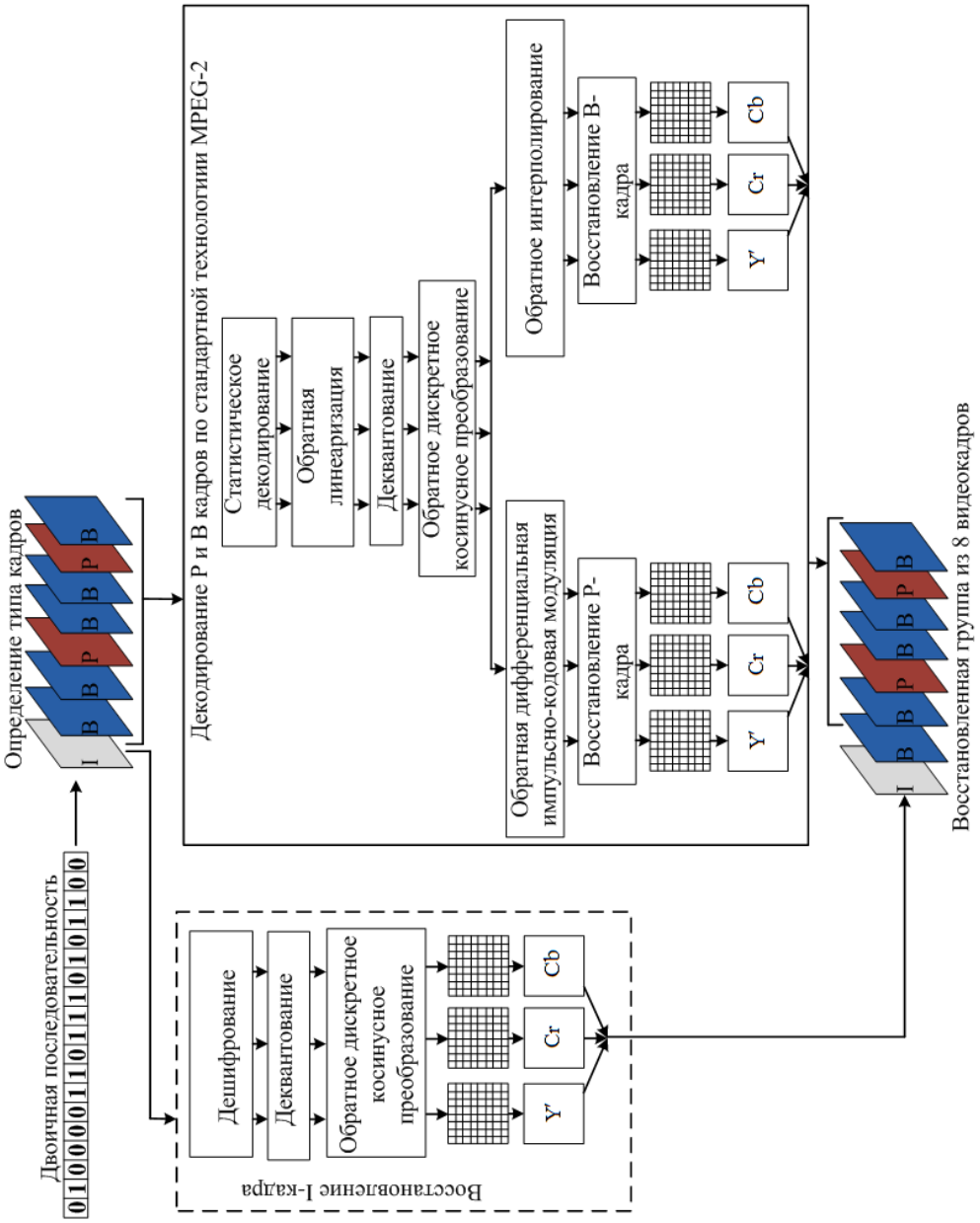


Рис. 12. Структура метода декодирования закрытого видеопотока

6. Декодирование незначимых $S_{незн}^{(\xi,\gamma)}$ структурных единиц, которое включает в себя такие этапы:

6.1. Обратная линеаризация трансформант ДКП блоков составляющей яркости и цветности $B(Y)_{m,n}^{(\xi,\gamma)}$, $B(Cr)_{m,n}^{(\xi,\gamma)}$ и $B(Cb)_{m,n}^{(\xi,\gamma)}$ незначимых структурных единиц.

6.2. Деквантование трансформант ДКП незначимых блоков.

7. Обратное ДКП значимых и незначимых блоков составляющей яркости и цветности $B(Y)_{m,n}^{(\xi,\gamma)}$, $B(Cr)_{m,n}^{(\xi,\gamma)}$ и $B(Cb)_{m,n}^{(\xi,\gamma)}$.

8. Построение композиции структурных единиц $S^{(\xi,\gamma)}$ базового видеокадра K_I, которое включает в себя следующие этапы:

8.1. Декодирование служебной информации для формирования структурных единиц $S^{(\xi,\gamma)}$.

8.2. Формирование композиций макроблоков $M(Y)^{(\xi,\gamma)}$, $M(C_r)^{(\xi,\gamma)}$ и $M(C_b)^{(\xi,\gamma)}$.

8.3. Формирование видеоизображения из блоков $B(Y)_{m,n}^{(\xi,\gamma)}$, $B(Cr)_{m,n}^{(\xi,\gamma)}$ и $B(Cb)_{m,n}^{(\xi,\gamma)}$.

9. Преобразование цифровых плоскостей видеоизображения I-кадра из формата YUV в формат RGB (формирование одного видеоизображения из 3-х цифровых плоскостей YCrCb).

10. Обратная дифференциальная импульсно-кодовая модуляция для восстановления P-кадров.

11. Обратное интерполирование для восстановления B-кадров.

12. Преобразование цифровых плоскостей видеоизображений P и B-кадров из формата YUV в формат RGB.

13. Формирование группы видеокадров из восстановленных I, P и B-кадров.

14. Формирование восстановленной видеопоследовательности из групп видеокадров.

После выделения базового кадра из битового потока группы видеокадров происходит определение значимых $S_{зн}^{(\xi,\gamma)}$ и незначимых $S_{незн}^{(\xi,\gamma)}$ структурных единиц (с учетом класса семантической насыщенности).

Для обеспечения безопасности и улучшения защиты динамического видеоинформационного ресурса с формированием нескольких каналов обработки и передачи видеопотока предлагается также применение открытого канала передачи кодового представления видеоизображений с встроенным информационным контейнером (стегановкладками).

Выводы

Предлагается усовершенствовать метод обработки (компрессии) видеоизображений на основе дифференциальной обработки трансформированного представления кадров согласно классификации семантически значимых фрагментов и кодирование с адаптацией параметров компрессии в зависимости от значения класса насыщенности и наличия встроенного контейнера [11].

Метод компрессии будет включать в себя три базовых механизма:

- дифференциальное представление видеоизображений (фрагментов, макроблоков) по комбинированной схеме;
- обеспечения дополнительного уменьшения объемов сжатых видеокадров и сохранение заданного уровня целостности за счет учета класса семантической насыщенности изображения и адаптации параметров компрессии;

- реализации технологии одномерного кодирования по блочной схеме.

Предлагается для адаптации степени компрессии и взаимооднозначного восстановления в зависимости от класса семантической насыщенности фрагментов изображений K_i применять стратегию дифференциального определения (выбора) параметров сжатия трансформант преобразования. Это обеспечит повышение эффективности компрессии за счет регуляризации яркостной составляющей изображений с обеспечением целостности информации, а именно сохранением исходной семантической составляющей видеоизображения.

Стратегия выбора параметров компрессии будет учитывать метод формирования параметров метода, матрицы квантования и механизма их адаптации в зависимости от класса семантической насыщенности видеокадра (фрагментов). Для каждого класса семантической насыщенности будет выполнен либо расчет параметров компрессии (коэффициентов матриц квантизации) либо будет осуществляться выбор параметров, определенных ранее эмпирическим путем на основе экспериментальных данных. Процесс компрессии представляет собой преобразование трансформанты путем поэлементной обработки в соответствие с выбранной стратегией.

В общем виде стратегия дифференциального выбора параметров компрессии будет иметь вид:

$$Q = F_q\{ X; M; F_P(X; M); F_K(X; M; P_i) \} = \{Q_1; Q_2; Q_3 \},$$

где $F_q\{ X; M; F_P(X; M); F_K(X; M; P_i) \}$ - функционал задающий построение матрицы параметров компрессии и квантования;

Q_1, Q_2, Q_3 – соответствующие K_i классу насыщенности матрицы параметров метода сжатия, $i=1,2,3$, определенные на основе решающего правила формирования.

Структурная схема формирования кодового значения в процессе компрессии видеоизображений с учетом класса семантической насыщенности приведена на рис. 13.

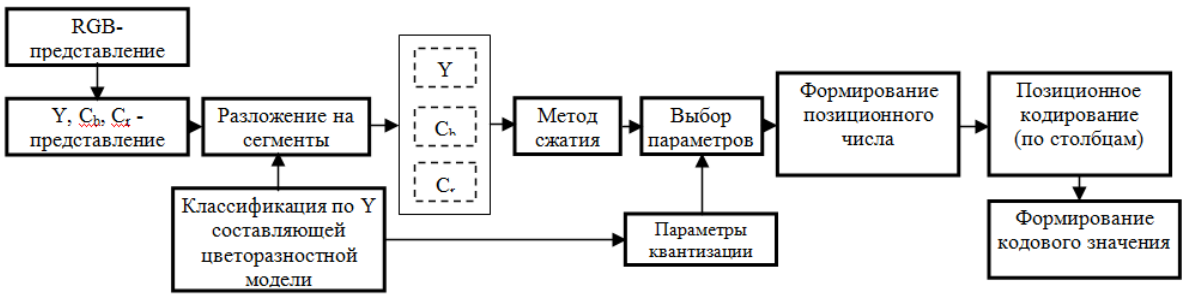


Рис. 13. Структурная схема формирования кодового значения в процессе компрессии видеоизображений

Таким образом, сформулированы методологические основы защиты

видеоинформационного ресурса в инфокоммуникационной составляющей критической инфраструктуры направленные на разработку технологию обработки видеоинформационного ресурса с целью обеспечения безопасности видеоинформационного ресурса, улучшения защиты самого ресурса и служебной информации видеопотока с разработкой комплекса методов.

Литература

1. Баранник В.В. Кодирование трансформированных изображений в инфокоммуникационных системах / В.В. Баранник, В.П. Поляков - Х.: ХУПС, 2010. – 234 с.
2. Баранник В.В. Технология двухкомпонентного кодирования видовых изображений для средств телекоммуникаций / В.В.Баранник, А.В.Власов, А.Н.Додух // Сучасна спеціальна техніка, вип. 4 (31). – 2012. – с. 70 – 79.
3. Баранник В.В. Методология двухкаскадного маскирования изображений в системах инфотелекоммуникаций / А.В. Власов, В.В. Баранник, А.В.Ширяев // АСУ и приборы автоматики. – 2013. – Вип. 162. – С. 50 – 55.
4. Богуш В.М. Інформаційна безпека держави /В.М. Богуш, О.К. Юдин. – К.: МК–Прес, 2005. – 432 с.
5. Быстрые алгоритмы в цифровой обработке изображений / [Т.С. Хуанг, Дж.О. Эклунд, Г.Дж. Нуссбаумер и др.]; под ред. Т.С. Хуанга; пер. с англ. – М. : Радио и связь, 1984. – 224 с.
6. Власов А.В. Метод кодирования видеоизображений с маскированием для повышения безопасности видеоинформационных ресурсов. / А.В. Власов, А.В. Ширяев // Радиоэлектронные и компьютерные системы. – 2013. – № 3. – С. 65 – 73.
7. Власов А.В. Количественная оценка качества маскирования изображений. / В.В. Баранник, А.В. Власов // IV Міжнародна науково – практична конференція [″Обробка сигналів і негауссівських процесів″], Черкаси, 22 – 24 травня 2013р. – С. 37 – 39.
8. Гуржий П.Н. Декодирование сжатых видеоданных в инфокоммуникационных системах объективного контроля // Сучасна спеціальна техніка. – 2014. – 1. – С. 22-30.
9. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2005. – 1073 с.
10. Крук Б.И. Телекоммуникационные системы и сети. Том 1, 2, 3 / Б.И. Крук, В.Н. Попантонопуло, В.П. Шувалов. – М.: Горячая линия-Телеком, 2003. – 647 с.
11. Патент на корисну модель № 91198, Україна, МПК G06F 15/00, G06F 17/00 Пристрій для динамічного кодування та захисту інформаційного ресурсу в інфокомунікаційних системах / Третяк В.Ф., Бараннік В.В., Власов А.В., Бойко Ю.П. та ін. - № u201400644; заяв. 23.01.2014; опубл. 25.06.2014; Бюл. № 12. - 6 с.
12. Ding Z. GPU accelerated interactive space-time video matting / Z. Ding, H. Chen, Y. Gua, Q. Peng // In Computer Graphics International. – 2010. – P. 163-168.
13. Gopinath R.A. On cosine-modulated wavelet orthogonal bases / R.A. Gopinath, C.S. Burrus // IEEE Trans. Image Proc. – 1995. – V. 4. – № 2. – P. 162-177.
14. Milyaev S. Image binarization for end-to-end text understanding in natural images / S. Milyaev, O. Barinova, T. Novikova, V. Lempitsky, P. Kohli // ICDAR. – 2013. – P. 35-42.
15. Vlasov A.V. Estimation of quality methods disguise images for detection edge contours // Science-Based Technologies. – 2013. – № 2 (18). – pp. 193 – 197.

МЕТОД КРИПТОКОМПРЕССИОННОГО ПРЕДСТАВЛЕНИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ АДАПТИВНОГО ОБОБЩЕННОГО ПОЗИЦИОННОГО КОДИРОВАНИЯ ДЛЯ БИНОМИНАЛЬНОГО ПРОСТРАНСТВА

Баранник В.В., Сидченко С.А.

Введение

Современное развитие государства в различных сферах деятельности (политическая, социальная, экономическая, управленческая) неразрывно связано с обеспечением должного уровня информатизации. В последнее время наибольшее применение получили системы сбора информации с использованием дистанционных инфокоммуникационных технологий (системы видеонаблюдения и мониторинга); корпоративные и локальные информационно-коммуникационные системы и сети кризисной инфраструктуры государства; телевизионные системы и системы видеоконференцсвязи; системы связи и передачи информации, в особенности мобильной радиосвязи и космические телекоммуникационные системы.

Видеоинформация приобретает значение важного ресурса, влияющего на национальную безопасность, безопасность коммерческих структур и соблюдения прав личности, и определяет уровень экономического развития государства, его оборонный потенциал, формирование общественного мнения. По данным компании Cisco доминирующими приложениями останутся видеосервисы и передача мультимедийных файлов. Так только по прогнозам в мировом Интернет-трафике к 2020 г. доля видео составит 79 % (в 2015 г. этот показатель составлял 63 %). В мире ежемесячно будет транслироваться три триллиона минут видео, что эквивалентно пяти миллионам лет видео в месяц.

В тоже время для внедрения видеоинформационного ресурса (ВИР) в жизнедеятельность общества и государства присутствуют следующие проблемные ограничения: *Первое проблемное ограничение (на законодательном уровне).* Любое видеонаблюдение приводит к получению персональных данных [1], как непосредственно, так и опосредованно. А любые персональные данные могут быть отнесены к конфиденциальной информации на основании законов или человеком, о котором эти данные собираются. При этом законодательством Украины не установлен четкий перечень сведений о физическом лице, которые являются персональными данными [1]. Современное состояние правового регулирования видеонаблюдения в Украине является ненадлежащим, многие вопросы регулируются подзаконными нормативными актами (внутренними приказами), которые должны детализировать отдельные положения закона, а в законе в свою очередь они не отражены [1]. Способы и методы защиты видео- и фотоматериалов на законодательном уровне не определены, хотя

необходимость обеспечения надлежащего хранения закреплена в ряде нормативно-правовых актов, в которых затрагиваются вопросы применения систем видеонаблюдения. При этом, резолюция Парламентской Ассамблеи Совета Европы по осуществлению видеонаблюдения в общественных местах рекомендует на законодательном уровне закрепить практику кодирования (шифрования) данных видеонаблюдения для защиты их от несанкционированного доступа и модификации, что может помочь гарантировать их достоверность для уголовных расследований [2].

Сейчас в Украине обеспечение защиты ВИР с длительной актуальностью по времени проводится в основном на основе организационных мероприятий, а для ВИР с ограниченной актуальностью по времени либо отсутствует, либо проводится на основе механизмов разграничения доступа.

Второе проблемное ограничение. Присутствие антропогенных и техногенных угроз требует обеспечение конфиденциальности ВИР, которое приводит к увеличению временных затрат на обработку и доставку видеоданных, т.е. происходит снижение их доступности. В свою очередь, поддержание доступности ВИР обеспечивается при снижении ее конфиденциальности.

Третье проблемное ограничение связано с ограниченностью пропускной способности телекоммуникационных систем, в особенности построенных на основе дистанционных средств. При этом для обеспечения конфиденциальности при поддержании заданной оперативности необходимо снизить объем полезной информации, что приведет к уменьшению достоверности, т.е. потере целостности.

Отсюда ***актуальной научно-прикладной проблемой*** является повышение категорий информационной безопасности видеоинформации, которая обрабатывается и передается с использованием беспроводных телекоммуникационных технологий в условиях реального времени.

Анализ последних исследований и публикаций. На сегодняшний день наиболее популярным решением для обеспечения конфиденциальности информации является шифрование. Однако существенным недостатком криптографических алгоритмов является то, что подавляющее большинство методов являются универсальными, т.е. не учитывают особенности источников информации. За последние годы было предложено множество специализированных алгоритмов шифрования для решения проблемы защиты цифровых изображений и потока видеоинформации. Один из первых широко распространенных подходов заключался в перестановке строк или столбцов кадров видеопотока [3]. Большинство таких алгоритмов не обеспечивают достаточную криптостойкость и были подвергнуты вскрытию. Вместе с тем, в настоящее время активно ведутся разработки в области визуальной

криптографии, впервые предложенные М. Наором и А. Шамиром [4]. Однако предложенные подходы применяются к исходным видеоданным, но на практике исходный объем ВИР может быть большим, что приведет к потере оперативности при обработке данных в реальном режиме времени.

В случае обеспечения конфиденциальности изображений следует учитывать, что:

1) для видеоинформации кардинально меняется вопрос, связанный с ее зашитой, а именно встает вопрос с резким разделением информационной защиты на семантический и синтаксический уровень;

2) видеоинформация – особый источник информации, имеющий аналоговую природу и до 90 % обусловленный психофизическими особенностями ее восприятия зрительной системой;

3) наличие многомерных связей;

4) отсутствие методологически обоснованного математического аппарата позволяющего устанавливать взаимосвязи между качественной и количественной сторонами видеоинформации;

5) наличие значительного количества показателей оценки качества и количества видеоинформации и видеоданных.

Поэтому возникла необходимость в разработке принципиально новой технологии, которая одновременно обеспечивает повышение оперативности доведения видеоинформации и ее защиту на основе методов семантической и синтаксической обработки изображений. В работах [5–9] была предложена технология криптокомпрессионного представления (ККП) изображений, предназначенная для сокрытия семантического содержания изображения с учетом как статистических, так и структурных особенностей источника информации. На основе разработанной технологии в работах [10, 11] предложен метод ККП изображений на основе статической схемы обобщенного позиционного кодирования в двумерном базисе. Одним из недостатков такого подхода является построение информационной составляющей (ИС) ККП на основе одинакового количества исходных элементов фрагмента видеоданных. Длина большинства кодограмм ИС ККП в битном представлении при таком подходе является значительно меньшей, чем выделяется для хранения кодового слова. Это приводит к появлению большого количества незначимых нулевых элементов в битных последовательностях ИС ККП [12, 13]. Этот недостаток влияет на объем ИС ККП и на выходные статистические характеристики ИС ККП. Поэтому в [12–16] был предложен метод ККП изображений на основе плавающей схемы в базисе по верхним границам. Для дополнительного сокращения избыточности и повышения криптостойкости представления кода ИС ККП предлагается внести изменение в структурную схему построения систем ККП изображений.

Целью исследований является разработка метода криптокомпрессионного представления изображений на основе адаптивного обобщенного позиционного кодирования для биноминального пространства для дополнительного сокращения избыточности и повышения криптостойкости видеоинформационных ресурсов.

Основной материал

Особенностью метода является выполнение двухкаскадного обобщенного позиционного кодирования в двоичном структурном пространстве, которое заключается в:

1) кодировании двоичного представления элементов фрагмента изображения;

2) формировании кода с учетом структурных особенностей исходного фрагмента изображения.

Рассмотрим подробней этапы двухкаскадного обобщенного позиционного кодирования.

Первоначально кадр изображения представляет собой массив пикселей размерностью $I^{(row)} \times I^{(col)}$, где $I^{(row)}$ – количество строк в изображении, а $I^{(col)}$ – количество столбцов. Количество δ градаций яркости элементов цветовой компоненты изображения характеризует максимально возможное значение яркости A. Исходный кадр изображения перед обработкой разбивается на сегменты A_φ размерностью $m \times n$, где m – количество строк фрагмента изображения, а n – количество столбцов.

Индекс φ сегмента A_φ в кадре определяется с помощью координатных переменных ψ и ζ, что задается таким выражением:

$$\varphi = (\psi - 1) \times \zeta_{max} + \zeta, \quad \varphi = \overline{1, \Phi},$$

где ψ – координата сегмента A_φ в кадре по вертикали;

ζ – координата сегмента A_φ в кадре по горизонтали;

ζ_{max} – максимальное значение координатной переменной по горизонтали;

Φ – количество сегментов A_φ в кадре.

Максимальное значение координатной переменной по вертикали ψ_{max} и по горизонтали ζ_{max} определяется из соотношения размера кадра и сегмента:

$$\psi_{max} = \frac{I^{(row)}}{m}, \qquad \zeta_{max} = \frac{I^{(col)}}{n}.$$

Сегмент A_φ представляет собой двухмерный массив. Данный массив содержит элементы $a(i, j)_\varphi$, $i = \overline{1, m}$, $j = \overline{1, n}$, которые содержат информацию о яркости. Размеры $m \times n$ массива выбираются кратными степени 2, т.е

$m, n \in 2, 4, 8, 16$. Обычно значения сторон массива принимаются равными $m = n$. В кадре исходные элементы $a(i, j)_\varphi$ в i-й строке j-го столбца сегмента A_φ могут принимать значение в диапазоне $[0; 255]$, т.е. $\delta = 255 = 2^8 - 1$.

На ***первом*** этапе для учета структурных особенностей исходного фрагмента изображения определяются максимальные значения g_i для каждой строки и максимальные значения g_j для каждого столбца на основе формул:

$$g_i = \max(a_{i,\,j}) + 1,\ j = \overline{1; n}, \tag{1}$$

$$g_j = \max(a_{i,\,j}) + 1,\ i = \overline{1; m}. \tag{2}$$

Последовательность максимальных значений g_i для каждой строки и максимальных значений g_j для каждого столбца образуют вектора $G^{(row)} = \{g_i\}$ и $G^{(col)} = \{g_j\}$ соответственно. После вычисления максимальных значений g_i для каждой строки и максимальных значений g_j для каждого столбца фрагмента изображения A_φ происходит определение необходимости формирования структурного кода N_ℓ по значениям яркости $a(i, j)_\varphi$ и минимаксной системе оснований на базе плавающей схемы обобщенного позиционного кодирования. Для этого предлагается ввести двоичный признак G_φ, который указывает на проведение дополнительного преобразования двоичного представления отдельного элемента $a_{i,\,j}$ фрагмента изображения согласно функционала:

$$G_\varphi = F_G(g_i,\, g_j),\ i = \overline{1; m},\ \ j = \overline{1; n}. \tag{3}$$

При значении двоичного признака $G_\varphi = 1$ на втором этапе будет осуществляться биноминальное кодирование двоичного представления отдельного элемента $a_{i,\,j}$ фрагмента изображения.

При значении двоичного признака $G_\varphi = 0$ второй этап проводиться не будет. В этом случае процесс формирования ИС ККП изображений в двухмерном базисе на основе плавающей схемы задается выражением:

$$N = \sum_{\tau=1}^{Q} a_\tau V_\tau, \tag{4}$$

где V_τ – значение весового коэффициента, которое определяется формулой:

$$V_\tau = \begin{cases} \prod\limits_{\xi=\tau+1}^{Q} s'_\xi = \prod\limits_{\xi=\tau+1}^{Q} \min\left(G^{(row)}\left(\xi - m\frac{\xi-1}{m}\right),\ G^{(col)}\left(\frac{\xi-1}{m} + 1\right)\right), & \tau < Q; \\ 1, & \tau = Q, \end{cases} \tag{5}$$

где Q – плавающее количество элементов исходного фрагмента изображения, принимающих участие в формировании кода в двухмерном базисе на основе плавающей схемы с учетом проверки на переполнение кодового слова;

ξ – линейная координата элемента $a_{i,j}$ при сканировании столбцов фрагмента изображения сверху вниз, начиная с левого столбца. Переход от двухмерной координаты элемента $a_{i,j}$ к линейной определяется зависимостью:

$$\xi = i + (j-1) \times m. \qquad (6)$$

Обратное преобразование координат задается формулами:

$$i = \xi - m\left[\frac{\xi - 1}{m}\right], \; j = \left[\frac{\xi - 1}{m}\right] + 1. \qquad (7)$$

Количество Q элементов исходного фрагмента изображения, принимающих участие в формировании кода в двухмерном базисе на основе плавающей схемы с учетом проверки на переполнение кодового слова, не превышает общее количество элементов в фрагменте изображения:

$$Q \le mn. \qquad (8)$$

Для контроля переполнения кодового слова при формировании кода ИС ККП N введем дополнительную величину T_Q, равную накопленному произведению оснований для Q элементов, принимающих участие в формировании кода, которая определяется по формуле:

$$T_Q = \prod_{\xi=1}^{Q} s_\xi = \prod_{\xi=1}^{Q} \min\left(G^{(row)}(\xi - m\frac{\xi-1}{m}),\; G^{(col)}(\frac{\xi-1}{m}+1)\right). \qquad (9)$$

Переполнения кодового слова не произойдет, если выполняется неравенство

$$T_Q \le 2^M - 1, \qquad (10)$$

где $2^M - 1$ – наибольшее число, которое может храниться в кодовом слове длиной $м$ элементов.

Действительно, поскольку выполняется неравенство $T_Q \ge N$, тогда

$$N \le 2^M - 1.$$

Максимальное количество элементов $Q_{пр}$, принимающих участие в формировании кода ИС ККП, определяется как значение аргумента, при котором величина T_Q достигает максимума при условии выполнения неравенства (10) и рассчитывается по формуле

$$Q_{пр} = \arg\max_{Q}(T_Q) = \arg\max_{Q}(\prod_{\xi=1}^{Q} s'_\xi) = \\ = \arg\max_{Q}(\prod_{\xi=1}^{Q} \min\ (G^{(row)}(\xi - m\frac{\xi-1}{m}),\ G^{(col)}(\frac{\xi-1}{m}+1))) \quad . \quad (11)$$

С учетом соотношения (11) выражение (4) для определения кода принимает вид

$$N = \sum_{\tau=1}^{Q_{пр}} a_\tau V_\tau\ , \quad (12)$$

и выражение (5) для определения весового коэффициента преобразуется в

$$V_\tau = \begin{cases} \prod_{\xi=\tau+1}^{Q_{пр}} \min\ (G^{(row)}(\xi - m\frac{\xi-1}{m}),\ G^{(col)}(\frac{\xi-1}{m}+1)), & \tau < Q_{пр} < mn; \\ 1, & \tau = Q_{пр} \le mn. \end{cases} \quad (13)$$

Из анализа выражений (11) – (13) при условии, что каждый элемент фрагмента изображения a_τ принимает максимальное значение равное 255, а его основание s_τ соответственно равно 256, с учетом соотношений (9) и (10) при длине кодового слова M, равной 64 бита, максимальное количество элементов $Q_{пр}$, которые могут принимать в формировании кода ИС ККП, равно 8. Это значение и есть минимальным количеством элементов, принимающих участие в формировании кода-номера.

Информационная составляющая кодограммы изображений на основе плавающей схемы системы обобщенного позиционного кодирования вычисляется за три шага, так, что:

1. На первом шаге (заключающемся в подготовке исходных данных и определении служебных составляющих):

– исходное изображение разбивается на фрагменты размерностью $m \times n$;

– определяется система оснований $S'(m \times n) = \min\ (G^{(row)}(i), G^{(col)}(j))$, $i = \overline{1, m}$, $j = \overline{1, n}$, исходного фрагмента изображения на основе выражений (1) и (2);

– преобразовывается исходный фрагмент изображения на основе выражения (6) из двумерной матрицы $A = \{a_{i,\,j}\}$, $i = \overline{1, m}$, $j = \overline{1, n}$, в одномерный вектор $A = \{a_\tau\}$, который в дальнейшем будет рассматриваться, как одномерное структурное число;

– расширяется система оснований $S'(m \times n) = \min\ (G^{(row)}(i), G^{(col)}(j))$ до мощности исходного фрагмента изображения в одномерном векторном виде $S'(m \times n) = \{s'_\xi\}$ на основе выражения:

$$s'_{\xi} = \min\ (G^{(row)}(\xi - m\frac{\xi-1}{m}),\ G^{(col)}(\frac{\xi-1}{m}+1)).$$

2. На втором шаге рассчитывается максимальное количество элементов $Q_{пр}$ одномерного структурного числа, принимающих участие в формировании ИС ККП, из соотношения

$$\begin{aligned} Q_{пр} &= \arg\max_{Q}(T_Q) = \arg\max_{Q}(\prod_{\xi=1}^{Q} s'_{\xi}) = \\ &= \arg\max_{Q}(\prod_{\xi=1}^{Q} \min\ (G^{(row)}(\xi - m\frac{\xi-1}{m}),\ G^{(col)}(\frac{\xi-1}{m}+1))) \end{aligned}, \quad (14)$$

при котором выполняется следующее условие

$$\prod_{\xi=1}^{Q} s'_{\xi} = \prod_{\xi=1}^{Q} \min\ (G^{(row)}(\xi - m\frac{\xi-1}{m}),\ G^{(col)}(\frac{\xi-1}{m}+1)) \le 2^{M} - 1. \quad (15)$$

3. На третьем шаге непосредственно формируется ИС ККП на основе выражений (4) и (5). Значение кода ИС является интегрированным, и формируется с учетом служебных данных по оператору $f\ (A;\ G)$, где:

$$\begin{aligned} f(A;G) = N &= \sum_{\tau=1}^{Q_{пр}} a_{\tau} V_{\tau} = \sum_{\tau=1}^{Q_{пр}} a_{\tau} \prod_{\xi=\tau+1}^{Q_{пр}} s'_{\xi} = \\ &= \sum_{\tau=1}^{Q_{пр}} (a_{\tau} \times \min\ (G^{(row)}(\xi - m[\frac{\tau-1}{m}),\ G^{(col)}(\frac{\xi-1}{m}+1))) \end{aligned}. \quad (16)$$

Формирование кода ИС ККП возможно и на основе рекуррентной схемы путем добавления очередного элемента одномерного структурного числа. Первоначальное значение кода N_1 определяется значением первого элемента a_1, т.е.

$$N_1 = a_1. \quad (17)$$

Процесс формирования кода задается следующими выражениями

$$\begin{aligned} N_{\tau} &= N_{\tau-1} s'_{\tau} + a_{\tau} = \\ &= N_{\tau-1} \times \min\ (G^{(row)}(\tau - m[\frac{\tau-1}{m}]),\ G^{(col)}([\frac{\xi-1}{m}]+1)) + a_{\tau} \end{aligned}, \quad (18)$$

где N_{τ}, $N_{\tau-1}$ – промежуточное значение кода для τ-го и $(\tau-1)$-го элементов.

Для исключения переполнения кодового слова перед каждым добавлением к коду ИС ККП $N_{\tau-1}$ очередного элемента a_{τ} проводится проверка на переполнение кодового слова, которая с учетом соотношения $\tau = Q$, определяется с помощью дополнительной величины T_{τ} на основе выражения (9) с учетом выполнения неравенства (10).

Действительно, поскольку выполняется неравенство $T_{\tau} \ge N_{\tau}$, тогда

$$N_\tau \le 2^M - 1.$$

Процесс формирования кода ИС ККП N заканчивается тогда, когда будет обработан последний Q-й элемент:

$$N = N_Q = N_{Q-1} s'_Q + a_Q = N_{Q-1} \times \min\left(g_{Q-m\left[\frac{Q-1}{m}\right]}, g_{\left[\frac{Q-1}{m}\right]+1}\right) + a_Q, \quad (19)$$

$$\text{при } T_Q \le 2^M - 1 \text{ и } Q \le mn.$$

В данном случаи Q-й элемент будет равен максимальному количеству элементов $Q_{пр}$ одномерного структурного числа, принимающих участие в формировании кода ИС ККП, рассчитанному по формуле (14).

Далее на ***втором*** этапе при значении двоичного признака $G_\varphi = 1$, выполняется биноминальное кодирование двоичного представления отдельного элемента $a_{i,j}$ фрагмента изображения. Для чего элемент фрагмента изображения рассматривается как одномерное плавающее структурное число. Плавающая схема формирования структурного кода подразумевает переменное количество ν кодируемых разрядов $b(a_{i,j})^{(\gamma)}$ двоичной последовательности элемента $a_{i,j}$.

Вычисление количества серий единиц $\eta_{i,j}$ для ν кодируемых разрядов $b(a_{i,j})^{(\gamma)}$ двоичной последовательности элемента $a_{i,j}$ выполняется, начиная со старшего разряда согласно следующему алгоритму:

– на нулевом шаге $\gamma = 0$ начальные значения разряда и длины серии приравнивают нулю: $b(a_{i,j})^{(0)} = 0$, $\eta_{i,j}^{(0)} = 0$;

– на γ-м шаге число серий увеличивается на 1:

$$\eta_{i,j}^{(\gamma)} = \eta_{i,j}^{(\gamma-1)} + 1, \text{ если } b(a_{i,j})^{(\gamma)} > b(a_{i,j})^{(\gamma-1)};$$

– в противном случае $\eta_{i,j}^{(\gamma)} = \eta_{i,j}^{(\gamma-1)}$, если $b(a_{i,j})^{(\gamma)} \le b(a_{i,j})^{(\gamma-1)}$;

– для конечного шага при $\gamma = \nu$ получаем искомое значение количества серий единиц $\eta_{i,j} = \eta_{i,j}^{(\nu)}$ для элемента $a_{i,j}$.

Формирование структурного представления $y_{i,j}$ двоичных данных для отдельного элемента $a_{i,j}$ задается функционалом $f_{bin}(a_{i,j}, \nu, \eta_{i,j})$:

$$y_{i,j} = f_{bin}(a_{i,j}, \nu, \eta_{i,j}), \quad (20)$$

где ν – количество кодируемых разрядов отдельного элемента $a_{i,j}$ фрагмента изображения,

$\eta_{i,j}$ – количество серий единиц отдельного элемента $a_{i,j}$ фрагмента изображения.

Полученное структурное представление $y_{i,j}$ характеризуется количеством $W_{i,j}$ допустимых структурных чисел для элемента фрагмента изображения, которое определяется по формуле:

$$W_{i,j} = \frac{(\nu+1)!}{(2\eta_{i,j})!(\nu+1-2\eta_{i,j})!}.$$

Формирование структурного представления $y_{i,j}$ двоичных данных для отдельного элемента $a_{i,j}$ фрагмента изображения производятся на основе выражения:

$$y_{i,j} = \sum_{\gamma=1}^{\nu} b(a_{i,j})^{(\gamma)} \times p_{i,j}^{(\gamma)}, \qquad (21)$$

где $b(a_{i,j})^{(\gamma)}$ – значение γ-го разряда элемента $a_{i,j}$;

$p_{i,j}^{(\gamma)}$ – весовой коэффициент отдельного элемента $a_{i,j}$ фрагмента изображения, зависящий от значений γ и $\eta_{i,j}$.

Расчет весовых коэффициентов проводится на основе рекуррентных выражений [17], позволяющих вычислить значение весового коэффициента $p_{i,j}^{(\gamma)}$ разряда $b_{i,j}^{(\gamma)}$ для элемента $a_{i,j}$ через весовой коэффициент $p_{i,j}^{(\gamma-1)}$ предыдущего разряда $b(a_{i,j})^{(\gamma)}$ элемента $a_{i,j}$. При этом возможны четыре варианта зависимости между разрядами отдельного элемента $a_{i,j}$ фрагмента изображения:

1) $|b(a_{i,j})^{(\gamma-2)} - b(a_{i,j})^{(\gamma-1)}| = 1$ и $|b(a_{i,j})^{(\gamma-1)} - b(a_{i,j})^{(\gamma)}| = 1$:

$$p_{i,j}^{(\gamma)} = p_{i,j}^{(\gamma-1)} \frac{\beta_{i,j}^{(\gamma-1)}+1}{\nu-\gamma+2}; \qquad (22)$$

2) $|b(a_{i,j})^{(\gamma-2)} - b(a_{i,j})^{(\gamma-1)}| = 1$ и $|b(a_{i,j})^{(\gamma-1)} - b(a_{i,j})^{(\gamma)}| = 0$:

$$p_{i,j}^{(\gamma)} = p_{i,j}^{(\gamma-1)} \left(\frac{(\beta_{i,j}^{(\gamma-1)}+1)\beta_{i,j}^{(\gamma)}}{(\nu-\gamma+2-\beta_{i,j}^{(\gamma-1)})(\nu-\gamma+2)} \right); \qquad (23)$$

3) $|b(a_{i,j})^{(\gamma-2)} - b(a_{i,j})^{(\gamma-1)}| = 0$ и $|b(a_{i,j})^{(\gamma-1)} - b(a_{i,j})^{(\gamma)}| = 1$:

$$p_{i,j}^{(\gamma)} = p_{i,j}^{(\gamma-1)} \frac{(\nu-\gamma-\beta_{i,j}^{(\gamma-1)}+3)(\nu-\gamma+2-\beta_{i,j}^{(\gamma-1)})}{(\beta_{i,j}^{(\gamma-1)})(\nu-\gamma+2)}; \qquad (24)$$

4) $|b(a_{i,j})^{(\gamma-2)} - b(a_{i,j})^{(\gamma-1)}| = 0$ и $|b(a_{i,j})^{(\gamma-1)} - b(a_{i,j})^{(\gamma)}| = 0$:

$$p_{i,\,j}^{(\gamma)} = p_{i,\,j}^{(\gamma-1)} \frac{\nu - \gamma - \beta_{i,\,j}^{(\gamma-1)} + 3}{\nu - \gamma + 2}, \qquad (25)$$

где ν – количество разрядов в обрабатываемом отдельном элементе $a_{i,\,j}$; $\beta_{i,\,j}^{(\gamma)}$ – рекуррентный параметр, равный количеству двоичных перепадов (переходов между «0» и «1») для последовательности, состоящей из $(\nu - \gamma + 1)$ необработанных разрядов отдельного элемента $a_{i,\,j}$ фрагмента изображения:

$$\beta_{i,\,j}^{(\gamma)} = \beta_{i,\,j}^{(\gamma-1)} - |\, b(a_{i,\,j})^{(\gamma)} - b(a_{i,\,j})^{(\gamma-1)} \,| \,. \qquad (26)$$

Для начального шага обработки $(\gamma = 1)$ принимаются следующие значения разряда $b_{i,\,j}^{(0)} = 0$ и рекуррентного параметра $\beta_{i,\,j}^{(0)} = 2\eta_{i,\,j}$. Весовой коэффициент $p_{i,\,j}^{(\gamma)}$ на первом шаге обработки для двух случаев значения первого разряда отдельного элемента $a_{i,\,j}$ фрагмента изображения определяется как:

1) $b(a_{i,\,j})^{(1)} = 1$:

$$p_{i,\,j}^{(1)} = W_{i,\,j} \frac{\nu + 1 - 2\eta_{i,\,j}}{\nu + 1}; \qquad (27)$$

2) $b(a_{i,\,j})^{(1)} = 0$:

$$p_{i,\,j}^{(1)} = W_{i,\,j} \frac{2\eta_{i,\,j}}{\nu + 1}. \qquad (28)$$

С учетом выражений (27) и (28) выражение (21) для вычисления структурного представления $y_{i,\,j}$ двоичных данных для отдельного элемента $a_{i,\,j}$ фрагмента изображения принимает вид:

$$y_{i,\,j} = b(a_{i,\,j})^{(1)} \times W_{i,\,j} \times \frac{\nu + 1 - 2\eta_{i,\,j}}{\nu + 1} + \overline{b(a_{i,\,j})^{(1)}} \times W_{i,\,j} \times \frac{2\eta_{i,\,j}}{\nu + 1} + \\ + \left(b(a_{i,\,j})^{(2)}; \ldots; b(a_{i,\,j})^{(\nu)}\right) \cdot \left(p_{i,\,j}^{(2)}; \ldots; p_{i,\,j}^{(\nu)}\right), \qquad (29)$$

где $\overline{b(a_{i,\,j})^{(\gamma)}}$ – инвертированное значение γ-го разряда элемента $a_{i,\,j}$;

$B \cdot P$ – выполнение скалярного умножения вектора $B = \{b(a_{i,\,j})^{(2)}; \ldots; b(a_{i,\,j})^{(\nu)}\}$ разрядов элемента $a_{i,\,j}$ фрагмента изображения на вектор $P = \{p_{i,\,j}^{(2)}; \ldots; p_{i,\,j}^{(\nu)}\}$ весовых коэффициентов.

В связи с тем, что согласно (22) – (25), (27) и (28) весовой коэффициент $p_{i,\,j}^{(\gamma)}$ отдельного элемента $a_{i,\,j}$ фрагмента изображения определяется рекуррентно, выражение (29) приводится к виду:

$$y_{i,\,j}=\frac{(\nu+1)!}{(\nu+1-2\eta_{i,\,j})!\times(2\eta_{i,\,j})!}\times(b(a_{i,j})^{(1)}\times\frac{\nu+1-2\eta_{i,\,j}}{\nu+1}+\overline{b(a_{i,j})^{(1)}}\times\frac{2\eta_{i,\,j}}{\nu+1}+$$

$$+\overline{b(a_{i,j})^{(2)}}\times b(a_{i,j})^{(1)}\times\frac{\nu+1-2\eta_{i,\,j}}{\nu+1}\times\frac{2\eta_{i,\,j}}{\nu}+b(a_{i,j})^{(2)}\times\overline{b(a_{i,j})^{(1)}}\times\frac{2\eta_{i,\,j}}{\nu+1}\times$$

$$\times\frac{(\nu-2\eta_{i,\,j}+1)(\nu-2\eta_{i,\,j})}{2\eta_{i,\,j}\times\nu}+b(a_{i,j})^{(2)}\times b(a_{i,j})^{(1)}\times\frac{\nu+1-2\eta_{i,\,j}}{\nu+1}\times$$

$$\times\frac{2\eta_{i,\,j}(2\eta_{i,\,j}-1)}{(\nu+1-2\eta_{i,\,j})\times\nu}+\overline{b(a_{i,j})^{(2)}}\times\overline{b(a_{i,j})^{(1)}}\times\frac{2\eta_{i,\,j}}{\nu+1}\times\frac{\nu-2\eta_{i,\,j}+2}{\nu}+\bullet\bullet\bullet+$$

$$+(\overline{b(a_{i,j})^{(\gamma-2)}}\times b(a_{i,j})^{(\gamma-1)}\times\overline{b(a_{i,j})^{(\gamma)}}+b(a_{i,j})^{(\gamma-2)}\times\overline{b(a_{i,j})^{(\gamma-1)}}\times b(a_{i,j})^{(\gamma)})\times$$

$$\times p_{i,\,j}^{(\gamma-1)}\times\frac{\beta_{i,\,j}^{(\gamma-1)}+1}{\nu-\gamma+2}+(\overline{b(a_{i,j})^{(\gamma-2)}}\times b(a_{i,j})^{(\gamma-1)}\times b(a_{i,j})^{(\gamma)}+b(a_{i,j})^{(\gamma-2)}\times$$

$$\times\overline{b(a_{i,j})^{(\gamma-1)}}\times\overline{b(a_{i,j})^{(\gamma)}})\times p_{i,\,j}^{(\gamma-1)}\times\times\frac{(\beta_{i,\,j}^{(\gamma-1)}+1)}{(\nu-\gamma+2-\beta_{i,\,j}^{(\gamma-1)})}\times\frac{\beta_{i,\,j}^{(\gamma)}}{(\nu-\gamma+2)}+$$

$$+(b(a_{i,j})^{(\gamma-2)}\times\times b(a_{i,j})^{(\gamma-1)}\times\overline{b(a_{i,j})^{(\gamma)}}+\overline{b(a_{i,j})^{(\gamma-2)}}\times\overline{b(a_{i,j})^{(\gamma-1)}}\times b(a_{i,j})^{(\gamma)})\times$$

$$\times p_{i,\,j}^{(\gamma-1)}\times b(a_{i,j})^{(\gamma-1)}\times\frac{(\nu-\gamma-\beta_{i,\,j}^{(\gamma-1)}+3)(\nu-\gamma+2-\beta_{i,\,j}^{(\gamma-1)})}{\beta_{i,\,j}^{(\gamma-1)}(\nu-\gamma+2)}+(b(a_{i,j})^{(\gamma-2)}\times$$

$$\times b(a_{i,j})^{(\gamma)}+\overline{b(a_{i,j})^{(\gamma-2)}}\times\overline{b(a_{i,j})^{(\gamma-1)}}\times\overline{b(a_{i,j})^{(\gamma)}})\times p_{i,\,j}^{(\gamma-1)}\times\frac{\nu-\gamma-\beta_{i,\,j}^{(\gamma-1)}+3}{\nu-\gamma+2}+$$

$$+(\overline{b(a_{i,j})^{(\nu-2)}}\times b(a_{i,j})^{(\nu-1)}\times\overline{b(a_{i,j})^{(\nu)}}+b(a_{i,j})^{(\nu-2)}\times\overline{b(a_{i,j})^{(\nu-1)}}\times b(a_{i,j})^{(\nu)})\times$$

$$\times\bullet\bullet\bullet\times p_{i,\,j}^{(\nu-1)}\times\frac{\beta_{i,\,j}^{(\nu-1)}+1}{2}+(\overline{b(a_{i,j})^{(\nu-2)}}\times b(a_{i,j})^{(\nu-1)}\times b(a_{i,j})^{(\nu)}+b(a_{i,j})^{(\nu-2)}\times$$

$$\times\overline{b(a_{i,j})^{(\nu-1)}}\times\overline{b(a_{i,j})^{(\nu)}})\times p_{i,\,j}^{(\nu-1)}\times\frac{(\beta_{i,\,j}^{(\gamma-1)}+1)\beta_{i,\,j}^{(\nu-1)}}{2(2-\beta_{i,\,j}^{(\gamma-1)})}+(b(a_{i,j})^{(\nu-2)}\times b(a_{i,j})^{(\nu-1)}\times$$

$$\times\overline{b(a_{i,j})^{(\nu)}}+\overline{b(a_{i,j})^{(\nu-2)}}\times\times\overline{b(a_{i,j})^{(\nu-1)}}\times b(a_{i,j})^{(\nu)})\times p_{i,\,j}^{(\nu-1)}\times$$

$$\times\frac{(3-\beta_{i,\,j}^{(\gamma-1)})(2-\beta_{i,\,j}^{(\nu-1)})}{2\beta_{i,\,j}^{(\gamma-1)}}+(b(a_{i,j})^{(\nu-2)}\times b(a_{i,j})^{(\nu-1)}\times b(a_{i,j})^{(\nu)}+$$

$$+\overline{b(a_{i,j})^{(\nu-2)}}\times\overline{b(a_{i,j})^{(\nu-1)}}\times\overline{b(a_{i,j})^{(\nu)}})\times p_{i,\,j}^{(\nu-1)}\times\frac{3-\beta_{i,\,j}^{(\nu-1)}}{2}),\quad(31)$$

где $b(a_{i,j})^{(\nu)}$ – значение ν-го (младшего) разряда элемента $a_{i,j}$;

$b(a_{i,j})^{(\nu-1)}$ – значение $(\nu-1)$-го (предпоследнего) разряда элемента $a_{i,j}$.

После формирования структурного представления двоичных данных образуется биноминальное структурное число $Y(G(g_i, g_j, A_\varphi)$. Для этого числа (i, j)-е координаты определяют местоположение элемента в локальном фрагменте изображения.

На ***третьем*** этапе для сокращения длины кода, полученных структурных представлений $y_{i,j}$ двоичной последовательности отдельных элементов $a_{i,j}$ фрагмента изображения, предлагается уменьшить значения оснований на основании значения двоичного признака G_φ. Для чего величина $s'_{i,j}$ рассчитывается по одной из следующих формул:

$$s'_{i,j} = \min(G^{(row)}(i);\ G^{(col)}(j)),\ i=\overline{1;m},\ j=\overline{1,n},\ G_\varphi = 0, \qquad (32)$$

$$s'_{i,j} = W_{i,j},\ i=\overline{1;m},\ j=\overline{1,n},\ G_\varphi = 1. \qquad (33)$$

С учетом выражений (32) и (33) величина $s'_{i,j}$ для фрагмента изображения вычисляется по следующей формуле:

$$s'_{i,j} = \min\ (G^{(row)}(i);\ G^{(col)}(j)) \times \overline{G_\varphi} + \frac{(\nu+1)!}{(2\eta_{i,j})!(\nu+1-2\eta_{i,j})!}) \times G_\varphi. \qquad (34)$$

После чего на ***четвертом*** этапе для построения ИС ККП изображения на основе плавающей схемы необходимо выполнить линеаризацию биноминального структурного числа $Y(G(g_i, g_j, A_\varphi)$.

Структурное число $Y(G(g_i, g_j, A_\varphi)$ в исходном виде представляет собой двумерную матрицу биноминального представления исходного фрагмента изображения $A=\{a_{i,j}\}$, $i=\overline{1,m}$, $j=\overline{1,n}$, которая преобразовывается в одномерный вектор

$$Y(G(g_i, g_j, A_\varphi) = \{y_{i,j}\} = \{y_\tau\}_{\tau=\overline{1,mn}} = \{y_{m(j-1)+i}\},\ i=\overline{1;m},\ j=\overline{1,n}. \qquad (35)$$

Для удобства проведения расчетов и для определения взаимнооднозначного соответствия элементов фрагмента изображения с основаниями предлагается расширить систему оснований до мощности исходного фрагмента изображения в одномерном векторном виде. Для этого воспользуемся формулой

$$S'^{(m\times n)} = \{s'_\tau\} = \{s'_{m(j-1)+i}\},\ \tau=\overline{1,mn}. \qquad (36)$$

Для контроля переполнения кодового слова при формировании кода ИС ККП N введем дополнительную величину T_Q, равную накопленному

произведению оснований для Q элементов, принимающих участие в формировании кода ИС ККП, которая определяется по формуле

$$T_Q = \prod_{\xi=1}^{Q} s'_\xi . \qquad (37)$$

Переполнения кодового слова не произойдет, если выполняется неравенство

$$T_Q \le 2^M - 1. \qquad (38)$$

Учитывая выражение (34) величина T_Q для фрагмента изображения определяется следующей формулой:

$$T_Q = \prod_{\xi=1}^{Q} \min\ (G^{(row)}(\xi - m[\frac{\xi-1}{m}]);\ G^{(col)}([\frac{\xi-1}{m}]+1)) \times \overline{G_\varphi} + \\ + \frac{(\nu+1)!}{(2\eta_{i,j})!(\nu+1-2\eta_{i,j})!}) \times G_\varphi) \qquad . \qquad (39)$$

Максимальное количество элементов $Q_{пр}$, принимающих участие в формировании кода ИС ККП, определяется как значение аргумента, при котором величина T_Q достигает максимума при условии выполнения неравенства (38) и рассчитывается по формуле

$$Q_{пр} = \arg\max_Q(\ T_Q) = \arg\max_Q(\ \prod_{\xi=1}^{Q} s'_\xi), \qquad (40)$$

которая при использовании выражения (39) примет вид:

$$Q_{пр} = \arg\max_Q\ (\prod_{\xi=1}^{Q} \min\ (G^{(row)}(\xi - m[\frac{\xi-1}{m}]);\ G^{(col)}([\frac{\xi-1}{m}]+1))\ ) \times \\ \times \overline{G_\varphi} + \frac{(\nu+1)!}{(2\eta_{i,j})!(\nu+1-2\eta_{i,j})!}) \times G_\varphi\ ) \qquad , (41)$$

где $G^{(row)}(\xi - m[\frac{\xi-1}{m}])$, $G^{(col)}([\frac{\xi-1}{m}]+1)$ – максимальное значение элемента $a_{i,j}$ в строке и столбце фрагмента изображения, соответственно.

Пятый этап состоит во втором каскаде формирования ИС ККП изображений (представления данных в системе обобщенного позиционного кодирования) в базисе по верхним границам на основе плавающей схемы задается выражением:

$$N = \sum_{\tau=1}^{Q} y_\tau V_\tau . \qquad (42)$$

В данном выражении весовой коэффициент V_τ определяется формулой:

$$V_\tau = \begin{cases} \prod_{\xi=\tau+1}^{Q} s'_\tau, & \rightarrow \quad \tau < Q; \\ 1, & \rightarrow \quad \tau = Q. \end{cases} \qquad (43)$$

где Q – плавающее количество элементов одномерного структурного числа, принимающих участие в формировании кода в двухмерном базисе с учетом особенностей структурного кода двоичной последовательности на основе плавающей схемы с учетом проверки на переполнение кодового слова. Это плавающее количество ограничено следующим значением:

$$Q \le mn . \qquad (44)$$

Формулу (43) вычисления весовых коэффициентов V_τ на основе выражений (34) и (7) преобразуется к следующему выражению:

$$V_\tau = \begin{cases} \prod_{\xi=\tau+1}^{Q} \min(G^{(row)}(\xi - m[\frac{\xi-1}{m}]); G^{(col)}([\frac{\xi-1}{m}]+1)\overline{G_\varphi} + \frac{(\nu+1)!}{(2\eta_{i,j})!(\nu+1-2\eta_{i,j})!}) \times G_\varphi), & \tau < Q; \\ 1, & \tau = Q. \end{cases} \qquad (45)$$

Для значения двоичного признака $G_\varphi = 1$ представляется возможным развертывание формулы (42) вычисления кода N в следующее выражение:

$$N = \sum_{\xi=1}^{Q_{пр}} (\frac{(\nu+1)!}{(\nu+1-2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})! \times (2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})!} \times$$

$$\times (b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(1)} \times \frac{\nu+1-2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}{\nu+1} + \overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(1)}} \times$$

$$\times \frac{2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}{\nu+1} + \overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(2)}} \times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(1)} \times$$

$$\times \frac{\nu+1-2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}{\nu+1} \times \frac{2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}{\nu} + b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(2)} \times$$

$$\times \overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(1)}} \times \frac{2\eta_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}{\nu+1} \times$$

$$\times \frac{(\nu - 2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} + 1)(\nu - 2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})}{2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} \times \nu} + b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(2)} \times$$

$$\times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(1)} \times \frac{\nu + 1 - 2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}}{\nu + 1} \times$$

$$\times \frac{2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}(2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} - 1)}{(\nu + 1 - 2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}) \times \nu} + \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(2)}} \times$$

$$\times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(1)}} \times \frac{2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}}{\nu + 1} \times \frac{\nu - 2\eta_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} + 2}{\nu} +$$

$$+ \dots + (\overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-2)}} \times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-1)} \times$$

$$\times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma)}} + b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-2)} \times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-1)}} \times$$

$$\times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma)}) \times p^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} \times \frac{\beta^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} + 1}{\nu - \gamma + 2} +$$

$$+ (\overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-2)}} \times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-1)} \times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma)} +$$

$$+ b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-2)} \times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-1)}} \times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma)}}) \times$$

$$\times p^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} \times \frac{(\beta^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} + 1)\beta^{(\gamma)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}}{(\nu - \gamma + 2 - \beta^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})(\nu - \gamma + 2)} +$$

$$+ (b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-2)} \times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-1)} \times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma)}} +$$

$$+ \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-2)}} \times \overline{b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma-1)}} \times b(a_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\gamma)}) \times$$

$$\times p^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} \times \frac{(\nu - \gamma - \beta^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} + 3)(\nu - \gamma + 2 - \beta^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})}{\beta^{(\gamma-1)}_{\xi - m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}(\nu - \gamma + 2)} +$$

$$+(b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\gamma-2)}\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\gamma-1)}\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\gamma)}+$$
$$+\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\gamma-2)}}\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\gamma-1)}}\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\gamma)}})\times$$
$$\times p^{(\gamma-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}\times\frac{\nu-\gamma-\beta^{(\gamma-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}+3}{\nu-\gamma+2}+\bullet\bullet\bullet+\overline{(b(a_{\xi-m[\frac{\xi-1}{m}],\frac{\xi-1}{m}})^{(\nu-2)}}\times$$
$$\times b(a_{\xi-m\left[\frac{\xi-1}{m}\right],\frac{\xi-1}{m}})^{(\nu-1)}\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu)}}+b^{(\nu-2)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}\times$$
$$\times\overline{b^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu)})\times p^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}\times$$
$$\times\frac{\beta^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}+1}{2}+\overline{(b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-2)}}\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-1)}\times$$
$$\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu)}+b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-2)}\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-1)}}\times$$
$$\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu)}})\times p^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}\times$$
$$\times\frac{(\beta^{(\gamma-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}+1)\beta^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}{2(2-\beta^{(\gamma-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})}+(b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-2)}\times$$
$$\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-1)}\times\overline{b^{(\nu)}_{\xi-m[\frac{\xi-1}{m}]+1,[\frac{\xi-1}{m}]+1}}+\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-2)}}\times$$
$$\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-1)}}\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu)})\times p^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}\times$$
$$\times\frac{(3-\beta^{(\gamma-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})(2-\beta^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})}{2\beta^{(\gamma-1)}_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1}}+(b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-2)}\times$$
$$\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-1)}}\times\overline{b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu)}})\times b(a_{\xi-m[\frac{\xi-1}{m}],[\frac{\xi-1}{m}]+1})^{(\nu-1)}\times$$

$$\times b(a_{\xi-m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\nu)} + \overline{b(a_{\xi-m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})^{(\nu-2)}} \times p^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1} \times$$

$$\times \frac{3-\beta^{(\nu-1)}_{\xi-m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1}}{2}) \times V_{\xi} \quad . \quad (46)$$

Здесь количество $Q_{пр}$ кодируемых во втором каскаде элементов определяется выражением:

$$Q_{пр} = \arg\max_{Q} \left(\prod_{\xi=1}^{Q} \frac{(\nu+1)!}{(2\eta_{\xi-m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})!(\nu+1-2\eta_{\xi-m[\frac{\xi-1}{m}],\,[\frac{\xi-1}{m}]+1})!}\right).$$

Выводы

1. Для понижения динамического диапазона значений исходных элементов фрагмента изображения и сокращения разрядов на их представление в ККП предлагается организовать двухкаскадную обработку на основе обобщенного позиционного кодирования с учетом двоичного признакового пространства, которая заключается в:

– биноминальном кодировании двоичного представления элементов фрагмента изображения;

– формировании ККП изображений с учетом полученных структурных особенностей кодированного фрагмента изображения на основе плавающей схемы обобщенного позиционного кодирования в базисе по верхним границам двоичного структурного пространства по количеству серий единиц.

Данный подход приведет к увеличению количества кодируемых элементов, принимающих участие в формировании кодов ИС ККП изображений на основе плавающей схемы обобщенного позиционного кодирования. Это обеспечивает криптостойкость за счет увеличения неопределенности при формировании ККП и повышения оперативности за счет уменьшения ИС ККП.

2. Кодирование двоичного представления элементов фрагмента изображения для повышения оперативности выполнения преобразования предлагается организовать на основе блока байтовой замены с входом по значению элемента исходного изображения и выходом со значениями его кода и количества серий единиц. Блок замены предлагается формировать на этапе инициализации выполнения преобразования (или заранее) на основе матрицы размерностью 2×256 элементов, подсчитанной для всех 256 возможных вариантов значений элементов исходного фрагмента изображения в диапазоне $[0; 255]$. Входом в таком блоке замены будет номер столбца, равный значению элемента исходного изображения, а для нулевого значению входом будет 266-ой столбец (или номер столбца,

равный значению элемента исходного изображения, увеличенный на один). Выходом будет с одной из строк данного столбца – количество серий единиц исходного элемента изображения, а со второй строки – значение его кода. Количество серий единиц для 8-битных исходных данных может принимать значение 0 для нулевого значения исходного элемента видеоданных и от 1 до 4 для остальных элементов. При этом количество серий единиц, равное 1, наблюдается у 36 исходных элементах, равное 2 – у 126 элементах, равное 3 – у 84 элементах, равное 4 – у 9 элементах. При разных значениях входных элементов на выходе:

– один из 9 кодов структурного представления будет соответствовать 4-м входным значениям исходных данных (при кодировании нулевого входного значения в одной группе с исходными значениями с количеством серий единиц, равным 4, значение одного из 10 кодов структурного представления будет одинаковым для 4-х входных элементов);

– один из 27 кодов структурного представления будет соответствовать 3-м входным значениям исходных данных;

– один из 47 кодов структурного представления будет соответствовать 2-м входным значениям исходных данных;

– только 42 кода структурного представления будут соответствовать одному входному значению исходных данных.

Биноминальное кодирование двоичного представления элементов фрагмента изображения приводит к изменению структуры видеоданных на синтаксическом уровне, связанном с уменьшением количества разрядов для хранения значений структурного представления двоичных данных для отдельного элемента. Однако на семантическом уровне представления изображения наблюдается незначительная корреляция между исходным элементом и его структурным представлением. При условии наличия априорной информации про количество разрядов, выделяемых для хранения каждого структурного представления отдельного элемента, появляется возможность построения изображения с ошибками и потерей мелких объектов.

3. Для повышения криптостойкости кодирования предлагается в блоке замены произвести случайную перестановку кодов структурного представления элемента в пределах групп одной серии единиц. Данный подход позволит изменить структуру видеоданных на синтаксическом и семантическом уровне. Блок байтовой замены, построенный таким образом, может выступать в роли долгосрочного ключевого элемента (по аналогии с S-блоками алгоритма ГОСТ 28147-89), быть одинаковым в одной группе пользователей, храниться в секрете и использоваться на протяжении определенного времени (его передача между пользователями не требуется).

4. Отойдя от принципа прямого математического построения кодов структурного представления элемента можно пойти дальше по пути

совершенствования построения блока байтовой замены. Вместо деления на группы по количеству серий единиц предлагается значения исходных элементов перераспределить между группами или случайным образом разбить на четыре группы (нумерация которых может быть тоже случайной в диапазоне [0; 3]). Один из возможных вариантов, формирующий полные группы, предусматривает следующее распределение количества элементов по группам:

– по 32 элемента в двух группах, что потребует для их хранения по 5 бит для каждого кода структурного представления;

– 64 элемента в третьей группе с выделением 6 бит для хранения каждого кода структурного представления;

– 128 элементов в последней группе с выделением 7 бит для хранения каждого кода структурного представления.

Предложенный блок байтовой замены обеспечивает:

– перемешивание значений исходных элементов;

– рассеивание на битовом уровне за счет перераспределения элементов исходной битовой последовательности между номеров группы и кодом структурного представления элемента с неравномерно уменьшенным количеством разрядов для его хранения;

– изменение структуры изображения на синтаксическом и семантическом уровнях;

– уменьшение динамического диапазона для кодов структурного представления элемента в зависимости от номера группы, что при выполнении второго каскада гарантированно обеспечит формирование кода ИС ККП на переменном количестве элементов;

– при записи выходной кодограммы двоичного представления элементов фрагмента изображения в разнотипном динамическом диапазоне, определяемом номером группы и, соответственно, количеством разрядов для их хранения, обеспечивается неопределенность определения начала и конца битовой последовательности для каждого отдельного кода структурного представления. При этом значения номеров группы для каждого кода выступают в качестве ключевого элемента. Информационная составляющая ККП в таком случае может быть сформирована без выполнения второго каскада преобразования, что значительно повысит оперативность формирования кодограммы с обеспечением ее криптостойкости.

5. Служебная составляющая ККП изображений формируется на основе двоичного признакового пространства и представляет собой биноминальный код значений количества серий единиц для всех элементов фрагмента изображений. Это обеспечит повышение криптостойкости служебных данных за счет отсутствия прямой взаимосвязи, которая наблюдалась между значениями системы оснований и исходными значениями фрагмента изображения, что позволяло восстановить

изображение с локализированной ошибкой, базируясь только на значениях системы оснований (в варианте построения ККП изображения без криптологического преобразования служебной составляющей). А так как система служебных данных для всего изображения, являющаяся ключевым элементом для декодирования (кодирования) ИС ККП, организуется по двум подходам и несет в себе разный физический смысл, то это создает криптоаналитику больше неопределенности при дешифровке служебных данных ККП.

Література

1. Соколан Т. С. Адміністративно-правове регулювання застосування відеоспостереження правоохоронними органами України / Тетяна Сергіївна Соколан // Дисертація на здобуття наукового ступеня кандидата юридичних наук за спеціальністю 12.00.07. – Київ : Київський національний університет імені Тараса Шевченка. – 2016. – 210 с.
2. Resolution of PACE 1604 (2008) “Video surveillance of public areas” [Електронний ресурс]. – Режим доступу: http://assembly.coe.int/nw/xml/XRef/ Xref-XML2HTML-en.asp?fileid=17633&lang=en.
3. Володин А. А. Обработка видео в системах телевизионного наблюдения / А. А. Володин, В. Г. Митько, Е. Н. Спинко // Вопросы защиты информации. – 2002. – № 4 (59). – С. 34 - 47.
4. Visual cryptography / M. Naor and A. Shamir // In EUROCRYPT’94. – Springer-Verlag Berlin, 1995 [Електронний ресурс]. – Режим доступу: http://www.fe.infn.it/u/filimanto/scienza/webkrypto/ visualdecryption.pdf.
5. Barannik V. Methodology of creation of cryptographic transformations on the basis of methods excluding redundancy / V. Barannik, S. Sidchenko, V. Larin // 10th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET'2010; Lviv-Slavske; Ukraine; 23 February 2010 - 27 February 2010. – p. 312.
6. Barannik V. Methodology compression of videoinformation in the cryptographic systems / V. Barannik, S. Sidchenko, V. Larin // Science-based technologies. – 2011. – Vol. 11. No. 3-4, doi.org/10.18372/2310-5461.11.5260 (eng).
7. Баранник В. В. Синтез комбинированных криптокомпрессионных систем для обеспечения безопасности видеоинформации в инфокоммуникациях / В. В. Баранник, С. А. Сидченко, И. М. Тупица // Автоматизированные системы управления и приборы автоматики. – Х.: ХНУРЭ. – 2014. – Вып. 169. – С. 39 - 44.
8. Barannik V. The methodological base of cryptocompression presentation of videoinformation resources / V. V. Barannik, S. A. Sidchenko, V. V. Larin // 12th International Conference: The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM 2013, Lviv; Ukraine; 19 February 2013 - 23 February 2013. – pp. 27 - 28.
9. Баранник В. В. Методология позиционирования полиадических кодовых конструкций на основе классифицирующих признаков в системе криптокомпрессионного представления / В. В. Баранник, С. А. Сидченко, И. М. Тупица, Н. А. Королева// Інформаційно-керуючі системи на залізничному транспорті. – 2015. – № 4. – С. 56 - 60., doi.org/10.18664/ikszt.v0i4.53977 (rus).
10. Баранник В. В. Метод дешифрируемо-стойкого представления изображений / В. В. Баранник, С. А. Сидченко, В. В. Ларин // Сучасна спеціальна техніка. – 2011. – №1 (24). – С. 24 - 29.

11. Barannik V.V. The Decoded-proof Presentation of Images on the Basis of the Polyadycal Encoding Systems / V. V. Barannik, S. A. Sidchenko, V. V. Larin // XIth International Conference CADSM 2011, The Experience of Designing and Application of CAD Systems in Microelectronics, Lviv-Polyana, Ukraine, Lviv Polytechnic National University, February 23 – 25, 2011. – P. 182.

12. Сидченко С. А. Способ представления изображений стойких к дешифрированию на основе плавающей схемы кодирования / С. А. Сидченко // Системи озброєння і військова техніка. – 2011. – Вип. 3 (27). – С. 68 – 70.

13. Barannik V. Methodology constructions of floating chart of decoded-proof presentation of images / V. Barannik, S. Sidchenko // 11th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET'2012; Lviv - Slavske; Ukraine; 21 February 2012 - 24 February 2012. – p. 437.

14. Barannik V. The method of crypto-semantic presentation of images based on the floating scheme in the basis of the upper boundaries / V. Barannik, I. Tupitsya, S. Sidchenko, R. Tarnopolov // 2nd International Scientific-Practical Conference Problems of Infocommunications Science and Technology, PIC S and T 2015; Kharkiv; Ukraine; 13 October 2015 - 15 October 2015. – pp. 248 - 250, doi.org/10.1109/infocommst.2015.7357326 (eng).

15. Баранник В. В. Метод криптосемантического представления изображений на основе плавающей схемы в базисе по верхним границам / В. В. Баранник, С. А. Сидченко, И. М. Тупица // Радиоэлектроника и информатика. – 2015. – № 4. – С. 9 - 12.

16. Barannik V. The application for internal restructuring the data in the entropy coding process to enhance the information resource security / V. Barannik, I. Tupitsya, S. Shulgin, S. Sidchenko, V. Larin // 2016 IEEE East-West Design & Test Symposium (EWDTS), Yerevan, 2016. – pp. 1 - 4, doi.org/10.1109/ewdts.2016.7807749 (eng).

17. Наукоемкие технологии в инфокоммуникациях: обработка и защита информации: коллективная монография / под ред. В.В. Баранник, В.М. Безрук. – Х. : СМИТ, 2013. – 398 с.

МЕТОДЫ ВЫЯВЛЕНИЯ СУГГЕСТИВНЫХ ВОЗДЕЙСТВИЙ НА ПОДСОЗНАНИЕ ЧЕЛОВЕКА В ТЕКСТОВЫХ СООБЩЕНИЯХ В УСЛОВИЯХ ИНФОРМАЦИОННО-ПСИХОЛОГИЧЕСКОГО ПРОТИВОБОРСТВА

Баранник В.В., Беликова Т.В.

Введение

Анализ военных конфликтов начала XXI века свидетельствует о появлении новых форм и методов вооруженной борьбы между государствами для достижения соответствующих политических целей и разрешения межгосударственных противоречий. На смену классическим формам вооруженной борьбы пришли так называемые “гибридные войны”. Они имеют скрытый характер и проводятся, преимущественно, в политической, экономической, информационной и других сферах. Сутью таких войн является смещение центра усилий с физического уничтожения противника в рамках масштабной войны к применению средств так называемой “мягкой силы” против страны-противника с целью дезинтеграции, изменения ее руководства и включения в сферу своего влияния. Составной частью “гибридные войны” являются информационные и информационно-психологические операции, которые проводятся с целью манипуляции массовым сознанием с использованием всех видов информационно-психологических воздействий, включая и на подсознание человека.

Сегодня понятие “манипуляция сознанием” подразумевает внедрение в сознание идей, мыслей и представлений путем распространения специально подготовленной по форме и содержанию информации. Сама манипуляция сознанием свое широкое распространение получила не только и не столько в военной сфере, а и в политике и экономике (рекламной сфере и сфере услуг). Манипуляция сознанием используется и в системе образования и дошкольного воспитания детей. Оно может иметь как позитивные, так и негативные последствия и быть направленным на сознательную и подсознательную сферу человека. Поскольку манипуляция - это вид духовного и психологического воздействия, мишенью которого является психика человеческой личности, то для достижения успеха, манипуляция должна оставаться незамеченной. Успех гарантирован, когда объект манипуляции, верит, что все происходящее естественно и неизбежно, и сам факт манипуляции не отражен в его памяти.

Такое воздействие на подсознание требует значительного мастерства и знаний. Поскольку манипуляция общественным сознанием стала технологией, существуют профессиональные работники, владеющие этой технологией или какой-то ее частью, которые относятся к людям не как к личностям, а как к объектам, особого рода вещам. С увеличением

количества информации, циркулирующей в информационном пространстве, для осуществления таких суггестивных воздействий и противодействий им требуется большое количество специалистов и автоматизация процессов для повышения их оперативности и увеличения сферы применения.

Отсюда актуальной научно-прикладной задачей является автоматизированное выявление в текстовой информации суггестивных воздействий на подсознание человека, нейтрализация таких негативных воздействий и составление информационных материалов с заданным видом суггестивного воздействия.

Анализ последних исследований и публикаций. Для анализа текстовой информации сегодня разработан и продолжается разрабатываться целый ряд подходов и методов [1–6], реализованных на основе программных средств анализа и лингвистической обработки текстов [7].

Большинство методов реализуют подходы к анализу семантической структуры текста и его логической сегментации [1–4]:

– статистический подход для анализа позволяет получить информацию о структуре текста только на основе вхождения в него отдельных слов, ключевыми из которых будут считаться те слова, число который в тексте выше заданного количества;

– семантические методы позволяют определить предметное содержание теста, его тематическую направленность, а также определить связи между отдельными частями текста и текста в целом. Семантически связанными считаются такие предложения либо абзацы, в которых есть одинаковые ключевые слова либо слова с одинаковым значением;

– лингвистические подходы основаны на синтаксических и морфологических методах. Они позволяют привести текстовые формы слов документа к словарным формам;

– контент-анализ позволяет определить частоту появления в тексте определенных характеристик, которые интересуют исследователя, а также позволяет делать некоторые выводы о намерениях создателя этого текста либо возможных реакциях адресата.

Данные походы реализованы во множестве программных средств [7]. Наибольшую популярность получила система TextAnalyst (http://www.analyst.ru), которая позволяет построить семантическую сеть понятий, выделенных в обрабатываемом тексте, со ссылками на контекст. Присутствует функция смыслового поиска фрагментов текста с учетом спрятанных в тексте смысловых связей со словами запроса. Представлены возможности анализа текста путем построения иерархического дерева тем (подтем), которые рассматриваются в тексте, и реферирования документа.

Альтернативный подход к анализу текстовой информации был предложен А.П. Журавлевым. Данный подход основан на определении

фонетического значения слов русского языка (семантического дифференциала) [5, 6]. В своих работах он представил экспериментальные данные лингвистической теории содержательности звуковой формы в русском языке.

Реализуя подходы обнаружения и анализа фонетических значений слов с использованием семантического дифференциала, можно провести анализ текстовых документов и выступлений, определить их направленность и осуществить корректирование соответственно заданным характеристикам влияния [8]. В целом, реализация технологии анализа текстов и выступлений позволяет оценивать “степень подготовленности” к эффективному восприятию и скрытую направленность информационно-психологического влияния.

Данный подход был реализован в российском программном комплексе ВААЛ (http://www.vaal.ru), который позволяет прогнозировать эффект неосознаваемого влияния текстов на массовую аудиторию, анализировать тексты с точки зрения такого влияния, составлять тексты с заданным вектором влияния и выявлять индивидуальные психологические качества авторов текста. Одним из самых больших недостатков системы ВААЛ является отсутствие описания ее математической базы и запретом поставки на экспорт и для коммерческого использования данного продукта в полном объеме. Поэтому необходимо создавать собственную информационно-аналитическую систему комплексного анализа текстовых документов, позволяющую определить степень суггестивного влияния на подсознание человека.

В [8–10] в формализованном виде рассмотрена система информационно-психологического противоборства и предложен теоретический подход к созданию системы комплексного анализа воздействия информации на подсознание человека. Данный теоретический подход позволяет определить в тексте документа отрезки текста, которые соответствуют определенным тематикам, а также выделить из них ключевые компоненты с выделением суггестивной направленности текста в целом. В работах [11–12] были предложены некоторые отдельные методы выявления суггестивных воздействий на подсознание человека, которые являются составными элементами системы комплексного анализа.

Целью исследований является разработка методов выявления суггестивных воздействий на подсознание человека в отдельных словах и текстовых сообщениях в условиях информационно-психологического противоборства на основе семантического дифференциала, фонетического и звукоцветового анализа.

1. Подходы к анализу слов для определения суггестивной направленности воздействия на подсознание человека

Метод анализа слов на основе семантического дифференциала

Семантический дифференциал – это метод анализа слов и текстовых документов на основе определения признакового аспекта по 25-ти биполярными шкалам [5, 6]. Каждая шкала представляет собой пару антонимов. В таблице 1 приведен перечень 25-ти шкал, по которым определяется семантический дифференциал.

Шкала отклонений семантического дифференциала представлена на рис. 1. Центральное значение шкалы – 3,0. Это нейтральное значение, которое не может выделить ни один признаковый аспект, так же нейтральной зоной считается зона от 2,5 до 3,5, значения в этих пределах считаются незначительными колебаниями. Все что выходит за пределы колебаний, можно считать отклонением от нормы. Именно зоны существенного отклонения говорят нам о том, к какому признаковому аспекту можно отнести слово.

Рис. 1. Шкала отклонений семантического дифференциала

Оценки являются вероятностными, то есть, подтверждены случайными колебаниям. При этом сами признаковые аспекты не стоит согласовывать со значением слова. Это обусловлено тем, что оценка дается по содержательности звуковой формы, а не по значению слова.

Предполагается, что в русском и украинском языках произносятся все буквы в слове. Однако сами буквы при написании не учитывают всех психологически важных особенностей звуков. Одна буква напрямую не может отразить мягкую и твердую согласную. Хотя сочетанием уже двух букв (текущей и последующей за ней) эта особенность уже может быть учтена. Поэтому обрабатываются не сами буквы, а именно звукобуквы, учитывающие особенности произношения. Всего в русском языке 46 звукобукв.

Определить семантическую составляющую для слова можно несколькими способами. Первым и более простым вариантом расчета является определение средней значимости всех звуков слова.

Таблица 1

25 пар антонимов для расчета семантического дифференциала

№ шкалы	Признаковый аспект		№ шкалы	Признаковый аспект	
	антоним 1	антоним 2		антоним 1	антоним 2
1	Хороший	Плохой	14	Веселый	Грустный
2	Большой	Маленький	15	Безопасный	Страшный
3	Нежный	Грубый	16	Величественный	Низменный
4	Женственный	Мужественный	17	Яркий	Тусклый
5	Светлый	Темный	18	Округлый	Угловатый
6	Активный	Пассивный	19	Радостный	Печальный
7	Простой	Сложный	20	Громкий	Тихий
8	Сильный	Слабый	21	Длинный	Короткий
9	Горячий	Холодный	22	Храбрый	Трусливый
10	Быстрый	Медленный	23	Добрый	Злой
11	Красивый	Отталкивающий	24	Могучий	Хилый
12	Гладкий	Шероховатый	25	Подвижный	Медлительный
13	Легкий	Тяжелый			

Определение средней значимости всех звуков слова проводится по формуле:

$$F = \frac{f_1 + f_2 + \ldots + f_n}{n}, \qquad (1)$$

где f_n – фонетическое значение отдельного звука (буквы) в слове;

n – количество звуков (букв) в слове.

Фонетическое значение звуков – установленная вероятностная величина. На основе проведения экспериментов, эта величина может быть изменена на более приемлемую. Фонетическое значение для каждого звука устанавливается отдельно в зависимости от шкалы, по которой будет произведен анализ. В табл. 2 приведена часть фонетических значений звукобукв для трёх шкал.

Но данный подход определения семантической составляющей для слова по формуле (1) не дает точного представления о значимости слова, так как не все звуки в слове равноправны. Психологи считают, что для человека первый звук в слове имеет куда большее значение, чем остальные, по их утверждению он в 4 раза заметнее. Выделяется в слове и ударный звук, хотя и не так как первый, только в 2 раза. Это говорит о том,

что при расчете суммарной фонетической составляющей всех звуков слова, вес первого звука нужно увеличить в 4 раза, а ударного в 2 раза.

Таблица 2

Фрагмент таблицы фонетического значения звукобукв

Признаковая шкала для расчета семантического дифференциала	А	Б	В	Г	Д
хороший – плохой	1,5	2,4	2,9	3,2	2,4
светлый – темный	2,2	3,2	3,0	3,3	3,2
красивый – отталкивающий	2,0	2,6	3,0	2,8	2,4

Но помимо расположения букв в слове, не менее важную роль играет встречаемость буквы в словах. Редко встречаемые буквы (например, Ф, Х) более заметны в словах, чем часто встречаемые (например, А, О, Т, Н). Следовательно, при расчете значимости слова, нужно брать во внимание встречаемость букв в словах. Коэффициент встречаемости (частотность) определяется как количество раз, которое буква встречается, на тысячу звукобукв. Звукобуквы в свою очередь еще делятся на ударные и безударные.

Отсюда следует, что информативность (заметность) звука находится в обратной зависимости от его частотности (встречаемости). То есть наименее информативный звук с максимальной частотностью, а остальные во столько раз информативнее, во сколько раз их частотность меньше максимальной для звуков данного слова.

Следовательно, при расчете фонетической составляющей звукового комплекса нужно увеличить вес средних оценок не только для первого и ударного звуков, но также и для всех звуков, кроме звука с максимальной частотой. Иначе говоря, необходимо сначала дописать каждому звуку (букве) свой вес в зависимости от положения в слове, а только после этого вычислять среднее арифметическое.

Таким образом, определение коэффициента каждого звука в слове проводится по формуле:

$$k_i = \frac{P_{max}}{P_i}, \qquad (2)$$

где k_i – коэффициент i-го звука в слове;

P_{max} – максимальная частотность звука в данном слове;

P_i – табличное значение частотности звукобуквы.

Исходя из вышеизложенного для первого звука необходимо увеличить коэффициент в 4 раза, т.е. выражение (2) примет вид

$$k_1 = 4k_i = 4\frac{P_{max}}{P_i}, \qquad (3)$$

а для ударного звука необходимо увеличить коэффициент в 2 раза, т.е. выражение (2) примет вид:

$$k_{уд} = 2k_i = 2\frac{P_{max}}{P_i}. \quad (4)$$

С учетом коэффициентов каждого звука, рассчитанных на основе выражений (2)–(4), для расчета фонетической составляющей слова используется следующая формула:

$$F = \frac{\sum_{i=1}^{n} f_i k_i}{\sum_{i=1}^{n} k_i}, \quad (5)$$

где F – фонетическая составляющая слова;

f_i – фонетическое значение кожного i-го звука (буквы) слова;

k_i – коэффициент для каждого i-го звука (звука);

n – количество звуков (букв) в слове.

Расчет семантического дифференциала для слова предлагается производить на основе следующих этапов (последовательности действий):

1) минимальной единицей для расчетов выступает слово, так как важны не только буквы, а и место их расположения в слове, что позволяет из набора букв получить звукобуквы. Поэтому на первом (подготовительном) этапе анализа производится перевод слова или последовательности букв в набор звукобукв;

2) на втором этапе для каждой звукобуквы выбираются табличные значение их частотности и значимости исходя из анализируемого признакового аспекта. Для этого используются 2 таблицы. Первая – таблица частотности звукобукв, вторая – таблица значимости звукобукв. Из таблицы частотности всегда выбираются одинаковые значение независимо от признака, по которому производится анализ. Таблица значимости имеет 25-ть разных наборов для каждой звукобуквы в зависимости от признакового аспекта;

3) на третьем этапе рассчитываются коэффициенты каждого звука в слове по формуле (2);

4) после чего полученные коэффициенты корректируются в зависимости от заметности букв в слове с помощью формул (3) для первого буквы слова и (4) для ударного звука. При этом, определение ударного звука в слове может быть реализовано с помощью специально сформированных словарей;

5) на пятом этапе происходит увеличение фонетической значимости для каждой звукобуквы в зависимости от коэффициента, для чего каждое фонетическое значение, полученное на втором этапе, перемножаем на соответствующий коэффициент;

6) финальным этапом является расчет значения семантического дифференциала для слова по формуле (5).

Метод фонетического анализа семантической составляющей слова

Фонетический анализ текста чем-то напоминает семантический дифференциал, результат сводится к выводу оценки слова по шкалам [5, 6]. Но в данном методе оценка производится по 20-ти однополярным шкалам в отличие от семантического дифференциала с 25-ю биполярными шкалами. Шкалы в свою очередь представляют 20-ть различных признаков, в таблице 3 представлен перечень этих признаков. Такие признаковые шкалы лучше подходят для характеристики какого-либо текста.

Таблица 3

20 признаков для фонетического анализа

№ признака	Признак для анализа	№ признака	Признак для анализа	№ признака	Признак для анализа	№ признака	Признак для анализа
1	прекрасный	6	печальный	11	тоскливый	16	возвышенный
2	бодрый	7	яркий	12	радостный	17	медлительный
3	светлый	8	темный	13	стремительный	18	тихий
4	нежный	9	сильный	14	угрюмый	19	суровый
5	минорный	10	устрашающий	15	тяжелый	20	зловещий

Методика анализа базируется на том, что человек привык в разговорной речи к некой частотности звуков и как установили психологи он определяет эту частотность довольно правильно. Соответственно, какое либо значительное отклонение от этой частотности должно быть замечено подсознанием человека. Определив, какие звуки преобладают в тексте и, дав им некоторые признаковые описания, можно судить о том какое психоэмоциональное воздействие окажет тот или иной текст на человека.

Отклонение значимости слова от среднего значения делает его выразительным или склонным к тому или иному признаку. Поэтому для оценки значимости слова необходима таблица отклонений значимости звуков. Фрагмент такой таблицы отклонений значимости звукобукв приведен в таблице 4. Значимость слова можно получить путем вычитания отклонений значимости звукобуквы от среднего значения шкалы 3,0:

$$k_{знач} = 3.0 - k_i,$$

где k_i – значение отклонения значимости i-ой звукобукв, определенной на основе таблицы.

Например, для буквы А по шкале “хороший – плохой” это будет $3.0-1.5=+1.5$. Это означает что отклонение уходит ближе к признаку “хороший”. Отрицательное же значение, будет значить, что отклонение происходит в сторону признака “плохой”.

Таблица 4

Фрагмент таблицы отклонения звукобукв

Признаковая шкала для фонетического анализа	Буква				
	А	Б	В	Г	Д
хороший – плохой	+1,5	+0,6	+0,1	-0,8	+0,6
светлый – темный	+0,8	-0,8	0,0	-0.7	-0,8
красивый – отталкивающий	+1,0	+0,4	0,0	+0,2	+0,6

Расчет фонетической составляющей для слова основан на следующих этапах:

1) первый подготовительный этап заключается в переводе слова или же последовательности букв в набор звукобукв. В отличии от семантического дифференциала, последовательность звукобукв не важна;

2) на втором этапе для каждой звукобуквы выбираются табличные значение их частотности и значимости в соответствии с анализируемым признаковым аспектом. Для этого используются 2 таблицы. Первая – таблица частотности звукобукв $f_{норм.}$, из которой всегда выбираются одинаковые значение независимо от признака по которому производится анализ. Вторая – таблица отклонения значимости звукобукв от нормы, которая имеет 20-ть разных наборов для каждой звукобуквы в зависимости от признакового аспекта;

3) на третьем этапе проводится определения частоты вхождения каждой звукобуквы в слово $f_{тек.}$ на основе подсчета общего числа звукобукв в анализируемом слове и количества каждой звукобуквы и операции деления числа вхождения звукобуквы на общее число звукобукв в слове:

$$f_{тек.} = \frac{n_i}{n_{общ.}},$$

где $n_{общ.}$ – количество звукобукв в слове;

n_i – частотность i-ой звукобуквы в слове;

5) следующим этапом является определение отклонения частотности звукобукв от нормы. При этом, нормальная частотность показывает, сколько раз должна встретиться определенная буква в обычном тексте. Но, как правило, если взять несколько различных текстов, то частотность не

будет точно совпадать с табличным значением. Из этого следует, что нормальная частотность подвержена колебаниям. Границы колебаний определяются по теории вероятностей. За единицу, при измерении размаха колебаний принимают величину σ. В теории вероятностей считается, что нормальные колебания какой-либо случайно величины не должны превышать $\pm 2\sigma$. Пока величина колеблется в этих приделах, можно считать, что она как бы "привязана" к средней точке колебания и далеко от этой точки не отклонится. Но если колебания превысят $\pm 2\sigma$, значит, они ненормальны. Поэтому для расчета отклонения частотности звукобукв от нормы будет применяться следующая формула:

$$k_{\text{отклон.}} = \frac{f_{\text{тек.}} - f_{\text{норм.}}}{\sqrt{f_{\text{норм.}}(1 - f_{\text{норм.}})} / n_{\text{общ.}}},$$

где $f_{\text{тек.}}$ – частота звукобуквы в анализируемом слове;

$f_{\text{норм.}}$ – нормальная частота звукобуквы в речи;

6) завершающим этапом расчета фонетической значимости слова является суммарное значение вкладов каждой звукобуквы в общий звуковой тон текста. Вклад звукобувы – это перемножение величины отклонения частотности от нормы на величину отклонения значимости от нейтральной точки, но только тех звукобукв, чье отклонение существенно от нормы:

$$F = \sum_{i=1}^{n_{\text{общ.}}} k_{\text{отклон.i}} k_{\text{знач.i}} = \sum_{i=1}^{n_{\text{общ.}}} k_{\text{отклон.i}} (3 - k_i) =$$

$$= \sum_{i=1}^{n_{\text{общ.}}} \frac{f_{\text{тек.i}} - f_{\text{норм.i}}}{\sqrt{f_{\text{норм.i}}(1 - f_{\text{норм.i}})} / n_{\text{общ.}}} (3 - k_i).$$

Метод звукоцветового анализа семантической составляющей слова

Как показывают исследования, звуки, которые используются в речи, у большинства людей ассоциируются с определенным цветом или же оттенком [5, 6, 13]. Но надо знать точное соответствие между звуками и цветом. Ведь каждый человек может воспринимать звуки по разному. Вполне точно мнения сходятся по поводу трех гласных: А – красная, Е – зеленая, И – синяя. Несколько разнообразно описывается буква О как бело-желтую. Это и не удивительно, что самые точные оценки даются основным опорным буквам языка (по мнению лингвистов). По итогам проведенных экспериментов, была составлена таблица наиболее единогласных результатов (табл. 5) [5, 6].

Из анализа данных в табл. 5 видно, что буквы Ё, Я, Ю, Й не имеют основных цветов, а связываются лишь с их оттенками и встречаются

довольно редко. Поэтому предложено их приплюсовывать к основным гласным Е, А, У, И соответственно.

Таблица 5

Цветовые ассоциации звуков

Звукобуква	Цветовая ассоциация
А	густо-красный, темно-красный
Я	ярко-красный, алый
О	светло-жёлтый, белый
Е	зеленый
Ё	желто-зеленый
И	синий
Й	синий, оттенки синего
У	тёмно-синий, тёмный сине-зелёный, тёмно-лиловый
Ю	голубой, оттенки циан
Ы	темный, темно-коричневый, чёрный
Э	серая, не ясная буква, исключена из анализа

Эксперимент показал, что с согласными буквами работа усложнилась, а результаты не дали однозначных данных. Следовательно, было принято решение, что согласным буквам не соответствуют какие либо определенные цвета и ими можно пренебречь, а данные не включать в таблицу.

Исходя из того, что между гласными буквами (звуками) и цветом есть некое соответствие, то можно “подсветить” текст или слово, то есть дать ему некую окраску, что в конечном итоге увеличит эмоциональное восприятие от прочитанного. Для определения такой цветовой картины, необходимо в слове подсчитать количество только тех букв, для которых присвоены цвета. Для определения частотности каждой буквы в слове также необходимо знать общее число букв в слове. Полученные частотности необходимо сопоставить с нормальными (среднестатистический показатель для языка) и вычислить нормированные разности этих частотностей, чтобы установить, случайно или нет частотности отличаются от нормальных и как именно отличаются.

Расчет звукоцветовой оценки для слова предполагает следующую последовательность действий:

1) первым этапом является подсчет количества букв в анализируемом слове (тексте);

2) далее нужно подсчитать количество отдельных букв в слове (тексте), у которых есть свой цвет по таблице соответствий букв и цветов (табл. 5);

3) подсчитав количество букв, необходимо определить частотность тех букв, которые соответствуют таблице путем деления их количества на общее количество букв в слове (тексте);

4) далее необходимо сопоставить полученные частотности с нормальными долями букв в текстах русского (украинского) языка. Для этого делим текущую частоту вхождение буквы в слове (тексте) на нормальную частоту;

5) имея результаты отношений долей звукобукв к нормальным значения, можем определить какие звукобуквы превышают норму. Выстраиваем значение по убыванию и отсекаем те значения, которые ниже необходимого порога;

6) теперь имея звукобуквы, которые превышают норму в слове (тексте), можем по данным из таблицы 5 определить цвет слова, сопоставив буквы с цветами.

Данный метод в данном материале отдельно для статического и динамического анализа текстовых документов рассматривать не будем в силу его узконаправленности и необходимости дополнительных исследований, хотя все ниже изложенные подходы к анализу текстов на основе методов тестирования для выявления суггестивных воздействий характерны и для него.

2. Подходы к анализу текстов на основе методов тестирования для выявления суггестивных воздействий

Подходы к статическому анализу текста для выявления суггестивных воздействий

Методы выявления суггестии ориентированы на анализ слов, что дает понимание, каким образом оно воспринимается человеком. Но они могут быть адаптированы и для анализа всего текста в целом. Первый подход заключается в анализе каждого слова в отдельности и определения среднего значения для всех слов. Недостатком является то, что анализируется каждое отдельное слово, а конечный результат не связан с рядом стоящими словами. А это говорит о том, что любой текст, составленный из этого набора слов, будет иметь одинаковую оценку. Но ведь используя одни и те же слова, можно составить текст абсолютно по-разному и с разным посланием.

Следовательно, необходимо каким-то образом зафиксировать слова текста в том порядке, в котором их расположил автор и никак иначе. Это даст уникальную оценку именно для такого упорядоченного набора слов.

Поэтому второй подход предполагает для такой явной фиксации либо же зависимости слов приведения этого набора отдельных слов в единую, неразрывную строку и проведение анализа этой полученной строки, как единого слова.

Каждый из этих подходов имеет свои плюсы и минусы, поэтому стоит рассматривать оба варианта. При этом анализ, во время которого анализируется весь текст целиком будем называть статическим.

Рассмотрим подход к статическому анализу текста по словам. Он предполагает следующие этапы:

1) поскольку анализ производится по словам, то анализируемый текст в данном случае необходимо разбить на отдельные слова;

2) каждое слово в отдельности нужно проанализировать тем методом выявления суггестивного воздействия, который выбран для анализа текста: семантический дифференциал, фонетический или звукоцветовой анализ;

3) после получения всех оценок по каждому слову, выводится средняя оценка для всех результатов путем вычисления среднего арифметического. Таким образом, мы имеем оценку, которая является средним значением, для каждого слова из которого составлен текст.

Далее рассмотрим подход, при котором анализ текста производится как единое целое, который состоит из следующих этапов:

1) весь анализируемый текст переводится в одну неразрывную строку, как очень длинное слово путем откидывания пробелов, знаков препинания, перевода чисел в текстовую форму написания. Таким образом, связывается конец одного слова с началом следующего;

2) полученная строка анализируется требуемым методом (семантический дифференциал, фонетический или звукоцветовой анализ), по аналогии с анализом одного слова. Это даст уникальную оценку для такой последовательности слов в текстовом документе.

Подходы к динамическому анализу текста для выявления суггестивных воздействий

Исходя из того, что статический анализ всего текста дает нам представление того, какое воздействие он окажет на человека при полном его прочтении, следует то, что мы получаем некие усредненные показатели для всех частей этого текста. То есть, нельзя сказать, что каждая из частей текста в отдельности имеет такой же показатель воздействия, как и весь текст целиком. Можно предположить, что текст построен таким образом, что каждая его часть, имеет свое отличительное воздействие на подсознание. Это становиться важным фактом в том случае, когда человек читает либо слышит только отдельную часть текста. В таком случае статический анализ всего текста будет не информативен, ведь всего один абзац может оказывать сильное воздействие, в то время как большая часть текста будет нейтральной, что в конечном итоге смажет результаты анализа. Более эффективным решением такое задачи может стать динамический анализ.

Динамический анализ подразумевает под собой разбиение текста на части и анализ каждой из этих частей в отдельности. Таким образом, можно проследить динамику изменения суггестивного воздействия на человека от начала и до конца текста или же оценить только конкретный его участок. Такой анализ так же может показать, какие части текста

имеют смысл урезать, а какие оставить, при необходимости сокращения текста. Таким подходом могут пользоваться рекламные компании, сокращая эфирное время рекламы, но, не урезая ее необходимого воздействия на подсознание.

Как правило, большинство текстов состоит из абзацев, как промежуточной единицей между фразой и главой, что служит в свою очередь для группировки однородных единиц изложения. Поэтому можно сделать вывод, что динамику лучше прослеживать по абзацам, так как обычно абзацем выражается некая общая мысль. Но бывает и так, что текст не разбит на абзацы, либо они не удовлетворяют нашему представлению о размере единого блока для анализа. В таком случае, текст может быть разбит на некие установленные заранее блоки текста. Это может быть либо разбиение по количеству символов на блок, либо же слов или предложений. Минусом такого разбиения будет то, что блоки текста, могут разрывать, связанные общей мыслю части текста либо предложения.

Анализ текста динамическим методом предполагает следующую последовательность этапов:

1) исходный текст необходимо разбить на некие блоки текста, которые требуется проанализировать. Удобнее и целесообразнее всего разбивать по абзацам;

2) необходимо определить какой блок текста будет анализироваться;

3) поскольку этот подход по своей сути является статическим, за исключением того, что статически анализируется не весь текст, а выбранная его часть, то на данном этапе применяется один из вариантов статического подхода к анализу выбранного фрагмента;

4) при необходимости можно вернуться к пункту 2 алгоритма и выбрать иной участок текста для повторного анализа. Таким образом, на основе динамического метода можно дать представление о том какими суггестивными воздействиями обладает каждая часть текста.

Обоснование подходов динамического анализа текстов накопительным итогом для выявления суггестивных воздействий

Как показывает статистика, около половины людей не дочитывают книги до конца [14]. Причиной этому может быть либо отсутствие времени, но также и отсутствие заинтересованности, либо же отсутствие неких эмоций, которые ожидались во время прочтения. Это ставит перед нами задачу, определить, как же можно построить текст так, чтоб он вызвал заинтересованность у читающего и не оттолкнул его на протяжении чтения. Для такого рода задачи, казалось бы, хорошо подходит динамический метод анализа, он даст представление о каждом участке текста. Но если взять во внимание тот факт, что читая текст от начала, человек не воспринимает каждый отрывок или абзац как несвязанный отрывок, а как бы накапливает ту информацию, которую

получил сейчас, с той, которая была абзацем ранее. То можно сделать вывод, что по мере освоения, каждая последующая полученная информация, будет дополнять уже имеющуюся информацию. А это значит то, что при анализе последующей части текста, необходимо так же учитывать и предыдущие его части, так как каждый последующий отрезок будет дополнять общую картину представления либо влияния на сознание человека.

Подход к динамическому анализу текста накопительным итогом предполагает следующие основные этапы:

1) исходный текст необходимо разбить на некие блоки текста, что позволит, двигаясь последовательно по блокам, анализировать накопленную информацию, так же как бы это делал человек читая текст. Удобнее и целесообразнее всего разбивать по абзацам;

2) поскольку данный подход подразумевает под собой то, что человек начиная читать текст, двигается от начальной точки чтения далее по тексту, тем самым пополняя информацию о прочитанном. То необходимо определить это самое начало, от которого следует производить анализ и накопление информации. Для этого предлагается выбрать блок текста, который будет проанализирован первым. Это не обязательно начало текста, ведь человек может начать чтение и не с самого начала, а, например, со второй главы;

3) этот подход хоть и отличается от динамического тем что анализирует блоки текста путем их накопления, но весь накопленный текст все так же анализируется с помощью статического метода. Так же как и в динамическом подходе, можно применить один из вариантов статического анализа по словам или строкой. Алгоритм от это не изменится;

4) для определения, какое суггестивное воздействие произведет следующий участок текста с уже накопленными предыдущими, необходимо к накопленным блокам текста, добавить следующий и вернуться к пункту 3 алгоритма произведя повторный анализ с вновь накопленной информацией.

3. Методы динамического анализа текстов для выявления суггестивной направленности

Метод анализа текстов на основе семантического дифференциала накопительным итогом

Метод анализа текстов на основе семантического дифференциала накопительным итогом предполагает следующие этапы:

1) исходный текст разбивается на необходимые блоки;

2) определяется первый блок текста, с которого будет производиться накопление информации;

3) выбранный блок текста представляется в виде одного слова;

4) для каждой буквы слова рассчитывается коэффициент по формуле (2);

5) полученные коэффициенты корректируются в зависимости от заметности букв в слове используя формул (3) и (4);

6) увеличивается фонетическая значимость для каждой звукобуквы на основе умножения каждого фонетического значения звукобуквы на соответствующий коэффициент;

7) проводится расчет значения семантического дифференциала по формуле (5);

8) после анализа всех блоков накопленного текста, необходимо подсчитать их среднее значение. Результатом будет значение семантического дифференциала накопленного текста;

9) при необходимости можно добавить еще один блок текста к уже накопленному, перейти к этапу 3 и продолжить анализ.

Метод фонетического анализа семантической составляющей накопительным итогом

Метод фонетического анализа семантической составляющей текста накопительным итогом предполагает следующие этапы:

1) исходный текст разбивается на необходимые блоки;

2) определяется первый блок текста, с которого будет производиться накопление;

3) следующим этапом является анализ накопленного текста методом фонетического анализа семантической составляющей по словам. Для этого определяется фонетическое значение для каждого слова из накопленного текста;

4) для каждой буквы слова подсчитывается общее число звукобукв и подсчитывается количество каждой звукобуквы. Это необходимо для определения частоты вхождения каждой звукобуквы в слово. На основе полученных данных определяется частота путем деления числа вхождения звукобуквы на общее число звукобукв в слове;

5) следующим этапом является определение отклонения частотности звукобукв от нормы;

6) далее необходимо посчитать вклад каждой звукобуквы в общий тон текста;

7) завершающим этапом расчета фонетической значимости слова является суммарное значение вкладов каждой звукобуквы в общий звуковой тон текста;

8) после анализа всех слов накопленного текста, необходимо посчитать среднее значение для всех значений слов. Результатом будет фонетическое значение накопленного текста;

9) при необходимости можно добавить еще один блок текста к уже накопленному и произвести анализ повторно.

4. Метод составления текста с заданной суггестивной направленностью контекста

На основе методов выявление суггестивного воздействия можно составить текст, таким образом, чтоб он имел заданное направление воздействия на подсознание человека. Но на составление такого текста может уйти довольно большое количество времени, если не использовать автоматизированные системы. Поскольку направленность текста зависит от использованных в текстах слов, то сам принцип задания или же корректировки направленности текста сводится к некому набору слов, которые лучше всего подходят под требуемую направленность. То есть можно простой заменой слов, как правило, это подбор синонимов, изменить суггестивное воздействие текста.

Самой важной частью такого составление текста, является таблица синонимов слов, которые могут быть использованы вместо имеющихся слов в тексте. Но сами синонимы мало что смогут нам сказать, так как, не зная, что даст каждая замена слова, придется постоянно производить анализ текста на предмет изменение его направленности. Таким образом, следует то, что в пару к синонимам, нам еще понадобится добавить значение суггестивного воздействия для каждого из синонимов. Это даст нам, либо машине, понятие того, что каждое из слов может принести в общую картину текста.

Исходя из данного предложения можно предложить такой алгоритм изменения суггестивного воздействия текста:

1) текст, который подвергается корректировке, необходимо разбить на слова, поскольку каждое из слов теоретически может быть заменено на более подходящее;

2) после разбиения на слова, необходимо выполнить проверку по базе знаний на наличие более удовлетворяющих поставленному требованию слов. Для этого необходимо из базы знаний получить все возможные синонимы для текущего слова. Все слова уже должны иметь свои коэффициенты суггестивного воздействия, которые необходимы для принятия решения в пользу того или другого синонима;

3) перебирая все синонимы для текущего слова, необходимо проверять какой из синонимов имеет большее воздействие в заданном направлении, нежели имеющееся слово;

4) на основе полученного числа удовлетворяемых слов, нужно произвести выборку именно того слова, которое лучше всего впишется в текст, а также даст лучший результат по направленности воздействия на подсознание;

5) после анализа и замены всех необходимых слов, можно произвести комплексный анализ всего текста, на предмет достижения необходимого результата, а в случае если результат не достигнут, можно повторно произвести замену.

Данный алгоритм хорошо подходит для автоматизированного анализа, поскольку участие человека в этом процессе необходимо как минимум для принятия решения, какое же слово лучше всего удовлетворяет требованием для замены. Машина, конечно, может выбрать наилучшее слово по максимальному коэффициенту, либо же, некоторым случайным образом из набора удовлетворяющих слов. Но результат такой выборки будет крайне непрактичным и текст может стать нечитабельным даже для автора, который все равно должен будет скорректировать текст после машины.

Для решения задачи автоматизированной корректировки текста, можно предложить следующий подход.

Поскольку вся сложность сводится именно к корректной выборке слова для замены, то именно на этом и стоит сосредоточить внимание. Необходимо решить, что является показателем успешного выполнения задачи. Первое, это конечно же то суггестивное воздействие, которого мы хотим добиться, ведь ради этого и производится анализ. Вторым, но не менее важным является читаемость текста, а именно не правильность форм слова (это может сделать и человек после корректировки машины, либо же машина с дополнительной базой знаний), а значение слова. Значение слова, которое будет подобрано для замены очень важно, так как поставив слово, которое не подходит под контекст текста, можно либо изменить смысл фразы, либо же это приведет к хаотичному набору слов. Если для первого данные у нас уже имеются, их всего лишь стоит дать машине, то для второго показателя, таких данных нет. Нужно дать машине понять какое слово можно поставить, а какое будет в контексте неуместным.

Для многих случайных процессов характерно некоторое влияние предшествующих событий на последующие. Такие процессы называют марковскими. В таких процессах вероятность находится в данном состоянии в данный момент можно вывести из сведений о предшествующем состоянии. Такой подход используется на данный момент в так называемых генераторах текста, которые используют цепи Маркова. Цепью Маркова первого порядка называется одна из форм марковских процессов, для которой каждое конкретное состояние зависит только от непосредственно предшествующего. Для такой цепи число состояний конечно, а вероятности, соответствующие переходам из одного состояния в другое, называют стационарными, имея в виду то, что они не зависят от времени. Из этого следует, что можно определить необходимое слово по той базе знаний, которая может дать вероятность появления его после некоторых слов.

Говоря другими словами, для того чтоб определить можно ли поставить слово на место другого, нужно знать связь этого слова со стоящим перед ним словом. А именно вероятность их появления вместе. Если вероятность появления такой пары слов достаточно высока, то можно

предположить, что данное слово подходит больше, нежели слово с меньшей вероятностью. Так как человеку будет более привычное сочетание слов, которое он чаще встречает.

Такую вероятность появления слова можно определить только путем анализа текстов, что даст базу знаний, основанную на проанализированном материале. Из этого следует, что вероятность появление слова – величина вероятностная. А поскольку зависит она от конкретного проанализированного материала, то можно создать базу для специализированных или узконаправленных текстов, что даст лучший подбор слов для замены.

Сформировать такую базу знаний можно на основе следующего подхода. В исходном тексте определяются слова, и сохраняется последовательность, какие слова следуют за какими. Затем на основании этих данных строятся связи слов. А именно собирается группа слов, которая может располагаться после текущего слова и так для каждого слова в тексте. Но в отличии от генераторов текста, где каждое последующее слово подбирается в зависимости от того какое слово стоит перед ним, в данном случае, важно то какое слово мы заменяем, а именно в первую очередь подбирается набор возможных синонимов, а только после этого определяется, имеют ли они связь с впереди стоящим словом. Поэтому более удобной записью построение связей будет обратная связь. Это означает, что необходимо собрать набор слов, перед которыми может располагаться данное слово. Для слова, которое может быть заменено, существует таблица синонимов с соответствующими коэффициентами для требуемого анализа. Каждый из синонимов содержит в свою очередь свою таблицу связей со словами, после которых может встречаться данное слово, со значением, показывающим частоту встречаемости.

Таким образом, имея такую базу знаний, принятие решения о выборе подходящего синонима автоматическим способом можно осуществить на основе следующих этапов:

1) синонимы, которые были отобраны как более подходящая замена для текущего слова по их суггестивному показателю, еще не могу в полной мере служить полноценной заменой, так как их значения могут быть недостаточно подходящими под контекст предложения. Поэтому для избегания подобной ситуации каждый синоним имеет таблицу связей со словами, а именно с теми словами, перед которыми это слово будет корректно применить, ну или, по крайней мере, такое сочетание слов встречалось ранее и дальнейшее использование подобного словосочетание не должно вызвать непонимание читающего текст. С помощью этих связей необходимо сначала проверить имеет ли слово связь с впередистоящим словом в тексте и за отсутствием таковой отсеять это слово из числа возможных синонимов на замену;

2) вторым этапом является непосредственно сама выборка подходящего слова для замены. Для этого необходимо найти синоним, связь которого с впередистоящим словом имеет набольшую частоту встречаемости. Это значит, что данная связка слов в текстах встречается чаще остальных комбинаций. Хотя для подобной выборки можно оставить некую зону колебаний, для того чтоб результат был несколько разнообразен.

Выводы

Разработанные методы могут использоваться для решения большого круга задач с целью выявления суггестивных воздействий на подсознание человека в отдельных словах, текстах (документах) целиком и их разных структурных (составных) элементах. Реализация данных методов позволит:

– оценивать эмоциональное влияние отдельных слов на подсознание человека;

– оценивать эмоциональное влияние фонетической структуры текстов на подсознание человека;

– оценивать уровень агрессивности текстов на основе анализа позитивного и негативного влияний отдельных слов на содержательное значение текста в целом;

– оценивать звукоцветовые характеристики текстов;

– задавать характеристики желаемого влияния и создавать (корректировать) структуры соответствующей направленности;

– настраивать тексты на лексически определенные социальные и профессиональные группы людей.

Литература

1. Герасимов Б.М. Извлечение информационных фраз из первичных электронных документов в информационно-поисковых системах / Б.М. Герасимов, О.Ю. Сергеев, И.Ю. Субач // Управляющие системы и машины. - 2006. - №1. - С. 26 - 29.

2. Рыбаков Ф.И. Автоматическое индексирование на естественном языке / Ф.И. Рыбаков, Е.А. Руднев, В.А. Петухов. – М.: Энергия, 1980. – 160 с.

3. Скороходько Е.Ф. Лінгвістичні основи автоматизації інформаційного пошуку / Е.Ф. Скороходько. – К.: Вища школа, 1970. – 242 с.

4. Сэлтон Г.А. Автоматическая обработка, хранение и поиск информации / Г.А. Сэлтон. – М.: Сов. радио, 1973. – 560 с.

5. Журавлев А.П. Фонетическое значение / А.П. Журавлев. – Л.: ЛГУ, 1974.

6. Журавлев А.П. Звук и смысл : кн. для внеклас. чтения учащихся ст. классов / А.П. Журавлев. – 2-е изд. испр. и доп. - М. : Просвещение, 1991. – 160 с.

7. Программы анализа и лингвистической обработки текстов [Електронний ресурс]. – Режим доступу: http://www.rvb.ru/soft/catalogue/index.html.

8. Сидченко С.А. Система анализа воздействия информации на подсознание человека в условиях информационно-психологического противоборства / С.А. Сидченко, К.И. Хударковский, В.Л. Петров // Теорія та методика навчання

математики, фізики, інформатики: Збірник наукових праць. Випуск V: В 3-х томах. – Кривий Ріг: Видавничий відділ НМетАУ, 2005. – Т.3: Теорія та методика навчання інформатики. – С. 303 - 314.

9. Сідченко С.О. Методика комплексного аналізу документу / С.О. Сідченко, С.В. Залкін, В.В. Бєлімов // Системи обробки інформації. – 2007. – Вип. 9 (67). – С. 109 - 113.

10. Sidchenko S. A. Method of complex information and psychological document analysis / S. A. Sidchenko , T. V. Saprykina // Наукоємні технології. – 2014. – № 1 (21). – С. 79 – 83.

11. Сидченко С.А. Тестирование семантической составляющей для выявления суггестивного воздействия / С.А. Сидченко, Т.В. Сапрыкина, В.А. Школяренко // Автоматизированные системы управления и приборы автоматики. – 2013. – Вип. 165. – С. 111 – 117.

12. Сидченко С.А. Метод составления текста с заданной суггестивной направленностью контекста / С.А. Сидченко, Т.В. Сапрыкина, В.А. Школяренко // Системи обробки інформації. – Х.: ХУПС. – 2014. – Вип. 4 (120). – С. 96 – 101.

13. Скрябин А. Цветная музыка стиха / А. Скрябин [Электронный ресурс]. – Режим доступа: http://ru.wikipedia.org/wiki/Скрябин,_Александр_Николаевич.

14. Самые читающие страны мира : справка [Электронный ресурс]. – Режим доступа: http://ria.ru/spravka/20080611/110842173.html.

ФОРМАЛІЗАЦІЯ ВИЗНАЧЕННЯ РІВНЯ ГАРАНТІЙ АВТОМАТИЗОВАНОЇ СИСТЕМИ ВІД НЕСАНКЦІОНОВАНОГО ДОСТУПУ

Бучик С.С., Юдін О.К., Нетребко Р.В.

Вступ

Стрімке зростання новітніх технологій, а також розвиток інфраструктури інформаційно-комунікаційних мереж державного та загального призначення призвело до створення інтегрованого інформаційного простору держави та всього суспільства. Інформаційні технології знаходять усе ширше застосування в таких сферах, як: державні системи управління, фінансовий обіг і ринок цінних паперів, розвинута система електронних платежів, система послуг зв'язку та телебачення, системи управління транспортом, високотехнологічні виробництва (особливо атомні, хімічні тощо) і т. ін. Будь-яке несанкціоноване та протиправне втручання в інформаційний простір наведених сфер життєдіяльності держави й суспільства може призвести до тяжких та не передбачуваних наслідків [1].

Досліджуючи нормативно-правову базу (НПБ) України в галузі захисту автоматизованих систем (АС) від несанкціонованого доступу (НСД), слід відмітити наступні документи: НД ТЗІ 2.5-004-99 «Критерії оцінки захищеності інформації в комп'ютерних системах від несанкціонованого доступу» затверджено наказом Департаменту спеціальних телекомунікаційних систем та захисту інформації Служби безпеки України від 28.04.1999 р. № 22 [2] із змінами згідно наказу Адміністрації Держспецзв'язку від 28.12.2012 № 806 та НД ТЗІ 2.7-010-09 «Методичні вказівки з оцінювання рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації від несанкціонованого доступу» затверджено наказом Адміністрації Державної служби спеціального зв'язку та захисту інформації України від 24 липня 2009 року № 172 [3]. Саме дані документи стали основою для проведення аналізу та подальшого дослідження. В розрізі продовження дослідження є актуальним здійснити аналіз теоретичних основ визначення рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації автоматизованих систем від несанкціонованого доступу, що дозволить автоматизувати процес визначення та зменшить витрати часу та матеріальних (людських) ресурсів, які витрачаються на такий процес.

Аналіз останніх досліджень і публікацій. Аналіз останніх досліджень і публікацій показав, що питання розвитку систем захисту, їх створення, організації та дослідження процесів їх функціонування відображенні в працях вітчизняних і закордонних вчених, серед яких Горбенко І.Д., Корченко О.Г., Задірака В.К., Конахович Г.Ф.,

Грайворонський М.В., Новіков О.М., Шаньгин В.Ф. і багато інших. Дані праці показують основні теоретичні положення з захисту інформації, методологічні та науково-теоретичні основи побудови систем захисту, оцінки їх ефективності та принципів вибору параметрів для оцінки ефективності. Тематиці визначення рівня гарантій АС від несанкціонованого доступу (НСД) присвячено небагато робіт, що на думку авторів пов'язано з тим, що процес визначення здійснюється експертною комісією та фактично єдиним документом, якій визначає підхід до визначення гарантій є НД ТЗІ 2.7-010-09. Дана робота є продовженням роботи теоретичні основи визначення стандартних ФПЗ на основі нормативно-правової бази (НПБ) [4] та тематики робіт над якими працює група науковців. Тому наразі стоїть питання можливості оцінки запропонованого рівня гарантій автоматизовано на основі документів НД ТЗІ 2.7-010-09 [3] та НД ТЗІ 2.5-004-99 [2]. Раніше на основі роботи [4] авторами була розроблена інформаційна система визначення стандартного (нестандартного) ФПЗ [5], але в даній інформаційній системі не здійснюється визначення рівня гарантій. Тому актуальним стоїть питання впровадження автоматизованої оцінки рівня гарантій.

Основний матеріал

Автоматизована система являє собою організаційно-технічну систему, що об'єднує обчислювальну систему, фізичне середовище, персонал і оброблювану інформацію. Прийнято розрізняти два основних напрями технічного захисту інформації (ТЗІ) в АС – це захист АС і оброблюваної інформації від несанкціонованого доступу і захист інформації від витоку технічними каналами [6]. Забезпечення безпеки інформації АС розглядається як набір функціональних послуг. Кожна послуга являє собою набір функцій, що дозволяють протистояти деякій множині загроз. Тому оцінка захищеності АС визначається з рівня реалізації послуг.

Крім функціональних критеріїв, що дозволяють оцінити наявність послуг безпеки в АС, критерії містять рівні гарантій, які дозволяють оцінити коректність реалізації послуг. Рівні гарантій включають вимоги до архітектури комплексу засобів захисту (КЗЗ), середовища розробки, послідовності розробки, випробування КЗЗ, середовища функціонування і експлуатаційної документації (рис. 1). Вводиться сім рівнів гарантій, ці рівні є ієрархічними. Ієрархія рівнів гарантій відбиває поступово наростаючу міру упевненості в тому, що послуги, які надаються, дозволяють протистояти певним загрозам, а механізми, що їх реалізують, в свою чергу, коректно реалізовані, і можуть забезпечити очікуваний споживачем рівень захищеності інформації під час експлуатації АС [2].

Гарантії забезпечуються як в процесі розробки, так і в процесі оцінки. В процесі розробки гарантії забезпечуються діями розробника

щодо забезпечення коректності розробки. В процесі оцінки гарантії забезпечуються шляхом перевірки додержання розробником вимог, аналізу документації, процедур розробки і постачання, а також іншими діями експертів, які проводять оцінку.

Критерії гарантій, розглядаються в нормативному документі НД ТЗІ 2.5-004-99, містять в собі вимоги до архітектури КЗЗ, середовища розробки, послідовності розробки, середовища функціонування, експлуатаційної документації та випробувань КЗЗ. Згідно з вимогами цього документа, оцінювання відповідності реалізованих засобів та заходів захисту встановленим вимогам та нормам здійснюється шляхом проведення експертизи [3].

Рис. 1. Вимоги критерій гарантій

Розглянемо вимоги до рівня гарантій більш детально. Оцінювання рівня гарантій функціональних послуг безпеки (ФПБ) передбачає виконання таких дій:

– першим етапом є ознайомлення з оцінювальною АС, збирання та аналіз документів, що характеризують організацію процесу розроблення, виробництва та постачання замовнику оцінюваної АС;

– другим етапом є розроблення програми перевірки дотримання вимог до рівня гарантій у процесі розроблення, виробництва та постачання замовнику оцінюваної АС;

– третім етапом є розроблення методики перевірки дотримання вимог до рівня гарантій у процесі розроблення, виробництва та постачання замовнику оцінюваної АС;

– четвертим є виконання оцінювання рівня гарантій згідно з розробленими програмою та методикою;

– п’ятим є аналіз та документування результатів оцінювання рівня гарантій [3].

Тому виходячи з даного порядку можна представити роботу експерта при подані на експертизу розробником або заявником АС разом з проектною, супровідною та експлуатаційною документацією та з визначеними в проектній документації функціональними специфікаціями АС наступною діаграмою (рис. 2).

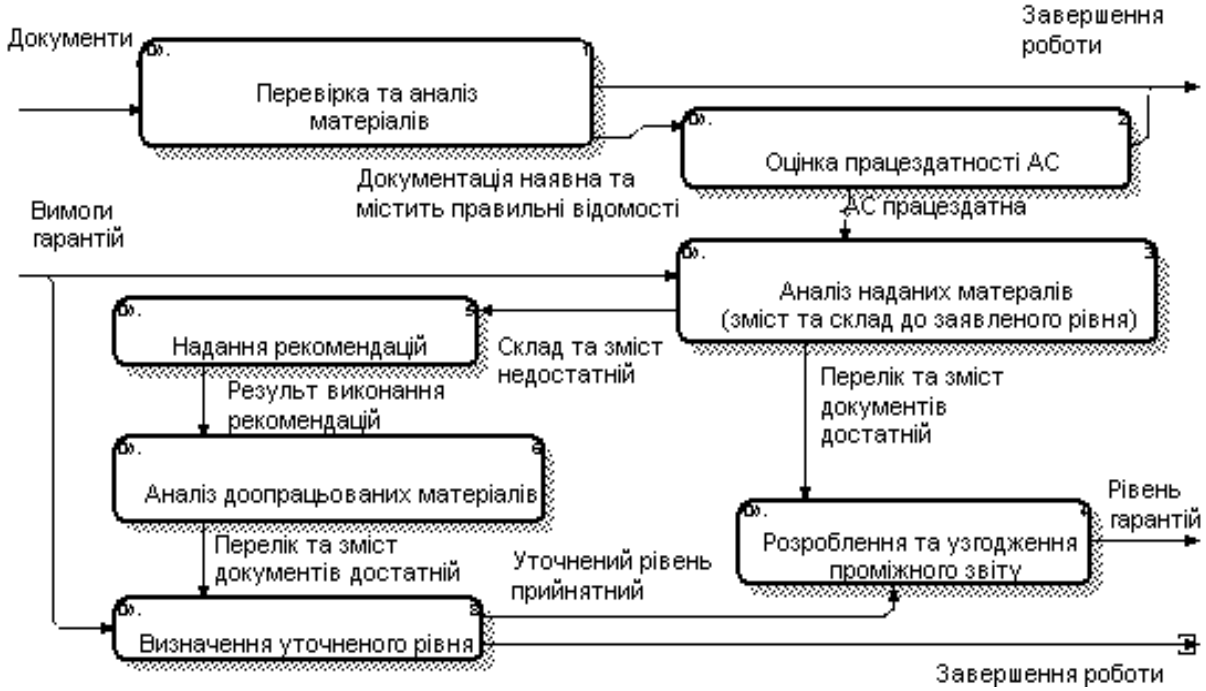


Рис. 2. Порядок роботи експерта

Виконуючи перший етап експерти перевіряють документацію на наявність функціональних специфікацій АС, якщо вони відсутні, експерт оцінює систему до вимог рівня гарантій Г-1.

Робота експерта відповідно до діаграми, яка представлена на рис. 2, полягає в наступному:

– проводиться перевірка факту надання експлуатаційної документації на АС та наявності в ній відомостей, необхідних для виконання подальших робіт з оцінювання працездатності АС. У випадку відсутності документації подальші роботи мають припинятися;

– якщо перший процес відбувся успішно проводиться оцінювання працездатності наданого АС та визначення його придатності для виконання подальших робіт. У випадку виявлення непрацездатності АС подальші роботи мають припинятися;

– після чого проводиться аналіз, згідно з методичними рекомендаціями, наданих розробником або заявником матеріалів з метою попереднього визначення відповідності їх складу та змісту вимогам до рівня гарантій, визначеного розробником АС або заявником відповідно до заявленого рівня гарантій, та надання рекомендацій та/або пропозицій

щодо їх доопрацювання та/або надання додаткових матеріалів у випадку виявлення певних невідповідностей;

– при деяких невідповідностях що аналізувались у попередньому процесі проводиться повторний аналіз складу та змісту доопрацьованих або додатково наданих розробником (заявником) матеріалів з метою визначення їх достатності для проведення подальших перевірок на відповідність вимогам до заявленого рівня гарантій або необхідності зниження заявленого рівня гарантій до такого, який випливає зі складу та змісту наданих матеріалів;

– потім відбувається прийняття рішення про прийнятність уточненого рівня гарантій, відповідність вимогам до якого має бути підтверджено в процесі експертизи, та про продовження робіт;

– якщо даний рівень гарантій вибраний проводиться документування отриманих результатів у погодженому з розробником (заявником) проміжному звіті з обов'язковою фіксацією уточненого рівня гарантій як такого, відповідність вимогам до якого має бути підтверджено в процесі експертизи [3].

Основною роботою експерта є збір та оцінка всіх необхідних документів перелік яких представлено в табл.1. Де позначення "+" – вимога до змісту документа з'являється або підвищується; "=" – вимога до змісту документа зберігається [3].

Таблиця 1

Рекомендований склад матеріалів (документів)

№ з/п	Тип документа	Заявлений рівень гарантій коректності реалізації ФПБ						
		Г-1	Г-2	Г-3	Г-4	Г-5	Г-6	Г-7
1	Технічне завдання (що містить функціональні специфікації КЗЗ ОЕ)	+	+	+	+	=	=	=
2	Ескізний проект (що містить проект архітектури)	+	=	+	=	+	+	=
3	Технічний проект (що містить детальний проект)	+	+	=	+	=	=	+
4	Робочий проект (реалізація)			+	=	+	=	+
5	Опис результатів аналізу відповідності між політикою безпеки та моделлю політики безпеки КЗЗ ОЕ		+	=	+	=	=	=
6	Опис результатів аналізу відповідності між моделлю політики безпеки КЗЗ ОЕ та проектом архітектури		+	=	=	+	+	=
7	Опис результатів аналізу відповідності між проектом архітектури та детальним проектом		+	=	=	=	+	+
8	Опис результатів аналізу відповідності між детальним проектом та реалізацією			+	=	=	=	+

Продовження таблиці 1

№ з/п	Тип документа	Заявлений рівень гарантій коректності реалізації ФПБ						
		Г-1	Г-2	Г-3	Г-4	Г-5	Г-6	Г-7
9	Опис методик діяльності розробника протягом життєвого циклу ОЕ	+	=	=	=	=	=	=
10	Документація використовуваних при розробленні інструментальних засобів			+	=	=	=	=
11	Опис методик забезпечення безпеки в процесі розроблення та виробництва ОЕ				+	=	+	=
12	Документація з керування конфігурацією ОЕ	+	=	=	+	=	=	=
13	Опис процедур безпечної інсталяції, генерації та запуску ОЕ	+	=	=	=	=	=	=
14	Опис процедур постачання ОЕ замовнику			+	=	=	+	=
15	Опис послуг безпеки, що реалізуються КЗЗ оцінюваного ОЕ	+	=	=	=	=	=	=
16	Настанови адміністратору з послуг безпеки	+	=	=	=	=	=	=
17	Настанови користувачу з послуг безпеки	+	=	=	=	=	=	=
18	Програма та методика випробувань функціональних послуг безпеки	+	=	=	=	=	=	=
19	Протоколи випробувань функціональних послуг безпеки	+	+	=	=	=	=	=
20	Опис результатів аналізу стійкості КЗЗ до атак з боку розробника				+	=	+	=

Вимоги до складу наданих документів визначенні залежно від заявленого рівня гарантій в АС. Назви документів не є обов’язковими, допускається об’єднання схожих за змістом документів в один.

Рекомендації щодо змісту документів, які надаються експерту викладені в нормативному документі НД ТЗІ 2.7-010-09 «Методичні вказівки з оцінювання рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації від несанкціонованого доступу» в додатку А.

Відповідно до переліку та змісту представлених документів визначається подальше опрацювання документів та визначення вимог критеріїв гарантій. Вимоги представлені в нормативному документі НД ТЗІ 2.5-004-99. Критерії гарантій включають вимоги до архітектури КЗЗ, середовища розробки, послідовності розробки, середовища функціонування, документації і випробувань КЗЗ. Виходячи з даних

критеріїв можна представити коефіцієнт гарантії позначенням $-K^{gar}$. Вимоги до архітектури представимо наступним коефіцієнтом – K^{arh}, середовища розробки – K^{sr}, послідовності розробки $-K^{pr}$, середовища функціонування $-K^{sf}$, документації – K^{d} і випробувань $-K^{v}$.

Введення даних коефіцієнтів дає можливість створити загальний математичний опис вимог гарантій, представлений наступною множиною:

$$K^{gar} = \{K^{arh}, K^{sr}, K^{pr}, K^{sf}, K^{d}, K^{v}\}. \quad (1)$$

Для того щоб АС одержала певний рівень гарантій, повинні бути задоволені всі вимоги, визначені для заданого рівня в кожному з розділів документа [3]. Розглянемо визначення вимог гарантій на прикладі архітектури табл. 2.

Таблиця 2

Вимоги до архітектури КЗЗ

Вимоги	Г1	Г2	Г3	Г4	Г5	Г6	Г7
КЗЗ повинен реалізовувати політику безпеки. Всі його компоненти повинні бути чітко визначені	+	=	=	=	=	=	=
КЗЗ повинен складатися з добре визначених і максимально незалежних компонентів. Кожний з компонентів повинен бути спроектований відповідно до принципу мінімуму повноважень	-	-	+	=	=	=	=
Критичні для безпеки компоненти КЗЗ повинні бути захищені від не критичних для безпеки за рахунок використання механізмів захисту, які надаються програмно-апаратними засобами більш низького рівня	-	-	-	+	=	=	=
З боку Розробника мають бути вжиті зусилля, спрямовані на виключення з КЗЗ компонентів, що не є критичними для безпеки. Мають бути наведені підстави для включення до КЗЗ будь-якого елемента, який не має відношення до захисту	-	-	-	-	+	=	=
Розробка ПЗ переважно має бути спрямована на мінімізацію складності КЗЗ. КЗЗ має бути спроектований і структурований так, щоб використовувати повний і концептуально простий механізм захисту з точно визначеною семантикою. Цей механізм повинен відігравати центральну роль в реалізації внутрішньої структури КЗЗ. Під час розробки КЗЗ значною мірою повинні бути задіяні такі підходи як модульність побудови і приховання (локалізація) даних	-	-	-	-	+	=	=

В таблиці використовуються такі позначення: “-” – вимога відсутня; “+” – вимога з’являється; “=“ – вимога зберігається [2].

Введемо позначення вимог перейшовши до умовних позначень. Умови, які стосуються вимог архітектури, отримали буквене позначення "a" з цифровим індексом (1...5), до середовища розробки отримали буквене позначення "r" з цифровим індексом (1...7), послідовності розробки отримали буквене позначення "pr" з цифровим індексом (1...5), середовища функціонування отримали буквене позначення "sf" з цифровим індексом (1...3), документації отримали буквене позначення "d" з цифровим індексом (1...5) і випробувань отримали буквене позначення "v" з цифровим індексом (1...4).

Після введення умовних позначень, табл. 2 буде мати наступний вигляд результат представлений в табл. 3.

Таблиця 3

Вимоги до архітектури КЗЗ (з умовними позначеннями)

Архітектура (вимоги)	Умовні позначення
КЗЗ повинен реалізовувати політику безпеки. Всі його компоненти повинні бути чітко визначені	a_1
КЗЗ повинен складатися з добре визначених і максимально незалежних компонентів. Кожний з компонентів повинен бути спроектований відповідно до принципу мінімуму повноважень	a_2
Критичні для безпеки компоненти КЗЗ повинні бути захищені від не критичних для безпеки за рахунок використання механізмів захисту, які надаються програмно-апаратними засобами більш низького рівня	a_3
З боку Розробника мають бути вжиті зусилля, спрямовані на виключення з КЗЗ компонентів, що не є критичними для безпеки. Мають бути наведені підстави для включення до КЗЗ будь-якого елемента, який не має відношення до захисту	a_4
Розробка ПЗ переважно має бути спрямована на мінімізацію складності КЗЗ. КЗЗ має бути спроектований і структурований так, щоб використовувати повний і концептуально простий механізм захисту з точно визначеною семантикою. Цей механізм повинен відігравати центральну роль в реалізації внутрішньої структури КЗЗ. Під час розробки КЗЗ значною мірою повинні бути задіяні такі підходи як модульність побудови і приховання (локалізація) даних	a_5

Підставивши умовні позначення слід створити матрицю знань, яку потім можна буде використати при створенні програмного продукту, який визначатиме рівень гарантій.

Матрицю знань створено таким чином: вимогам, які виконуються у визначеній АС надаємо значення 1, а тим умовам, які не виконуються – 0. Для вимог архітектури відповідно до семи рівнів гарантій матриця буде мати наступний вигляд (табл. 4).

Таблиця 4

Матриця знань

K^{arh}, Рівень гарантій	a_1	a_2	a_3	a_4	a_5
$K^{arh}{}_1$	1	0	0	0	0
$K^{arh}{}_2$	1	0	0	0	0
$K^{arh}{}_3$	1	1	0	0	0
$K^{arh}{}_4$	1	1	1	0	0
$K^{arh}{}_5$	1	1	1	1	1
$K^{arh}{}_6$	1	1	1	1	1
$K^{arh}{}_7$	1	1	1	1	1

Відповідно до матриці (табл. 4) отримаємо наступні логічні рівняння.

$$K^{arh}{}_1 = K^{arh}{}_2 = a_1, \tag{2}$$

$$K^{arh}{}_3 = a_1 \wedge a_2, \tag{3}$$

$$K^{arh}{}_4 = a_1 \wedge a_2 \wedge a_3, \tag{4}$$

$$K^{arh}{}_5 = K^{arh}{}_6 = K^{arh}{}_7 = a_1 \wedge a_2 \wedge a_3 \wedge a_4 \wedge a_5. \tag{5}$$

Якщо один із наведених вище логічних виразів (2-5) приймає значення "1", то обирають відповідну вимогу, якщо значення "0" – вимога відкидається.

Отож, дані рівняння відображають формалізацію основ до вимог архітектури по визначенню рівня гарантій.

Відповідно до НД ТЗІ 2.7-004-99 формалізуємо усі представлені вимоги до наступних логічних рівнянь.

Вимоги архітектури K^{arh} визначаються за допомогою наступних рівнянь.

$$K^{arh} = \begin{cases} K^{arh}{}_1 = K^{arh}{}_2 = a & (6) \\ K^{arh}{}_3 = a_1 \wedge a_2 & (7) \\ K^{arh}{}_4 = a_1 \wedge a_2 \wedge a_3 & (8) \\ K^{arh}{}_5 = K^{arh}{}_6 = K^{arh}{}_7 = a_1 \wedge a_2 \wedge a_3 \wedge a_4 \wedge a_5. & (9) \end{cases}$$

Вимоги середовища розробки K^{sr} визначаються за допомогою наступних рівнянь.

$$К^{sr} = \begin{cases} К^{sr}{}_1 = К^{sr}{}_2 = r_1 \wedge r_4 & (10) \\ К^{sr}{}_3 = r_1 \wedge r_2 \wedge r_4 & (11) \\ К^{sr}{}_4 = К^{sr}{}_5 = r_1 \wedge r_2 \wedge r_3 \wedge r_4 \wedge r_5 \wedge r_6 & (12) \\ К^{sr}{}_6 = К^{sr}{}_7 = r_1 \wedge r_2 \wedge r_3 \wedge r_4 \wedge r_5 \wedge r_6 \wedge r_7 . & (13) \end{cases}$$

Вимоги послідовність розробки $К^{pr}$ визначаються за допомогою наступних рівнянь.

$$К^{pr} = \begin{cases} К^{pr}{}_1 = pr_1 \wedge pr_3 \wedge pr_4 & (14) \\ К^{pr}{}_2 = К^{sr}{}_3 = К^{sr}{}_4 = К^{sr}{}_5 = К^{sr}{}_6 = К^{sr}{}_7 = pr_1 \wedge pr_2 \wedge pr_3 \wedge pr_4 & (15) \end{cases}$$

Вимоги послідовність розробки K^{sf} визначаються за допомогою наступних рівнянь.

$$К^{sf} = \begin{cases} К^{sf}{}_1 = К^{sf}{}_2 = sf_1 & (16) \\ К^{sf}{}_3 = К^{sf}{}_4 = К^{sf}{}_5 = sf_1 \wedge sf_2 . & (17) \\ К^{sf}{}_6 = К^{sf}{}_7 = sf_1 \wedge sf_2 \wedge sf_3 & (18) \end{cases}$$

Вимоги документації $К^{d}$ визначаються за допомогою наступних рівнянь.

$$К^{d} = К^{d}{}_1 = К^{d}{}_2 = К^{sf}{}_3 = К^{sf}{}_4 = К^{sf}{}_5 = К^{sf}{}_6 = \quad (19)$$
$$= К^{sf}{}_7 = d_1 \wedge d_2 \wedge d_3 \wedge d_4 \wedge d_5$$

Вимоги випробування КЗЗ $К^{v}$ визначаються за допомогою наступних рівнянь.

$$К^{v} = \begin{cases} К^{v}{}_1 = v_1 \wedge v_2 & (20) \\ К^{v}{}_2 = К^{v}{}_3 = v_1 \wedge v_2 \wedge v_3 & (21) \\ К^{v}{}_4 = К^{v}{}_5 = К^{v}{}_6 = К^{v}{}_7 = v_1 \wedge v_2 \wedge v_3 \wedge v_4 & (22) \end{cases} .$$

Визначивши вимоги гарантій та їх математичний опис, з’являється можливість визначити один із рівнів гарантій.

Основні результати. Основними результатами автори вважають здійснення формалізації визначення рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації автоматизованих систем від несанкціонованого доступу. Запропоновані авторами теоретичні основи дають можливість в подальшому розробити експертну систему, яка допоможе експертам визначати гарантії безпеки автоматизованої системи від несанкціонованого доступу автоматизовано.

Висновки

Таким чином, у роботі запропоновано, показано та проаналізовано теоретичні основи визначення рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації АС від НСД. Висвітлено необхідні НД ТЗІ, які регламентують порядок оцінки та визначення рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації АС від НСД в Україні. На основі нормативних документів, здійснено формалізацію визначення (узгодження) рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації АС від НСД. Запропоновані авторами теоретичні основи дають можливість в подальшому розробити експертну систему, яка визначатиме рівні гарантій АС від НСД. Це полегшить роботу експертів щодо визначення гарантій безпеки АС від НСД, зменшить витрачений на це ресурс часу. У сукупності з автоматизацією процесу визначення стандартних (нестандартних) профілів захищеності АС від НСД, реалізація викладених в статті результатів надасть можливість в подальшому удосконалити інформаційну систему розроблену групою авторів з урахуванням гарантій безпеки.

Література

1. Юдін О. К. Державні інформаційні ресурси. Методологія побудови класифікатора загроз: монографія / О. К. Юдін, С. С. Бучик. – К.: НАУ, 2015. – 214 с.

2. Критерії оцінки захищеності інформації в комп'ютерних системах від несанкціонованого доступу: НД ТЗІ 2.5-004-99 [Електронний ресурс]. – Режим доступу: http://www.dstszi.gov.ua/dstszi/doccatalog/document?id=41649.

3. Методичні вказівки з оцінювання рівня гарантій коректності реалізації функціональних послуг безпеки в засобах захисту інформації від несанкціонованого доступу: НД ТЗІ 2.7-010-09 [Електронний ресурс]. – Режим доступу: http://www.dsszzi.gov.ua/dsszzi/doccatalog/document?id=103247.

4. Юдін О. К. Теоретичні основи визначення стандартних функціональних профілів захищеності автоматизованої системи від несанкціонованого доступу / О. К. Юдін, С. С. Бучик, С. В. Мельник // Наукоємні технології. – 2016. – № 2 (30). – С.195 – 205, doi.org/10.18372/2310-5461.30.10564.

5. А. с. 66492 Україна. Комп'ютерна програма. Інформаційна система визначення функціонального профілю захищеності автоматизованої системи від несанкціонованого доступу / Бучик С. С., Мельник С. В. (Україна). – № 67055; заявл. 10.05.16; опубл. 28.10.16, Бюл. № 42.

6. Загальні положення щодо захисту інформації в комп'ютерних системах від несанкціонованого доступу: НД ТЗІ 1.1-002-99 [Електронний ресурс]. – Режим доступу: www.dsszzi.gov.ua/dsszzi/doccatalog/document?id=106340.

РОЗШИРЕНА КЛАСИФІКАЦІЯ КВАНТОВИХ МЕТОДІВ БЕЗПЕЧНОЇ КОМУНІКАЦІЇ

Гнатюк С.О., Жмурко Т.О., Поліщук Ю.Я., Сєйлова Н.А.

Вступ

Останні два десятиріччя бурхливо розвивається новий та неоднозначний у сприйнятті науковців мультидисциплінарний напрямок – *квантова криптографія* (КК). Як і всі напрямки науки, що змушують науковців сперечатись про їх доцільність, переваги та недоліки, КК досліджується багатьма науковими центрами та університетами, в наслідок чого з'являються нові та удосконалені протоколи, що забезпечують захист інформації, ґрунтуючись на непорушних постулатах квантової фізики, та у своїй більшості досягають теоретико-інформаційної стійкості. Однак, поява великої кількості нових протоколів та методів КК в деякій мірі значно ускладнює роботу науковців та дослідницьких центрів, оскільки їх класифікація проводиться частково та своєчасно не оновлюється, що ускладнює пошук і не дає змоги у повній мірі оцінити рівень існуючих досягнень для їх подальшого ефективного використання.

Аналіз існуючих досліджень і постановка завдання

На сьогодні існує декілька класифікацій методів та протоколів КК, проте вони або орієнтовані на узагальнення протоколів одного виду, наприклад, у роботі [1] представлена класифікація протоколів квантового розділення секрету (КРС, *quantum secret sharing*). У роботі [2] проводиться класифікація квантових протоколів шифрування з відкритим ключем (*quantum public-key encryption protocol*) за шістьма елементами кортежу та виділяють з 64 видів три основні типи протоколів. За аналогічним методом класифікують квантові протоколи симетричного шифрування (*quantum symmetric-key encryption protocols*) [3]. У [4] наведена класифікація протоколів квантової телепортації (КТ), а також представлено двосторонню КТ та двосторонній квантовий прямий безпечний зв'язок (КПБЗ). Часткова класифікація протоколів квантового розподілу ключів (КРК), КРС та систематизація деяких атак на КК проводилась у [5]. Робота [6] представляє класифікацію протоколів одного з напрямків квантової теорії ігор (КТІ). Класифікація Корченка-Василіу-Гнатюка [7] є однією з найбільш повних, однак вона є досить застарілою – не враховує протоколи КТ, КТІ тощо. Як правило науковці проводять окремо класифікацію різного роду атак на протоколи КК [8-10, 15], а найчастіше досліджують конкретний протокол і конкретну атаку на нього [9-16], тобто стійкість до певного класу атак не виділяють як окрему класифікаційну ознаку.

З огляду на це, **метою** роботи є розробка розширеної класифікації сучасних квантових методів безпечної комунікації за рахунок розширення номенклатури методів та базових ознак, а також огляд комерційних систем квантового розподілу ключів.

Основна частина

До квантових методів безпечної комунікації відносяться: КРК, КПБЗ, КРС, квантовий потоковий шифр (КПШ), квантовий цифровий підпис (КЦП), квантова стеганографія (КС), КТ та КТІ.

Квантовий розподіл ключів (*quantum key distribution*) – найбільш розвинутий та досліджений напрямок КК, який у сучасних дослідженнях [17] прийнято розділяти на:

– *DV-QKD (discrete-variable)* – КРК з дискретними змінними – як носій інформації використовується фаза фотонів/поляризація, детектором є лічильники фотонів, діапазон передачі – 100 км, в обов'язкові компоненти має входити активне охолоджуюче обладнання. Головна перевага протоколів в тому, що за відсутності помилок, Аліса і Боб (легітимні користувачі) відразу розділяють ідеальний секретний ключ. Проте головними і суттєвими недоліками є відсутність джерел одиночних фотонів і низька ефективність їх детекторів, що спричиняє пошуки інших варіантів реалізації протоколів КРК. Протоколи DV-QKD можуть бути реалізовані з декількома джерелами, але вимагають використання методів обрахунку кількості фотонів (*photon-counting*). До них відносять такі протоколи [18]: BB84, SARG04, The six-state protocol, Singapore protocol, B92, протоколи типу GV, модифікаціями якого є протоколи Koashi-Imoto та Guo-Shi, а також типу N09, до якого відносять оригінальний протокол N09 та Sun-Wen protocol.

– *CV-QKD (continuous-variable)* – КРК з безперервними змінними – як носій інформації використовуються амплітудно-фазове поле, детектор – когерентний, діапазон передачі – 25 км, компоненти використовуються стандартні. Лічильники фотонів замінені на стандартні p-i-n фотодіоди, які є більш швидкими та більш ефективними. Крім того вони використовують когерентні методи виявлення (*coherent detection techniques*), що широко використовуються в класичних оптичних комунікаціях. Згідно [19-20] вони поділяються на P&M (*Prepare-and-Measure*) та E-B (*Entanglement-base*).

– *DFS-QKD (differential phase shift).* На відміну від DV-QKD та CV-QKD протоколів замість того, щоб бути закодованими в кубіти, ключова інформація кодується в фазу послідовних імпульсів слабкого когерентного світла. У протоколі DPS, Аліса кодує логічні біти у фази імпульсів. Якщо фаза модулюється «0», то Аліса посилає логічний нуль, а якщо фаза між двома імпульсами π, то вона кодує логічну одиницю. Якщо відносна фаза між двома імпульсами дорівнює «0», то Боб виявить «0», і аналогічно, якщо фаза між двома імпульсами π, то він отримає логічну «1».

Історично перший протокол КРК – *BB84* [21], запропонований у 1894 році Ч. Беннетом та Ж. Брассаром, основними задачами BB84 є генерація та розподіл ключів шифрування між двома абонентами, що з'єднані квантовим та класичним каналами зв'язку, У протоколі BB84 використовуються 4 поляризовані стани фотонів (0°, 45°, 90°, 135°), які

передаються квантовим каналом зв'язку. Пошук та виправлення помилок виконується з використанням відкритого класичного каналу, який не повинен бути конфіденційним, тільки аутентифікованим. Для виявлення факту дій зловмисника використовується процедура контролю помилок, а для забезпечення безумовної стійкості використовується класична процедура підсилення секретності (*privacy amplification*).

«Six states» протокол [22] передбачає використання чотирьох станів, аналогічних протоколу BB84, і додатково вводяться ще два можливих напрямки поляризації – правоциркулярний та лівоциркулярний. Такі зміни з одного боку зменшують кількість інформації, що може бути отримана зловмисником, а з іншого боку ефективність протоколу також зменшується (до 33%). Також запропоновано узагальнення протоколу з шістьома станами на багаторівневі квантові системи. Даний протокол має дещо більшу інформаційну місткість та значно більшу стійкість до атаки «перехоплення – повторної посилки» кудитів.

Протокол 4+2 [23] є перехідним між BB84 та B92. У ньому використовуються чотири квантових стани для кодування: «0» та «1» у двох базисах. Стани в кожному базисі вибираються неортогональними, крім того, стани в різних базисах також мають бути попарно неортогональними. Для протоколу 4+2 характерна менша кількість помилок відносно протоколу BB84 для кубітів і менша кількість корисної інформації, що може отримати зловмисник, але одночасно відбувається й зменшення відносної ефективності протоколу.

У протоколі *GV* [24] кодування «0» та «1» виконується за допомогою двох ортогональних станів. Кожен з цих двох станів є суперпозицією двох локалізованих нормалізованих хвильових пакетів. Для захисту проти атаки «перехоплення – повторної посилки» використовується випадковий час відправлення пакетів. Модифікований варіант протоколу Гольденберга-Вайдмана – це протокол *Koashi-Imoto* [25], удосконалений тим, що замість випадкового часу відправлення пакетів використовується асиметризація інтерферометра, тобто світло розбивається у нерівних пропорціях між довгим і коротким плечами інтерферометра.

Протокол *B92* [26] з використанням потужних імпульсів, але зловмисник може одержати більше інформації про ключ для заданого рівня створюваних їм помилок, ніж у протоколі BB84, тобто стійкість протоколу B92 нижче стійкості протоколу BB84. Ефективність протоколу становить 25%.

Протокол E91 [27], відноситься до КРК з використанням переплутаних станів. Під час передавання інформації за протоколом E91 перехоплення одного із фотонів пари не дає зловмиснику ніякої корисної інформації. Крім того, запропоновано узагальнення схеми Екерта на тривимірні та багатовимірні квантові системи, що значно збільшує інформаційну місткість протоколу.

Протоколи зі станами «приманки» (decoy states protocols) є удосконаленим варіантом протоколу BB84, у якому відправник, шляхом заміни підмножини імпульсів, вводить так звані приманки [28]. Даному типу протоколів характерний більш високий рівень безпеки, ніж у BB84. Крім того, такі протоколи відзначаються стійкістю проти PNS атаки. До явних переваг протоколів зі станами «приманки» також можна віднести і збільшення довжини каналу за рахунок лінійної залежності від втрат у каналі. Проте, без попередньої аутентифікації користувачів на таких протоколах не можливо побудувати завершене повноцінне рішення проблеми розподілення криптографічних ключів.

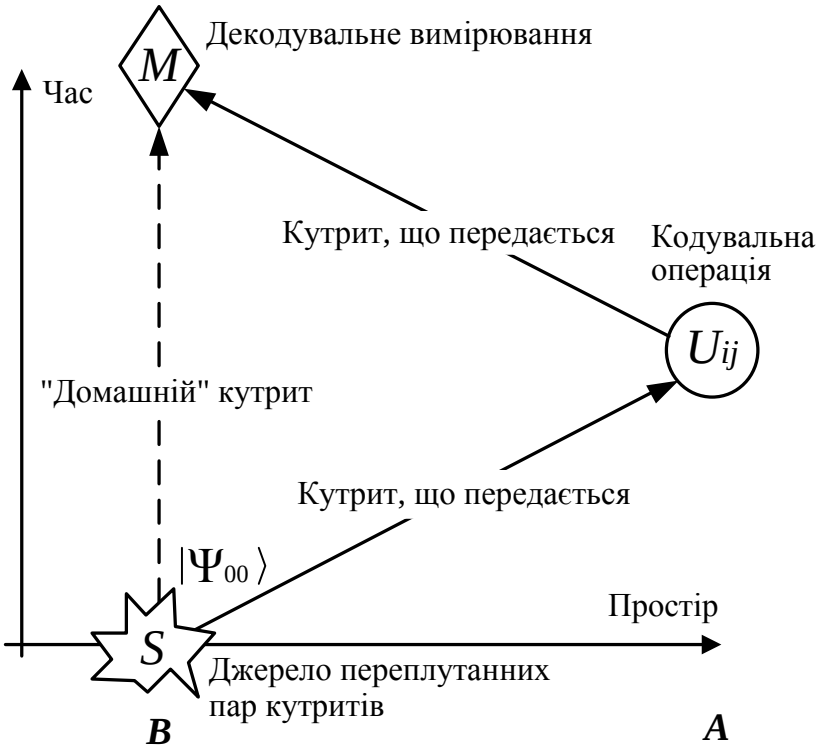
SARG04 protocol [29] на рівні квантової частини протокол еквівалентний BB84, проте замість джерел одиночних фотонів використовуються послаблені лазерні імпульси. Стійкий до PNS атаки.

COW (coherent one-way) protocol [30] – це новий протокол для практичної квантової криптографії, розроблений Н. Жізаном (N. Gisin) та ін. в 2004 р. У протоколі Аліса (передавач) посилає пару імпульсів, один порожній і один не порожній (містить середнє число фотонів 0,5). Біти кодуються в парі імпульсів, з значенням біту визначають по положенню не порожнього імпульсу: перше положення «0» і друге положення «1». Боб, приймач, використовує детектор, щоб розрізняти імпульси. Аліса і Боб також перевіряють узгодженість імпульсів. Боб випадковим чином вибирає невелику частину імпульсів, що не використовується як дані, щоб відправити на інтерферометр, який вимірює когерентність між суміжними кубітами. У зв'язку з цим заходом безпеки, зловмисник не може виконати PNS-атаку, так як видалення або блокування фотонів, неможливе без порушення системи. На відміну від інших протоколів інтерферометр використовується тільки для оцінки інформації на когерентність і не може призвести до помилок на ключі. Протокол не використовує посимвольний тип кодування (як BB84, B92, SARG04).

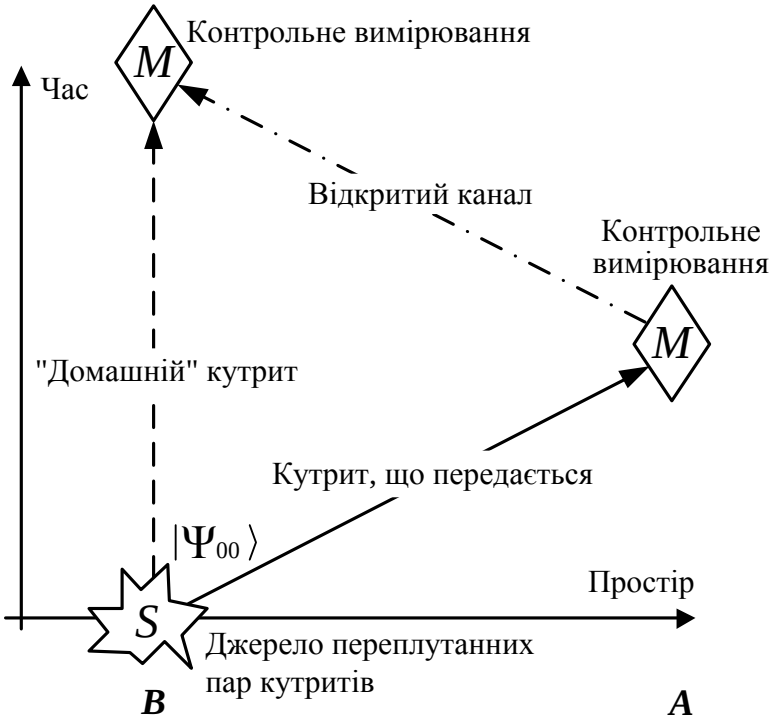
KMB09 [31] (розроблений Khan M.M., Murphy M., Beige A. у 2009 р.) – протокол у якому Аліса і Боб використовують два взаємно неупереджені базиси – один із них кодує «0», а інший кодує «1». Безпека схеми обумовлено швидкістю передачі мінімального індексу помилки (ITER) і квантового рівня помилок, що вноситься зловмисником. ITER значно збільшується для більш високих розмірностей фотонних станів. Це дозволяє мати більше шуму в лінії передачі, тим самим збільшуючи можливу відстань між Алісою та Бобом без необхідності проміжних вузлів.

Квантовий прямий безпечний зв'язок [32-34, 51-53] характерною особливістю даного методу є відсутність криптографічних перетворень, відповідно відсутня і проблема розподілу ключів шифрування. Протоколи КПБЗ можна поділити на такі типи: детерміністичний протокол Бострома-Фелбінгера (також відомий як пінг-понг (PP) протокол, рис. 1 [54]) та різні

його варіанти [33, 35, 51-53], протоколи з передаванням переплутаних кубітів блоками [36], протоколи з одиничними кубітами та протоколи з групами переплутаних кудитів.



а)



б)

Рис. 1. Схематичне відображення роботи пінг-понг протоколу з парами переплутаних кутритів у режимах: передавання повідомлення (а), контролю підслуховування (б)

Більшість запропонованих до теперішнього часу протоколів КПБЗ потребують передачі кубітів блоками. Це дозволяє виявити прослуховування квантового каналу до початку передачі самого повідомлення й таким способом гарантувати безпеку передачі – якщо прослуховування виявлене до передачі повідомлення, то легітимні сторони переривають сеанс і ніяка інформація не витікає до зловмисника. Але для зберігання таких блоків кубітів необхідна квантова пам'ять великого об'єму. Технологія квантової пам'яті активно розробляється, але поки ще далека від масового застосування в стандартному телекомунікаційному устаткуванні. Тому з погляду технічної реалізації перевагу мають протоколи, у яких передача здійснюється одиничними кубітами або невеликими їхніми групами (за один цикл протоколу). Таких протоколів запропоновано небагато, і вони мають тільки асимптотичну безпеку, тобто атака буде виявлена з високою ймовірністю, але до цього зловмисник зможе одержати деяку частину повідомлення. Отже, виникає проблема підсилення безпеки таких протоколів, тобто створення таких методів попередньої обробки передаваної інформації, які зроблять перехоплену зловмисником інформацію для нього некорисною.

Квантове розділення секрету. Переважна частина квантових протоколів розділення секрету (КПРС) використовує властивості переплутаних квантових станів [37-38]. У роботах [1, 51-53] наведено детальний огляд сучасного стану протоколів КРС та проведено їх класифікацію. Наведемо деякі відомості про КРС: у 1998 році був запропонований перший протокол КРС, який, аналогічно деяким протоколам квантового безпечного зв'язку, використовує ГХЦ-триплети (четвірки) кубітів. Цей протокол дозволяє відправнику розділити своє повідомлення між двома (трьома) абонентами таким чином, що вони зможуть його прочитати, тільки діючи спільно. *Напів-квантовий* КРС з використанням ГХЦ-триплетів (четвірок) кубітів. У цьому протоколі сторони, що приймають розділене повідомлення, мають доступ до квантового каналу, але обмежені деяким набором операцій та називаються «класичними», в тому сенсі, що вони не мають можливості готувати переплутані стани, виконувати будь-які квантові операції та вимірювання. Дані сторони можуть вимірювати кубіти у «класичному» $\{|0\rangle,|1\rangle\}$ базисі, впорядковувати їх за допомогою затримок вимірювань, приготувати кубіти у «класичному» базисі, а також відправляти або повертати кубіти без збурення їх станів. Сторона, що розділяє своє повідомлення, може виконувати будь-які квантові операції. Цей протокол має перевагу – обладнання у приймальних сторін буде дешевше, оскільки їм не потрібно дороге обладнання для створення багатокубітних переплутаних станів та вимірювань, наприклад, в ГХЦ-базисі. До напів-квантових КРС відносяться дві групи протоколів – це напів-квантові КРС, що ґрунтуються на рандомізації, та КРС, що ґрунтуються на «вимірюванні – повторній

відправці» кубітів. Також існують КРС з використанням одиничних фотонів, що готуються у двох взаємно незміщених базисах і посилаються блоками. Цей протокол дозволяє відправнику розділити своє повідомлення між двома абонентами (або більшою кількістю абонентів). Усі КРС є захищеними як проти зовнішнього зловмисника, так і проти нечесних дій учасників протоколу. На відміну від класичних схем розділення секрету, квантові та напів-квантові схеми дозволяють виявити підслуховування та не потребують шифрування повідомлень. Найзначнішим недоліком більшості протоколів КРС є потреба у наявності великої квантової пам'яті в усіх сторін, що поки знаходиться за межами можливостей сучасних технологій [1].

Квантовий потоковий шифр передбачає шифрування даних подібно до класичних потокових шифрів, але із застосуванням квантового шумового ефекту [39] і може використовуватись в оптичних комунікаційних мережах. КПШ базується на протоколі *Yuen 2000 (Y-00)* [40-41, 51-53]. Вихідними даними передавача у даній схемі є послідовність когерентних станів, що переносить інформацію про дані чи ключ. Теоретико-інформаційна стійкість протоколу Y-00 забезпечується рандомізацією, що базується на квантовому шумі, а також на додаткових математичних (обчислювальних) схемах. Ще однією перевагою КПШ є більша захищеність порівняно із звичайними потоковими шифрами завдяки квантовому шумовому ефекту і неможливості клонування квантових станів. Що стосується недоліків КПШ, то варто відмітити, перш за все, складність практичної реалізації системи [42].

Квантовою телепортацією називається передача квантового стану на відстань за допомогою роз'єднаної в просторі заплутаної пари і класичного каналу зв'язку, при якій стан руйнується в точці відправлення при проведенні вимірювання, після чого відтворюється в точці прийому**.** При цьому обов'язковою є передача інформації між джерелом і приймачем класичним, неквантовим каналом, яка може здійснюватися не швидше, ніж зі швидкістю світла. Протоколи КТ розділяють на два класи: *імовірнісні* та *детерміновані*. Експерименти та протоколи групи науковців під керівництвом Цайлінгера та Де-Мартіні – відносять до імовірнісних (для таких протоколів не важливо скільки фотонів загубиться при пересиланні, вони більш придатні для передачі на великі відстані). Детальне дослідження протоколів проведено у [43].

Квантова теорія ігор. Одним з напрямків у дослідженнях КК є вивчення різних стратегій передачі повідомлення одержувачу, для цього використовуються основи теорії ігор з елементами квантової фізики. У деяких випадках дії відправника і зловмисника розглядаються як асиметрична квантова гра двох гравців [44]. У КТІ суперпозиція використовується для моделювання невизначеності, коли неможливо знати, якою стратегією в цей момент часу скористається гравець. У

класичному математичному апараті теорії ймовірності такої можливості немає, там враховується лише ймовірність, з якою гравець може вибрати ту чи іншу стратегію. А використання суперпозиції для моделювання цієї невизначеності дозволить вирішувати завдання, які стандартна імовірнісна концепція описати не може. Загальна схема квантової гри наведена у [45] та полягає у наступному: побудова квантової гри починається з вибору початкового стану і визначення можливої заплутаності системи двох гравців. Далі шукають можливі квантові рішення (стратегії) гравців різних типів. Після того, як обидва гравці вибрали свою індивідуальну квантову стратегію, вводиться оператор розчеплення (розплутування) для підготовки вимірювання стану гравців. Оператори заплутування і розплутування залежать від додаткового параметра, який вимірює ступінь заплутаності системи. Очікуваний виграш у квантовій версії загальної гри двох гравців залежить від матриці виграшів і спільної ймовірності спостерігати чотири вимірюваних величини, які і є результатами гри. Квантовий стан заплутаності двох гравців зовсім не означає, що заплуталися самі гравці (або їх думки). Процес квантової декогеренції забороняє такі макроскопічні заплутані системи, створені з мікроскопічних квантових частинок.

До КТІ [46] входять такі протоколи як квантове вручення біту (*quantum bit commitment*), квантове підкидання монети (*quantum coin tossing*) та квантова рулетка (*quantum gambling*). Поняття класичного вручення біту (математична версія відправки заклеєного конверту) є загальним примітивом для розробки секретних криптографічних протоколів. Дані, що відправляються від Аліси до Боба перебувають у стані замкнутості (локінгу) і можуть бути розшифровані тільки при вручені Бобу ключа від Аліси. *Квантове вручення біту* [47], як і квантові протоколи розподілу ключів, забезпечує теоретико-інформаційну (безумовну) стійкість. Проте, спираючись лише на властивості квантового каналу, експериментальне квантове вручення бітів неможливе. Ця проблема була вирішена науковцями у 2013 р. шляхом поєднання квантової фізики і теорії відносності, при цьому відбулася передача безумовно стійкого повідомлення (між Женевою та Сінгапуром за 50 мс).

Квантове підкидання монети [47] використовує квантові монети, які, на відміну від звичайних, можуть перебувати у нескінченній кількості станів. Також, можливі деякі змішані стани, які пояснюються явищем суперпозиції у квантовій механіці. Використання змішаної позиції гарантує одній зі сторін постійний виграш. Якщо Аліса використовує змішаний стан, то у будь-якому випадку, не залежно від дій Боба з монетою, і за умови попередньої домовленості про кінцеве число ходів, вона буде вигравати, оскільки у кінці гри зможе перевести свою монету зі змішаного стану в чистий (абсолютний).

Квантова рулетка [46], завдяки використанню непорушних постулатів квантової механіки, унеможливлює шахрайство, яке можливе у класичній версії гри «вибір виграшної коробки». Наприклад, Аліса приховує м'яч у будь-якій коробці, а Боб відгадує його місцезнаходження. Якщо грали віддалено, Аліса може легко збрехати про виграш. Або, якщо коробки були з Бобом, він міг обдурити, стверджуючи, що він знайшов м'яч. Використовуючи квантову рулетку шахрайство неможливе з огляду на те, що опоненти можуть завжди з'ясувати, який вибір зробив інший, коли гра закінчена (завдяки суперпозиції своїх дій). Хоча зазначені протоколи є лише примітивами безпеки, проте на їх основі можуть бути побудовані більш складні протоколи, здатні повністю змінити уявлення про безпеку в інформаційно-комунікаційних системах. Окрім того, також проведено багато досліджень та багато класичних ігор переведено у квантовий простір, наприклад, «prisoners' dilemma», «the battle of the sexes», «the Monty Hall problem», «rock-scissors-paper», «quantum tic-tac-toe» [48, 50].

З урахуванням зазначених квантових методів безпечної комунікації, а також нової базової ознаки (стійкість до певного типу кібератак), розширена класифікація матиме вигляд представлений на рис. 2. Точками на перетині ліній позначено уразливість до певного виду атаки. Кібератаки, позначені: C – coherent attack – когерентні атаки; NC – non-coherent attack – некогерентні атаки; MiM – man-in-the-middle – атака «людина посередині»; DoS – denial of servise attack – атака «відмова в обслуговувані»; TC – timing channel attack – часова незбалансованість детектора; FS – заміна існуючого квантового каналу на кращий; PBS – photon beam splitting attack – атака поділу пучка фотонів; PNS – photon number splitting attack – атака поділу числа фотонів; TH – trojan horse attack – атака типу «Троянський кінь». Зазначені атаки детально описані у роботі [10].

Комерційні системи квантової криптографії

Як зазначалось, значна частина теоретичних та практичних досліджень у галузі КК присвячена розробці та вдосконаленню протоколів КРК, єдиному на сьогоднішній день напряму КК, який досягає теоретико-інформаційної (безумовної) стійкості [8] та вже пройшов шлях від теоретичних досліджень до практичної реалізації.

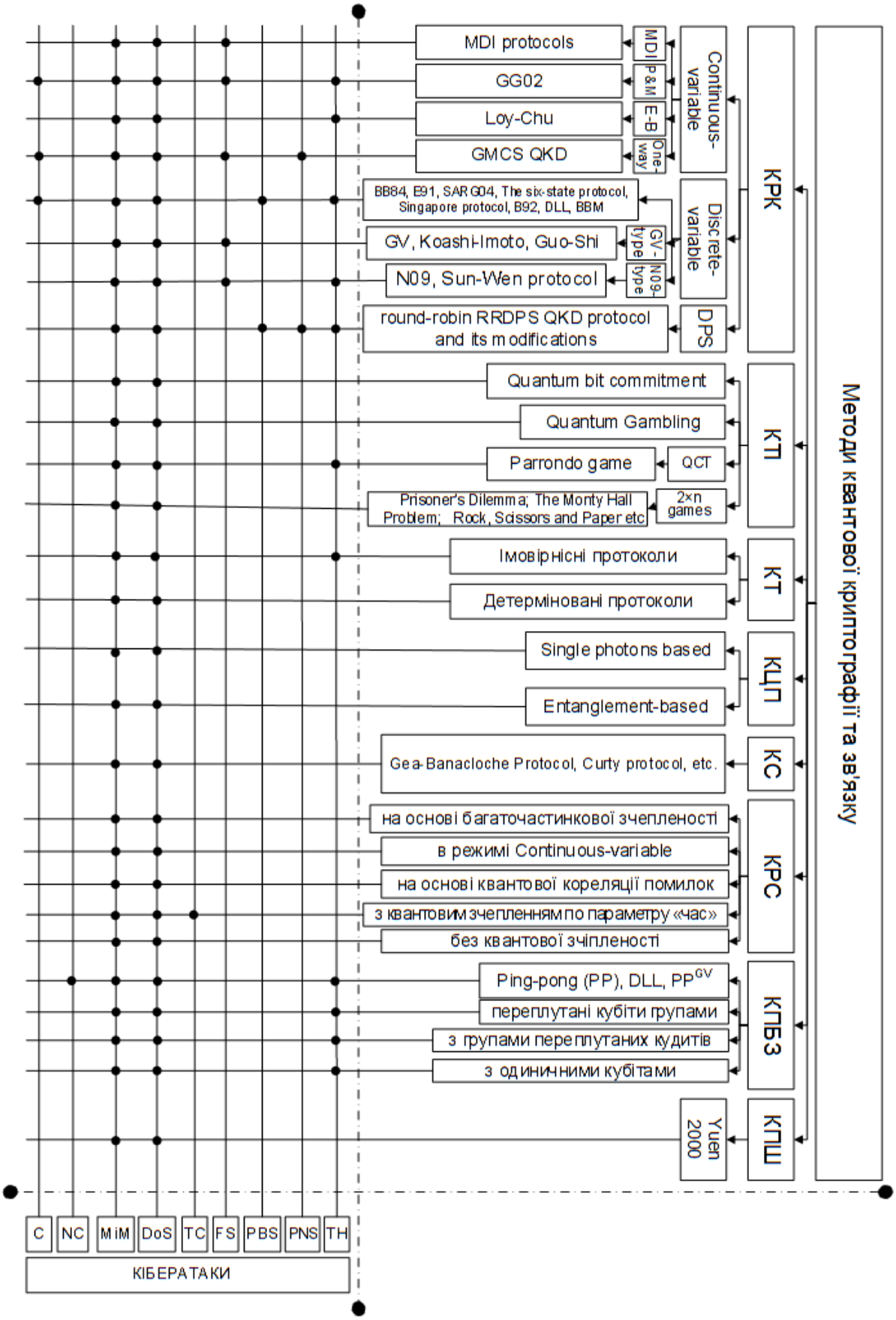


Рис. 2. Розширена класифікація квантових методів безпечної комунікації (криптографії та зв'язку)

Основними виробниками систем КК, а саме систем КРК, які вже декілька років представлені на ринку комерційними системами [55-63] є: id Quantique, Inc. (Швейцарія), MagiQ Technologies, Inc. (США), SmartQuantum, Inc. (Франция), QuintessenceLabs, Pty Ltd (Австралия), D-Wave Systems (Канада), Toshiba Research Europe Ltd (Великобританія), QinetiQ (Великобританія), NEC (Японія). На рис. 3-10 представлені деякі комерційні системи КРК.

Рис. 3. Генератор квантових ключів Q-Key maker [55]

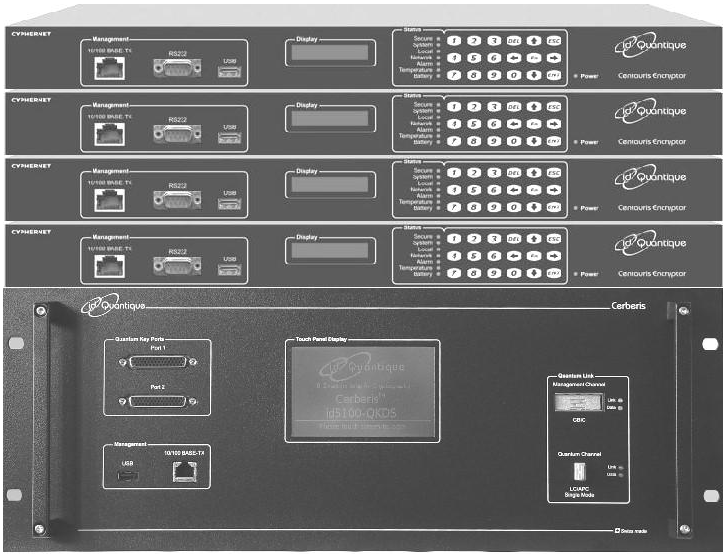


Рис. 4. CN8000 Multilink Encryption [56]

Рис. 5. Комерційна система Cerberis QKD, фірми id Quantique [58]

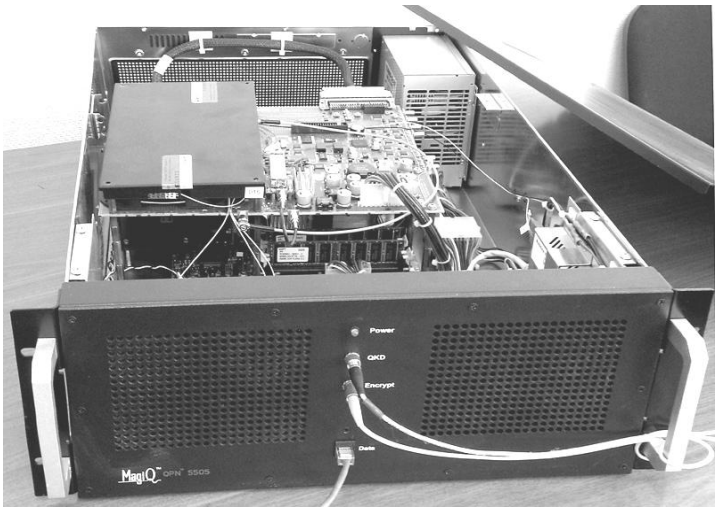


Рис. 6. Квантова криптосистема, розроблена MagiQ Technologies, Inc. [59]

Рис. 7. Квантова криптосистема, розроблена Toshiba Research Europe Ltd [60]

Рис. 8. Квантова криптосистема Clavis^{2}QKD [62]

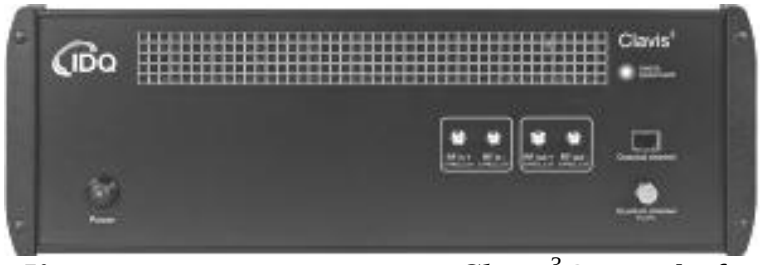


Рис. 9. Квантова криптосистема Clavis^{3}QKD Platform [63]

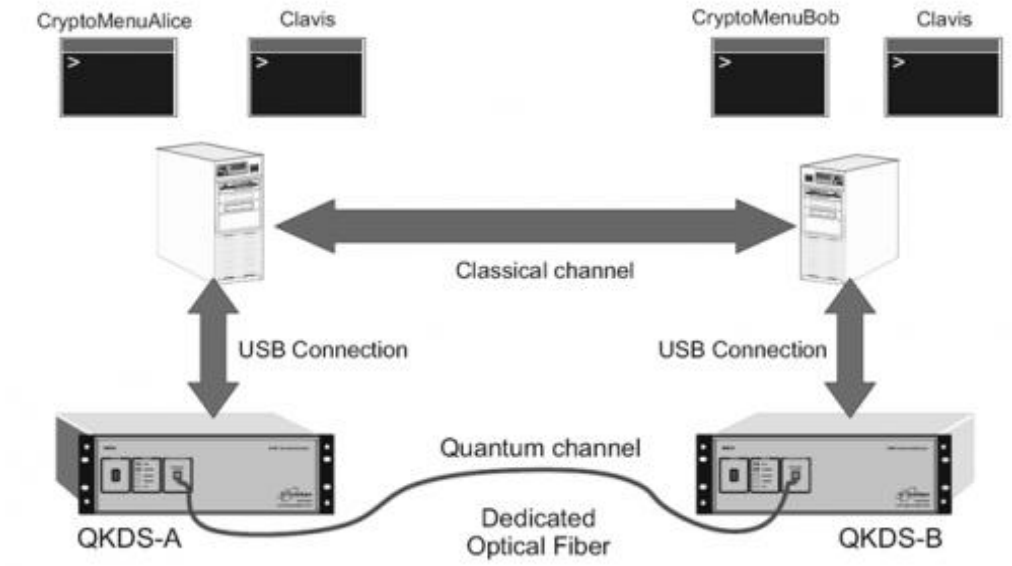


Рис. 10. Схема реалізації роботи криптосистеми Clavis^{2}QKD [62, 63]

Окрім того, на сьогодні системи КРК [63], впроваджені у таких банках Європи (Швейцарія): Notenstein Private Bank, Swissquoute Bank, Hyposwiss Private Bank, Global Bank. А також захист інформації засобами КРК проводиться у таких комерційних компаніях: Battelle's System Heralds (США), Bloombase Extends Enterprise (США), Colt (Великобританія), Cygate (Швеція, Фінляндія).

Висновки

Таким чином, у цій роботі проведено аналіз сучасних квантових методів безпечної комунікації – визначено їх переваги і недоліки, відомі класифікації. На підставі часткових узагальнень теоретичних положень та практичних досягнень у галузі квантової криптографії та зв'язку, розроблено розширену класифікацію. За рахунок розширення номенклатури методів та базових ознак, ця класифікація дає можливість виявити низку проблем у цій галузі та дозволяє розширити можливості щодо вибору відповідних методів для побудови сучасних квантових систем захисту інформації.

Література

1. Лимарь И.В., Василиу Е.В. Классификация квантовых технологий разделения секрета / Захист інформації. – Том 16. – №3. – 2014 – С. 201 - 214.

2. Chenmiao W., Li Ya. A complete Classifi-cation of Quantum Public-key Encryption Protocols Available from: http://arxiv.org/pdf/ 1507.03765v2.pdf

3. Chong X., Li Y., Yong P. Dongqing Chen The Classification of Quantum Symmetric-Key Encryption Protocols. Available from: http://arxiv.org/pdf/1006.4216.pdf

4. Hassanpour S., Houshmand M. Bidirectional quantum teleportation and secure direct communication via entanglement swapping. Available from http://arxiv.org/ftp/arxiv/papers/1411/1411.0206.pdf

5. Xiaoqing T. Introduction to Quantum Crypto-graphy (part of Theory and Practice of Cryptography and Network Security Protocols and Technol.) Available from http://cdn.intechopen.com/pdfs-wm/43793.pdf

6. D'ariano G.M. Quantum bit commitment: a complete classification of protocols. Available from http://www.qubit.it/research/publications/0209150.pdf

7. Korchenko O., Vasiliu E., Gnatyuk S. Modern quantum technologies of information security, Aviation. – Vilnius: Technika, 2010. – Vol. 14, No. 2. – p. 58 - 69.

8. Кузнецова А.В. Стратегии атак на квантовые протоколы защиты информации. – Цифрові технології. – 2013. – № 14. – С. 134 - 137.

9. Scarani V., Ribordy G., Gisin N. Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations. Available from: http://www.unige.ch/gap/quantum/ _media/publications:bib: prl57901.pdf

10. Корченко А.Г. Методы перехвата информации в информационно-коммуникационных системах на основе квантових технологий / А.Г. Корченко, Е.В. Василиу, Т.А. Жмурко, С.А. Гнатюк // Монография. – Х. : Цифрова друкарня. – 2013. – № 1. – С. 98 - 110.

11. Waks E. Security of Quantum Key Distribution with Entangled Photons Against Individual Attacks / E. Waks, A. Zeevi, Y. Yamamoto // Physical Review A. – 2002. – V. 65, issue 5. – 052310.

12. Василиу Е.В. Анализ атаки на пинг-понг протокол с триплетами Гринбергера-Хорна-Цайлингера / Е.В. Василиу // Наукові праці ОНАЗ ім. О.С. Попова. – 2008. – № 1. – С. 15 - 24.

13. Василиу Е.В. Анализ атаки пассивного перехвата на пинг-понг протокол с полностью перепутанными парами кутритов / Е.В. Василиу, Р.С. Мамедов // Восточноевропейский журнал передовых технологий. – 2009. – № 4/2 (40). – С. 4 - 11.

14. Василиу Е.В. Анализ атаки двух злоумышленников на протокол квантовой прямой безопасной связи / Е.В. Василиу, С.В. Николаенко // Труды Северо-Кавказ. фил. МТУСИ. – 2013. – С. 324 - 330.

15. Suzuki S., Rodney V. Classification of Quantum Repeater Attacks. Available from: www.internetsociety.org/sites/default/files/01_2_3.pdf

16. Jain N., Anisimova E., Khan I. et al. Trojan-horse attacks threaten the security of practical quantum cryptography. Available from http://iopscience.iop. org/1367-2630/16/12 /123 030

17. Scarani V. et al. The security of practical quantum key distribution.

18. Shukla C., Banerjee A., Pathak A. Secure Quantum Communication with Orthogonal States Available from: http://arxiv.org/pdf/1407.3412.pdf

19. Cutolo A., Mignani A.G., Tajani A. Photonics for safety and security, World Scientific Publishing Co. Pte. Ltd. – Singapore. – 2014. – 422 p.

20. Cerf N.J., Leuchs G., Polzik E.S. Quantum information with CV of atoms and Light, Imperial College Press. – 2007. – 604 p.

21. Bennett C.H., Brassard G. Quantum crypto-graphy: public key distribution and coin tossing // Proceedings of the IEEE Intern. Conf. on Comp., Syst. and Signal Proces. – Bangalore, India. – 1984. – P. 175 - 179.

22. Bruss D. Optimal Eavesdropping in Quantum Cryptography with Six States // Physical Review Letters. – 1998. – V. 81, № 14. – P. 3018 - 3021.

23. Huttner B., Imoto N., Gisin N., Mor T. Quantum Cryptography with Coherent States // Physical Review A. – 1995. – V. 51, № 3. – P. 1863 - 1869.

24. Goldenberg L., Vaidman L. Quantum Cryptography Based On Orthogonal States // Physical Review Letters. – 1995. – V. 75, № 7. – P. 1239 - 1243.

25. Koashi M., Imoto N. Quantum Cryptography Based on Split Transmission of One-Bit Information in Two Steps, Phys. Rev.Let. – 1997. – V. 79, № 12. – P. 2383 - 2386.

26. Bennett C.H. Quantum cryptography using any two non-orthogonal states // Physical Review Letters. – 1992. – V. 68, № 21. – P. 3121 – 3124.

27. Ekert A. Quantum cryptography based on Bell's theorem, Phys. Rev. Let. – 1991. – V. 67, № 6. – P.661 - 663.

28. Wang X.-B. Comment on Decoy State Quantum Key Distribution // arXiv:quant-ph/0501143

29. Scarani V., Acin A., Ribordy G., Gisin N., Phys. Rev. Lett. 92, 2004, 057901.

30. Gisin N. et al., Rev. Mod. Phys., 74, 2002. – P.145 - 195.

31. Khan M.M., Murphy M., Beige A. High error-rate quantum key distribution for long-distance communication Available from: http://iopscience. iop.org/ article/10.1088/1367-2630/11/6/063043/pdf

32. Chuan W., Fu Guo D., Gui Lu L. Multistep quantum secure direct communication using multiparticle Greenberg-Horne-Zeilinger state. – Optics Communications. – 2005. – V. 253. – P. 15 - 19.

33. Bostrom K., Felbinger T. Deterministic secure direct communication using entanglement // Physical Review Letters. – 2002. – V. 89, № 18. – 187902.

34. Cai Q.-Y., Li B.-W. Improving the capacity of the Bostrom – Felbinger protocol // Physical Review A. – 2004. – V. 69, № 5. – 054301.

35. Василиу Е.В., Василиу Л.Н. Пинг-понг протокол с трех- и четырехкубитными состояниями Гринбергера-Хорна-Цайлингера // Труды Одесского политех. ун-та. – 2008. – Вып. 1(29). – С. 171 - 176.

36. Wang Ch., Deng F.-G., Li Y.-S. et al. Quantum secure direct communication with high dimension quantum superdense coding // Physical Review A. – 2005. – V. 71, № 4. – 044305.

37. Нильсен М., Чанг И. Квантовые вычисления и квантовая информация. – М.: Мир, 2006. – 824 с.

38. Стин Э. Квантовые вычисления. Перевод с английского И.Д. Пасынкова: НИЦ «Регулярная и хаотическая динамика», М. – Ижевск, 2000. – 111 с.

39. Yan F.-L., Gao T., Li Yu.-Ch. Quantum secret sharing protocol between multiparty and multiparty with single photons and unitary transformations // Chinese Phys.Let. – 2008. – V. 25, № 4. – P. 1187 – 1190

40. Yuen H. P. KCQ: A New Approach to Quantum Cryptography I. General Principles and Key Generation // arXiv:quant-ph/0311061

41. Nair R., Yuen H.P. On the Security of the Y-00 Direct Encryption Protocol // arXiv:quant-ph/0702093v2

42. Hirota O., Kurosawa K. An immunity against correlation attack on quantum stream cipher by Yuen 2000 protocol // arXiv:quant-ph/0604036v1

43. Hassanpour S., Houshmand M. Bidirectional quantum teleportation and secure direct communication via entanglement swapping http://arxiv.org/ abs/1411.0206

44. Eisert J., Wilkens M., Lewenstein M. Quantum games and quantum strategies, http://journals.aps. org/prl/abstract/10.1103/PhysRevLett.83.3077

45. Старобогатов Р. Квантовые стратегии предотвращения финансово-экономических кризисов http://econf.rae.ru/pdf/2012/05/1282.pdf

46. Жмурко Т.О. Протоколи квантової теорії ігор / Т.О. Жмурко // Політ. Сучасні проблеми науки: міжнар. наук.-практ. конф. молодих учених і студентів, м. Київ, 2-3 квітня 2014 р., НАУ. – С. 6.

47. Nayak A., Sikora J., Tunzel L. Quantum and classical coin-flipping protocols based on bit-commitment and their point games http://arxiv.org/abs/1504.04217

48. Flitney A.P., Abbott D. An introduction to quan-tum game theory http://arxiv.org/pdf/quant-ph/0208069.pdf

49. Gnatyuk S., Zhmurko T., Falat P. Efficiency increasing method for quantum secure direct communication protocols // Proc. of the EEE 8th Intern. Conf. on «Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications» (IDAACS'2015), Warsaw, Poland, September 24-26, 2015: Vol. 1. – p. 468 - 472.

50. Leaw J.N., Cheong S.A. Strategic Insights From Playing the Quantum Tic-Tac-Toe.

51. Корченко О.Г., Василіу Є.В., Гнатюк С.О. Сучасні квантові технології захисту інформації // Захист інформації. – 2010. – № 1. – С. 77 - 89.

52. Korchenko O.G., Vasiliu Ye.V., Gnatyuk S.O. Modern directions of quantum cryptography // Proc. of the fourth world congress «Aviation in the XXI-st century» – «Safety in Aviation and Space Technologies». – Kyiv, 2010. – V. 1. – P. 17.1 - 17.4.

53. Korchenko O., Vasiliu Ye., Gnatyuk S. et al. Quantum Secure Telecommunication Systems, Telecommunications Networks – Current Status and Future Trends (ed. by J.H. Ortiz), InTech, 2012. – p. 211 - 236.

54. Василиу Е.В. Безопасные системы передачи конфиденциальной информации на основе протоколов квантовой криптографии : монография / Е.В. Василиу, В.Я. Мильчевич, С.В. Николаенко, А.В. Мильчевич. – Харьков: Цифровая типография. – 2013.– № 1. – 168 с.

55. Bovino F.A. Practical Quantum Cryptography: The Q-KeyMaker / Bovino F.A., Giardina M. // ArXiv.org [Online] – 2011. – Mode of access: http://arxiv.org/abs/1104.2475.

56. Cerberis Encryption Solution [Electronic resource] : Layer 2 Encryption with Quantum Key Distribution. – Electronic data. – Geneva : ID Quantique SA, 2010. – Mode of access: http://www.idquantique.com/products/cerberis.htm.

57. Google and NASA Launch Quantum Computing AI Lab [Electronic resource] : MIT Technology Review. – Electronic data. – Big Sandy : MIT, 2013. – Mode of access: http://www.technologyreview.com/news/514846/google-and-nasa-launch-quantum-computing-ai-lab.

58. ID Quantique SA, Cerberis Encryption Solution: Layer 2 Encryption with Quantum Key Distribution [Online] Available: http://www.idquantique.com/products/cerberis.htm.

59. QPN-8505 Security Gateway: Data Sheet [Electronic data] (1 file: 143 754 bytes). – MagiQ Technologies Inc, 2007. – Mode of access: http://www.magiqtech.com/MagiQ/Products_files/8505_Data_Sheet.pdf

60. Quantum Key Distribution System [Electronic resource] : / Toshiba Research Europe Ltd., Cambridge Research Laboratory. – Electronic data. – Tokyo, Japan : Toshiba Corporation, 2010. – Mode of access: http://www.toshiba-europe.com/research/crl/qig/quantumkeyserver.html.

61. Toshiba QKD system [Electronic resource]. – Electronic data. – Toshiba Research Europe Ltd, 2015. – Mode of access: http://www.toshiba.eu/eu/Cambridge-Research-Laboratory/Quantum-Information-Group/Quantum-Key-Distribution / Toshiba-QKD-system/.

62. Clavis2 QKD Platform For R&D [Online] Available: http://www.idquantique.com/photon-counting/clavis2-qkd-platform/

63. Clavis3 The new QKD research platform [Online] Available: http://www.idquantique.com/wordpress/wp-content/uploads/Clavis_3_Datasheet_260116.pdf

64. Gnatyuk S.O. Contemporary Commercial Quantum Information Security Systems / S.O. Gnatyuk, T.O. Zhmurko, M.O. Riabyi // Proceedings of the CSE-2013. – 2013. – P. 74-77.

ОЦІНКА ВРАЗЛИВОСТІ ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ ТА МЕРЕЖ ВІД КІБЕРАТАК

Ларін В.В., Ширяев А.В., Медведев Д.О.

Вступ

Створення технологій виявлення впливів на інформацію, в тому числі у відкритих мережах – це природна захисна реакція на появу нової зброї. Економічну і науково-технічну політику підключення держави до світових відкритих мереж слід розглядати через призму інформаційної безпеки. Будучи відкритою, орієнтованою на дотримання законних прав громадян на інформацію та інтелектуальную власність, ця політика повинна передбачати захист мережевого обладнання на території країни від проникнення в нього елементів інформаційної зброї. Це особливо важливо сьогодні, коли здійснюються масові закупівлі іноземних інформаційних технологій.

Зрозуміло, що без підключення до світового інформаційного простору країну очікує економічне животіння. Оперативний доступ до інформаційних і обчислювальних ресурсів, що підтримуються мережею Internet, зрозуміло, слід сприймати як фактор подолання міжнародної ізоляції і внутрішньої дезінтеграції, як умови зміцнення державності, інститутів громадянського суспільства, розвитку соціальної інфраструктури.

Стратегія кібербезпеки України має наступні цілі:

- разом з основними ризиками і проблемами виявити економічні та геополітичні можливості;
- порівняти між собою ступінь підготовленості і політичної уваги до проблеми безпеки Інтернету в третіх країнах;
- позначити основні і найважливіші проблеми, що вимагають вирішення;
- оцінити поточні та заплановані заходи, а також відзначити ті проблемні зони, до яких державі слід приділити більше уваги.

У середовищі, де постійно з’являються і еволюціонують кіберзагрози, держави при зустрічі з новими, глобальними загрозами отримають більшу вигоду від гнучких, оперативних стратегій кібербезпеки. Транскордонний характер загроз змушує країни вступати в тісну міжнародну взаємодію. Співпраця на європейському рівні необхідна не тільки для ефективної підготовки до кібератак, а й для своєчасної реакції на них. Комплексна державна стратегія кібербезпеки – перший крок на цьому шляху.

Для реалізації стратегій кібербезпеки приватний і державний сектори повинні працювати в тісній співпраці. Співробітництво повинно здійснюватися за допомогою обміну інформацією, передовими практиками

(наприклад, в сфері управління інцидентами), а також навчаннями на державному та загальносвітовому рівнях.

1. Загальна класифікація та характеристики атак на відмову

Атака на відмову в обслуговуванні – напад на комп'ютерну систему з наміром зробити комп'ютерні ресурси недоступними користувачам, для яких комп'ютерна система була призначена.

Типовий сценарій атаки на відмову полягає в наступному: нападник ініціює початок атаки, після чого одне або декілька джерел атаки починають надсилати зловмисні пакети через мережу Інтернет, що призводить до блокування послуги для клієнтів. Як правило при атаці нападник не використовує свій власний комп’ютер, тому джерела атаки, насправді, являються агентами – машинами третіх осіб, що були взяті під контроль атакуючим. Атака може блокувати ключовий ресурс шляхом використання певних слабкостей в програмному забезпеченні (ПЗ) жертви (атаки вразливості) або шляхом пересилки великих об’ємів трафіку, який жертва повинна обробляти (атаки переповнення). Атаки вразливості зазвичай використовують пакети спеціального типу або змісту для використання конкретних вразливостей.

Як правило для успішного використання вразливості досить декількох пакетів, тому такі атаки мають низький об’єм трафіку. Ці дві властивості (спеціальний вид пакетів і низькі об’єми) спрощують протидію таким атакам. На рис. 1 зображені характеристики атак на відмову.

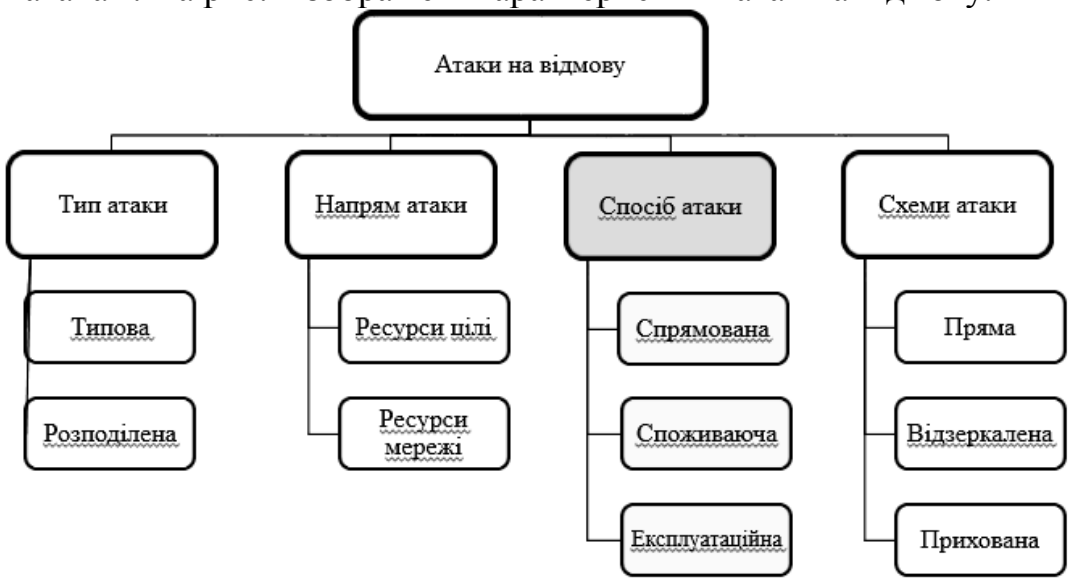


Рис. 1. Класифікація атак на відмову

На сьогоднішній момент існує досить багато різних видів атак на відмову, кожна з яких використовує певну особливість побудови мережі або вразливості ПЗ. Наприклад, атаки можуть здійснюватися шляхом безпосередньої пересилки великої кількості пакетів (протокол датаграм користувачів - User Datagram Protocol (UDP), протокол міжмережевих

управляючих повідомлень - Internet Control Message Protocol (ICMP), використання проміжних вузлів (Smurf, Fraggle), передачі занадто довгих пакетів ("Ping of Death"), некоректних пакетів ("Land") або великої кількості трудоємних запитів. Зауважимо, що протягом останнього часу відбувається розвиток цього напряму діяльності та поява нових видів і способів атак. З останніх тенденцій можна відзначити появу атак погіршення якості та низькочастотних атак і безумовно, цей процес буде продовжуватися, потребуючи нових досліджень та розробки нових методів протидії. Основні існуючі класи атак досить добре вивчені. Наприклад, атаки класифіковані згідно з протоколами, по яким вони здійснюються. Виділені наступні атаки: синхронізація - synchronization (SYN flood), протокол управління передачі – transmission control protocol (TCP reset), ICMP flood, UDP flood, система доменних імен - Domain Name System (DNS request), загальний інтерфейс шлюзу - Common Gateway Interface (CGI request), "Mail bomb", протокол визначення адреси – Address Resolution Protocol (ARP storm) і атаки на алгоритмічну складність.

Виділяють три типа атак по техніці здійснення (рис. 2):

– цільова (використовують недоліки в протоколах, програмах);

– споживаюча (завантажують ресурси системи);

– експлуатаційна (використовують вразливості, помилки кода).

Найбільш поширеним видом кібератаки на цільову систему можна вважати розподілені Distributed Denial of service (DDoS) атаки на сервер додатків або на інше мережеве обладнання. Наприклад, великомасштабна DDoS-атака за допомогою бот-мережі може заблокувати доступ в кіберпростір на рівні країни.

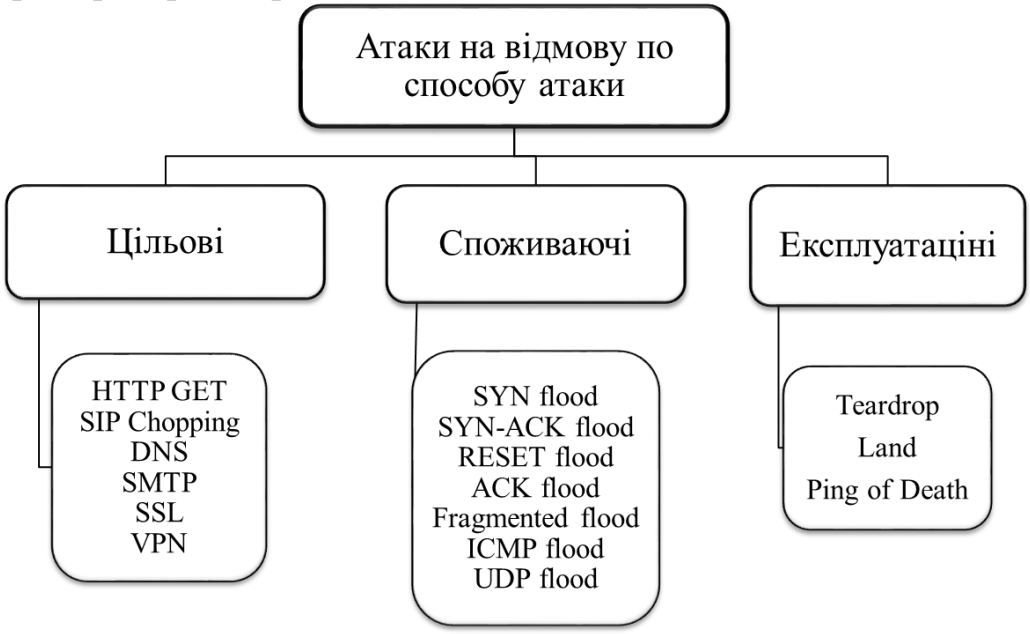


Рис. 2. Види атак на відмову по способу атаки

2. Аналіз розподілених атак на відмову

Розподілена атака на відмову в обслуговуванні – це реальна і зростаюча загроза, з якою стикаються компанії в усьому світі. Ці атаки реалізуються великою кількістю програмних агентів, розміщених на хостах, які зловмисник скомпрометував раніше. Реалізація цих атак може призвести не тільки до виходу з ладу окремих хостів і служб, а й повністю або тимчасово зупинити мережі. У зв’язку з критичністю і нетривалістю даного класу атак, побудова ефективних засобів захисту від них являє собою складну науково-технічну проблему. На рівні маршрутизаторів захист від DDoS-атак вже досить успішно реалізували компанії Cisco Systems.

В обчислювальній техніці, атаки на відмову в обслуговуванні або розподілені атаки на відмову в обслуговуванні є спробою зробити машини або мережевий ресурс недоступним для можливих користувачів. Мотиви і цілі з DoS-атаки можуть відрізнятися, але в загальному випадку складаються із зусиль одного або декількох людей тимчасово або на невизначений термін перервати або призупинити надання мережевих послуг.

Відмови в обслуговуванні можуть також призвести до проблем у «гілці» мережі, в якій знаходиться фактична жертва нападників. Наприклад, пропускна здатність маршрутизатора між Інтернетом та локальною мережею може споживатися атакою, що призводить до збитків не тільки для потенційної жертви, але й для всієї мережі. Якщо атака здійснюється на досить великому масштабі, то через неправильно налаштоване обладнання мережевої інфраструктури, про що зловмисник міг і не знати, може бути порушена робота всієї мережі (рис. 3).

Одними з найпрогресивніших досліджень в області захисту від DDoS-атак є роботи професорів Peng T. Ong, Jaydip Sen та доктора Ashish Gupta. Структура атаки майже завжди є дуже складною, що не дозволяє в багатьох випадках відстежити нападника. Зв'язок між майстром і демонами може бути непомітним, так що стає важко знайти головний комп'ютер. Хоча деякі докази можуть існувати на одному або декількох комп'ютерах в мережі DDoS з місцем розташування майстра. Демони, як правило, автоматизовані так, що вони не є необхідними для постійного діалогу, котрий відбудеться між майстром і рештою мережі. Насправді, такі методи, які зазвичай використовуються, свідомо маскують особу і місцезнаходження господаря в мережі DDoS. Ці методи роблять процес аналізу атаки, блокування атакуючого трафіку і відстежування його до джерела надзвичайно важким.

У більшості випадків, системні адміністратори заражених систем навіть не знають, що демони були встановлені в системі. Навіть якщо вони знайшли і знищили ПЗ DDoS, вони не можуть допомогти іншим користувачам визначити, чи є десь в системі ще розміщене подібне ПЗ.

Популярними системами для експлуатації є Web-сервери, електронна пошта, інші сервери, так як ці системи можуть мати велику кількість відкритих портів, великий обсяг трафіку, і навряд чи будуть швидко виведені з ладу, навіть при атаці на них.

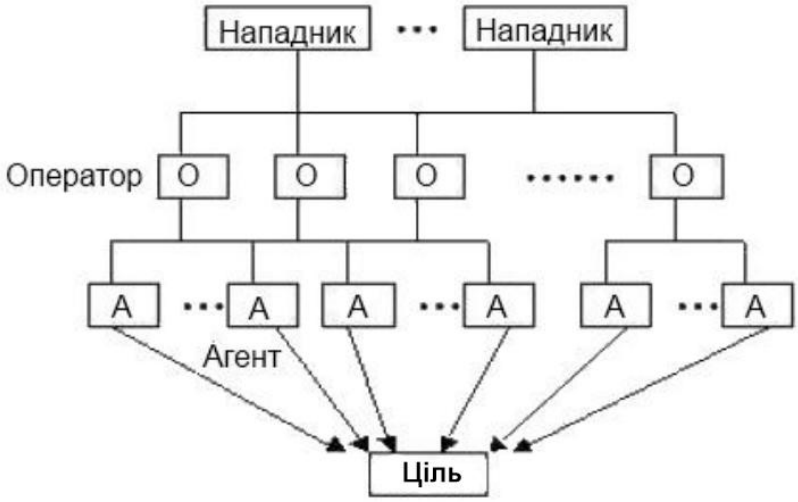


Рис. 3. Концептуальна схема DDoS-атаки

3. Типові атаки на відмову, що можуть використовуються в телекомунікаційних системах та мережах

Для початку розглянемо типові атаки на відмову, які здійснюються з одного комп'ютера.

Прямі атаки. Ці атаки використовують певні особливості побудови протоколів мережі Інтернет. Часто такого роду атаки використовуються в якості елементів для конструювання складних розподілених нападів. Приклади таких атак наведемо нижче.

ICMP атаки. В атаках цього типу використовуються особливості реалізації протокола ICMP. Оскільки цей протокол є внутрішнім механізмом підтримки працездатності Internet protocol (IP-мереж), то він природно викликав підвищену увагу нападників. Виділяють декілька типів. Перший тип – атака на розрив з'єднання (connection-reset attacks) використовує ICMP-повідомлення, що сигналізують про фатальну помилку і спричиняють до припинення TCP-з'єднання. Другий тип полягає в звуженні пропускної здатності (throughput-reduction attacks), при цьому використовуються ICMP-повідомлення, що вимагають від джерела зменшення пропускного каналу. Це службові пакети, що надсилаються мережею в разі її перевантаження. Нарешті третій тип пов'язаний з пакетами ICMP, які відповідають за допустиму фрагментацію пакетів. Надсилаючи сфальшований пакет на адресу джерела нападник добивається ситуації коли великий файл надсилається мізерними шматочками.

Віддзеркалені атаки. В основі здійснення віддзеркалених атак лежить використання фальшування адреси та особливостей протоколів TCP-

з’єднання, а саме – момент коли один комп’ютер відповідає на запит іншого. Схема виконання атаки наступна: пакет-запит фальшується від імені жертви та надсилається на велику кількість машин. Потім всі ці машини відповідають жертві пакетами відповідями, завантажуючи його канал беззмістовним (хоча й законним) трафіком. Як випливає з опису схеми найбільш сильно ефект віддзеркалення проявляється при здійсненні поглинаючих атак.

Приховані атаки. Дослідники відмічають обмеженість існуючих механізмів захисту проти атак такого типу, але не пропонують алгоритму ефективної протидії.

В чому ж полягає така атака? Атакуючий комп’ютер «А» встановлює з’єднання з жертвою та надсилає велику кількість пакетів, на короткий час завантажуючи канал. Отримавши переповнення каналу протокол TCP вмикає механізм RTO (retransmission timeout) певний час не приймаючи пакети від «А». Оцінивши цей час, протягом нього нападник не надсилає пакетів. Після того як час RTO пройшов і до того як з’єднання буде закрите (оскільки воно не активне) нападник надсилає наступну порцію пакетів. В результаті підтримується порожнє з’єднання, яке завантажує пам’ять і ресурси машини. Такі атаки також отримали назву пульсуючих атаки (pulse attacks). Крім того, що один атакуючий комп’ютер може підтримувати декілька десятків таких з’єднань, атаку може бути запущено з декількох десятків тисяч комп’ютерів.

4. Розподілені атаки на відмову, що використовуються в телекомунікаційних системах та мережах

Розподілена атака на відмову (Distributed denial-of-service (DdoS)) це атака на відмову, що реалізується з кількох підконтрольних машин (агентів). При найбільш розповсюдженому сценарії всі машини, задіяні в схемі одночасно починають надсилати пакети жертві з максимальною інтенсивністю. Велика кількість агентів дозволяє швидко завантажити ресурси жертви як основні так і резервні.

Типова DdoS атака складається з двох етапів. На першому етапі відбувається пошук вразливих систем в мережі Інтернет та встановлення на них інструментів атаки. Цей етап також відомий як перетворення комп’ютерів на «зомбі». На другому етапі атакуючий дає команду своїм «зомбі» через захищений канал на здійснення атаки проти вибраної жертви. Зауважимо, що пакети трафіка атаки можуть використовувати фальшиву IP адресу джерела, щоб ускладнити для жертви ідентифікацію атакуючих комп’ютерів. Кількість керованих агентів при здійсненні розподіленої атаки на відмову може коливатися від кількох десятків до 100000 скомпроментованих машин.

5. Види, принцип дії та характеристика розподілених атак на відмову типу "flood-atack"

DDoS-атаки завжди пов'язані з низкою систем. Типовий сценарій DDoS-атаки може відбуватись приблизно за наступними кроками:

– зловмисник знаходить одну або декілька систем в Інтернеті, які можна скомпрометувати і експлуатувати. Це зазвичай здійснюється за допомогою вкраденого облікового запису в системі з великим числом користувачів через неуважних адміністраторів чи користувачів, переважно із з'єднанням з високою пропускною здатністю до Інтернету;

– в зломану систему завантажуються будь-яка кількість таких інструментів, як сканери, детектори операційної системи (ОС), руткіти, а також програми DoS/DDoS. Ця система стає майстром DDoS. Майстер за допомогою ПЗ дозволяє знайти ряд інших систем, які можна експлуатувати. Зловмисник сканує великі діапазони IP мережевих адресних блоків, щоб знайти системи, що мають вразливі місця в безпеці. Автоматизовані інструменти, що використовуються не є частиною інструментарію DDoS, але є способом обміну всередині груп злочинних хакерів. Ці зламані системи є початковими жертвами нападу DDoS. Згодом ці системи будуть експлуатуватись демонами DDoS, які здійснюватимуть фактичний напад;

– зловмисник має список систем, якими він може керувати, і в яких системах є демони DDoS. Фактичний напад відбувається, коли зловмисник запускає програму в головній системі, що спілкується з демонами DDoS, щоб почати атаку.

Атака "відмова в обслуговуванні" характеризується явною спробою нападників відключити законних користувачів мережі або ресурсу від використання його доступних сервісів. Є дві основних форми DoS атак: ті, які є сервісами злому (crash services) і сервіси флуду (flood services).

Атака DoS може бути здійсненна у ряді напрямків. П'ятьма основними типами атак є:

– споживання обчислювальних ресурсів, таких як пропускна здатність, дисковий простір або процесорний час;

– перешкоджання доступу до інформації про конфігурацію мережі, наприклад, інформацію про маршрутизацію;

– перешкоджання доступу до інформації про стан, наприклад, небажане скидання сеансів TCP;

– перешкоджання доступу до фізичних компонентів мережі;

– перешкоджання масової комунікації між передбачуваними користувачами і жертвами, так що вони вже не можуть спілкуватися.

Рис. 4 ілюструє тип широкосмугової атаки, котрий називається відбиваючою розподіленою відмовою в обслуговуванні (DRDOS-атакою). Метою DRDOS-атаки є приховати джерела трафіку атаки за допомогою третіх осіб (маршрутизаторів або веб-серверів) для передачі трафіку атаки

до жертви. Ці безневинні треті особи називаються відбивачами. Будь-яка машина, яка відповідає на вхідний пакет може стати потенційним відбивачем.

Напад DRDOS складається з трьох етапів. Перший етап являє собою типову DDoS-атаку. На другому етапі, після того як нападник отримав контроль над певною кількістю «зомбі», замість інструктажу «зомбі» для відправки трафіку атаки жертвам безпосередньо, «зомбі» наказано вислати третім особам фальшиві пакети з IP-адресом жертви в якості джерела IP-адресу. На третьому етапі, треті особи будуть надсилати відповідь до жертви, яка являє собою атаку DDoS. У порівнянні з традиційною DDoS-атакою, трафік від нападу DRDOS збільшений за допомогою третіх осіб. Це робить напад більш ширшим а, отже, і процес зупинення атаки буде більш важким. Крім того, джерела IP-адрес атаки є безневинними третіми особами. Це робить процес простеження джерела атаки вкрай складним. Нарешті, DRDOS атаки мають здатність посилювати трафік атаки, що робить атаку ще більш потужною.

Розглянемо види розподілених атак на відмову типу "flood-atack".

HTTP (HyperText Transfer Protocol) flood – обладнання телекомунікаційної системи налаштовується за допомогою веб-інтерфейсу. Розгорнутий на обладнанні HTTP-сервер є вразливим до атак. Більшість захищеного ПЗ залишає порт 80 відкритим для вільного проходження HTTP трафіку. Атака пов’язана з навантаженням веб-сервера звичайними HTTP-запитами, відбувається перевантаження мережі та зростання кількості відмов.

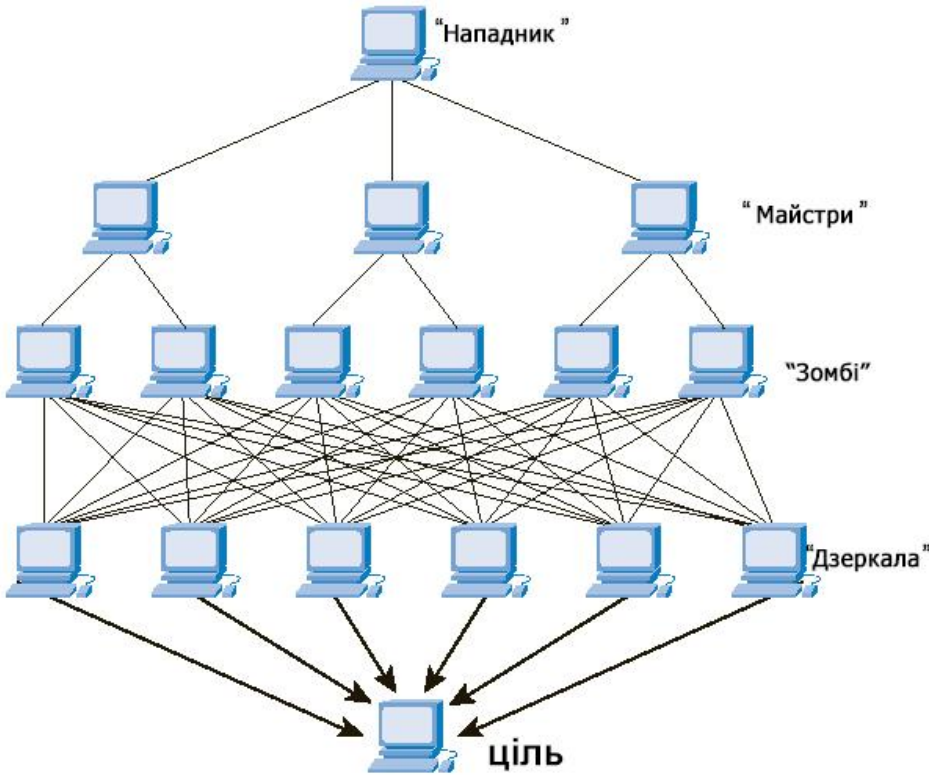


Рис. 4. Відбиваюча розподілена атака на відмову

SYN flood-атака використовує уразливість TCP, а саме трьох ступінчатий запит, тому сервер повинен виділити велику структуру даних для всіх вхідних SYN пакетів, незалежно від його достовірності. Під час SYN flood-атаки, атакуючий посилає SYN пакети з вихідними IP-адресами, які не існують або не використовується. Коли сервер заносить інформацію запиту в стек пам'яті, він буде чекати підтвердження від клієнта, який відправив запит. У той час як запит очікує підтвердження, він буде залишатися в стеку пам'яті. Оскільки IP-адреса джерела, використовуваного в ході SYN flood-атаки може виявитися помилковою, сервер не отримає пакета з підтвердженням запитів.

ICMP flood – спрямована на мережеве обладнання, виникає при завантаженні мережі атакуємої системи пакетами типу ICMP-ECHO. Ці пакети сфальсифіковані для максимального зменшення пропускної здатності каналу передачі даних.

Smurf-атака типу ICMP flood, - найбільш небезпечний різновид атаки, працює на основі ICMP. Може проводитися за допомогої звичайної команди "ping", яка використовується для визначення доступності будь-якого хоста посилкою пакета. Пакет може бути відправлений за адресою широкомовного транслюємого запиту по всій мережі. Як наслідок, на широкомовний пакет дадуть відповідь кілька тисяч машин, то комп’ютер-ініціатор може не впоратись з обробкою ECHO-відповідей.

6. Технічні засоби захисту інформації та аналіз методів кіберзахисту

Загалом технічними засобами захист інформації забезпечують, коли:

– джерело і носій інформації локалізовані в межах об'єкта захисту і містять механічні перешкоди від контакту з ними зловмисника чи дистанційного впливу на них полів його технічних засобів добування інформації;

– співвідношення енергії носія і перешкод на виході приймача каналу витоку таке, що зловмисникові не вдасться зняти з носія інформацію належної якості;

– зловмисник не може знайти джерело чи носій інформації;

– замість справжньої зловмисник приймає несправжню інформацію, котру він оцінює як справжню.

Ці варіанти реалізують такими методами захисту що зображено на рис. 5.

Способи захисту на основі інженерних конструкцій у поєднанні з технічними засобами охорони також поширені. Разом вони утворюють так званий фізичний захист. Але цей термін не можна вважати вдалим, оскільки інші методи захисту інформації за допомогою технічних засобів також ґрунтуються на законах фізики. З огляду на те, що основу розглянутого методу становлять інженерні конструкції і технічні засоби

охорони, доцільно визначити його як інженерний захист і технічна охорона об'єктів (ІЗТОО).

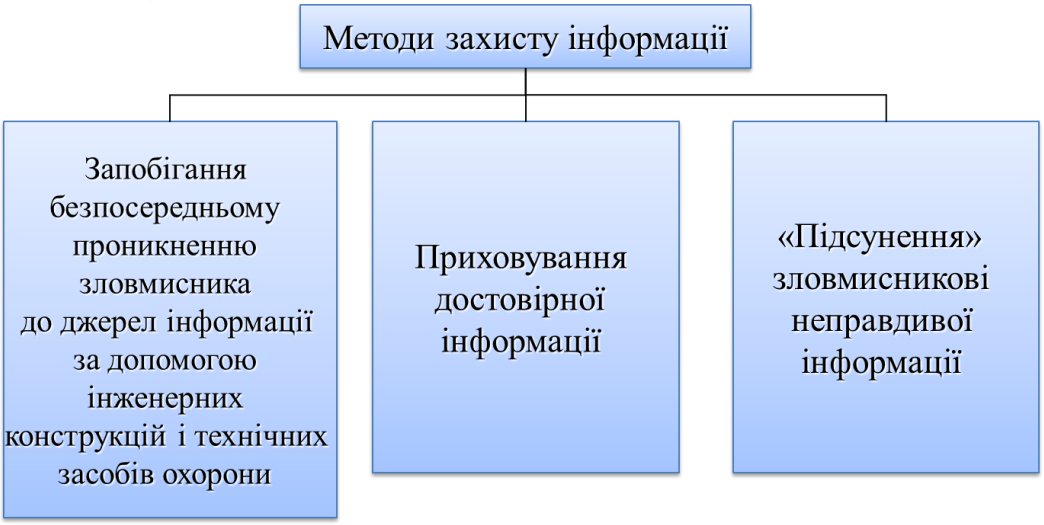


Рис. 5. Методи захисту інформації

Канали витоку інформації по фізичних принципах можна класифікувати на такі групи:

– акустичні (включаючи вібраційні і акустоперетворювальні);

– візуально-оптичні (спостереження, фотографування);

– електромагнітні (у тому числі магнітні, електричні і параметричні);

– матеріально-речові (папір, фото, магнітні носії, відходи і т.п.);

– комп'ютерний метод знімання (віруси, закладки, в тому числі «back door», «троянські коні», логічні бомби);

– перехоплення при передачі по каналах зв'язку (передача даних, аудіо- та відеоінформації).

7. Типи основних кібератак

Кібератаки мають наступну класифікацію зображену на рис. 6, які залежать від відповідної характеристики атаки.

Дана класифікація кібератак дозволяє визначити напрямки подальших досліджень щодо розробки методів та побудови ефективних систем захисту інформації. Розглянемо кібератаки по об’єкту атаки, їх класифікація зображена на рис. 7.

Рис. 6. Класифікація кібератак

Рис. 7. Класифікація кібератак по об’єкту атаки

8. Міжмережевий екран, як засіб кіберзахисту телекомунікаційних систем

Коли користувач виходить в Інтернет, незалежно від того, через кабель чи через Wi-Fi, він може не уявляти, що наражає свій комп’ютер або смартфон на небезпеку. Існує статистика за якою, незахищений комп’ютер (без оновлень і фаєрволу) – може підхопити вірус в середньому протягом п’яти хвилин після виходу в інтернет.

Міжмережевий екран (його іноді ще називають "брандмауер" або "файервол") - це програма-сторож. Міжмережевий екран – це, по суті, фільтр, який стежить за вихідним і вхідним трафіком – тобто, за всім, що Ваш комп'ютер отримує і надсилає, та повідомляє Вас про підозрілі дії, щоб Ви приймали рішення, дозволяти їх чи ні.

Міжмережеві екрани можуть працювати на різних рівнях протоколів моделі OSI. На мережевому рівні виконується фільтрація вхідних і вихідних пакетів по IP-адресам (наприклад, не пропускаються пакети з мережі Internet, які направлені на ті сервери, доступ до яких зовні заборонено). На транспортному рівні фільтрація відбувається ще й за номерами портів TCP і прапорців, що містяться в пакетах (наприклад, запити на встановлення з'єднання). На прикладному рівні виконується аналіз прикладних протоколів (FTP, HTTP, SMTP) і контроль за змістом потоків даних (заборона внутрішнім абонентам на отримання будь-яких типів файлів: рекламної інформації або виконуваних програмних модулів).

Однак, слід пам’ятати, що брандмауер не може захистити Вас, якщо Ви самі завантажуєте шкідливу програму чи вірус, а також не зможе захистити від проникнення через вразливості ПЗ. Щоб вирішити цю проблему необхідно обов’язково використовувати антивірус та оновлювати ПЗ.

Існують два види міжмережевих екранів:

– програмні;

– апаратні.

Переваги програмних екранів:

– низька вартість;

– високі функціональні можливості;

– висока гнучкість при конфігуруванні і модернізації.

Недоліки:

– злом системи може бути проведений через уразливості ОС;

– висока вартість повного вирішення, яка включає: вартість файрвольного ПО, вартість ОС, вартість додаткового ПЗ (БД, інтерпретатори, драйвера), вартість апаратної платформи.

– нижче надійність, так як в апаратних платформах є рухомі і магнітні елементи;

– можливий доступ до системи через зовнішні порти та приводи;

– потрібно більш високий рівень кваліфікації обслуговуючого персоналу.

У апаратних міжмережевих екранів переваги та недоліки зворотні програмним.

Покоління міжмережевих екранів:

- екрани з пакетною фільтрацією;
- екрани рівня з'єднання;
- екрани рівня програми;

– екрани посередники;
– екрани з повною пакетної перевіркою.

9. Міжмережеві екрани з пакетною фільтрацією

Міжмережеві екрани з пакетною фільтрацією можуть також бути програмними пакетами, що базуються на ОС загального призначення (таких як Windows NT і Unix) або на апаратних платформах міжмережевих екранів. Міжмережевий екран має декілька інтерфейсів, по одному на кожну з мереж, до яких підключений екран. Аналогічно міжмережевих екранів прикладного рівня, доставка трафіку з однієї мережі в іншу визначається набором правил політики. Якщо правило не дозволяє явним чином певний трафік, то відповідні пакети будуть відхилені або анульовані міжмережевим екраном.

Правила політики посилюються за допомогою використання фільтрів пакетів. Фільтри вивчають пакети і визначають, чи є трафік дозволеним, згідно з правилами політики і станом протоколу (перевірка з урахуванням стану). Якщо протокол додатка функціонує через TCP, визначити стан відносно просто, так як TCP сам по собі підтримує стан. Це означає, що коли протокол знаходиться в певному стані, дозволена передача тільки певних пакетів.

Міжмережеві екрани з фільтрацією пакетів не використовують модулі доступу для кожного протоколу і тому можуть використовуватися з будь-яким протоколом, працюючим через IP. Деякі протоколи вимагають розпізнавання міжмережевим екраном виконуваних ним дій.

Як правило, міжмережеві екрани з фільтрацією пакетів мають можливість підтримки більшого обсягу трафіку, тому в них відсутнє навантаження, створювана додатковими процедурами налаштування та обчислення, що мають місце в програмних модулях доступу.

Завдання, що виконуються екранами з пакетною фільтрацією:
– обмеження трафіку на підставі інформації 3 і 4 рівнів OSI;
– поділ мережі на зони з різним рівнем довіри.

Приклад створення пакетних фільтрів для міжмережевих екранів з пакетною фільтрацією наведено на рис. 8.

10. Міжмережеві екрани рівня з'єднання

Наведено приклад того, як працюють екрани рівня з'єднання. Скажімо, комп'ютер «А» знаходиться в мережі, захищеній брандмауером рівня з'єднання, і хоче переглянути веб-сторінку на комп'ютері «В», який за межами брандмауера. Комп'ютер «А» надсилає запит на веб-сторінку до комп'ютеру «В», який перехоплюється і записується брандмауером перед передачею. Комп'ютер «В» приймає запит, який прийшов з адреси брандмауера, і починає посилати дані веб-сторінки назад через Інтернет. Коли він досягає брандмауеру, його порівнюють з проханням комп'ютера

«А», щоб побачити, чи IP-адреса та порт збігаються, тоді дані або передаються або блокуються.

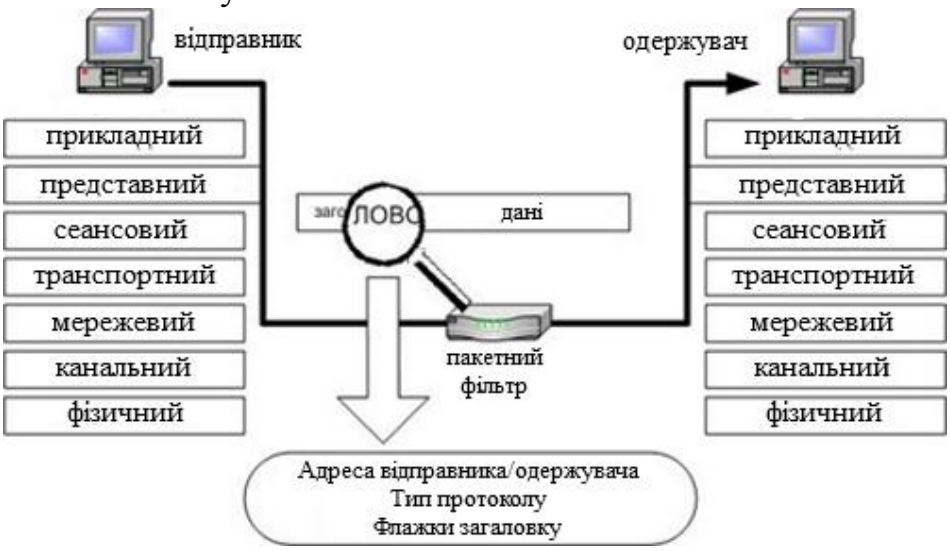


Рис. 8. Міжмережеві екрани з пакетною фільтрацією

ПЗ або апаратні брандмауери, що використовують екрани рівня з'єднання будуть також включати в себе деякий спосіб спільного використання Інтернет, так як це є частиною функції цього типу брандмауера. Як ви можете здогадатися, кабельні / DSL домашні маршрутизатори використовують цей метод в першу чергу. Більш конкретно, вони використовують переклад мережевих адрес, який являє собою поєднання функцій екрану рівня з'єднання з спільним доступом до Інтернету.

Основні завдання міжмережевих екранів рівня з'єднання:

– відстеження коректності встановлених сесій протоколів 4-го рівня;

– захист від DoS (Denial of Service, відмова в обслуговуванні) атак - обмеження на кількість одночасних сесій, час життя сесії, кількості встановлюваних сесій за проміжок часу і т.д. (рис. 9)

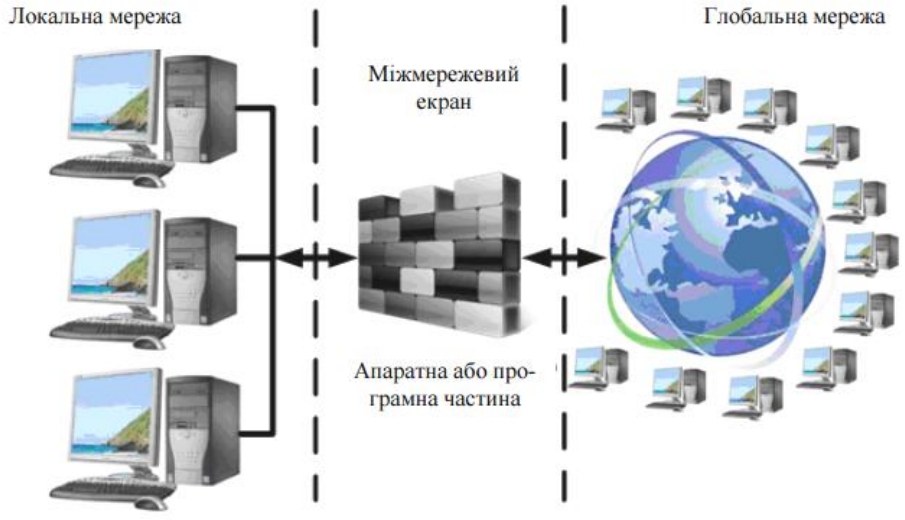


Рис. 9. Міжмережеві екрани рівня з'єднання

Плюси:

– відносно високий рівень захисту (в порівнянні з пакетними файрволом);

– «прозорість» для додатків;

– відносно висока продуктивність (у порівнянні з екранами рівня додатків);

– масштабованість;

– низька вартість (практично така ж як у пакетних файрволів).

Недоліком є те, що не аналізується інформація протоколів більш високого рівня.

11. Міжмережеві екрани прикладного рівня

Application-level gateway (ALG) - шлюз прикладного рівня - компонент NAT-маршрутизатора (Network Address Translation), який розуміє який-небудь прикладний протокол, і при проходженні через нього пакетів цього протоколу модифікує їх таким чином, що користувачі за Network Address Translation - «перетворення мережевих адресів» (NAT) можуть користуватися протоколом.

NAT-маршрутизатор ретранслює пакети, що надійшли зсередини локальної мережі і направляє їх в зовнішню мережу, використовуючи свій зовнішній IP-адрес, як адрес відправника. Також може підмінятися порт. Але деякі мережеві протоколи у змісті своїх пакетів передають і намагаються використовувати локальний IP-адрес або порт відправника. Звичайно ж, після проходження NAT і підміни локального IP на зовнішній ці параметри стають невірними - а значить, віддалена сторона не може налагодити з'єднання.

Application-level gateway - «шлюз прикладного рівня» (ALG), ідентифікувавши пакет як такий, що відноситься до даного протоколу, підставляє в якості IP-адреси і порту свої адресу і порт.

ALG схожий на проксі-сервер; зазвичай поняттям «проксі-сервер» називають сервер, який виконує додаткові операції на зразок кешування, в той час як завдання ALG - забезпечити, щоб клієнти могли користуватися протоколом.

Завдання міжмережевих екранів прикладного рівня (рис. 10) наступні:

– відстеження коректності роботи протоколів більш високих рівнів, наприклад відстеження команд протоколів HTTP, FTP, SMTP, POP3, SNMP і термінація сесії у разі неправильного порядку команд, можливість блокування певних команд та ін.;

– динамічне відкриття необхідних портів для складних протоколів;

– обмеження доступу до URL ресурсів на основі бази, масок та аналізу сторінок, блокування завантаження або отримання поштою певних файлів;

– блокування певних Java і ActiveX аплетів і скриптів.

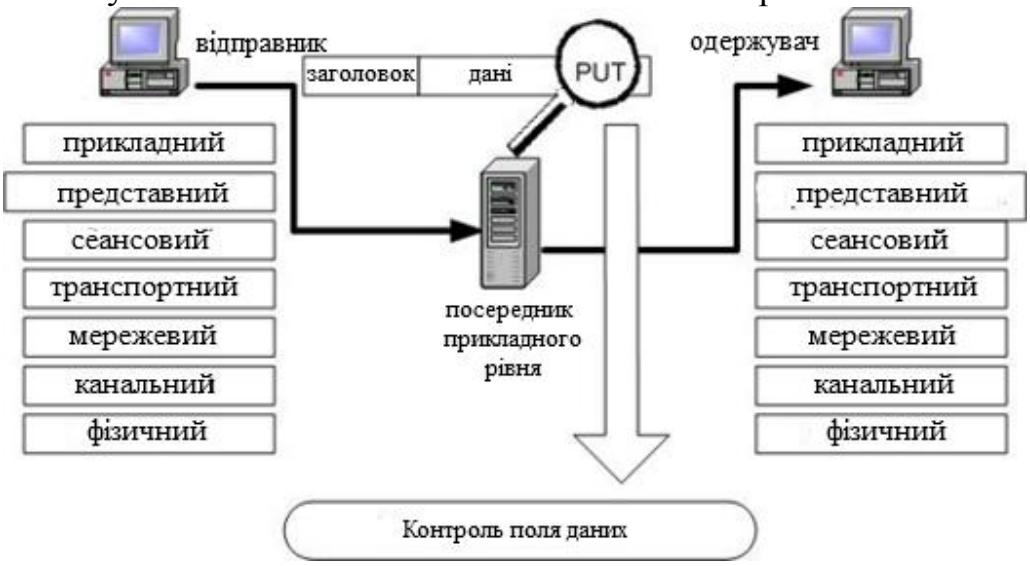


Рис. 10. Міжмережеві екрани прикладного рівня

12. Міжмережеві екрани посередники (Proxy)

Можуть застосовуватися для протоколів 7-го або 4-го рівнів, тобто працювати з екранами рівнів програми або з'єднання.

Особливості використання екранів посередників:

– використовується програма-посередник;

– спочатку встановлюється сесія між хостом-джерелом і програмою-посередником, а потім програма-посередник встановлює сесію до хосту-призначення;

– проксі-програма виконує дії (заборонити або дозволити) на основі політик, встановлених користувачем;

– кожен протокол 7-го або 4-го рівня має свою власну проксі-програму.

Головним плюсом даного механізму – є високий рівень безпеки для підтримуваного протоколу. Мінуси:

– порушується модель клієнт-сервер - кожне з'єднання «клієнт-сервер» вимагає двох з’єднань: одне від клієнта до міжмережевого екрану й інше від міжмережевого екрану до клієнта;

– низька продуктивність;

– висока вартість.

Загальний принцип функціонування екрану посередника схематично показаний на рис. 11.

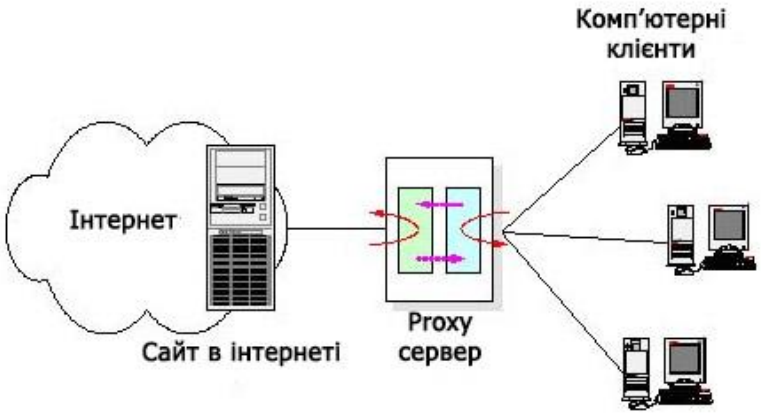


Рис. 11. Принцип функціонування екрану посередника

Висновки

В результаті аналізу видів кібератак було визначено, що найбільш доцільно захищати телекомунікаційні системи та мережі, від атак на відмову, які виконуються через відправку спеціальних пакетів по протоколах TCP.

Для захисту від цих кібератак реалізуються такі методи захисту:

– запобігання безпосередньому проникненню зловмисника до джерел інформації за допомогою інженерних конструкцій і технічних засобів охорони;

– приховування достовірної інформації;

– «підсунення» зловмисникові неправдивої інформації.

Для реалізації цих методів існують такі засоби: міжмережеві екрани, антивіруси, віртуальні приватні мережі та механізми виявлення та запобігання вторгненню.

Література

1. Юдін О.К. Кодування в інформаційно-комунікаційних мережах: Монографія.- К.: Книжкове видавництво НАУ, 2007. – 302 с.

2. Ларін В.В., Комолов Д.С., Ялівець К.В., Гаврилов Д.С. Метод захисту низькочастотних складових в алгоритмі кодування JPEG. - Системи обробки інформації. – 2015. - № 9 (134). С. 121 – 123.

3. Larin V., Krasnikov P., Gavrulov D. The analysis of template method of video processing. Proceedings of 2015 1st International Conference on Advanced Information and Communication Technologies-2015 (AICT'2015), Lviv, Ukraine, October 29 – November 1, 2015. – P. 87 – 89.

4. Захист інформації в автоматизованих системах управління: навч. посіб. / Ю.В. Стасєв, О.А. Смірнов, В.В. Бараннік; за ред. Ю.В Стасєва. – Х.: ХУПС. – 2015 – 264 с.

МЕТОД АВТЕНТИФІКАЦІЇ У БЕЗДРОТОВИХ МЕРЕЖАХ НА ОСНОВІ МОДЕЛІ ДОВІРИ

Лужецький В. А., Войтович О. П., Шулятицька О. О.

Вступ

Бездротові сенсорні мережі (БСМ) все активніше проникають у різні сфери діяльності людини. Сенсорні мережі, як частина Інтернету речей (Internet of Things), займають важливе місце в мережах зв'язку [1, 2]. Проте багато досліджень показують проблеми безпеки, що виникають при експлуатації цих систем [3, 4], зокрема: небезпечні комунікації, слабка автентифікація, відсутність керування доступом, слабкі криптографічні алгоритми, обмін бібліотеками та оновленнями тощо. Більшість цих проблем зумовлено, зокрема такою важливою особливістю бездротових сенсорних мереж як робота від автономних джерел живлення, і питання про тривалість роботи такої мережі стоїть на першому місці. Відповідно метою є покращення безпеки за рахунок аналізу та систематизації відомих методів передачі даних та автентифікації у БСМ, а також розробка методу автентифікації вузлів, який би дозволив при підвищенні рівня безпеки не витрачати додатково велику кількість ресурсів.

Поставлена мета передбачає вирішення таких задач: аналіз протоколів, які використовуються в сенсорних мережах, протоколів маршрутизації, в тому числі енергоефективних; аналіз протоколів автентифікації, що використовуються в сучасних БСМ; розробка методу автентифікації, що дозволить реалізувати захист передаваної інформації при економії енергоресурсів, моделювання запропонованого методу.

1. Класифікація БСМ

Бездротові сенсорні мережі будуються з сенсорів (мотів), які складаються з первинних перетворювачів (температури, освітлення, шуму, тиску тощо), мікроконтролерів, та пристроїв приймання і передавання сигналів, забезпечуючи значну площу покриття системи при малій потужності передавачів.

Класифікація БСМ наведена на рис. 1.

Розрізняють сенсорні мережі за типом живлення: з автономним живленням (безпроводні), підключені до мережі живлення (проводні) та відновлювальні (наприклад від сонячних батарей) [5].

За протоколом передавання даних БСМ можуть поділятися на: ZigBee, Bluetooth, WiFi, 6LoWPAN та UWB. Найбільш відомим з протоколів є протокол ZigBee WSN [6]. Для того, щоб розробити стандартний стек протоколів для бездротових сенсорних мереж ZigBee Alliance, використовуваних раніше, розроблений стандарт IEEE 802.15.4, який описує фізичний рівень і рівень доступу до середовища для бездротових мережі передачі даних на короткі відстані (до 75 м) з низьким

енергоспоживанням, але з високим ступенем надійності. Стандарт IEEE 802.15.4 є базовою основою для протоколу ZigBee.

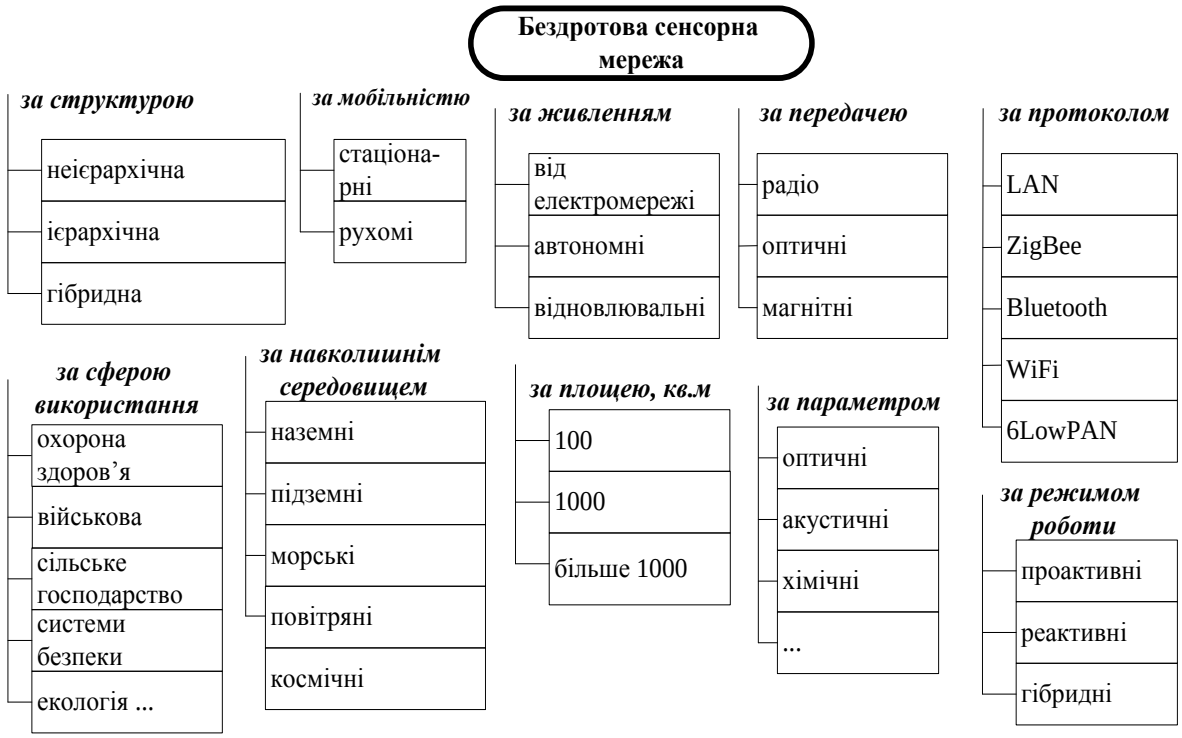


Рис. 1. Класифікація бездротових сенсорних мереж

БСМ також реалізуються з використанням технології бездротового зв'язку Bluetooth [7]. Ці мережі складаються з пристрою, що приймає та майстер-пристроїв (ролі можуть бути об'єднані), здатні передавати дані в обох синхронних і асинхронних режимах. Спеціально для БСМ були запропонована енергоекономна версія в специфікації v.4.0 бездротової технології Bluetooth, що отримала скорочення BLE.

Для реалізації БСМ використовується набір стандартів IEEE 802.11 (Wi-Fi) зв'язку. Протокол IEEE 802.11s може бути використаний для ієрархічної організації вузькоспеціалізованих бездротових мереж з мобільними і статичними вузлами (mesh-network) [8].

БСМ реалізуються також на основі технології бездротового зв'язку на коротких відстанях при низькій енергії UWB (Ultra-Wide Band) [1].

Концепція 6LoWPAN (взаємодія по IPv6 над малопотужними бездротовими персональними мережами) отримана від ідеї про те, що IP є більш надійним протоколом, зокрема, для маленьких пристроїв, які мають малу потужність і обмежені можливості обробки. В протоколі 6LoWPAN визначені механізми інкапсуляції і ущільнення заголовків, які дозволяють пакетам IPv6 бути відправленими та отриманими поверх протоколу IEEE 802.15.4. Основними споживачами цих мереж є додатки, яким необхідне бездротове підключення до Інтернет на більш низьких швидкостях передачі даних та для пристроїв з обмеженим можливостями [9].

За мобільністю розрізняють стаціонарні і рухомі сенсорні мережі. Залежно від сфери застосування можливо використовувати мобільні (роботи) сенсорні мережі і стаціонарні. Мобільні сенсорні мережі є окремим випадком MANET [10].

За типом архітектури розрізняють такі сенсорні мережі:

- неієрархічна мережа однотипних сенсорів. Сенсори здійснюють розподілену обробку даних, при цьому передача даних здійснюється з використанням багатократних ретрансляцій (маршрутизацій). Один або декілька сенсорів мають бездротовий доступ до зовнішнього шлюзу, який у свою чергу пов'язаний з іншими мережами (дротовими, стільниковими тощо);

- ієрархічна кластеризована мережа неоднотипних сенсорів. Однотипні відео-, аудіо, і скалярні сенсори передають дані до головного вузла (кластеру), який відповідає за виконання централізованої обробки даних. Головний вузол передає зібрану інформацію через зовнішній шлюз до концентратора пам'яті і на центральну станцію;

- гібридна мережа з однотипними і неоднотипними сенсорами. Кожен шар виконує різні функції. Обмежені ресурсом, однотипні сенсори з низьким енергоспоживанням відповідають за виконання простіших завдань (виявлення скалярних фізичних розмірів тощо), тоді як сенсорні пристрої з великою потужністю відповідальні за рішення комплексних задач (обробка, передача інформації). Зберігання і обробка даних може бути виконана розподіленим способом в кожному різному шарі [1].

За середовищем моніторингу мережі поділяються на: наземні, підземні, морські, повітряні тощо [1, 9], а за параметрами моніторингу на: акустичні, хімічні та інші. Залежно від середовища моніторингу в сенсорах використовують первинні перетворювачі і мікрокомп'ютери, які реєструють певні параметри (наприклад, рівень радіації на забрудненій території).

За режимом роботи виокремлюють: проактивні, реактивні, гібридні БСМ. Вузли проактивної мережі періодично вмикають свої передавачі і сенсори, знімають показання і передають їх на базову станцію, таким чином вони роблять «моментальне фото» свого оточення з деякою періодичністю і використовуються зазвичай для додатків, які потребують регулярного моніторингу деяких значень. Вузли реактивних мереж з деякою періодичністю знімають показання, але передають їх якщо отримані дані потрапляють у визначену область нормальних показників. В той же час відомості про неочікувані і різні зміни в показниках сенсорів або їх вихід за діапазон нормальних значень негайно передаються на базову станцію. Цей вид мережі призначений для роботи з додатками реального часу. Гібридні мережі – це комбінація двох вище перерахованих типів, де сенсорні вузли не тільки періодично відправляють зняті дані, але і реагують на різкі зміни в значеннях.

Одним з найважливіших параметрів, який впливає на застосування та розвиток тих чи інших механізмів безпеки, що можуть застосовуватись у БСМ, є споживані ресурси. У сенсорів з автономним джерелом живлення є суттєві обмеження на використання ресурсів, а отже і обмеження на використання механізмів захисту з ресурсоємними вимогами.

2. Аналіз протоколів маршрутизації в БСМ

Для визначення маршруту передачі інформації в БСМ від кінцевого вузла до вузла координатора, а також між кінцевими вузлами, використовуються спеціальні протоколи маршрутизації. Протоколи маршрутизації в БСМ вирішують наступні завдання [1]:

1. Самоорганізація вузлів мережі (самоконфігурування, самовідновлення та оптимізація).
2. Маршрутизація пакетів даних і адресація вузлів.
3. Мінімізація енергоспоживання вузлів мережі і збільшення загального часу життя всієї мережі.
4. Збір і агрегація даних.
5. Регулювання швидкості передачі і обробки даних в мережі.
6. Максимізація зони покриття мережі.
7. Забезпечення заданої якості обслуговування (QoS).
8. Захист від несанкціонованого доступу.

При виборі шляху передачі інформації в мережі як метрика в них можуть бути використані такі параметри [2]:

- довжина шляху (кількість ділянок прийому інформації);
- надійність;
- затримка;
- пропускна здатність;
- завантаження;
- вартість передачі трафіку і ін.

Протоколи маршрутизації БСМ відповідають за підтримку маршрутів у мережі і повинні гарантувати надійний зв'язок навіть в жорстких несприятливих умовах. Багато протоколів маршрутизації були спеціально розроблені для БСМ, де енергозбереження є суттєвою проблемою, на вирішення якої спрямовано протокол. Інші ж були розроблені для загального застосування в бездротових мережах, але знайшли своє застосування і в БСМ.

Існує велика кількість протоколів маршрутизації для БСМ, їх класифікація наведена на рис. 2. Залежно від використовуваного режиму роботи мережі, що обумовлює необхідність передачі інформації від вузлів, усі протоколи маршрутизації поділяю на проактивні (всі шляхи визначаються заздалегідь, до того як вони будуть потрібні), реактивні (шлях визначаються на вимогу) та гібридні (комбінація перших двох).

Рис. 2. Класифікація протоколів маршрутизації БСМ

Протоколи, що враховують структуру мережі, діляться на:

1) протоколи однорівневої (плоскої) (flat-based) маршрутизації – всі вузли БСМ мають однакову функціональність;

2) протоколи ієрархічної (hierarchical-based) маршрутизації – вузли мережі виконують різні функції, вони можуть бути і фізично різними;

3) протоколи маршрутизації на основі інформації про місцезнаходження вузла (location-based).

Робота протоколу маршрутизації може ґрунтуватися на різних операціях:

1) на основі шляху (multipath routing) – використовуються кілька маршрутів від джерела до точки призначення, що підвищує надійність з'єднання, але збільшує накладні витрати і енерговитрати;

2) на основі запитів (query-based) – вузол посилає запит на дані в мережу, і інший вузол, який має запитувані дані, відповідає на запит;

3) на основі «переговорів» (negotiation routing) між вузлами;

4) на основі якості обслуговування (QoS-based), що дозволяє забезпечити певний рівень послуг в мережі;

5) когерентні та некогерентні.

У протоколах, спрямованих на агрегацію даних, проміжні вузли, що розташовуються між джерелами інформації та базовою станцією, можуть здійснювати агрегацію даних і посилати базовій станції вже зведені дані. Цей процес дозволяє сенсорним вузлів економити енергію.

Усі протоколи маршрутизації також можна розділити на два види - в одних ініціатором з'єднання є джерело інформації, а в інших - одержувач.

Класифікація протоколів маршрутизації БСМ на основі типів вузлів показана на рис.3.

Отже, проаналізовано протоколи маршрутизації, за якими можна визначити оптимальний маршрут передачі даних. Оскільки дані треба не тільки передати, але й при цьому зберегти енергію сенсору, необхідно розглянути протоколи, які спрямовані на ефективне енергоспоживання.

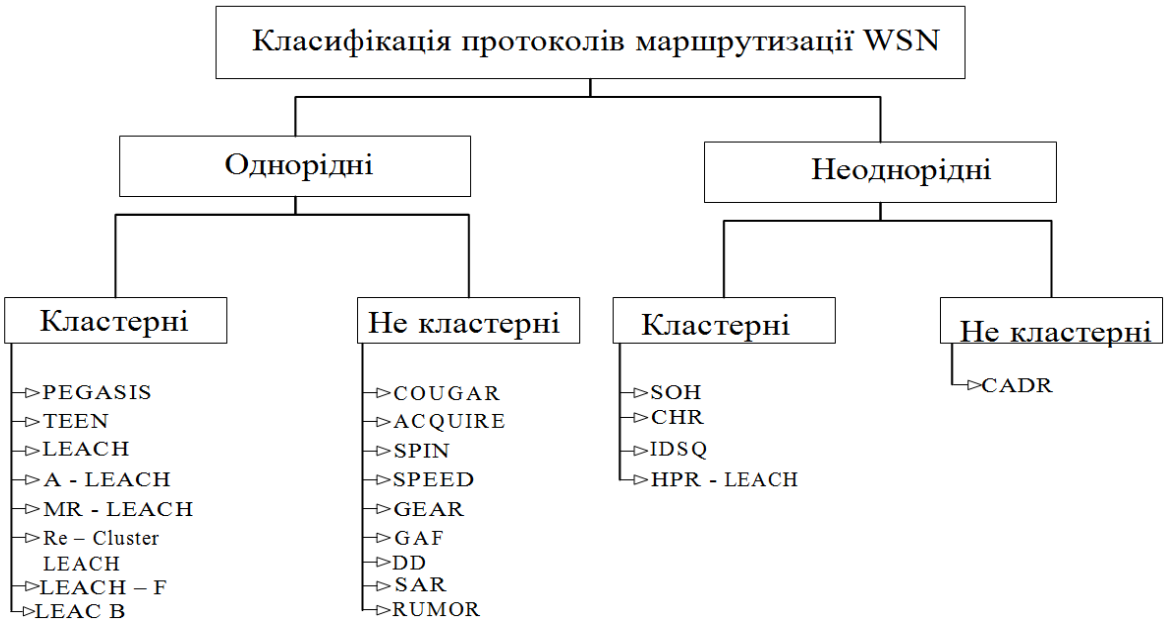


Рис. 3. Класифікація протоколів маршрутизації БСМ на основі типів вузлів

3. Характеристика енергоефективних протоколів БСМ

Існує велика кількість протоколів, які використовують БСМ, але невелика частина з них складає саме ті, які забезпечують енергоефективне споживання потужності вузла мережі. Характеристика енергоефективних протоколів наведена в табл. 1.

Таблиця 1

Характеристика енергоефективних протоколів

Протокол	**Характеристика**
An Energy Efficient Neighbour Node Discovery Method for Wireless Sensor Networks [10]	Точно виявляє сусідні вузли та керує живленням, значно скорочує використання енергії. Оскільки протокол маршрутизації динамічний, побудований на основі адміністративного значення відстані, призначеної для кожного шляху в мережі, це значно зменшує кінцеву затримку та накладні витрати протоколу маршрутизації.
An Energy-Efficient Unicast Routing Protocol for Wireless Sensor Networks [11]	Розроблений для уникнення заторів на конкретних вузлах при передачі даних та збалансованого споживання енергії вузла, що збільшує термін служби мережі. При прокладанні маршруту враховується залишкова енергія сенсорних вузлів, для уникнення незбалансованого споживання енергії вузлів.

Продовження таблиці 1

Протокол	Характеристика
A Review on Energy Efficient Routing in Wireless Sensor Networks [12]	Протокол є ефективним з точки зору використання енергії всієї мережі, оскільки окремі частини не будуть відокремлені через енергетичне виснаження проміжних вузлів. Враховує віддаленість вузлів один від одного, за допомогою GPS-приймачів або інших способів визначення позиції.
An Energy Aware WSN Geographic Routing Protocol [13]	Маршрутизація базується на двох параметрах: місцезнаходженні та енергетичному рівні вузлів. Кожен вузол знає про місцезнаходження і енергетичний рівень своїх сусідів. Це дозволяє продовжити термін служби сенсорів, а отже час життя мережі.
EE-DSR: Energy Efficient Dynamic Source Routing in Wireless Ad Hoc Networks Using Residual Energy [14]	Використовується як базова модель з меншою пропускною здатністю і меншим використанням енергії. Існуючі ефективні алгоритми маршрутизації енергії, а саме RMER (надійна маршрутизація мінімальною енергією) і RMECR (надійна маршрутизація мінімальна вартість енергії) посилюються і реалізуються в E-DSR.
Location based energy-efficient reliable routing protocol for wireless sensor networks [15]	Енергоефективний надійний протокол маршрутизації на основі положення вузла. Географічні протоколи маршрутизації є ефективними і зручними для оптимального споживання енергії та використання смуги пропускання. Більшість існуючих географічних протоколів маршрутизації використовують жадібні алгоритми маршрутизації для пересилання пакетів від джерела до місця призначення.

Аналіз базових енергоефективних протоколів показав, що вони забезпечують швидку та безпечну передачу при цьому зберігаючи рівень потужності, отже їх можна використовувати для енергоефективної автентифікації.

4. Алгоритми автентифікації

Щоб забезпечити необхідний рівень безпеки використовують стійкі методи автентифікації вузлів у мережі. На сьогоднішній день існує низка схем автентифікації в БСМ, основними з яких є:

- на основі протоколів шифрування;
- на основі ієрархії;
- на основі довірчого центру.

Характеристики методів автентифікації наведена в табл. 2.

Таким чином, у вище зазначених методів автентифікації є такі проблеми:

- можливість зламу мережі шляхом попереднього перегляду трафіку, зміни даних, підміни вузла, коли для зменшення енергоспоживання використовують слабкі методи захисту.

- використання складних криптографічних схем автентифікації призводить до значного споживання енергії, і як наслідок, батарея потребує частої підзарядки або заміни.

- вдала атака на довірений центр призводить до компрометації всієї мережі або до зупинки її функціонування.

Таблиця 2

Характеристики методів автентифікації

Метод автентифікації	Характеристика
An Improved Dynamic User Authentication Scheme for Wireless Sensor Networks [16]	Переваги 1. Динамічна схема автентифікації дозволяє авторизованим користувачам запитувати будь-які сенсорні вузли і має невелике обчислювальне навантаження. 2. Реєстрація і фаза зміни паролів відбувається через захищений канал.
	Недоліки 1. Не може протистояти атаці replay (перегравання) 2. Не може протистояти атаці (forgery) фальсифікації 3. Паролі можуть бути виявлені за допомогою будь-якого з вузлів. Немає механізму захисту паролів у фазах авторизації та автентифікації. 4. Великі витрати енергії
An Efficient Broadcast Authentication Scheme in Wireless Sensor Networks [17]	Переваги 1. Схема ефективно знижує потребу в зберіганні даних і передбачає повторні маніпуляції механізмів підписання додаткових повідомлень. 2. Метод базується на однонаправлених функціях з невеликою складністю обчислень, порівняно з асиметричними примітивами. 3. Відсутня синхронізації за часом 4. Не потрібна буферизація приймача. 5. Індивідуальна перевірка справжності. 6. Система миттєвої автентифікації повідомлень.
	Недоліки 1. Час підписання залежить від розміра відкритого ключа. 2. Призводить до швидкого зменшення енергії батареї.
An Enhanced Trust Center Based Authentication in ZigBee Networks [18]	Переваги 1. Проста і недорога автентифікація. 2. Енергоефективна для малих мереж. 3. Ефективна пам’ять.

Продовження таблиці 2

Метод автентифікації	Характеристика
An Enhanced Trust Center Based Authentication in ZigBee Networks [18]	Недоліки 1. При збільшенні мережі збільшуються витрати енергії на взаємодію вузлів. 2. Наявність третьої сторони. 3. Слабка стійкість мережевого ключа. 4. При атаці на довірений центр мережа не буде функціонувати.
Design of authentication protocol for LR-WPAN using pre-authentication mechanism [19]	Переваги 1. Кожен сенсор посилає тільки випадкові числа і хеш-значення. Довірений центр використовує шифрування при відсиланні мережевого ключа, що унеможливлює отримання загального секрету за рахунок підслуховування. 2. Кожен сенсор генерує випадкові числа, і використовує їх для розрахунку MacTag для кожного процесу автентифікації, чим запобігає атаці перенаправлення. 3. Кожен сенсор пов'язаний з перевірками справжності MacTag, щоб запобігти атаці модифікації повідомлення.
	Недоліки 1. При вдалій атаці на довірений центр компрометується вся мережа. 2. Кількість обмінів даними між довіреним центром та вузлом збільшується. 3. Може використовуватись тільки для мереж з постійним живленням.
Lightweight and provably secure user authentication with Anonymity for the global mobility network [20]	Переваги 1. Захист від повторного посилання. 2. Запобігання шахрайству. 3. Протистоїть атаці оффлайн вгадування пароля. 4. Протистоїть атаці «людини в середині»
	Недоліки 1. Складна криптографія. 2. Потребує багато енергії на криптографічні перетворення та складні обчислення, а отже застосовується тільки для мереж з постійним живленням.
Concept of Designing the Wireless Sensor Networks Based on Ant Intelligence [21]	Переваги 1. Використання для мереж із низьким електроживленням. 2. Дозволяє процес відновлення (збільшення) мережі
	Недоліки 1. Швидке зниження енергії батареї вузла, коли багато вузлів передають дані по одному і тому ж оптимальному маршруту.

Для усунення вказаних проблем доцільно використовувати методи енергоефективної маршрутизації з деякими змінами, що мають забезпечити як належний рівень захисту БСМ, так і скорочення споживання енергії.

5. Метод автентифікації на основі моделі довіри

Пропонується метод автентифікації на основі моделі довіри, який передбачає використання простої процедури автентифікації, що виключає існування довіреного центру та зменшує споживання енергії.

На рис. 4 схематично показана взаємодія вузлів БСМ підчас автентифікації за запропонованим методом. Тут сірим кольором позначено вузол-передавач, а чорним вузол-приймач.

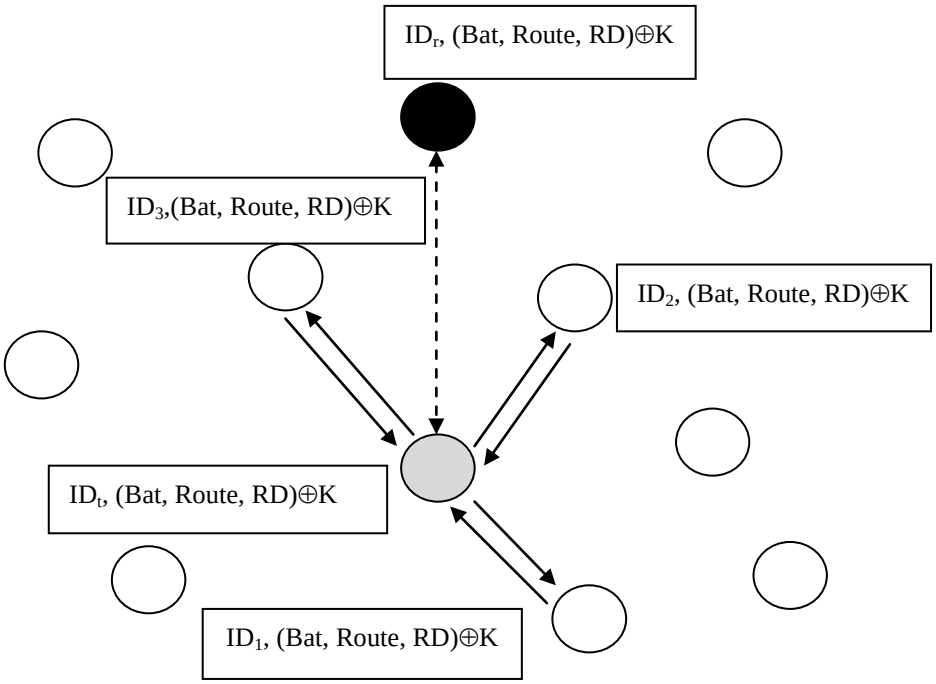


Рис. 4. Схема взаємодії вузлів БСМ

Автентифікація реалізується шляхом опитування як мінімум трьох найближчих вузлів про їх довіру до четвертого вузла. В енергозберігаючих протоколах маршрутизації [16-21], вузол відсилає запит сусіднім вузлам про готовність до передачі даних (рівень заряду батареї, характеристика маршруту, наприклад відстань між вузлами, тощо) для побудови оптимального маршруту передачі даних. У табл. 3 описані символи, використовувані в схемі.

Нехай, вузол має невелику базу даних, що містить як мінімум три ідентифікатори сусідніх вузлів для подальшої автентифікації і визначення рівня довіри до інших вузлів.

Метод автентифікації, що пропонується передбачає виконання пересилань даних, які показано на рис. 5.

Таблиця 3

Опис введених позначень

Позначення	Опис
ID	Ідентифікатор
K	Ключ
RD	Рівень довіри (1 чи 0)
SN	Вузол
Bat	Рівень заряду батареї
Route	Характеристика маршруту

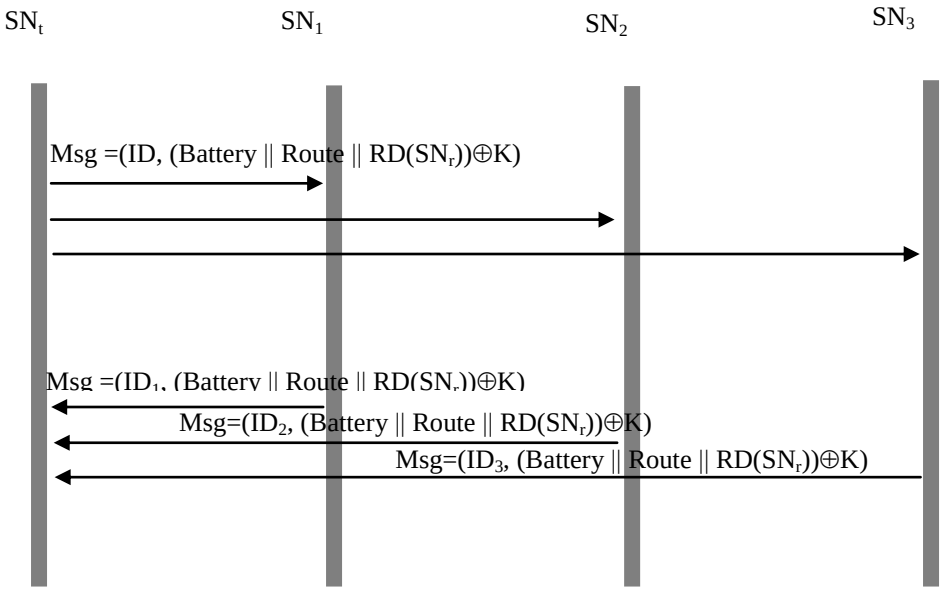


Рис. 5. Схема фази автентифікації

Схема автентифікації - це простий протоколом, який передбачає взаємодію вузла-передавача SN_t, вузла-приймача SN_r та вузлів сусідів SN_i.

Вузол-передавач SN_t, який хоче відправити дані до неавтентифікованого вузла-приймача SN_r, спочатку формує запит до щонайменше трьох довірених вузлів-сусідів про довіру до нього. Запитувані вузли дають позитивну чи негативну відповідь про довіру. Якщо два вузла з трьох відповіли позитивно, то вузол-приймач SN_r стає довіреним.

Таким чином, вузол-передавач SN_t формує висновок про можливість передавання даних через вузол-приймач SN_r. У разі недовіри сусідніх вузлів до вузла SN_r вузол-передавач SN_t опитує інші сусідні вузли з метою пошуку іншого маршруту.

Загальна процедура автентифікації вузлів мережі передбачає реалізацію двох етапів.

1) *Етап реєстрації*

На етапі реєстрації для вузлів в БСМ можна використовувати більш складні схеми автентифікації, якщо вони забезпечують більшу стійкість.

Найближчі вузли беруть участь в схемі автентифікації, і новому вузлу достатньо лише три рази успішно автентифікуватися, щоб бути визнаним довіреним. Оскільки така перевірка справжності виконується невелику кількість разів, економиться енергія як нового вузла, так і мережі в цілому. Як ідентифікатор вузла (або ключ для подальшого обміну даними), може бути використане значення таймера (або іншої випадкової величини). Це дозволить уникнути використання генераторів псевдовипадкових чисел.

2) *Етап автентифікації*

Випадок 1. Сусідні вузли довіряють один одному.

Вузол SN_t опитує сусідні вузли SN_1, SN_2, SN_3 про рівень заряду батареї та характеристику маршруту і формує оптимальний енергоефективний маршрут. Оскільки опитані вузли SN1, SN2, SN3 довіряють один одному і мають спільний ключ K (отриманий на етапі реєстрації) дані передаються передбаченим протоколом.

Випадок 2. Поява нового вузла.

Новий вузол (SN_r), який передбачає участь у передачі даних, повинен спочатку пройти фазу реєстрації для отримання ідентифікатора та спільного ключа, і тільки після цього він зможе взяти участь у роботі мережі.

Крок 1. Вузол SN_t формує запит, що містить такі параметри: ідентифікатор ID_t, ключ K, на якому шифруються всі інші параметри, рівень заряду батареї, характеристика маршруту і рівень довіри RD:

$$Msg = (ID, (Battery \,||\, Route \,||\, RD(SN_r)) \oplus K).$$

Цей запит розсилається трьом найближчим вузлам, щоб дізнатися про довіру до вузла SN_r. При цьому всі дані, що передаються і приймаються, зашифровані з використанням спільного ключа K.

Крок 2. Вузол SN_t отримує відповідь від вузлів SN_1, SN_2, SN_3, які містять ідентифікатори, рівні заряду батареї, характеристики маршруту і рівень довіри до SN_r, та найголовніше, його особистий ідентифікатор. Якщо два з трьох вузлів відповіли про довіряють SN_r, та передають однаковий ідентифікатор вузла, то новий вузол вноситься до бази довірених вузлів і може брати участь у передачі даних. У випадку, коли сусідні вузли дали негативну відповідь (рівень довіри нуль, або значення ідентифікатора нового вузла у різних відповідях не збіглися), то SN_r не має довіри (потенційний зловмисник), і вузол SN_t шукає інший вузол, який може передавати дані, і повторює схему автентифікації.

6. Моделювання

Моделювання відбувалося за допомогою емулятора Atarraya [22] – це емулятор керування подіями, який може бути використаний для навчання і дослідження управління топологією і протоколами БСМ.

Емулятор має можливість моделювання з використанням двох базових топологій: TC – Topology Constructions (топологія будування), TM – Topology Maintenance (топологія підтримання).

Протокол Dynamic Global Energy-based Topology Recreation (DGETRec) базується на тому, що при зниженні заряду батареї вузла до критичного значення, вся мережа змінює топологію передачі даних.

Протокол Energy Local Patching DSR (ELPDSR) базується на тому, що при зниженні заряду батареї вузла до критичного значення, вузол пробуджує свої дочірні вузли для передачі даних, і подальшого функціонування в мережі.

Експеримент полягав у тому, що при однакових параметрах модельованих сенсорних мереж (кількість вузлів – 100, однакові дані, що передаються у мережі, як за складом, так і за розміром тощо, використовуються різні протоколи (DGETRec та ELPDSR). Результати моделювання показали, що у випадку використання протоколу DGETRec, мережа швидко деградувала і перестала передавати дані, в той час як при використанні енергоефективного протоколу ELPDSR мережа продовжувала працювати (рис. 6, 7).

Отже, якщо продовжити термін електроживлення (потужності) кожного з вузлів, то і час служби самої мережі теж продовжиться на доволі великий час, тобто мережа стане більш економною, оскільки не потрібно буде змінювати батарею чи підживлювати заряд кожного з вузлів.

Висновки

Оскільки БСМ знайшли своє широке застосування в Інтернеті речей, який зараз дуже швидко розвивається, то постають дві найголовніші проблеми: безпека даних, що передаються, та швидке зниження енергії батареї вузла (з автономним живленням).

Для вирішення першої проблеми потрібно використовувати принципово нові криптографічні перетворення, які складають малоресурсну криптографію.

Вирішення другої проблеми забезпечується використанням протоколів енергоефективной автентифікації.

Запропонований метод автентифікації заснований на моделі довіри, і використовує строгу автентифікацію тільки на етапі реєстрації. Подальша автентифікація вузлів мережі здійснюється на основі моделі довіри через опитування сусідніх вузлів про рівень довіри до нового вузла, і на основі мажоритарного принципу робиться висновок про можливість передачі даних. Оскільки опитування відбувається в межах стандартного енергоефективного протоколу для пошуку маршруту, то на фазу автентифікації не витрачається додаткова енергія.

Експериментальні дослідження показали, що при використанні енергоефективних протоколів термін служби мережі збільшується.

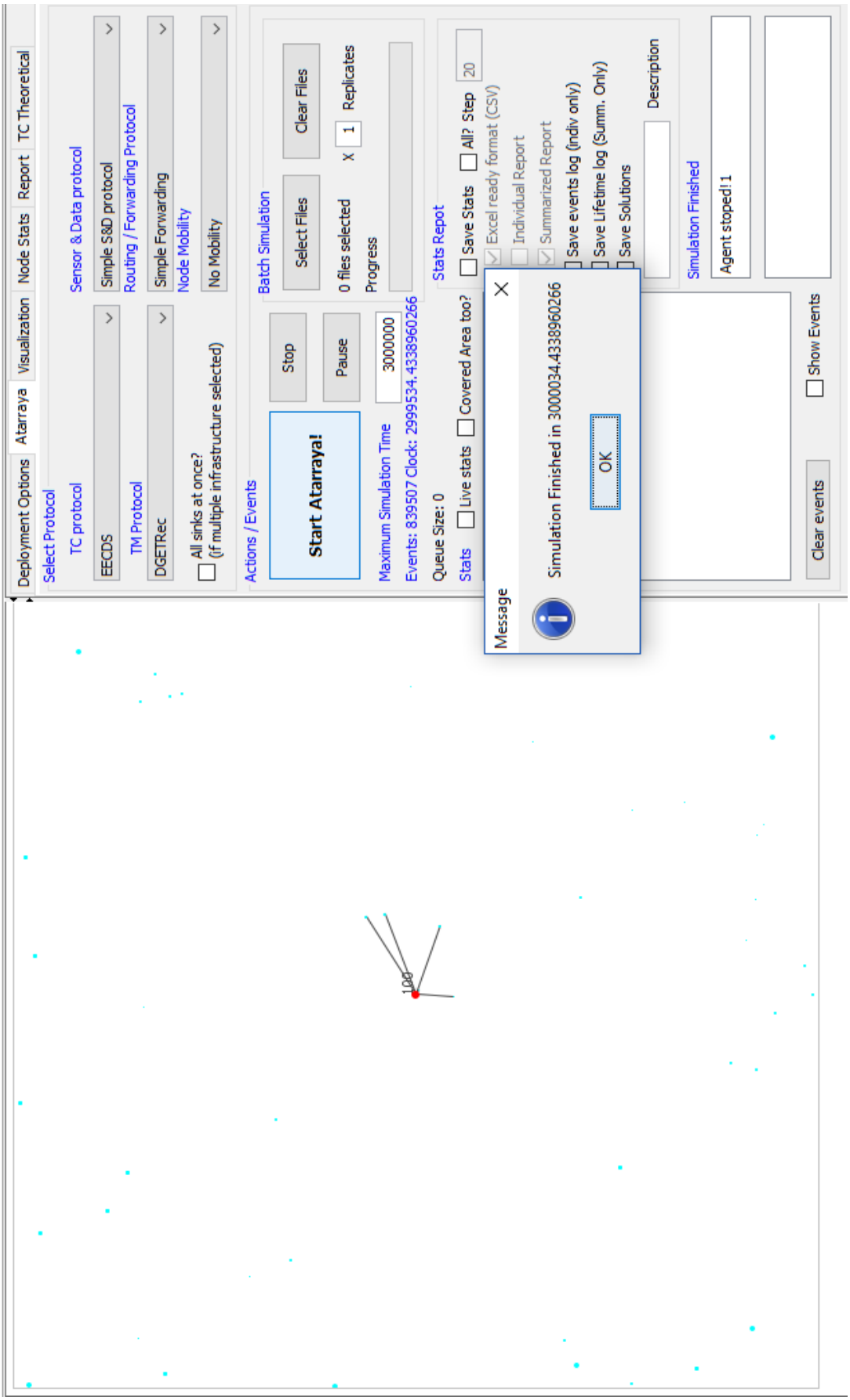


Рис. 6. Результати моделювання топології мережі протоколом DGETRec

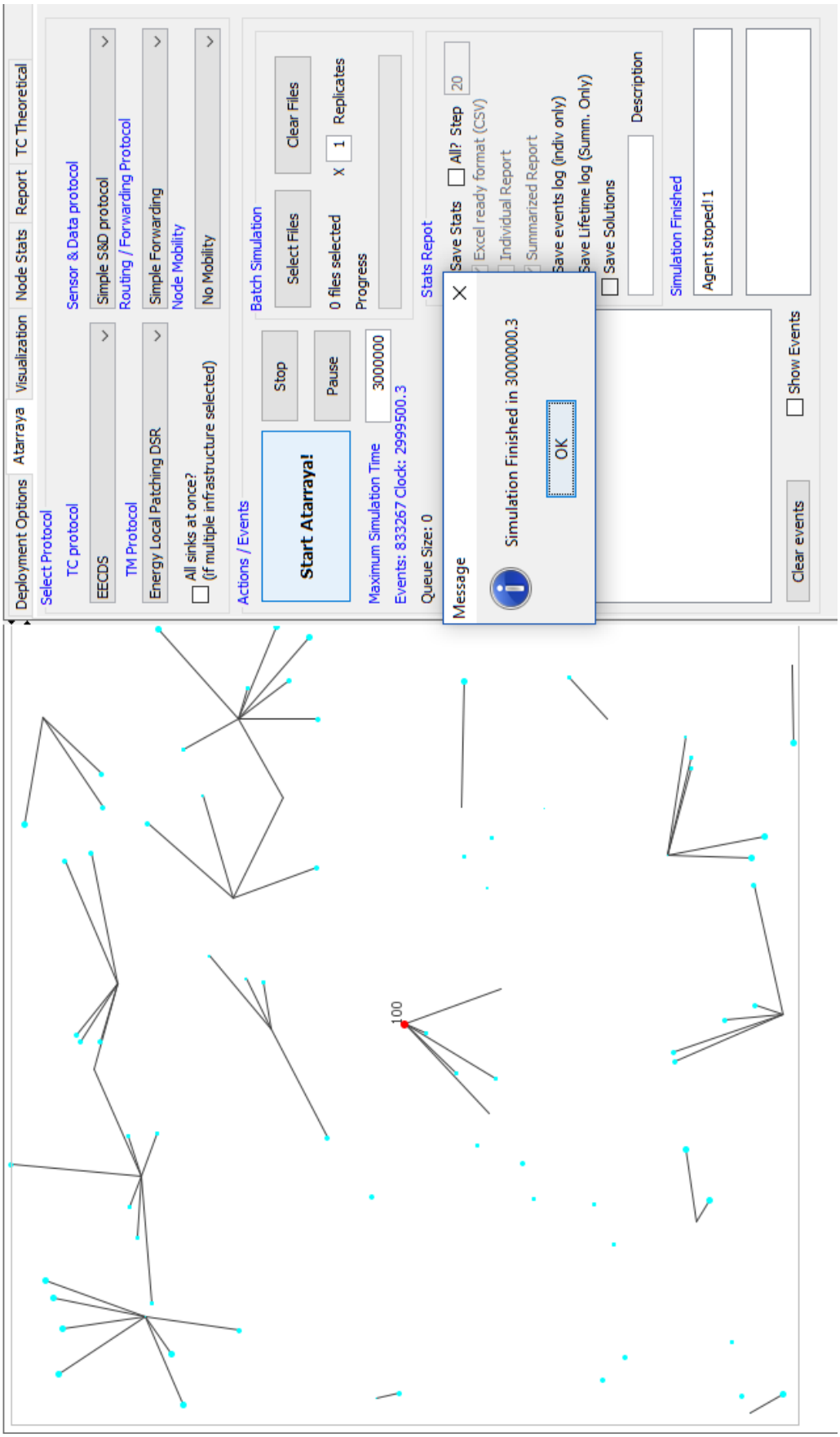


Рис. 7. Результати моделювання топології мережі протоколом Energy Local Patching DSR

Література

1. Roslyakov A., Vanyashin S., Scallops A., Samsonov M. Internet of Things. – Samara, 2014 (in Russion).
2. Yinbiao S., Lanctot P., Jianbin F. Internet of things: wireless sensor networks. – White Paper, International Electrotechnical Commission. – 2014.
3. IOActive Labs Research: Hacking Robots Before Skynet : http://blog.ioactive.com/2017/02/hacking-robots-before-skynet.html
4. Корченко, О. Г. Аналіз загроз та механізмів забезпечення інформаційної безпеки в сенсорних мережах / [Корченко О. Г., Алєксандер М. Б., Одарченко Р. С., Наджі А. А. А., Петренко О. Ю.] // Захист інформації. – 2016. – 18(1). – С. 48 - 56.
5. Voitovych O., Shulyatitska O., Malyushytskyy V. Simulation and security of sensor networks : Inżynier XXI wieku projectuje myśli przyszłości. monografia [pod red: Jacek Rysiński] - Bielsko-Biała: Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Białej, 2016 – pp. 367-373. ISBN: 978-83-65182-51-7.
6. Postolsky S.: Review of problem areas in the security of wireless sensor networks, attacks and mechanisms for their protection: http://book.itep.ru/4/41/zigbee.htm
7. Bluetooth: https://www.bluetooth.com/news-events/press-releases
8. Zhang Y., Jijun L.,Honglin Hu.: Wireless Mesh Network.
9. Prikhodko T. Problems of local networks at the data link and network layer osi model. – 2015.
10. Yatskiv V. Theoretical bases of creation and structural organization. Components of wireless sensor networks increased efficiency. Ternopil, 2016.
11. Karthikeyan V., Vinod A., Jeyakumar P. An Energy Efficient Neighbour Node Discovery Method for Wireless Sensor Networks //arXiv preprint arXiv:1402.3655. – 2014.
12. Chung Y. An energy-efficient unicast routing protocol for wireless sensor Networks //Tech. Int. J. Comput. Sci. Emerg. Tech. – 2011. – Т. 2. – С. 60 - 64.
13. Tiwari R., Saxena A. A review on energy efficient routing in wireless sensor networks //Journal of engineering trends and technology. – 2015.
14. Elrahim A. G. A. et al. An energy aware WSN geographic routing protocol //Universal Journal of Computer Science and Engineering Technology. – 2010. – Т. 1. – №. 2. – С. 105 - 111.
15. Ramesh V., Supriya K. S., Subbaiah P. Design of novel energy conservative preemptive dynamic source routing for MANET //Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on. – IEEE, 2014. – P. 1 - 7.
16. Alasem R., Reda A., Mansour M. Location based energy-efficient reliable routing protocol for wireless sensor networks //Recent Researches in Communications, Automation, Signal processing, Nanotechnology, Astronomy and Nuclear Physics, WSEAS Press, Cambridge, UK. – 2011.
17. Tseng H. R., Jan R. H., Yang W. An improved dynamic user authentication scheme for wireless sensor networks //Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE. – IEEE. 2007. – P. 986 - 990.
18. Chang S. M. et al. An efficient broadcast authentication scheme in wireless sensor networks //Proceedings of the 2006 ACM Symposium on Information, computer and communications security. – ACM. 2006. – P. 311 - 320.
19. Lee K. et al. An enhanced Trust Center based authentication in ZigBee networks //International Conference on Information Security and Assurance. – Springer Berlin Heidelberg, 2009. – P. 471 - 484.
20. Lee S. H., Kim J. H. Design of authentication protocol for LR-WPAN using pre-authentication mechanism //Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE. – IEEE. 2009. – P. 1 - 5.
21. Chen C. et al. Lightweight and provably secure user authentication with anonymity for the global mobility network //International Journal of Communication Systems. – 2011. – Т. 24. – №. 3. – P. 347 - 362.
22. Yatskiv V. et al. Concept of designing the wireless sensor networks based on ant intelligence //Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2015 IEEE 8th International Conference on. – IEEE, 2015. – Т. 2. – P. 863 - 866.
23. Atarraya - A Topology control simulator for Wireless Ad-hoc and Sensor Networks - http://www.csee.usf.edu/~mlabrador/Atarraya/

СИНТЕЗ НЕВИРОДЖЕНОГО КРИПТОГРАФІЧНОГО ПЕРЕТВОРЕННЯ НА ОСНОВІ ГРУПОВОГО ВИКОРИСТАННЯ ДВОРОЗРЯДНИХ МАТРИЧНИХ ОПЕРАЦІЙ

Рудницький В.М., Сисоєнко С.В., Миронець І.В.

Вступ

Наукові дослідження, направлені на підвищення якості систем криптографічного захисту інформації, завжди були і будуть актуальними та мають як теоретичну, так і практичну цінність в сучасних умовах.

В роботах [1-4] було доведено, що при кодуванні інформації кількома випадковими невиродженими операціями криптографічного перетворення інформації, з подальшим додаванням результатів кодування за модулем 2 дозволяє підвищити якість псевдовипадкової послідовності за рахунок того, що результат додавання буде виродженим. З цього можна зробити висновок, що даний підхід не може бути використаним для підвищення якості криптоалгоритмів, тому що втрачається можливість оберненого криптографічного перетворення.

Метою даного дослідження є визначення можливих шляхів вдосконалення систем криптографічного захисту інформації.

Виходячи з даних результатів подальші дослідження будуть направлені на пошук операцій, які замінять додавання за модулем 2, для отримання результуючого перетворення на основі декількох перетворень випадковими невиродженими операціями криптографічного перетворення інформації.

Основна частина

Подальші дослідження можуть проводитись в двох напрямках:

1. Синтез дворозрядних операцій криптографічного перетворення інформації, які забезпечать дані вимоги на основі операцій криптографічного перетворення інформації наведених в табл. 1 [1];
2. Використання операцій криптографічного перетворення інформації наведених в табл. 1 [1] для побудови результуючого перетворення.

При проведенні досліджень на першому етапі обмежимося побудовою результуючого криптографічного перетворення на основі кодування інформації двома випадковими невиродженими операціями.

Розглянемо перший напрямок дослідження.

При дослідженні можливості використання операцій криптографічного перетворення інформації для побудови результуючого перетворення обмежимося операціями, в яких присутнє додавання за модулем 2 та відсутня інверсія результатів операції.

Серед всіх операцій наведених в табл. 1 лише чотири операції відповідають даним вимогам.

Таблиця 1

Повна група дворозрядних операцій криптографічного перетворення інформації

№	операція	№	операція	№	операція
1	$F_{3,5}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$	9	$F_{3,9}=\begin{bmatrix}x_1\\x_1\oplus x_2\oplus 1\end{bmatrix}$	17	$F_{10,6}=\begin{bmatrix}x_2\oplus 1\\x_1\oplus x_2\end{bmatrix}$
2	$F_{6,5}=\begin{bmatrix}x_1\oplus x_2\\x_2\end{bmatrix}$	10	$F_{5,12}=\begin{bmatrix}x_2\\x_1\oplus 1\end{bmatrix}$	18	$F_{9,3}=\begin{bmatrix}x_1\oplus x_2\oplus 1\\x_1\end{bmatrix}$
3	$F_{3,6}=\begin{bmatrix}x_1\\x_1\oplus x_2\end{bmatrix}$	11	$F_{5,9}=\begin{bmatrix}x_2\\x_1\oplus x_2\oplus 1\end{bmatrix}$	19	$F_{12,10}=\begin{bmatrix}x_1\oplus 1\\x_2\oplus 1\end{bmatrix}$
4	$F_{5,3}=\begin{bmatrix}x_2\\x_1\end{bmatrix}$	12	$F_{6,12}=\begin{bmatrix}x_1\oplus x_2\\x_1\oplus 1\end{bmatrix}$	20	$F_{9,10}=\begin{bmatrix}x_1\oplus x_2\oplus 1\\x_2\oplus 1\end{bmatrix}$
5	$F_{5,6}=\begin{bmatrix}x_2\\x_1\oplus x_2\end{bmatrix}$	13	$F_{12,5}=\begin{bmatrix}x_1\oplus 1\\x_2\end{bmatrix}$	21	$F_{12,9}=\begin{bmatrix}x_1\oplus 1\\x_1\oplus x_2\oplus 1\end{bmatrix}$
6	$F_{6,3}=\begin{bmatrix}x_1\oplus x_2\\x_1\end{bmatrix}$	14	$F_{9,5}=\begin{bmatrix}x_1\oplus x_2\oplus 1\\x_2\end{bmatrix}$	22	$F_{10,12}=\begin{bmatrix}x_2\oplus 1\\x_1\oplus 1\end{bmatrix}$
7	$F_{3,10}=\begin{bmatrix}x_1\\x_2\oplus 1\end{bmatrix}$	15	$F_{12,6}=\begin{bmatrix}x_1\oplus 1\\x_1\oplus x_2\end{bmatrix}$	23	$F_{10,9}=\begin{bmatrix}x_2\oplus 1\\x_1\oplus x_2\oplus 1\end{bmatrix}$
8	$F_{6,10}=\begin{bmatrix}x_1\oplus x_2\\x_2\oplus 1\end{bmatrix}$	16	$F_{10,3}=\begin{bmatrix}x_2\oplus 1\\x_1\end{bmatrix}$	24	$F_{9,12}=\begin{bmatrix}x_1\oplus x_2\oplus 1\\x_1\oplus 1\end{bmatrix}$

Визначені моделі операцій наведено в табл. 2. Визначені алгоритми криптографічного перетворення інформації двох блоків змінних в табл. 3. Алгоритми побудови псевдовипадкової послідовності наведено табл. 4.

Таблиця 2

Операції криптографічного перетворення інформації двох блоків змінних

Номер операції	Модель операції	
	Пряме перетворення	Обернене перетворення
1	$F^{k}_{6,5}=\begin{bmatrix}z_1\oplus z_2\\z_2\end{bmatrix}$	$F^{d}_{6,5}=\begin{bmatrix}w_1\oplus w_2\\w_2\end{bmatrix}$
2	$F^{k}_{3,6}=\begin{bmatrix}z_1\\z_1\oplus z_2\end{bmatrix}$	$F^{d}_{3,6}=\begin{bmatrix}w_1\\w_1\oplus w_2\end{bmatrix}$
3	$F^{k}_{5,6}=\begin{bmatrix}z_2\\z_1\oplus z_2\end{bmatrix}$	$F^{d}_{6,3}=\begin{bmatrix}w_1\oplus w_2\\w_1\end{bmatrix}$
4	$F^{k}_{6,3}=\begin{bmatrix}z_1\oplus z_2\\z_1\end{bmatrix}$	$F^{d}_{5,6}=\begin{bmatrix}w_2\\w_1\oplus w_2\end{bmatrix}$

Таблиця 3

Алгоритм криптографічного перетворення інформації двох блоків змінних

Номер алгоритму	Модель операції	
	Пряме перетворення	Обернене перетворення
1	$G_{6,5}^{k}=\begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix}$	$G_{6,5}^{d}=\begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix}$
2	$G_{3,6}^{k}=\begin{bmatrix} F_1 \\ F_1 \oplus F_2 \end{bmatrix}$	$G_{3,6}^{d}=\begin{bmatrix} F_1 \\ F_1 \oplus F_2 \end{bmatrix}$
3	$G_{5,6}^{k}=\begin{bmatrix} F_2 \\ F_1 \oplus F_2 \end{bmatrix}$	$G_{6,5}^{d}=\begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix}$
4	$G_{6,3}^{k}=\begin{bmatrix} F_1 \oplus F_2 \\ F_1 \end{bmatrix}$	$G_{5,6}^{d}=\begin{bmatrix} F_2 \\ F_1 \oplus F_2 \end{bmatrix}$

Розглянемо результати розрахунку псевдовипадкової послідовності більш детально на прикладах.

Таблиця 4

Таблиця алгоритму побудови псевдовипадкової послідовності

Послі-довність	Кроки алгоритму побудови псевдовипадкових послідовностей відповідних блоків				
	1	2	3	4	5
1	$F_{1;1}^{k}(z_1)$	$F_{1;2}^{k}(z_2)$	$F_{1;3}^{k}(z_3)$	$F_{1;4}^{k}(z_4)$	$F_{1;5}^{k}(z_5)$
2	$F_{2;1}^{k}(z_1)$	$F_{2;2}^{k}(z_2)$	$F_{2;3}^{k}(z_3)$	$F_{2;4}^{k}(z_4)$	$F_{2;5}^{k}(z_5)$
Результат перетворення	$z_1\{z_{1.1};z_{1.2}\}$	$z_2\{z_{2.1};z_{2.2}\}$	$z_3\{z_{3.1};z_{3.2}\}$	$z_4\{z_{4.1};z_{4.2}\}$	$z_5\{z_{5.1};z_{5.2}\}$

Приклад 1. Розглянемо обробку перших двох блоків змінних перетворення за допомогою першого алгоритму криптографічного перетворення інформації

$$G_{6,5}^{k}=\begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix}. \quad (1)$$

Позначимо F_1 через $F_{1;1}^{k}(z_1)$, тоді

$$F_1=F_{1;1}^{k}(z_1)=F_{6,5}^{k}(z_1)=\begin{bmatrix} z_{1.1} \oplus z_{1.2} \\ z_{1.2} \end{bmatrix}. \quad (2)$$

Представимо F_2 через $F_{2;2}^{k}(z_2)$, тоді

$$F_2 = F^k_{2;2}(z_2) = F^k_{6,3}(z_2) = \begin{bmatrix} z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (3)$$

Підставивши в вираз (1) значення виразів (2) і (3), перший алгоритм криптографічного перетворення інформації двох блоків змінних буде наступним:

$$G^k_{6,5} = \begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix} = \begin{bmatrix} F^k_{1;1}(z_1) \oplus F^k_{2;2}(z_2) \\ F^k_{2;2}(z_2) \end{bmatrix} = \begin{bmatrix} F^k_{6,5}(z_1) \oplus F^k_{6,3}(z_2) \\ F^k_{6,3}(z_2) \end{bmatrix} =$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \\ z_{1.2} \oplus z_{2.1} \\ z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (4)$$

Далі перевіримо можливість використання даного перетворення для криптографії. Для цього необхідне існування оберненого перетворення.

Обернене матричне криптографічне перетворення знайдемо на основі методу синтезу операцій оберненого криптоперетворення наведеного в [5].

Сутність даного методу полягає в наступному: якщо операція криптографічного прямого перетворення без урахування групи операцій інверсії задана виразом

$$\vec{F}_k = \begin{pmatrix} a_{11}x_1 \oplus a_{12}x_2 \oplus \ldots \oplus a_{1n}x_n \\ a_{21}x_1 \oplus a_{22}x_2 \oplus \ldots \oplus a_{2n}x_n \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{n1}x_1 \oplus a_{n2}x_2 \oplus \ldots \oplus a_{n\,n}x_n \end{pmatrix}, \quad (5)$$

де $a_{ij} \in [0,1]$; $b_i \in [0,1]$; $x_1 \ldots x_n$ – операнди-розряди відповідно; $\oplus$ – операція «сума за mod 2», тоді операція криптографічного оберненого перетворення буде задана виразом

$$\vec{F}_d = \begin{pmatrix} b_{11}y_1 \oplus b_{12}y_2 \oplus \ldots \oplus b_{1n}y_n \\ b_{21}y_1 \oplus b_{22}y_2 \oplus \ldots \oplus b_{2n}y_n \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ b_{n1}y_1 \oplus b_{n2}y_2 \oplus \ldots \oplus b_{n\,n}y_n \end{pmatrix}, \quad (6)$$

де b_i – коефіцієнти матриці оберненого перетворення, y_i – операнди-розряди інформації, які отримані в результаті застосування операції прямого перетворення ($y_i = \vec{F}_k(x_i)$) відповідно.

Тоді результатом виконання операції оберненого перетворення повинен бути вираз, що має вигляд:

$$\vec{F}_r = \begin{pmatrix} a_{11}x_1 & & & & & & & & & & \\ & a_{22}x_2 & & & & & & & & & \\ . & . & . & . & . & . & . & . & . & . & . \\ & & & & & & & & & & a_{n\,n}x_n \end{pmatrix}, \quad (7)$$

де $\vec{F}_r$ – еталонна матриця або матриця-результат, $x_1 ... x_n$ – початкові операнди-розряди інформації; $a_{ij} = 1$ при $i = j$, тому що потрібно забезпечити невиродженість перетворення, тобто повинна виконуватись умова $a_{11} \cdot a_{22} - a_{12} \cdot a_{21} \neq 0$, а також відсутні перестановки рядків матриці.

Використавши даний метод отримаємо:

$$G^d_{6,5} = \begin{bmatrix} w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \\ w_{1.2} \oplus w_{2.2} \\ w_{2.2} \\ w_{2.1} \oplus w_{2.2} \end{bmatrix}. \quad (8)$$

Перевіримо коректність отриманої операції оберненого перетворення, підставивши в вираз (8) значення виразу (4). Отримаємо:

$$G^d_{6,5} = \begin{bmatrix} w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \\ w_{1.2} \oplus w_{2.2} \\ w_{2.2} \\ w_{2.1} \oplus w_{2.2} \end{bmatrix} =$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \\ z_{1.2} \oplus z_{2.1} \oplus z_{2.1} \\ z_{2.1} \\ z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \end{bmatrix} = \begin{bmatrix} z_{1.1} \\ z_{1.2} \\ z_{2.1} \\ z_{2.2} \end{bmatrix}. \quad (9)$$

В результаті отримано коректну операцію оберненого криптографічного перетворення, яку можна реалізувати як операцію обробки чотирьох бітів інформації при паралельній реалізації, або обробки двох блоків інформації з наступною обробкою кожного блоку окремо при послідовній реалізації. При обробці блоків інформації необхідно використовувати обернені операції. Для даного прикладу алгоритми прямого та оберненого перетворення співпадають, також співпадають прямі та обернені операції криптоперетворення.

Слід відмітити, що при реалізації даної операції були задіяні обидві псевдовипадкові послідовності, а це повинно привести до підвищення якості загального результату перетворення.

Приклад 2. Позначимо F_1 через $F^k_{1;2}(z_2)$, тоді

$$F_1 = F^k_{1;2}(z_2) = F^k_{3,6}(z_2) = \begin{bmatrix} z_{2.1} \\ z_{2.1} \oplus z_{2.2} \end{bmatrix}. \quad (10)$$

Представимо F_2 через $F^k_{2;1}(z_1)$, тоді

$$F_2 = F^k_{2;1}(z_1) = F^k_{5,6}(z_1) = \begin{bmatrix} z_{1.2} \\ z_{1.1} \oplus z_{1.2} \end{bmatrix}. \quad (11)$$

Підставивши значення виразів (10) і (11) в вираз (1), отримаємо наступне значення:

$$G^k_{6,5} = \begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix} = \begin{bmatrix} F^k_{1;2}(z_2) \oplus F^k_{2;1}(z_1) \\ F^k_{2;1}(z_1) \end{bmatrix} = \begin{bmatrix} F^k_{3,6}(z_2) \oplus F^k_{5,6}(z_2) \\ F^k_{5,6}(z_1) \end{bmatrix}$$

$$= \begin{bmatrix} z_{2.1} \oplus z_{1.2} \\ z_{2.1} \oplus z_{2.2} \oplus z_{1.1} \oplus z_{1.2} \\ z_{1.2} \\ z_{1.1} \oplus z_{1.2} \end{bmatrix}. \quad (12)$$

Перевіримо можливість використання даного перетворення для криптографії.

$$G^d_{6,5} = \begin{bmatrix} w_{2.2} \oplus w_{2.1} \\ w_{2.1} \\ w_{1.1} \oplus w_{2.1} \\ w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \end{bmatrix}. \quad (13)$$

Перевіримо коректність отриманої операції оберненого перетворення підставивши в вираз (13) значення виразу (12).

$$G^d_{6,5} = \begin{bmatrix} w_{2.2} \oplus w_{2.1} \\ w_{2.1} \\ w_{1.1} \oplus w_{2.1} \\ w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \end{bmatrix} =$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{1.2} \\ z_{1.2} \\ z_{1.2} \oplus z_{2.1} \\ z_{1.2} \oplus z_{2.1} \oplus z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{1.2} \oplus z_{1.1} \oplus z_{1.2} \end{bmatrix} = \begin{bmatrix} z_{1.1} \\ z_{1.2} \\ z_{2.1} \\ z_{2.2} \end{bmatrix}. \quad (14)$$

При реалізації даної операції були задіяні обидві псевдовипадкові послідовності, що підвищує якість загального результату перетворення.

Приклад 3. Позначимо F_1 через $F^k_{1;1}(z_1)$, тоді

$$F_1 = F^k_{1;1}(z_1) = F^k_{6,5}(z_1) = \begin{bmatrix} z_{1.1} \oplus z_{1.2} \\ z_{1.2} \end{bmatrix}. \quad (15)$$

Позначимо F_2 через $F_{2;1}^k(z_1)$, тоді

$$F_2 = F_{2;1}^k(z_1) = F_{5,6}^k(z_1) = \begin{bmatrix} z_{1.2} \\ z_{1.1} \oplus z_{1.2} \end{bmatrix}. \quad (16)$$

Перший алгоритм криптографічного перетворення інформації виразу (1) після підстановки значень виразів (15), (16) буде наступним:

$$G_{6,5}^k = \begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix} = \begin{bmatrix} F_{1;1}^k(z_1) \oplus F_{2;1}^k(z_1) \\ F_{2;1}^k(z_1) \end{bmatrix} = \begin{bmatrix} F_{6,5}^k(z_1) \oplus F_{5,6}^k(z_1) \\ F_{5,6}^k(z_1) \end{bmatrix}$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{1.2} \\ z_{1.2} \oplus z_{1.1} \oplus z_{1.2} \\ z_{1.2} \\ z_{1.1} \oplus z_{1.2} \end{bmatrix} = \begin{bmatrix} z_{1.1} \\ z_{1.1} \\ z_{1.2} \\ z_{1.1} \oplus z_{1.2} \end{bmatrix}. \quad (17)$$

При реалізації даної операції підвищити якість загального результату перетворення інформації не можливо, оскільки отримано вироджений результат внаслідок втрати інформації другого блоку даних $z_2\{z_{2.1}; z_{2.2}\}$.

Приклад 4. Позначимо F_1 через $F_{1;2}^k(z_2)$, тоді

$$F_1 = F_{1;2}^k(z_2) = F_{3,6}^k(z_2) = \begin{bmatrix} z_{2.1} \\ z_{2.1} \oplus z_{2.2} \end{bmatrix}. \quad (18)$$

Представимо F_2 через $F_{2;2}^k(z_2)$, тоді

$$F_2 = F_{2;2}^k(z_2) = F_{6,3}^k(z_2) = \begin{bmatrix} z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (19)$$

Після підстановки значень (18) та (19) перший алгоритм криптографічного перетворення інформації двох блоків змінних (1) матиме вигляд:

$$G_{6,5}^k = \begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix} = \begin{bmatrix} F_{1;2}^k(z_2) \oplus F_{2;2}^k(z_2) \\ F_{2;2}^k(z_2) \end{bmatrix} = \begin{bmatrix} F_{3,6}^k(z_2) \oplus F_{6,3}^k(z_2) \\ F_{6,3}^k(z_2) \end{bmatrix} =$$

$$= \begin{bmatrix} z_{2.1} \oplus z_{2.1} \oplus z_{2.2} \\ z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \\ z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix} = \begin{bmatrix} z_{2.2} \\ z_{2.2} \\ z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (20)$$

Підвищення якості загального результату перетворення інформації не можливо, оскільки отримано вироджений результат внаслідок втрати інформації першого блоку даних $z_1\{z_{1.1}; z_{1.2}\}$.

Приклад 5. Позначимо F_1 через $F_{1;1}^k(z_1)$, тоді

$$F_1 = F^k_{1;1}(z_1) = F^k_{6,5}(z_1) = \begin{bmatrix} z_{1.1} \oplus z_{1.2} \\ z_{1.2} \end{bmatrix}. \quad (21)$$

Представимо F_2 через $F^k_{1;2}(z_2)$, тоді

$$F_2 = F^k_{1;2}(z_2) = F^k_{3,6}(z_2) = \begin{bmatrix} z_{2.1} \\ z_{2.1} \oplus z_{2.2} \end{bmatrix}. \quad (22)$$

Вираз (1) після підстановки значень виразів (21) та (22) набуде вигляду:

$$G^k_{6,5} = \begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix} = \begin{bmatrix} F^k_{1;1}(z_1) \oplus F^k_{1;2}(z_2) \\ F^k_{1;2}(z_2) \end{bmatrix} = \begin{bmatrix} F^k_{6,5}(z_1) \oplus F^k_{3,6}(z_2) \\ F^k_{3,6}(z_2) \end{bmatrix} =$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \\ z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \\ z_{2.1} \\ z_{2.1} \oplus z_{2.2} \end{bmatrix}. \quad (23)$$

Перевіримо можливість використання даного перетворення для криптографії:

$$G^d_{6,5} = \begin{bmatrix} w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \\ w_{1.2} \oplus w_{2.2} \\ w_{2.1} \\ w_{2.1} \oplus w_{2.2} \end{bmatrix}. \quad (24)$$

Перевіримо коректність отриманої операції оберненого перетворення підставивши в вираз (24) значення (23):

$$G^d_{6,5} = \begin{bmatrix} w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \\ w_{1.2} \oplus w_{2.2} \\ w_{2.1} \\ w_{2.1} \oplus w_{2.2} \end{bmatrix} =$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \oplus z_{2.1} \oplus z_{2.2} \\ z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \oplus z_{2.2} \\ z_{2.1} \\ z_{2.1} \oplus z_{2.1} \oplus z_{2.2} \end{bmatrix} = \begin{bmatrix} z_{1.1} \\ z_{1.2} \\ z_{2.1} \\ z_{2.2} \end{bmatrix}. \quad (25)$$

При реалізації даної операції підвищити якість загального результату перетворення інформації не можливо, оскільки використано лише інформацію, що знаходиться в першій послідовності. При втраті інформації другої послідовності отримаємо вироджений результат.

Приклад 6. Позначимо F_1 через $F^k_{2;1}(z_1)$, тоді

$$F_1 = F^k_{2;1}(z_1) = F^k_{5,6}(z_1) = \begin{bmatrix} z_{1.2} \\ z_{1.1} \oplus z_{1.2} \end{bmatrix}. \quad (26)$$

Представимо F_2 через $F^k_{2;2}(z_2)$, тоді

$$F_2 = F^k_{2;2}(z_2) = F^k_{6,3}(z_2) = \begin{bmatrix} z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (27)$$

Перший алгоритм криптографічного перетворення інформації двох блоків змінних (1) після підстановки значень (26), (27) буде наступним:

$$G^k_{6,5} = \begin{bmatrix} F_1 \oplus F_2 \\ F_2 \end{bmatrix} = \begin{bmatrix} F^k_{2;1}(z_1) \oplus F^k_{2;2}(z_2) \\ F^k_{2;2}(z_2) \end{bmatrix} = \begin{bmatrix} F^k_{5,6}(z_1) \oplus F^k_{6,3}(z_2) \\ F^k_{6,3}(z_2) \end{bmatrix}$$

$$= \begin{bmatrix} z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \\ z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \\ z_{2.1} \oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (28)$$

Перевіримо можливість існування оберненого перетворення:

$$G^d_{6,5} = \begin{bmatrix} w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \\ w_{1.1} \oplus w_{2.1} \\ w_{2.2} \\ w_{2.1} \oplus w_{2.2} \end{bmatrix}. \quad (29)$$

Перевіримо коректність отриманої операції оберненого перетворення, підставивши в вираз (29) значення виразу (28).

$$G^d_{6,5} = \begin{bmatrix} w_{1.1} \oplus w_{1.2} \oplus w_{2.1} \oplus w_{2.2} \\ w_{1.1} \oplus w_{2.1} \\ w_{2.2} \\ w_{2.1} \oplus w_{2.2} \end{bmatrix} =$$

$$= \begin{bmatrix} z_{1.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \oplus z_{2.1} \oplus z_{2.2} \\ z_{1.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{2.1} \oplus z_{2.2} \\ z_{2.1} \\ z_{2.1} \oplus z_{2.1} \oplus z_{2.2} \end{bmatrix} = \begin{bmatrix} z_{1.1} \\ z_{1.2} \\ z_{2.1} \\ z_{2.2} \end{bmatrix}. \quad (30)$$

При реалізації даної операції підвищити якість загального результату перетворення інформації не можливо, оскільки використано лише інформацію, що знаходиться в другій послідовності. При втраті інформації першої послідовності отримаємо вироджений результат.

Приклад 7. Розглянемо обробку другого і третього блоків змінних перетворення за допомогою другого алгоритму криптографічного перетворення інформації:

$$G_{3,6}^{k} = \begin{bmatrix} F_1 \\ F_1 \oplus F_2 \end{bmatrix}. \quad (31)$$

Позначимо F_1 через $F_{1;2}^{k}(z_2)$, тоді

$$F_1 = F_{1;2}^{k}(z_2) = F_{3,6}^{k}(z_2) = \begin{bmatrix} z_{2.1} \\ z_{2.1} \oplus z_{2.2} \end{bmatrix}. \quad (32)$$

Представимо F_2 через $F_{2;3}^{k}(z_3)$, тоді

$$F_2 = F_{2;3}^{k}(z_3) = F_{3,6}^{k}(z_3) = \begin{bmatrix} z_{3.1} \\ z_{3.1} \oplus z_{3.2} \end{bmatrix}. \quad (33)$$

Другий алгоритм криптографічного перетворення інформації другого та третього блоків змінних (31) після підстановки (32), (33) буде наступним:

$$G_{3,6}^{k} = \begin{bmatrix} F_1 \\ F_1 \oplus F_2 \end{bmatrix} = \begin{bmatrix} F_{1;2}^{k}(z_2) \\ F_{1;2}^{k}(z_2) \oplus F_{2;3}^{k}(z_3) \end{bmatrix} = \begin{bmatrix} F_{3,6}^{k}(z_2) \\ F_{3,6}^{k}(z_2) \oplus F_{3,6}^{k}(z_3) \end{bmatrix} =$$

$$= \begin{bmatrix} z_{2.1} \\ z_{2.1} \oplus z_{2.2} \\ z_{2.1} \oplus z_{3.1} \\ z_{2.1} \oplus z_{2.2} \oplus z_{3.1} \oplus z_{3.2} \end{bmatrix}. \quad (34)$$

Перевіримо можливість використання даного перетворення для криптографії. Для цього необхідне існування оберненого перетворення.

Використавши даний метод отримаємо:

$$G_{3,6}^{d} = \begin{bmatrix} w_{2.1} \\ w_{2.1} \oplus w_{2.2} \\ w_{2.1} \oplus w_{3.1} \\ w_{2.1} \oplus w_{2.2} \oplus w_{3.1} \oplus w_{3.2} \end{bmatrix}. \quad (35)$$

Перевіримо коректність отриманої операції оберненого перетворення. Для цього, підставивши в вираз (35) значення виразу (34), отримаємо:

$$G_{3,6}^{d}=\begin{bmatrix} w_{2.1} \\ w_{2.1}\oplus w_{2.2} \\ w_{2.1}\oplus w_{3.1} \\ w_{2.1}\oplus w_{2.2}\oplus w_{3.1}\oplus w_{3.2} \end{bmatrix}=$$

$$=\begin{bmatrix} z_{2.1} \\ z_{2.1}\oplus z_{2.1}\oplus z_{2.2} \\ z_{2.1}\oplus z_{2.1}\oplus z_{3.1} \\ z_{2.1}\oplus z_{2.1}\oplus z_{2.2}\oplus z_{2.1}\oplus z_{3.1}\oplus z_{2.1}\oplus z_{2.2}\oplus z_{3.1}\oplus z_{3.2} \end{bmatrix}=\begin{bmatrix} z_{2.1} \\ z_{2.2} \\ z_{3.1} \\ z_{3.2} \end{bmatrix}. \quad (36)$$

В результаті отримано коректну операцію оберненого криптографічного перетворення, яку можна реалізувати як операцію обробки чотирьох бітів інформації при паралельній реалізації, або обробки двох блоків інформації з наступною обробкою кожного блоку окремо при послідовній реалізації. Для даного прикладу алгоритми прямого та оберненого перетворення співпадають, також співпадають прямі та обернені операції криптоперетворення.

Слід відмітити, що при реалізації даної операції були задіяні обидві псевдовипадкові послідовності, а це повинно привести до підвищення якості загального результату перетворення.

Приклад 8. Позначимо F_1 через $F_{1;3}^{k}(z_3)$, тоді

$$F_1=F_{1;3}^{k}(z_3)=F_{5,6}^{k}(z_3)=\begin{bmatrix} z_{3.2} \\ z_{3.1}\oplus z_{3.2} \end{bmatrix}. \quad (37)$$

Представимо F_2 через $F_{2;2}^{k}(z_2)$, тоді

$$F_2=F_{2;2}^{k}(z_2)=F_{6,3}^{k}(z_2)=\begin{bmatrix} z_{2.1}\oplus z_{2.2} \\ z_{2.1} \end{bmatrix}. \quad (38)$$

Підставивши у вираз (31) значення виразів (37) і (38) отримаємо наступний вираз:

$$G_{3,6}^{k}=\begin{bmatrix} F_1 \\ F_1\oplus F_2 \end{bmatrix}=\begin{bmatrix} F_{1;3}^{k}(z_3) \\ F_{1;3}^{k}(z_3)\oplus F_{2;2}^{k}(z_2) \end{bmatrix}=\begin{bmatrix} F_{5,6}^{k}(z_3) \\ F_{5,6}^{k}(z_3)\oplus F_{6,3}^{k}(z_2) \end{bmatrix}=$$

$$=\begin{bmatrix} z_{3.2} \\ z_{3.1}\oplus z_{3.2} \\ z_{3.2}\oplus z_{2.1}\oplus z_{2.2} \\ z_{3.1}\oplus z_{3.2}\oplus z_{2.1} \end{bmatrix}. \quad (39)$$

Перевіримо можливість використання даного перетворення для криптографії. Для цього необхідне існування оберненого перетворення.

Використавши даний метод отримаємо:

$$G^{d}_{3,6} = \begin{bmatrix} w_{2.1} \oplus w_{2.1} \oplus w_{2.2} \oplus w_{3.2} \\ w_{2.1} \oplus w_{2.2} \oplus w_{3.1} \oplus w_{3.2} \\ w_{2.1} \oplus w_{2.2} \oplus w_{3.1} \oplus w_{3.1} \\ w_{2.1} \end{bmatrix}. \quad (40)$$

Перевіримо коректність отриманої операції оберненого перетворення, підставивши в вираз (40) значення виразу (39):

$$G^{d}_{3,6} = \begin{bmatrix} w_{2.1} \oplus w_{2.1} \oplus w_{2.2} \oplus w_{3.2} \\ w_{2.1} \oplus w_{2.2} \oplus w_{3.1} \oplus w_{3.2} \\ w_{2.1} \oplus w_{2.2} \oplus w_{3.1} \oplus w_{3.1} \\ w_{2.1} \end{bmatrix} =$$

$$= \begin{bmatrix} z_{3.2} \oplus z_{2.1} \oplus z_{3.1} \oplus z_{3.2} \oplus z_{3.1} \oplus z_{3.2} \oplus z_{3.2} \\ z_{3.1} \oplus z_{3.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{3.2} \oplus z_{2.1} \oplus z_{3.1} \oplus z_{3.2} \oplus z_{3.2} \\ z_{2.1} \oplus z_{2.2} \oplus z_{3.2} \oplus z_{3.1} \oplus z_{3.2} \oplus z_{2.1} \oplus z_{2.2} \oplus z_{3.2} \oplus z_{3.2} \\ z_{3.2} \end{bmatrix} = \begin{bmatrix} z_{2.1} \\ z_{2.2} \\ z_{3.1} \\ z_{3.2} \end{bmatrix}. \quad (41)$$

При реалізації даної операції були задіяні обидві псевдовипадкові послідовності, що підвищує якість загального результату перетворення.

Приклад 9. Розглянемо обробку третього і четвертого блоків змінних перетворення за допомогою третього алгоритму криптографічного перетворення інформації

$$G^{k}_{5,6} = \begin{bmatrix} F_2 \\ F_1 \oplus F_2 \end{bmatrix}. \quad (42)$$

Позначимо F_1 через $F^{k}_{1;3}(z_3)$, тоді

$$F_1 = F^{k}_{1;3}(z_3) = F^{k}_{5,6}(z_3) = \begin{bmatrix} z_{3.2} \\ z_{3.1} \oplus z_{3.2} \end{bmatrix}. \quad (43)$$

Представимо F_2 через $F^{k}_{2;4}(z_4)$, тоді

$$F_2 = F^{k}_{2;4}(z_4) = F^{k}_{6,3}(z_4) = \begin{bmatrix} z_{4.1} \oplus z_{4.2} \\ z_{4.1} \end{bmatrix}. \quad (44)$$

Третій алгоритм криптографічного перетворення інформації третього та четвертого блоків змінних виразу (42) після підстановки (43), (44) буде наступним:

$$G_{5,6}^{k}=\begin{bmatrix}F_2\\ F_1\oplus F_2\end{bmatrix}=\begin{bmatrix}F_{2;4}^{k}(z_4)\\ F_{1;3}^{k}(z_3)\oplus F_{2;4}^{k}(z_4)\end{bmatrix}=\begin{bmatrix}F_{6,3}^{k}(z_4)\\ F_{5,6}^{k}(z_3)\oplus F_{6,3}^{k}(z_4)\end{bmatrix}=$$

$$=\begin{bmatrix}z_{4.1}\oplus z_{4.2}\\ z_{4.1}\\ z_{3.2}\oplus z_{4.1}\oplus z_{4.2}\\ z_{3.1}\oplus z_{3.2}\oplus z_{4.1}\end{bmatrix}. \quad (45)$$

Перевіримо можливість використання даного перетворення для криптографії [5].

Використавши даний метод отримаємо:

$$G_{6,5}^{d}=\begin{bmatrix}w_{3.1}\oplus w_{3.2}\oplus w_{4.1}\oplus w_{4.2}\\ w_{3.1}\oplus w_{3.2}\oplus w_{3.2}\oplus w_{4.1}\\ w_{3.2}\\ w_{3.1}\oplus w_{3.2}\oplus w_{4.2}\end{bmatrix}. \quad (46)$$

Перевіримо коректність отриманої операції оберненого перетворення.

$$G_{6,5}^{d}=\begin{bmatrix}w_{3.1}\oplus w_{3.2}\oplus w_{4.1}\oplus w_{4.2}\\ w_{3.1}\oplus w_{3.2}\oplus w_{3.2}\oplus w_{4.1}\\ w_{3.2}\\ w_{3.1}\oplus w_{3.2}\oplus w_{4.2}\end{bmatrix}=$$

$$\begin{bmatrix}z_{4.1}\oplus z_{4.2}\oplus z_{4.1}\oplus z_{3.2}\oplus z_{4.1}\oplus z_{4.2}\oplus z_{3.1}\oplus z_{3.2}\oplus z_{4.1}\\ z_{4.1}\oplus z_{4.2}\oplus z_{3.2}\oplus z_{3.2}\oplus z_{3.2}\oplus z_{4.1}\oplus z_{4.2}\\ z_{4.1}\\ z_{4.1}\oplus z_{4.2}\oplus z_{4.1}\oplus z_{3.1}\oplus z_{3.2}\oplus z_{4.1}\oplus z_{3.1}\oplus z_{3.2}\oplus z_{4.1}\end{bmatrix}=\begin{bmatrix}z_{3.1}\\ z_{3.2}\\ z_{4.1}\\ z_{4.2}\end{bmatrix}. \quad (47)$$

В результаті отримано коректну операцію оберненого криптографічного перетворення.

Слід відмітити, що при реалізації даної операції були задіяні обидві псевдовипадкові послідовності, а це повинно привести до підвищення якості загального результату перетворення.

Приклад 10. Позначимо F_1 через $F_{1;4}^{k}(z_4)$, тоді

$$F_1=F_{1;4}^{k}(z_4)=F_{6,5}^{k}(z_4)=\begin{bmatrix}z_{4.1}\oplus z_{4.2}\\ z_{4.2}\end{bmatrix}. \quad (48)$$

Представимо F_2 через $F_{2;5}^{k}(z_5)$, тоді

$$F_2=F_{2;3}^{k}(z_3)=F_{3,6}^{k}(z_3)=\begin{bmatrix}z_{3.1}\\ z_{3.1}\oplus z_{3.2}\end{bmatrix}. \quad (49)$$

Вираз (42) після підстановки значень виразів (48), (49) буде наступним:

$$G_{5,6}^{k}=\begin{bmatrix}F_2\\ F_1\oplus F_2\end{bmatrix}=\begin{bmatrix}F_{2;3}^{k}(z_3)\\ F_{1;4}^{k}(z_4)\oplus F_{2;3}^{k}(z_3)\end{bmatrix}=\begin{bmatrix}F_{3,6}^{k}(z_3)\\ F_{6,5}^{k}(z_4)\oplus F_{3,6}^{k}(z_3)\end{bmatrix}=$$

$$=\begin{bmatrix}z_{3.1}\\ z_{3.1}\oplus z_{3.2}\\ z_{4.1}\oplus z_{4.2}\oplus z_{3.1}\\ z_{4.2}\oplus z_{3.1}\oplus z_{3.2}\end{bmatrix}. \qquad (50)$$

Перевіримо можливість використання даного перетворення для криптографії [5].

Використавши даний метод, отримаємо:

$$G_{6,5}^{d}=\begin{bmatrix}w_{3.1}\\ w_{3.1}\oplus w_{3.2}\\ w_{3.1}\oplus w_{3.2}\oplus w_{4.1}\oplus w_{4.2}\\ w_{3.2}\oplus w_{4.1}\end{bmatrix}. \qquad (51)$$

Перевіримо коректність отриманої операції оберненого перетворення, підставивши в вираз (51) значення виразу (50):

$$G_{6,5}^{d}=\begin{bmatrix}w_{3.1}\\ w_{3.1}\oplus w_{3.2}\\ w_{3.1}\oplus w_{3.2}\oplus w_{4.1}\oplus w_{4.2}\\ w_{3.2}\oplus w_{4.1}\end{bmatrix}=$$

$$=\begin{bmatrix}z_{3.1}\\ z_{3.1}\oplus z_{3.1}\oplus z_{3.2}\\ z_{3.1}\oplus z_{3.1}\oplus z_{3.2}\oplus z_{3.1}\oplus z_{4.1}\oplus z_{4.2}\oplus z_{3.1}\oplus z_{3.2}\oplus z_{4.2}\\ z_{3.1}\oplus z_{3.2}\oplus z_{3.1}\oplus z_{3.2}\oplus z_{4.2}\end{bmatrix}=\begin{bmatrix}z_{3.1}\\ z_{3.2}\\ z_{4.1}\\ z_{4.2}\end{bmatrix}. \qquad (52)$$

При реалізації даної операції були задіяні обидві псевдовипадкові послідовності, а це підвищує якість загального результату перетворення.

Приклад 11. Перевіримо обробку четвертого та п’ятого блоків змінних перетворення за допомогою четвертого алгоритму криптографічного перетворення інформації

$$G_{6,3}^{k}=\begin{bmatrix}F_1\oplus F_2\\ F_1\end{bmatrix}. \qquad (53)$$

Позначимо F_1 через $F_{1;5}^{k}(z_5)$, тоді

$$F_1=F_{1;5}^{k}(z_5)=F_{3,6}^{k}(z_5)=\begin{bmatrix}z_{5.1}\\ z_{5.1}\oplus z_{5.2}\end{bmatrix}. \qquad (54)$$

Позначимо F_2 через $F^k_{2;4}(z_4)$, тоді

$$F_2 = F^k_{2;4}(z_4) = F^k_{6,3}(z_4) = \begin{bmatrix} z_{4.1} \oplus z_{4.2} \\ z_{4.1} \end{bmatrix}. \quad (55)$$

Вираз (53) після підстановки значення виразів (54), (55) буде наступним:

$$G^k_{6,3} = \begin{bmatrix} F_1 \oplus F_2 \\ F_1 \end{bmatrix} = \begin{bmatrix} F^k_{1;5}(z_5) \oplus F^k_{2;4}(z_4) \\ F^k_{1;5}(z_5) \end{bmatrix} = \begin{bmatrix} F^k_{3,6}(z_5) \oplus F^k_{6,3}(z_4) \\ F^k_{3,6}(z_5) \end{bmatrix} =$$

$$= \begin{bmatrix} z_{5.1} \oplus z_{4.1} \oplus z_{4.2} \\ z_{5.1} \oplus z_{5.2} \oplus z_{4.1} \\ z_{5.1} \\ z_{5.1} \oplus z_{5.2} \end{bmatrix}. \quad (56)$$

Переведемо перевірку існування оберненого перетворення [5]:

$$G^d_{5,6} = \begin{bmatrix} w_{4.1} \oplus w_{4.1} \oplus w_{4.2} \oplus w_{5.2} \\ w_{4.1} \oplus w_{4.2} \oplus w_{5.1} \oplus w_{5.2} \\ w_{5.1} \\ w_{5.1} \oplus w_{5.2} \end{bmatrix}. \quad (57)$$

Перевіримо коректність отриманої операції оберненого перетворення, підставивши в вираз (57) значення виразу (56).

$$G^d_{5,6} = \begin{bmatrix} w_{4.1} \oplus w_{4.1} \oplus w_{4.2} \oplus w_{5.2} \\ w_{4.1} \oplus w_{4.2} \oplus w_{5.1} \oplus w_{5.2} \\ w_{5.1} \\ w_{5.1} \oplus w_{5.2} \end{bmatrix} = \quad (58)$$

$$= \begin{bmatrix} z_{4.1} \oplus z_{4.2} \oplus z_{5.1} \oplus z_{4.1} \oplus z_{4.2} \oplus z_{5.1} \oplus z_{4.1} \oplus z_{5.1} \oplus z_{5.2} \oplus z_{5.1} \oplus z_{5.2} \\ z_{4.1} \oplus z_{4.2} \oplus z_{5.1} \oplus z_{4.1} \oplus z_{5.1} \oplus z_{5.2} \oplus z_{5.1} \oplus z_{5.1} \oplus z_{5.2} \\ z_{5.1} \\ z_{5.1} \oplus z_{5.2} \oplus z_{5.1} \end{bmatrix} = \begin{bmatrix} z_{4.1} \\ z_{4.2} \\ z_{5.1} \\ z_{5.2} \end{bmatrix}$$

При реалізації даної операції були задіяні обидві псевдовипадкові послідовності, а це повинно привести до підвищення якості загального результату перетворення.

Приклад 12. Позначимо F_1 через $F^k_{1;4}(z_4)$, тоді

$$F_1 = F^k_{1;4}(z_4) = F^k_{6,5}(z_4) = \begin{bmatrix} z_{4.1} \oplus z_{4.2} \\ z_{4.2} \end{bmatrix}. \quad (59)$$

Позначимо F_2 через $F^k_{2;5}(z_5)$, тоді

$$F_2 = F^k_{2;5}(z_5) = F^k_{5,6}(z_5) = \begin{bmatrix} z_{5.2} \\ z_{5.1} \oplus z_{5.2} \end{bmatrix}. \quad (60)$$

Підставивши в вираз (53) значення виразів (59) та (60) отримаємо:

$$G_{6,3}^{k} = \begin{bmatrix} F_1 \oplus F_2 \\ F_1 \end{bmatrix} = \begin{bmatrix} F_{1;4}^{k}(z_4) \oplus F_{2;5}^{k}(z_5) \\ F_{1;4}^{k}(z_4) \end{bmatrix} = \begin{bmatrix} F_{6,5}^{k}(z_4) \oplus F_{5,6}^{k}(z_5) \\ F_{6,5}^{k}(z_4) \end{bmatrix} =$$

$$= \begin{bmatrix} z_{4.1} \oplus z_{4.2} \oplus z_{5.2} \\ z_{4.2} \oplus z_{5.1} \oplus z_{5.2} \\ z_{4.1} \oplus z_{4.2} \\ z_{4.2} \end{bmatrix}. \quad (61)$$

Використавши даний метод [5] отримаємо:

$$G_{5,6}^{d} = \begin{bmatrix} w_{4.1} \oplus w_{4.1} \oplus w_{5.1} \oplus w_{5.2} \\ w_{5.2} \\ w_{4.1} \oplus w_{4.2} \oplus w_{5.1} \oplus w_{5.2} \\ w_{4.1} \oplus w_{5.1} \oplus w_{5.2} \oplus w_{5.2} \end{bmatrix}. \quad (62)$$

Перевіримо коректність отриманої операції оберненого перетворення підставивши в вираз (62) значення виразу (61).

$$G_{5,6}^{d} = \begin{bmatrix} w_{4.1} \oplus w_{4.1} \oplus w_{5.1} \oplus w_{5.2} \\ w_{5.2} \\ w_{4.1} \oplus w_{4.2} \oplus w_{5.1} \oplus w_{5.2} \\ w_{4.1} \oplus w_{5.1} \oplus w_{5.2} \oplus w_{5.2} \end{bmatrix} =$$

$$\begin{bmatrix} z_{4.1} \oplus z_{4.2} \oplus_{5.2} \oplus z_{4.1} \oplus z_{4.2} \oplus z_{5.2} \oplus z_{4.1} \oplus z_{4.2} \oplus z_{4.2} \\ z_{5.2} \\ z_{4.1} \oplus z_{4.2} \oplus_{5.2} \oplus z_{4.2} \oplus z_{5.1} \oplus z_{5.2} \oplus z_{4.1} \oplus z_{4.2} \oplus z_{4.2} \\ z_{4.1} \oplus z_{4.2} \oplus_{5.2} \oplus z_{4.1} \oplus z_{4.2} \oplus z_{4.2} \oplus z_{4.2} \end{bmatrix} = \begin{bmatrix} z_{4.1} \\ z_{4.2} \\ z_{5.1} \\ z_{5.2} \end{bmatrix}. \quad (63)$$

В результаті отримано коректну операцію оберненого криптографічного перетворення, яку можна реалізувати як операцію обробки чотирьох бітів інформації при паралельній реалізації, або обробки двох блоків інформації з наступною обробкою кожного блоку окремо при послідовній реалізації.

Отримані результати перевірені на повній множині операцій та алгоритмів.

Висновки

Використання матричних операцій криптографічного перетворення в поєднанні з груповими операціями криптографічного перетворення забезпечує підвищення якості шифрування (отриманої псевдовипадкової послідовності), а також забезпечує можливість розшифрування інформації, так як забезпечує використання умови отримання невиродженого перетворення. Подальші дослідження будуть направлені на пошук алгоритмів побудови обернених перетворень меншої складності.

Література

1. Ланських Є.В. Оцінка якості псевдовипадкових послідовностей на основі використання операцій додавання за модулем два / Є.В. Ланських, С.В. Сисоєнко, М.О. Пустовіт // Наука і техніка Повітряних Сил Збройних Сил України. – 2015. – №4(21) – С. 147 - 150. – Режим доступу: http://nbuv.gov.ua/UJRN/Nitps_2015_4_36.

2. Фауре Е.В. Синтез і аналіз псевдовипадкових послідовностей на основі операцій криптографічного перетворення / Е.В. Фауре, С.В. Сисоєнко, Т.В. Миронюк // Системи управління, навігації та зв'язку. – Полтава, ПНТУ, 2015. – №4(36) – С. 85 - 87.

3. Рудницький В.М. Оцінка якості псевдовипадкових послідовностей на основі додавання за модулем / В.М. Рудницький, Е.В. Фауре, С.В. Сисоєнко // Вісник інженерної академії України. – Київ, 2016. – №3 – С. 219 - 221.

4. Фауре Е.В. Метод підвищення стійкості псевдовипадкових послідовностей до лінійного криптоаналізу/ Е.В. Фауре, С.В. Сисоєнко// Науковий прогрес та процес розвитку країни в аспекті євроінтеграції : зб. наук. праць «ΛΌГОΣ». СПЕЦВИПУСК. - 2016. - Т.1. - С. 119 - 122.

5. Рудницький В. М. Метод синтезу матричних моделей операцій криптографічного кодування та декодування інформації / В. М. Рудницький, В. Г. Бабенко, С. В. Рудницький // Збірник наукових праць Харківського університету Повітряних Сил – Вип. 4 (33). – Х. : ХУПС ім. І. Кожедуба, 2012. – С. 198 – 200.

МЕТОД СЕЛЕКЦИИ ЗНАЧИМЫХ СТРУКТУРНЫХ ЕДИНИЦ ВИДЕОКАДРА ДЛЯ КОДИРОВАНИЯ ВИДЕОДАННЫХ

Тарнаполов Р.В.

Введение

На сегодняшний день, в современных системах управления ведомственных структур перед разработчиками стоит вопрос внедрения технологий безопасности. Безопасность касается и систем видеоконференцсвязи, которая широко используется ведомственными структурами, для организации управления, в режиме реального времени.

Существующие технологии скрытия видеоинформационных ресурсов, которые обеспечивают необходимую конфиденциальность, при использовании технологии гарантированной стойкости. Однако они не учитывают целостность и доступность видеоинформации. Эти методы имеют существенный недостаток: их работа основана на закрытии всего потока, что приводит к некорректной работе системы видеоконференцсвязи, при использовании каналов связи с низкой пропускной способностью. Возникает проблема в достижении конфиденциальности в условиях обеспечения требуемой целостности и доступности видеоинформации. Для решения этой проблемы разрабатываются селективные методы шифрования.

Под селективным шифрованием понимается использование средств криптографической защиты изображений в процессе сжатия на разных этапах компрессии. Селективное шифрование используется во временной и пространственной областях. Его суть заключается в скрытии наиболее значимых компонент видеопотока. Эти компоненты формируются в процессе сжатия видеоданных. Существующие селективные методы имеют ряд недостатков. Селективный метод был предложен Тангом - это способ считывания квантованных коэффициентов ДКП случайным образом. Недостатком которого есть низкая компрессия видеоданных. Способ шифрования низкочастотных коэффициентов ДКП был предложен Ченгом и Ли, в котором шифруются только низкочастотные составляющие дискретного косинусного преобразования, т.е. коэффициенты, расположенные в верхнем левом углу каждой матрицы коэффициентов ДКП. Недостатком данного способа является возможность извлечь информацию из-за появления контуров объектов на изображении, информация о которых может содержаться в незашифрованных высокочастотных коэффициентах ДКП. Для устранения этих недостатков нужно применять другие селективные подходы шифрования. Предлагается повысить эффективность выявления значимых блоков, которые включают в себя значимые структурные единицы. Что повысить защищенность видеокадра и позволит уменьшить объем и время обработки шифрованных видеоданных также.

Таким образом, целью статьи является разработка и исследование метода селекции значимых структурных единиц видеокадра для кодирования видеоданных, в процессе сжатия.

Основная часть

Для селекции значимых структурных единиц $S_{зн}$ предлагается выявлять наиболее информативные, в плане структурного и семантического содержания, составляющие базового кадра. Поскольку наиболее полную информацию несет яркостная составляющая видеокадра K_I, то значимые структурные единицы предлагается выявлять на базе яркостных компонент. Поэтому принятие решения по закрытию структурной единицы предлагается осуществлять по результатам анализа информационной составляющей совокупности блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей.

Для определения энергетической насыщенности блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ предлагается ввести понятии блоков трех типов:

- слабонасыщенные блоки (блоки, в которых присутствуют равномерные участки изображения);
- средней насыщенности (блоки, в которых имеются незначительные отличия между пикселями, соответственно присутствуют плавные переходы контрастности);
- сильнонасыщенные блоки (блоки, в которых присутствуют резкие переходы яркости и контрастности изображения).

Определение энергетической насыщенности блоков предлагается осуществлять после ДКП. С помощью дискретного косинусного преобразования осуществляется переход от пространственно-временного представления видеокадра в пространственно-спектральное. Компоненты трансформанты ДКП являются интегральными характеристиками структурного содержания фрагмента изображения. Причем интегральные свойства компонент зависят от их положения в трансформанте (рис 1).

На рис. 1 представлено расположение низкочастотных и высокочастотных компонент трансформанты ДКП в блоках $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей макроблока. Из рис. 1 видно, что низкочастотные компоненты находятся в области первых пяти диагоналей, среднечастотные компоненты расположены в пределах пятой-девятой диагоналей, а высокочастотные компоненты – между девятой и тринадцатой диагоналями.

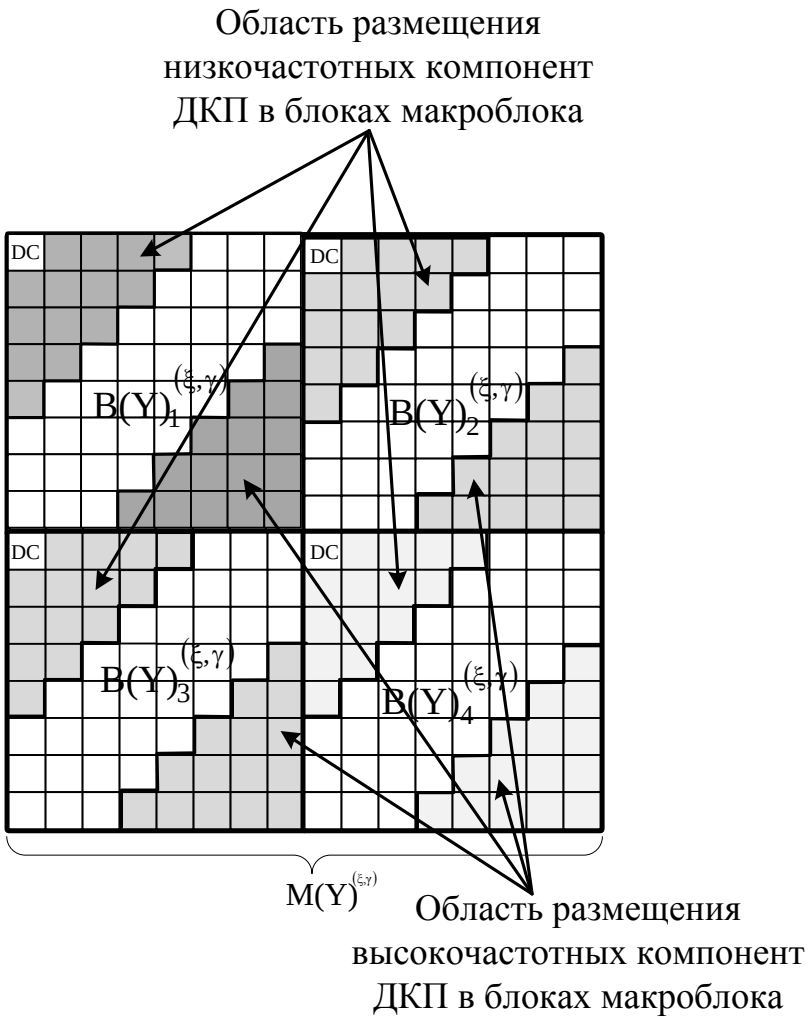


Рис. 1. Схема расположения низкочастотных и высокочастотных компонент трансформанты ДКП в блоках $B(Y)_{\varphi}^{(\xi,\gamma)}$ *яркостной составляющей макроблока*

Интегральная зависимость компонент трансформанты ДКП выглядит следующим образом:

1. Значение компоненты в верхнем левом углу трансформанты ДКП пропорциональны средней яркости изображения. Они характеризуют степень насыщенности блока изображения низкочастотными перепадами. К низкочастотным перепадам относят ступенчатые изменения уровня яркости или координаты цвета.

2. Компоненты в средней части трансформанты определяют степень насыщенности блока изображения линейными, равномерными изменениями уровня яркости.

3. Значения компонент в нижней правой области трансформанты ДКП характеризуют степень насыщенности высокочастотными перепадами блока изображения. К высокочастотным перепадам относят импульсные изменения значений элементов изображений.

Поэтому можно сделать вывод о том, что энергией блока называется величина, характеризующая наличие неоднородно визуальных контуров блока изображения.

Значения компонент изменяются по мере преобладания в изображении различных структурных особенностей.

Широкий класс изображений содержит в основном линейные, монотонные и ступенчатые структурные изменения уровня яркости. Импульсные изменения занимают меньшую площадь изображения. Кроме того, они могут быть вызваны шумами дискретизации. Поэтому наибольшие значения имеют компоненты, расположенные в верхней левой части трансформанты. Компоненты в нижней части трансформанты соответствуют высокочастотным изменениям и поэтому имеют меньшие значения.

На рис. 2 показано расположение компонент в трансформанте ДКП блока яркостной составляющей видеокадра.

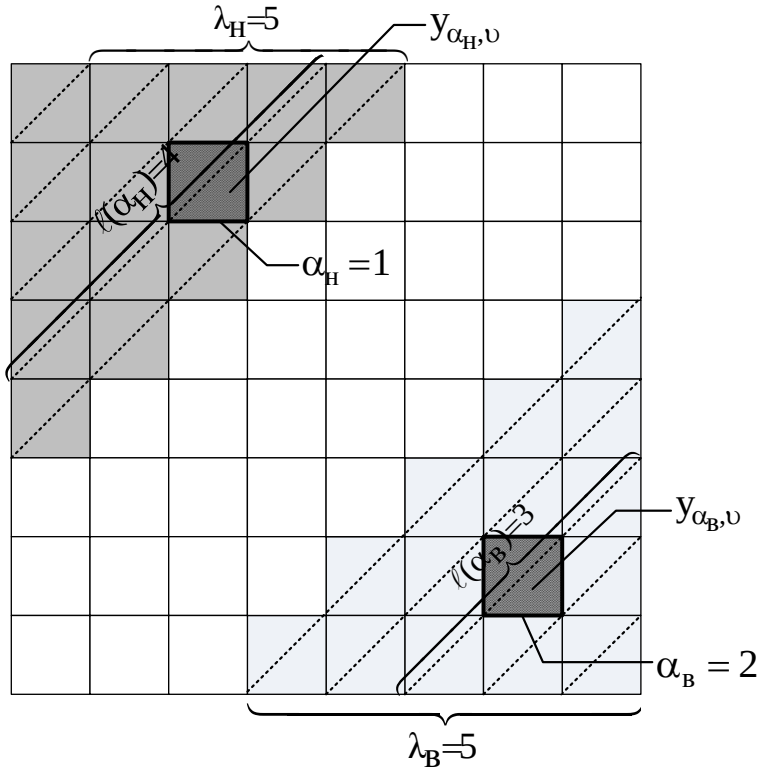


Рис. 2. Схема расположения компонент в трансформанте ДКП блока яркости видеокадра

Для трансформанты ДКП в сильнонасыщенных блоках изображениях характерны следующие особенности:

- значения компонент ДКП уменьшается по диагональному зигзагу слева - направо, сверху – вниз;
- компоненты ДКП с большими значениями сконцентрированы в относительно малой области трансформанты. Компоненты с минимальными значениями занимают большую площадь трансформанты;

- при большой площади изображения, имеющей мало изменяющуюся яркость, размер области трансформанты с большими значениями компонент имеет маленькую площадь.

Предлагается оценивать структурную и семантическую информативность структурной единицы с позиции спектральных характеристик. Очевидно, что чем больше площадь однородной яркостной площади, и чем меньше площадь, заполненная мелкими деталями, тем меньше степень структурной и семантической информативности обрабатываемого блока видеокадра. Наоборот, чем чаще яркостные перепады, и чем больше площадь, отводимая под мелкие детали и контурные перепады, тем выше структурная и семантическая информативность. В связи с чем, для оценки значимости структурных единиц предлагается использовать информацию, содержащуюся в спектральном представлении изображения.

Для определения блоков с выраженными яркостными ступенчатыми перепадами предлагается использовать информацию, содержащуюся в совокупности низкочастотных компонент.

Такую информацию предлагается оценивать с помощью показателя $Z(B_н)_φ^{(ξ,γ)}$ суммарных значений низкочастотных компонент, которые находятся в первых 4-х диагоналях $(1 \le \lambda_н \le 5)$. Показатель $Z(B_н)_φ^{(ξ,γ)}$ рассчитывается по следующей формуле:

$$Z(B_н)_φ^{(ξ,γ)} = \frac{\log_2 \sum_{\alpha_н=1}^{\lambda_н} \sum_{\upsilon=1}^{\ell(\alpha_н)} y_{\alpha_н,\upsilon}^2}{\sum_{\alpha_н=1}^{\lambda_н} \ell(\alpha_н)}, \qquad (1)$$

где $Z(B_н)_φ^{(ξ,γ)}$ – показатель, который определяет суммарное значение низкочастотных компонент ДКП блока $B(Y)_φ^{(ξ,γ)}$ яркости; $y_{\alpha_н,\upsilon}$ – значение компоненты трансформанты; $\lambda_н$ – количество диагоналей с низкочастотными компонентами в трансформанте; υ – индекс элемента внутри $\alpha_н$-ой диагонали; $\alpha_н$ – индекс низкочастотной $\lambda_н$-ой диагонали; $\ell(\alpha_н)$ – длина низкочастотной $\alpha_н$-ой диагонали.

Выражение (1) позволяет определить наличие значительных яркостных перепадов в блоках яркости. Такой подход не учитывает мелкую деталировку. Соответственно, он не позволяет с полной уверенностью определить блоки с сильной информационной насыщенностью.

В случае большой концентрации мелких деталей в блоке видеокадра K_I, увеличиваются значения высокочастотных компонент трансформанты

ДКП. Поэтому для более точного определения значимости структурных единиц с учетом концентрации мелких деталей в блоках яркостной составляющей предлагается дополнительно оценивать информацию на основе концентрации высокочастотной деталировки видеокадра.

Для этого необходимо оценивать показатель $Z(B_в)_φ^{(ξ,γ)}$ суммарных значений высокочастотных компонент 10-13 диагоналей ($10 \le \lambda_в \le 13$). Показатель $Z(B_в)_φ^{(ξ,γ)}$ определяется следующим образом:

$$Z(B_в)_φ^{(ξ,γ)} = \frac{\log_2 \sum_{\alpha_в=1}^{\lambda_в} \sum_{υ=1}^{\ell(\alpha_в)} y_{\alpha_в,υ}^2}{\sum_{\alpha_в=1}^{\lambda_в} \ell(\alpha_в)}, \qquad (2)$$

где $Z(B_в)_φ^{(ξ,γ)}$ – показатель, который определяет суммарное значение высокочастотных компонент ДКП блока $B(Y)_φ^{(ξ,γ)}$ яркости; $y_{\alpha_в,υ}$ – значение компоненты трансформанты; $\lambda_в$ – количество диагоналей с высокочастотными компонентами в трансформанте ДКП; $υ$ – индекс элемента внутри $\alpha_в$-ой диагонали; $\alpha_в$ – индекс высокочастотной $\lambda_в$-ой диагонали; $\ell(\alpha_в)$ – длина низкочастотной $\alpha_в$-ой диагонали.

Показатель $Z(B_н)_φ^{(ξ,γ)}$ суммарных значений низкочастотных компонент и показатель $Z(B_в)_φ^{(ξ,γ)}$ суммарных значений высокочастотных компонент, которые получены в результате расчетов (1), (2), позволяют классифицировать блоки яркостной составляющей видеокадра K_I по степени насыщенности.

Оценку значимости структурной единицы $S^{(ξ,γ)}$ предлагается осуществлять на основе энергетической значимости макроблока $M(Y)^{(ξ,γ)}$ яркостной составляющей. В свою очередь, оценку значимости макроблока $M(Y)^{(ξ,γ)}$ яркостной составляющей предлагается проводить на основе структурной и семантической насыщенности блока $B(Y)_φ^{(ξ,γ)}$. Для этого необходимо разработать метод, базирующийся на системе правил для принятия решения по энергетической значимости структурных единиц и макроблоков на основе информации о значимости блоков яркостной составляющей.

В основе правил лежит система сравнения показателя $Z(B_н)_φ^{(ξ,γ)}$ совокупности значений низкочастотных компонент с пороговыми

значениями $\delta_{min_н}$ и $\delta_{max_н}$. Будем считать, что $\delta_{max_н}$ – верхний предел для оценки показателя $Z(B_н)_{\varphi}^{(\xi,\gamma)}$ совокупности значений низкочастотных компонент блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей. $\delta_{min_н}$ – нижний предел для оценки показателя $Z(B_н)_{\varphi}^{(\xi,\gamma)}$ совокупности значений низкочастотных компонент блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей.

Предлагается проводить оценку энергетической значимости макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей базового видеокадра K_I. Макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей будет считаться энергетически значимым в двух случаях:

1. Если в состав макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей входит один и больше блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ с высокой степенью семантической и структурной насыщенности. Это можно описать следующим выражением:

$$M(Y)^{(\xi,\gamma)} = M(Y)_{зн}^{(\xi,\gamma)} \text{ и } M=1, \text{ если } Z(B_н)_{\varphi}^{(\xi,\gamma)} > \delta_{max_н}.$$

2. Если в состав макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей входят два $N_{sr} = 2$ и больше $N_{sr} > 2$ блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ со средней степенью семантической и структурной насыщенности, то есть выполняется неравенство:

$$(\delta_{min_н} \le Z(B_н)_{\varphi}^{(\xi,\gamma)} \le \delta_{max_н}),$$

тогда:

$$M(Y)^{(\xi,\gamma)} = M(Y)_{зн}^{(\xi,\gamma)} \text{ и } M=1 \text{ если } N_{sr} \ge 2,$$

$$N_{sr} = N_{sr} + 1, \text{ если } (\delta_{min_н} \le Z(B_н)_{\varphi}^{(\xi,\gamma)} \le \delta_{max_н}),$$

где N_{sr} – количество блоков со средней структурной и семантической насыщенности.

Остальные структурные единицы обрабатываются по стандартному алгоритму видеокомпрессии.

Структурная схема метода селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ с использованием информации по совокупности значений низкочастотных компонент трансформанты ДКП блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$, $\varphi = \overline{1,4}$ яркостной составляющей представлена на рис. 3.

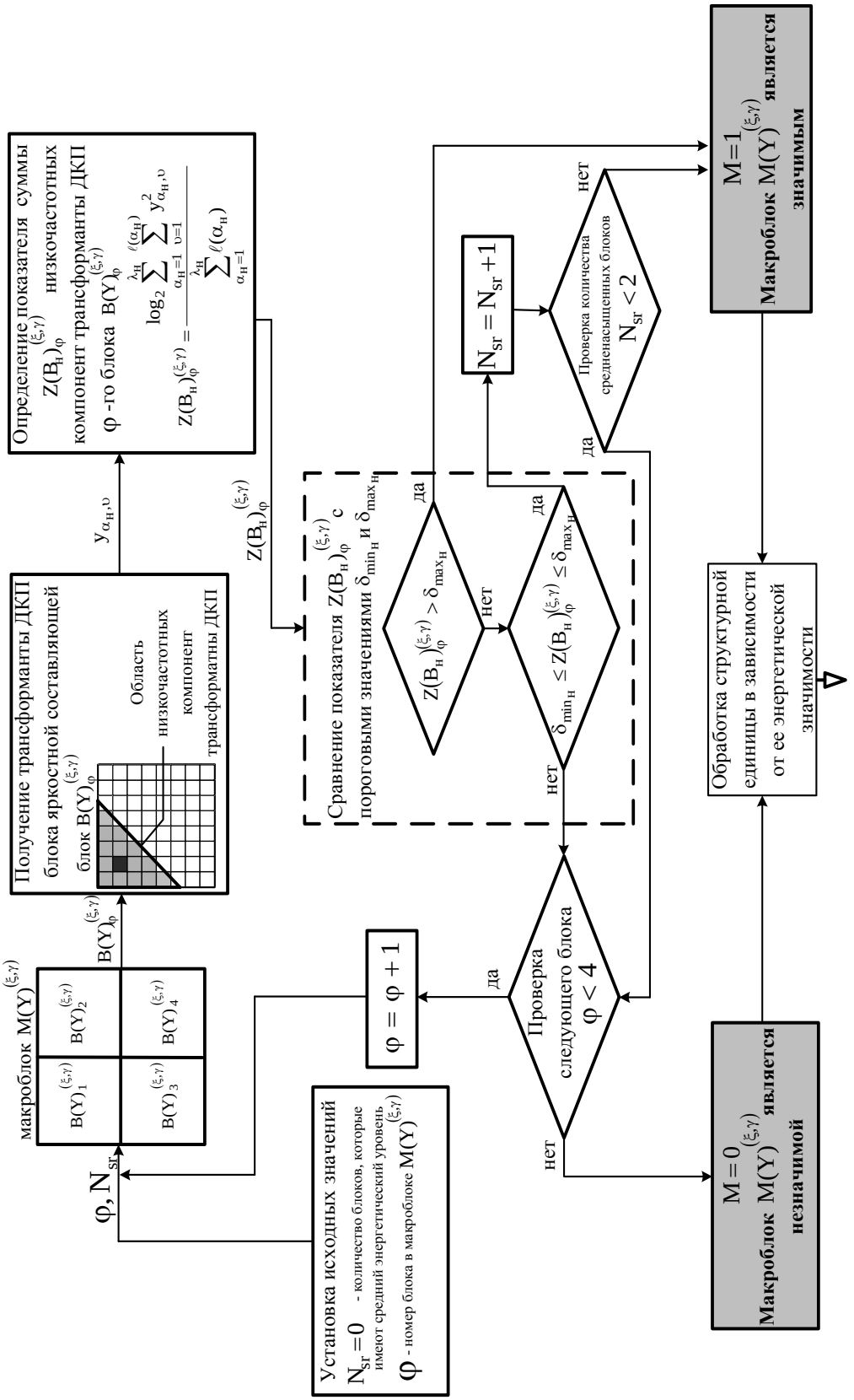


Рис. 3. Структурная схема метода селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ с использованием информации по совокупности значений низкочастотных компонент трансформанты ДКП блоков $B(Y)_\varphi^{(\xi,\gamma)}$ яркостной составляющей

Процесс выбора значимого макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей (рис. 2) происходит следующим образом:

1. В начале проверки значимого макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей переменная N_{sr} для подсчета средненасыщенных блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей, принимает значение $N_{sr}=0$, а переменная φ, которая определяет номер блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей для проверки, принимает значение $\varphi=1$.

2. После образования трансформант ДКП блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей их проверка осуществляется по очереди с 1-го по 4-й блок.

3. Для блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей производится расчет показателя $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ для совокупности значений низкочастотных компонент с учетом выражения (1).

4. Показатель $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ сравнивается с пороговыми значениями δ_{min_i} и δ_{max_i} для определения энергетической насыщенности блока $B(Y)_{\varphi}^{(\xi,\gamma)}$. Если значения показателя $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ для блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей превышает верхний порог $\delta_{max_н}$, то блок считается энергетически значимым по степени структурной и семантической насыщенности

$$Z(B_{i})_{\varphi}^{(\xi,\gamma)} > \delta_{max_i}\ .$$

В этом случае метка M принимает значение M=1, соответственно макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей будет считаться энергетически значимым. В результате чего алгоритм проверки останавливается.

6. Если показатель $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ суммарных значений низкочастотных компонент блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей находится между пороговыми значениями

$$\delta_{min_i} \le Z(B_{i})_{\varphi}^{(\xi,\gamma)} \le \delta_{max_i}\ ,$$

то блок $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей будет средненасыщенным, а переменная N_{sr} для подсчета средненасыщенных блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей, примет значение

$$N_{sr} = N_{sr} + 1.$$

Для того, чтобы считать, что макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей обладает высокой энергетической значимостью,

необходимо наличие двух и более блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей, входящих в его состав.

7. После чего проверяется количество средненасыщенных блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей. Если количество блоков со средней степенью семантической и структурной насыщенности больше или равно двум:

$$N_{sr} \geq 2,$$

то макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей будет считаться энергетически значимым. В этом случае также метка M принимает значение M=1, а макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей будет считаться энергетически значимым. После чего дальнейшая проверка блоков яркостной составляющей прекращается. Если в результате проверки всех 4-х блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей значение показателя $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ блоков оказалось меньше нижнего порогового значения

$$\delta_{min_{i}} > Z(B_{i})_{\varphi}^{(\xi,\gamma)},$$

или количество средненасыщенных блоков меньше

$$N_{sr} < 2,$$

то метка M принимает значение M=0. Соответственно, такой макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей будет считаться энергетически значимым.

В результате чего энергетическая значимость структурной единицы $S^{(\xi,\gamma)}$ определяется на основе энергетической значимости макроблока $M(Y)^{(\xi,\gamma)}$. Таким образом, структурная единица считается значимой $S^{(\xi,\gamma)} = S_{зн}^{(\xi,\gamma)}$ если в результате проверки макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей по информации о совокупности значений низкочастотных компонент трансформанты ДКП блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ метка приняла значение M=1.

На рис. 4 представлен алгоритм селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ с использованием информации по совокупности значений низкочастотных компонент трансформанты ДКП блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей.

Разработанный метод позволяет выявлять (селекционировать) значимые структурные единицы $S_{зн}$ базового видеокадра K_I на основе оценки показателя $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ по совокупности значений низкочастотных

компонент блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей с пороговыми значениями. В результате работы такого метода происходит выявления участков изображения базового видеокадра, которые обладают выраженными структурными переходами, текстурными и яркостными перепадами. Наряду с этим появляется значительный недостаток.

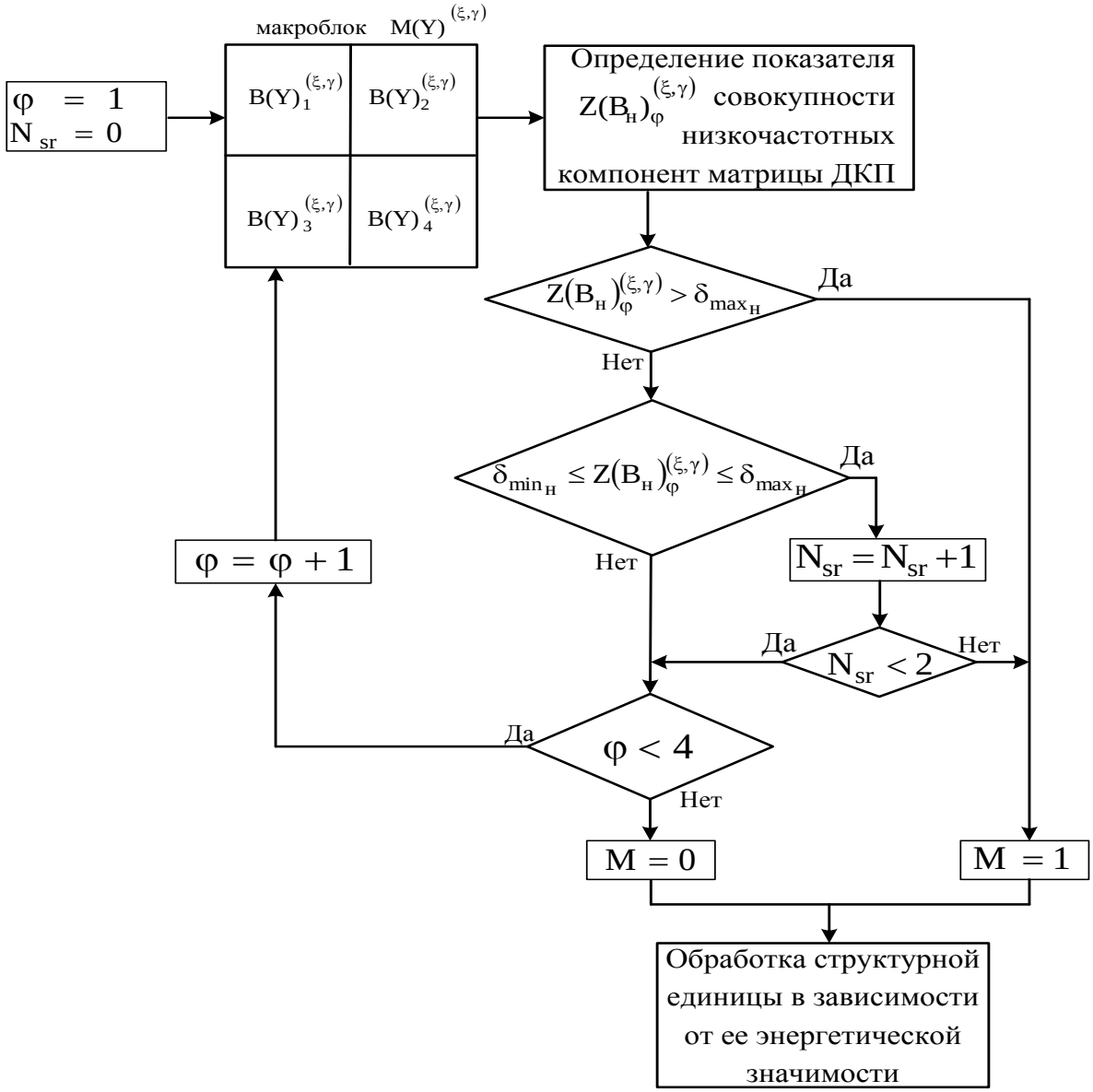


Рис. 4. Алгоритм значимых структурных единиц $S^{(\xi,\gamma)}$ *с использованием информации по совокупности значений низкочастотных компонент трансформанты ДКП блоков* $B(Y)_{\varphi}^{(\xi,\gamma)}$ *яркостной составляющей*

Метод выявления значимых структурных единиц $S_{зн}$, основанный на оценки только низкочастотного показателя $Z(B_i)_{\varphi}^{(\xi,\gamma)}$ блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей, также определяет как значимые структурные

единицы базового видеокадра, для которых характерны следующие визуальные особенности:

1. Фрагменты видеоизображения с выраженными текстурными перепадами. Такими фрагментами изображения могут быть текстурные перепады.

2. Фрагменты фоновых однородных областей видеокадра, имеющие высокую яркостную насыщенность, но в которых присутствуют контрастные незначимые мелкие детали.

Таким образом, в результате обработки видеоизображения возникают ошибки второго рода, когда незначимая область изображения будет идентифицирована как значимая. Поэтому в случае использования метода селекции видеоданных в ведомственных системах ВКС, основанного только на анализе низкочастотных компонент блоков яркостной составляющей базового видеокадра, будет формироваться избыточное количество структурных единиц $S_{зн}^{(\xi,\gamma)}$, которые закрываются. Это приводит к увеличению времени обработки и снижению интенсивности закрытых видеоданных.

Для снижения вероятности ошибки второго рода предлагается для более точной идентификации энергетически насыщенных структурных единиц $S_{зн}^{(\xi,\gamma)}$ дополнительно учитывать информацию по высокочастотным компонентам блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей. Это позволит отсекать структурные единицы, для которых $S_{незн}^{(\xi,\gamma)}$ характерны такие особенности:

1. Наличие однородных фрагментов изображения с высокой яркости и контрастности, в состав которых входят незначимые мелкие детали.

2. Наличие фрагментов изображения с выраженными текстурными перепадами, которые не являются значимыми.

Поэтому предлагается дополнительно разработать правило для оценки высокочастотной составляющей в сильно- и средненасыщенных блоках $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей на основе сравнения показателя $Z(B_{в})_{\varphi}^{(\xi,\gamma)}$ по совокупности значений высокочастотных компонент блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей с пороговым значением $\delta_{в}$.

Макроблок $M(Y)^{(\xi,\gamma)}$ яркостной составляющей будет считаться энергетически значимым $M(Y)^{(\xi,\gamma)} = M(Y)_{зн}^{(\xi,\gamma)}$ и M=1 в следующих случаях:

1. Если одновременно выполняются следующие условия: значения показателя $Z(B_{i})_{\varphi}^{(\xi,\gamma)}$ по низкочастотной составляющее для блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей превышает верхний порог $\delta_{max_{i}}$:

$$Z(B_{í})_{\varphi}^{(\xi,\gamma)} > \delta_{max_{í}},$$

и значения показателя $Z(B_{в})_{\varphi}^{(\xi,\gamma)}$ по высокочастотной составляющее для блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ превышает порог $\delta_{в}$. Это можно описать следующим выражением:

$$Z(B_{н})_{\varphi}^{(\xi,\gamma)} > \delta_{max_{н}} \text{ и } Z(B_{в})_{\varphi}^{(\xi,\gamma)} > \delta_{в}.$$

2. Если одновременно выполняются следующие условия: в состав макроблока $M(Y)^{(\xi,\gamma)}$ яркостной составляющей входят два $N_{sr} = 2$ и больше $N_{sr} > 2$ блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$, и значения показателя $Z(B_{í})_{\varphi}^{(\xi,\gamma)}$ по низкочастотной составляющей которых находятся в пределах пороговых значений $\delta_{min_{í}}$ и $\delta_{max_{í}}$:

$$(\delta_{min_{í}} \le Z(B_{í})_{\varphi}^{(\xi,\gamma)} \le \delta_{max_{í}}),$$

а значения показателя $Z(B_{в})_{\varphi}^{(\xi,\gamma)}$ по высокочастотной составляющее для блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ превышает порог $\delta_{в}$, т. е. выполняется неравенство:

$$Z(B_{в})_{\varphi}^{(\xi,\gamma)} > \delta_{в}.$$

Структурная схема метода селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ с использованием информации по совокупности значений низкочастотных и высокочастотных компонент трансформанты ДКП блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$, $\varphi = \overline{1,4}$ яркостной составляющей представлена на рис. 5.

На рис. 6 представлен алгоритм селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ с использованием информации по совокупности значений низкочастотных и высокочастотных компонент трансформанты ДКП блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$, $\varphi = \overline{1,4}$ яркостной составляющей.

Общее правило для определения энергетически значимой структурной единицы, где $S_{зн}^{(\xi,\gamma)} = S^{(\xi,\gamma)}$ если M=1, будет иметь следующий вид:

$$M = \begin{cases} 1, \rightarrow (Z(B_{н})_{\varphi}^{(\gamma,\xi)} > \delta_{max_{н}}) \vee (Z(B_{в})_{\varphi}^{(\gamma,\xi)} > \delta_{в}); \\ 1, \rightarrow (\delta_{min_{н}} \le Z(B_{н})_{\varphi}^{(\gamma,\xi)} \le \delta_{max_{н}}) \vee (Z(B_{в})_{\varphi}^{(\gamma,\xi)} > \delta_{в}) \vee (N_{sr} \ge 2). \end{cases}$$

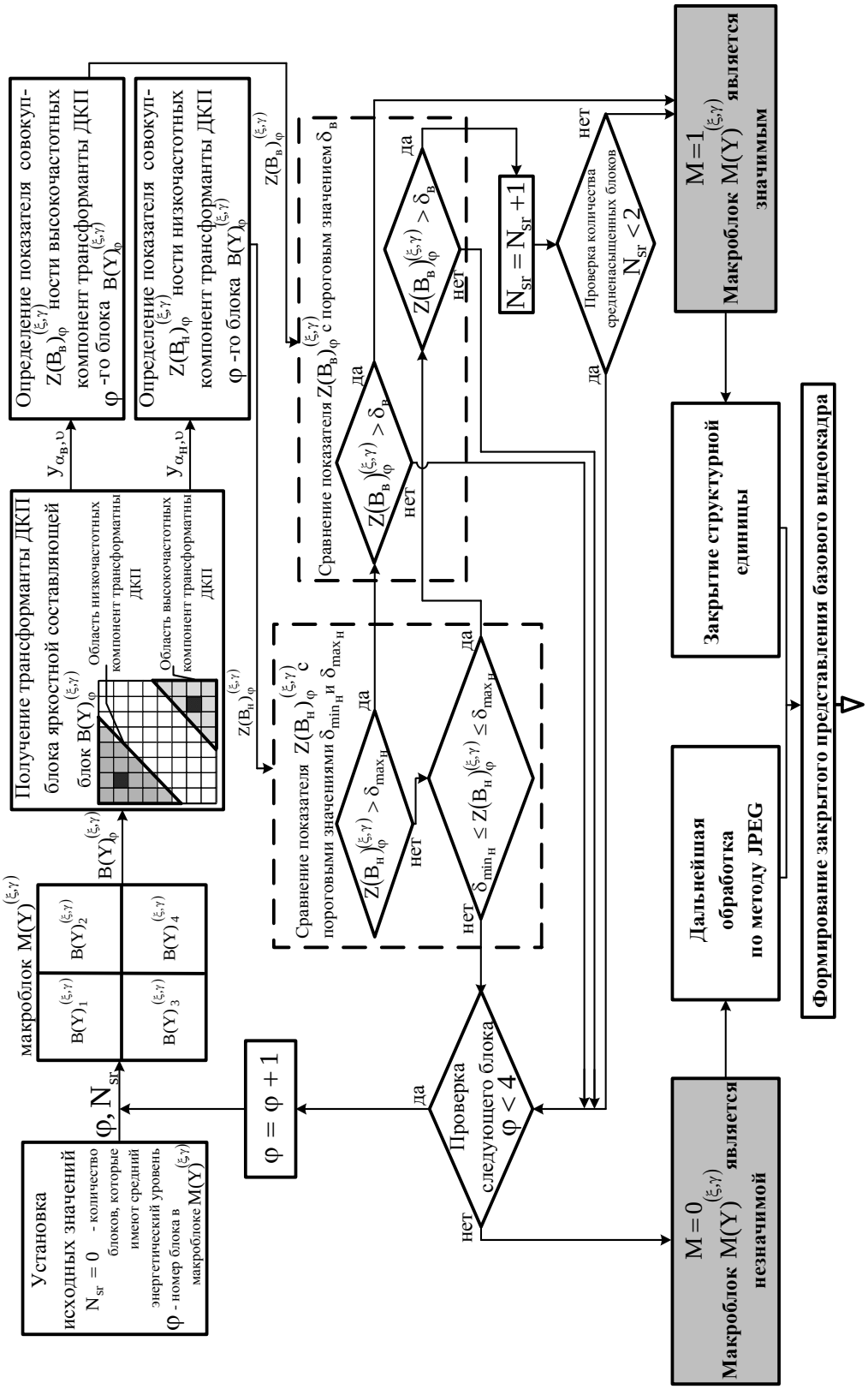


Рис. 5. Структурная схема метода селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ *с использованием информации по совокупности значений низкочастотных и высокочастотных компонент трансформанты ДКП блоков* $B(Y)_\varphi^{(\xi,\gamma)}$, $\varphi = \overline{1,4}$ *яркостной составляющей.*

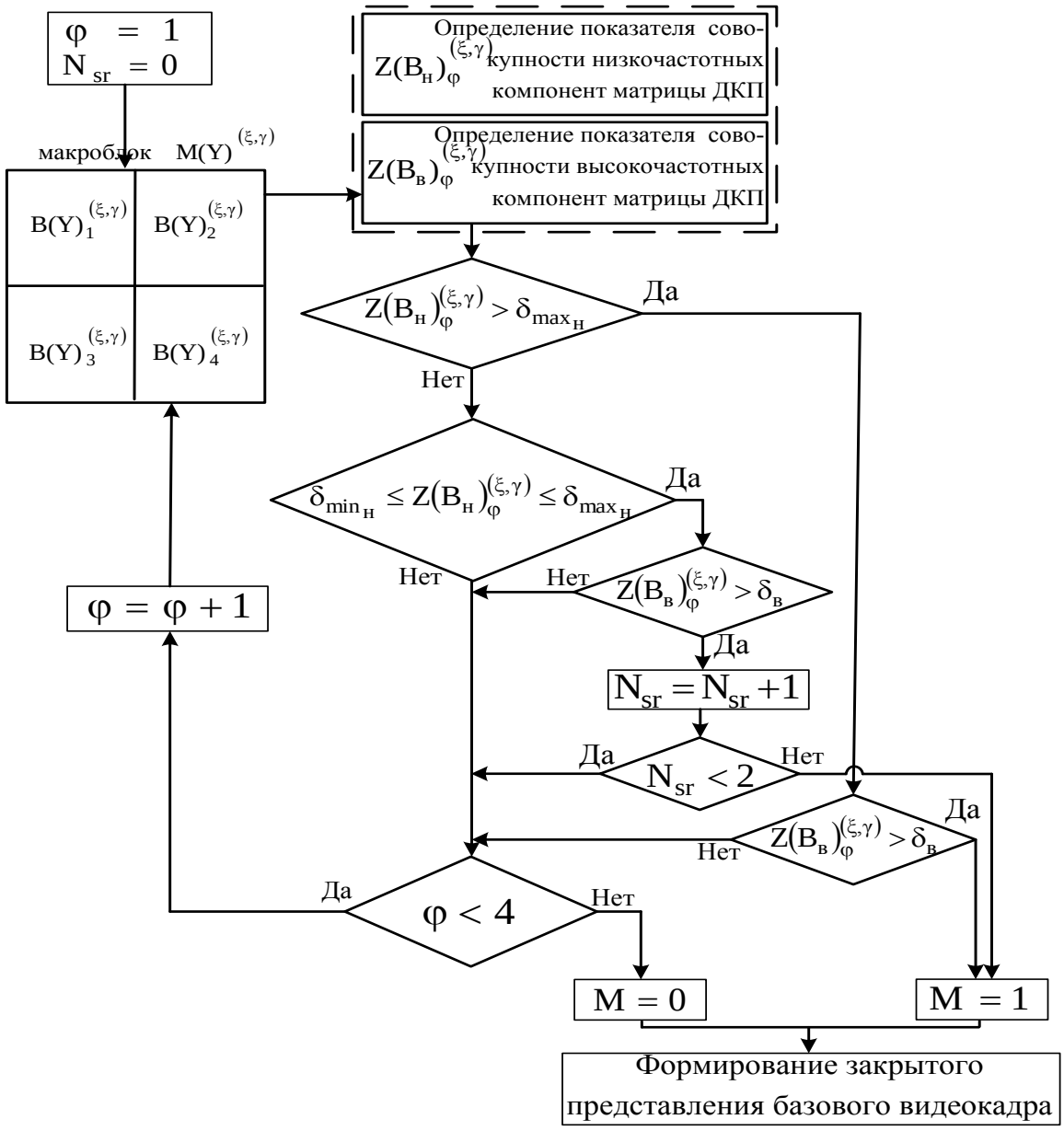


Рис. 6 Алгоритм селекции значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ *с использованием информации по совокупности значений низкочастотных и высокочастотных компонент трансформанты ДКП блоков* $B(Y)_{\varphi}^{(\xi,\gamma)}$, $\varphi=\overline{1,4}$ *яркостной составляющей.*

Выводы

Поэтому при использовании селективного подхода, основанного на закрытии значимых блоков, с одной стороны выполняются требования по обеспечению конфиденциальности и целостности видеоинформационного ресурса. Но с другой стороны, реализация такого подхода приводит к увеличению интенсивности передаваемых закрытых видеоданных, в результате чего снижается пропускная способность закрытого видеоканала.

Таким образом разработана система показателей (метрика) для выявления наиболее значимых блоков яркостной составляющей видеокадра по степени семантической и структурной насыщенности на основе оценки информации, содержащейся в суммарных значениях низкочастотных компонент и оценки информации суммарных значений высокочастотных компонент трансформанты ДКП.

Разработана методологическая база для определения энергетической значимости структурной единицы базового видеокадра, базирующаяся на системе правил для оценки структурной и семантической насыщенности блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей. Здесь учитываются как значение показателя по совокупности низкочастотных, так и значения показателя по совокупности высокочастотных компонент трансформанты ДКП блока $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей. Это позволяет:

1. Производить оценку блоков и макроблоков яркостной составляющей видеокадра по низкочастотным компонентам трансформанты ДКП для выявления участков изображения, которые обладают выраженными структурными переходами, текстурными и яркостными перепадами.

2. Производить оценку блоков и макроблоков яркостной составляющей видеокадра по высокочастотным компонентам трансформанты ДКП для выявления участков изображения, которые имеют выраженные текстурные перепады и в которых присутствуют контрастные незначимые мелкие детали.

3. Устранять ошибки второго рода и осуществлять выбор значимых структурных единиц $S_{зн}^{(\xi,\gamma)}$ с высоким уровнем определения структурной и семантической насыщенности блоков $B(Y)_{\varphi}^{(\xi,\gamma)}$ яркостной составляющей.

Основным отличием данного метода является проведение оценки информативности для структурных единиц в спектральной области на основе иерархии порогового взвешивания низкочастотных и высокочастотных составляющих. Это создает условие для закрытия видеопотока на основе технологии внутрикадровой селекции.

Литература

1. Д. Ватолин, А. Ратушняк, М.Смирнов, В. Юкин, Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: Диалог-Мифи, 2003. – 381с.

2. Ян Ричардсон. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. – Москва: Техносфера, 2005. - 368с.

3. Баранник В.В. Кодирование трансформированных изображений в инфокоммуникационных системах / В.В. Баранник, В.П. Поляков - Х.: ХУПС, 2010. – 212 с.

4. Баранник В.В. Методологические принципы представления апертур во множестве одномерных двухосновных позиционных чисел / В.В. Баранник, Д.С. Кальченко // АСУ и приборы автоматики. – 2011. – Вып. 155. – С. 15 – 22.

ПОСТРОЕНИЕ СИСТЕМ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ СЕТЕЙ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Толюпа С.В.

Введение

Потребности в обеспечении безопасности информационных систем связаны с тем, что существуе множество субьектов и структур, которые заинтерисованы в чужой информации и готовы платить за это высокую цену. В таких условиях всё бiльше распространяется аксиома, что защита информации должна по своим характеристикам быть соответствующей масштабам существующих киберугроз.

По результатам исследования видно, что уровень опасности в киберпространстве постоянно растет. Согласно полученным данным, по миру 91% компаний, которые участвовали в исследовании, сталкивались с киберугрозами за последние 12 месяцев. В Украине эта цифра выглядит еще более угрожающе – 97%. При этом следует отметить, что 48% респондентов заявили о росте числа кибератак.

Многие организации пострадали от киберпреступников: 45% украинских компаний заявили о потере конфиденциальных данных из-за вирусных атак. В мире этот показатель ниже, но не существенно – 32%.

В Украине наблюдается хороший уровень осведомленности обосновных киберугрозах, но при этом уровень инвестиций в области информационной безопасности ниже среднемировых показателей. По результатам исследования, 95% респондентов, которые недовольны финансированием, были бы не против увеличить объем инвестиций на 25% и более. При этом только 35% украинских компаний считают существующий уровень инвестиций в IT-безопасность достаточным. В мире этот показатель равен 55%.

В большинстве компаний более популярен реактивный подход к IT-безопасности. Это в полной мере относится и к инвестициям: компании начинают вкладывать деньги в систему защиты после того, как инцидент уже произошел.

Данные исследования свидетельствуют о том, что IT-специалисты как в мире, так и в Украине достаточно хорошо информированы об угрозах информационной безопасности. Но, к сожалению, за последний год практически во всех компаниях были инциденты, связанные с киберугрозами, и почти половина организаций столкнулась с проблемой потери конфиденциальной информации.

В большинстве украинских компаний считают, что для надежной защиты от киберугроз им предстоит еще многое сделать. Речь идет об увеличении числа специалистов службы IT-безопасности, большем объеме инвестиций и внедрении новейших решений и технологий.

Отход от этого правила приведёт к дополнительным убыткам. Для каждой информационной системы имеется оптимальный уровень защищённости, который не обходимо постоянно поддерживать.

Нет сомнений, что защита критически важных для информационных систем массивов соответствует международным, корпоративным, нормативным и методическим документам. Используются высокостоимостные технические средства и внедряються строго регламентированые организационные мероприятия. Однако нет ответа на самый важный вопрос – насколько предложенные и реализованные решения действительно хороши, какая их планируемая и реальная эффективность.

Такому положению, имеющемуся в информационной системе, но нежелательному в области информационной безопасности есть ряд причин:

– игнорирование системного похода к методологии анализа и синтеза построения систем безопасности (СБ);

– отсутствие механизмов полного и достоверного подтверждения качества таких системы;

– недостатки информативно-методического обеспечения информационной безопасности, прежде всего в области показателей и критериев.

Основная часть

Всем специалистам в области защиты информации известны основные постулаты, которые не утратили актуальность до сих пор:

– абсолютной защиты создать невозможно (самый защищённый компьютер находится в закрытом сейфе, отключённый от сети и неработающий);

– система защиты информации должна быть *комплексной*;

– СБ должна быть адаптированной к смене обстановки;

– СБ должна быть системою, а не простым набором хаотически собранных технических средств и организационных мероприятий, как это часто бывает на практике;

– системный поход к защите информации должен применятся, начиная с подготовки технического задания и оканчиваться оценкой эффективности и качества СБ в процессе её эксплуатации – *жизненного цикла* СБ.

Система безопасности должна иметь целевое назначение. Причём, чем более конкретно сформулирована цель защиты информации, детально выяснены имеющиеся ресурсы, и определён комплекс ограничений, тем в большей степени возможно получение положительного результата. Когда цель обеспечения информационной безопасности проста и принципиально достигается, то достаточно несложных по структуре СБ. Однако при

расширении круга решаемых проблем для обеспечения информационной безопасности, содержание целевого назначения системы на формализованном уровне определяет многомерный, векторный характер. При этом важность свойств отдельных элементов СБ понижается и на первый план выходят общесистемные задачи – определение оптимальной структуры и режимов функционирования системы, организация взаимодействия между элементами системы, учёт воздействий внешней среды и т.д. При целенаправленном объединении элементов в систему последняя требует специфических свойств, которые не имеются ни в одной из её элементов, частей.

При построении СБ ИС на первый план встаёт проблема – как создать такую систему защиты информации, которая бы смогла при минимальных затратах выполнять максимально задачи защиты информации. Эту проблему необходимо решать постепенно, начиная с главного этапа жизненного цикла – проектирования систем защиты информации в комплексе с особенностями объекта защиты.

Обоснованный выбор требуемого уровня защиты информации является системообразующей задачей, поскольку как занижение, так и завышение уровня неизбежно ведет к потерям. При этом, в последнее время роль данного вопроса резко возросла в связи с тем, что, во-первых, теперь в число защищаемых помимо военных, государственных и ведомственных, включены также секреты промышленные, коммерческие и даже личные, а во-вторых, самая информация все больше становится товаром.

На современном этапе развития информационных технологий (ИТ) обеспечения информационной безопасности (ИБ) в масштабах всей информационной системы пока еще невозможно в силу отсутствия на рынке реальных решений, позволяющих строить именно интегрированные системы безопасности. Это можно объяснить недостаточной зрелостью международных стандартов в области защиты информации, хотя движение з этом направлении прослеживается уже достаточно явно. С другой стороны, построение многокомпонентных, а тем более однокомпонентных СБ в большинстве случаев уже не является современным решением проблемы ИБ, особенно для крупных компаний. Поэтому, на наш взгляд, в настоящее время оптимальным решением является построение именно комплексных систем безопасности построенных на основе многоуровневой иерархической модели.

Известно, что в области IT давно и достаточно успешно применяется стековая модель описания сложных ИС, в которой система рассматривается в виде иерархии нескольких функционально-единообразных уровней. Поскольку любая СБ в конечном итоге должна «накладываться» на реальную ИС, для ее описания также целесообразно было бы использовать многоуровневую иерархическую модель.

Одним из главных преимуществ представления СБ в виде иерархии функционально-независимых уровней является существенное упрощение процесса проектирования системы, поскольку теперь проектирование одной многофункциональной и сложной системы можно разложить на несколько законченных этапов проектирования гораздо менее сложных систем для каждого уровня в отдельности и завершающего этапа контроля целостности системы защиты при переходе от уровня к уровню[1].

Проблема обеспечения целостности системы защиты в рамках предложенной модели безопасности принимает достаточно понятную и наглядную форму - это, как уже было сказано, обеспечение полноты реализации функций защиты на каждом уровне модели и обеспечение целостности функций защиты при переходе от уровня к уровню. Очевидно, что максимальная степень комплексности СБ достигается в том случае, когда применяемые технические средства, решения и методы обеспечивают защиту каждого уровня в соответствии с самыми жесткими требованиями и при этом все используемые с СБ технические средства проявляют свою функциональность на каждом уровне модели. Очевидно, что построить настолько «комплексную» СБ в принципе возможно только при неограниченных ресурсах проекта. Поэтому на практике необходимо найти разумный и, главное, обоснованный компромисс между «комплексностью» системы, т.е. ее функциональной наполненностью, и совокупной стоимостью ее построения и эксплуатации [2].

Для успешного использования современных информационных технологий необходимо эффективно управлять не только сетью, но и СБ, при этом на уровне ИС автономно должна работать система, реализующая управление составом событий информационной безопасности, планирование модульного состава СБ и аудит. Поскольку объект управления – СБ является весьма сложной организационно-технической системой, функционирующей в условиях неопределенности, противоречивости и неполноты знаний о состоянии информационной среды, управление такой системой должно быть основано на применении системного анализа, методов теории принятия решений и необходимой интеллектуальной поддержки.

Вместе с тем в области разработки методов и систем защиты информации в настоящее время практически отсутствуют исследования, направленные на обеспечение автоматизированной поддержки управления системой безопасности для решения проблемы обеспечения требуемого уровня защищенности информации в течение всего периода функционирования.

Одним из вариантов решения данной проблемы является использование методов интеллектуальной поддержки принятия решений управления СБ в сегменте информационной системы, что в свою очередь, требует разработки на основе принципов системного анализа и

общенаучных подходов методологических основ защитой информации, соответствующих моделей, методов, алгоритмов и программного обеспечения.

Таким образом, целью является разработка методологических основ управления защитой информации в сегменте информационной системы для решения научно-практической проблемы обеспечения требуемого уровня защищенности информации в течение жизненного цикла системы защиты информации в условиях неопределенности информационных воздействий с использованием интеллектуальной поддержки принятия решений.

Системы управления безопасностью, в которых реализованы интеллектуальные технологии идентификации кибератак и реагирования на события безопасности, являются продуктами зарубежных компаний, коды программ, используемые в них методы формирования командной управляющей информации неизвестны.

Анализ соответствующих зарубежных и отечественных публикаций позволил выявить растущую популярность средств оценки риска, программных комплексов анализа и управления рисками.

Анализируются представленные на рынке программные продукты для автоматизации управления рисками нарушения информационной безопасности. Показано, что недостатками этих систем являются: необходимость наличия экспертов высокой квалификации; трудности, возникающие при адаптации методов к потребностям конкретной организации; невозможность оценить эффективность конкретного комплекса средств защиты, применяемого на объекте защиты; требование наличия на предприятии достоверной статистики по инцидентам информационной безопасности.

Проведенный анализ существующих стандартов в области менеджмента информационной безопасности позволяет сделать вывод о том, что целью стандартов является формирование общих понятий и этапов управления. Вместе с тем, стандарты не формируют конкретных подходов к управлению безопасностью; они определяют функциональные требования в отношении средств защиты и не предлагают методик сравнительного анализа различных комплексов средств защиты в целях выявления наиболее рационального варианта построения системы безопасности.

Для реализации упреждающей стратегии защиты в СБ сегмента информационной системы обосновывается необходимость разработки практически применимых моделей и методов интеллектуальной поддержки планирования рационального модульного состава СБ, оценки и прогнозирования риска нарушения информационной безопасности и управления защитой информации в условиях неопределенности информационных воздействий.

Главным направлением поиска путей защиты информации является неуклонное повышение системности подхода к самой проблеме защиты информации. Понятие системности интерпретировалось прежде всего в том смысле, что защита информации заключается не только в создании соответствующих механизмов, а представляет собой регулярный процесс, осуществляемый на всех этапах жизненного цикла систем обработки данных при комплексном использовании всех имеющихся средств защиты. При этом все средства, методы и мероприятия, используемые для защиты информации, непременно и наиболее рационально объединяются в единый целостный механизм - систему защиты.

Основные трудности реализации систем защиты состоят в том, что они должны удовлетворять двум группам противоречивых требований. С одной стороны, должна быть обеспечена надежная защита находящейся в системе информации, что в более конкретном выражении формулируется в виде двух обобщенных задач: исключение случайной и преднамеренной выдачи информации посторонним лицам и разграничение доступа к устройствам и ресурсам системы всех пользователей, администрации и обслуживающего персонала. С другой стороны, системы защиты не должны создавать заметных неудобств в процессе работы с использованием ресурсов системы. В частности, должны быть гарантированы: полная свобода доступа каждого пользователя и независимость его работы в пределах предоставленных ему прав и полномочий. К сожалению, необходимость системного подхода к вопросам обеспечения безопасности информационных технологий пока еще не находит должного понимания у пользователей современных ИС.

Основываясь на принципах *системного анализа*, который представляет собой теорию и практику улучшающего вмешательства в проблемную ситуацию, предлагается вариант *декомпозиции* проблемы разрешения имеющихся противоречий в области обеспечения безопасности информации.

Следует отметить, что основной проблемой при построении управляющей системы является разработка *модели киберугроз*, что связано со специфичностью взаимодействия объекта управления – СБ с окружающей средой. В связи с этим необходимо рассмотреть *концепцию построения модели киберугроз* безопасности информации, базирующаяся на разрабатываемой классификационной схеме преднамеренных целенаправленных киберугроз информационной среде информационной системы. Показывается целесообразность построения *совокупности моделей*: функциональной, на основе описания последовательности действий злоумышленника (нарушителя) с помощью деревьев угроз, и пространственной графовой, систематизированных в формате интегральной структурной модели каналов несанкционированного доступа, утечки и деструктивных воздействий, позволяющей провести

всесторонний анализ реальных угроз, повысить адекватность модели угроз для конкретного объекта защиты.

При разработке принципов управления СБ в реальном времени, которые обеспечивают удержание требуемого уровня защищенности информации при функционировании СБ в рамках действующего плана в ситуации деструктивных информационных воздействий.

На основе системного подхода предлагается формализованное описание информационной системы, с помощью модели, отображающей семантику предметной области. Предлагается описание множества атак в виде кортежей

$$U^{внш} = \langle S^{внш}, A, З_c, З_x, П, O(C) \rangle \; ; \qquad (1)$$
$$U^{вн}_{l(m)} = \langle S^{k-1}_l, A, З_c, З_x, ПП, {}^{k}(C^{k}_m) \rangle ,$$

где $U^{внш}$ – удаленная кибератака на информационные активы сегмента информационной системы; $U^{вн}_{l(m)}$ – внутренняя кибератака на информационные активы уровня критичности k, обрабатываемые в сегментах C_m, когда нарушитель имеет учетную запись как пользователь с правом доступа к информации, уровень критичности которой не более (k-1), и пытается превысить свои привилегии; $S^{ввн}$ – внешний источник киберугрозы; S^{k-1}_l – внутренний источник киберугрозы; A – коммуникационное оборудование в канале связи; $З_c, З_x$ - сервисы безопасности на пути распространения атаки, сетевые и хостовые; П – протоколы, пакеты; O – объект доступа; C^k_m – сегмент, в котором обрабатывается информация, наивысший уровень критичности которой равен k; l, m– номера сегментов.

Приводится оценка числа путей распространения кибератак, анализируется возможность идентификации кибератаки по индикаторам аномальных событий на пути распространения. С помощью характеристического предиката вводится множество индикаторов

$$И = \{и_j : и_j - \text{индикатор сетевой, хостовый или периметровый}\}. \qquad (2)$$

Поскольку единственным эффективным способом идентифицировать кибератаку является анализ комбинаций аномальных событий, предлагается сопоставлять множеству возможных путей P распространения кибератак множество индикаторов

$$\tau_a \subseteq P \times И = \{(p_i, и_j) : p_i \in P \;\wedge\; и_j \in И\}, \qquad (3)$$

а вероятность того, что подозрительная активность является кибератакой, оценивать числом индикаторов на пути распространения. Сечение соответствия по $\tau_a(p_i)$ определяет набор индикаторов, соответствующий реализации кибератаки на данном пути.

Так как в системе управления реального времени предъявляются требования к времени вычислений информации, то для решения задачи управления в условиях неполноты, противоречивости и неопределенности данных о состоянии информационной среды целесообразно использовать механизм нечеткого логического вывода. Информацией, которая поступает на вход системы нечеткого логического вывода, являются входные переменные – число признаков аномальных событий. Эти переменные соответствуют реальным процессам в сети. Информация, которая формируется на выходе системы нечеткого логического вывода, соответствует выходной переменной, которая является вероятностью того, что совокупность аномальных событий в сети является кибератакой (вероятность кибератаки).

Вводятся лингвистические переменные: «число аномальных сетевых событий на пути распространения атаки», «число аномальных событий на хосте», «число аномальных событий на периметре», «вероятность того, что подозрительная активность в сети является кибератакой». В рассмотрение вводятся нечеткие множества A, B, C, D с функциями принадлежности $\mu_{\tilde{A}}$, $\mu_{\tilde{B}}$, $\mu_{\tilde{C}}$,$\mu_{\tilde{D}}$:

$$\begin{aligned} A &= \{\mu_{\tilde{A}}(x) \mid x:\ \mu_{\tilde{A}}(x) \in [0,1],\ x \in X\}, \\ B &= \{\mu_{\tilde{B}}(x) \mid x:\ \mu_{\tilde{B}}(x) \in [0,1],\ x \in X\}, \\ C &= \{\mu_{\tilde{C}}(x) \mid x:\ \mu_{\tilde{C}}(x) \in [0,1],\ x \in X\}. \\ D &= \{\mu_{\tilde{D}}(p) \mid p:\ \mu_{\tilde{D}}(p) \in [0,1],\ p \in [0,1]\}, \end{aligned} \quad (4)$$

где X – множество чисел индикаторов событий информационной безопасности.

Функции принадлежности лингвистических переменных для входных и выходной переменных, базы продукционных правил формируются на основе экспертных данных и результатов моделирования.

В условиях, когда управляющая система не обладает полной информацией о состоянии информационной среды, обосновывается необходимость разработки модели противодействия киберугрозам, в которой существует возможность выбора того управляющего воздействия, которое в наибольшей степени соответствует состоянию объекта управления. Формулируются принципы разработки модели противодействия угрозам, приводится формализованное описание метода принятия решений по выбору рационального варианта реагирования на события безопасности.

Процесс выбора рационального варианта реагирования на события безопасности описывается кортежем

$$\langle U_i, V_j, C(V_j), P_a, P(z_l), J, U^*(P_a)\rangle, \quad (5)$$

где U_i – вариант реагирования; V_j – исход; C_j– оценка ущерба; z – параметр неопределенности состояния среды; $P(z_l)$ – вероятность состояния среды;

J – целевая функция выбора; $U^*(P_a)$– рациональный вариант реагирования; P_a – вероятность кибератаки.

Модель выбора рационального варианта предлагается формировать в виде графа связи вариантов реагирования на события безопасности и исходов, а также с использованием функции реализации в табличной форме.

Анализ возможных вариантов реагирования $\{U_i\}$ на события безопасности показал, что число управляющих воздействий для каждой ситуации ограничено, $i \in [1,3]$. Поскольку выбор осуществляется в условиях возможного осуществления кибератаки, предлагается связывать систему предпочтений альтернатив с оценкой ущерба: отсутствие ущерба, ущерб одному пользователю, ущерб группе пользователей, ущерб от реализации кибератаки ($\{V_j\}, j \in [1,4]$).

Задается функционал, по которому осуществляется выбор рационального варианта реагирования:

$$J(U_i,z) = \sum_{l=1}^{s} C_j\left(V_j(U_i,z_l)\right)\cdot p(z_l), \qquad (6)$$

где $p(z_l) = \prod_{i=1}^{I} p_{ij}\left(V_j(U_i),P_a\right)$, вероятности p_{ij} наступления каждого j-го исхода при выборе i-го варианта реагирования предлагается рассчитывать как функции вероятности кибератак

$$p_{ij} = p_{ij}\left(V_j(U_i),P_a\right),\ \forall i: \sum_j p_{ij} = 1. \qquad (7)$$

Рациональное управляющее воздействие $U^*(P_a)$ определяется как

$$U^*(P_a) = U\left(\arg\ \min_i \left(J(U_i,z)\right)\right). \qquad (8)$$

На основе адаптированного для выбора рационального варианта реагирования метода принятия решений разрабатываются модели противодействия киберугрозам с учетом возможных путей их распространения: локальное сетевое вторжение, по радиоканалу через беспроводную точку доступа, удаленное вторжение через сети открытого доступа.

Для преодоления трудностей в слабоформализованных ситуациях более высокий качественный уровень управления в условиях реального времени предполагает обеспечение необходимой и достаточной интеллектуальной поддержки. Предлагаемая в работе структура построения системы интеллектуальной поддержки оперативного управления приведена на рис. 1.

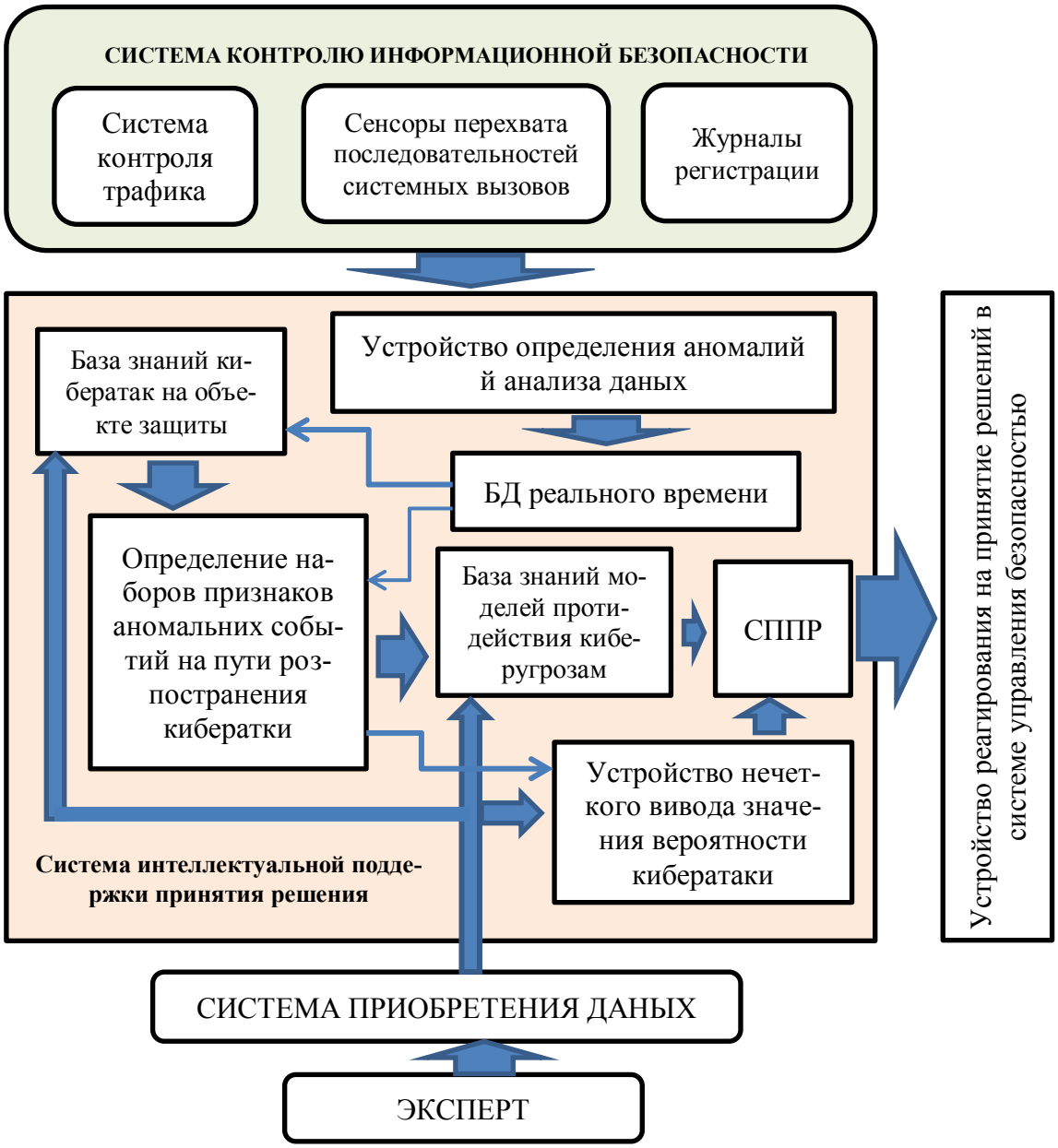


Рис.1. Структура системы интеллектуальной поддержки управления системой безопасности

В системе интеллектуальной поддержки управления СБ предлагается использовать интеллектуальные технологии: механизм нечеткого логического вывода для численной оценки вероятности кибератаки; организованное упорядочение информации о событиях в базе знаний; модели противодействия угрозам; принятие решений по выбору рационального варианта реагирования на события безопасности.

Ввиду необходимости максимальной структуризации разрабатываемой системы и принятия решений, предлагается типовая СБ , которую можно рассмотреть в виде следующих пяти функциональных уровней: физический уровень: физическая охрана помещений, в которых

обрабатывается или хранится конфиденциальная информация; организация контроля доступа сотрудников в данные помещения; ответственное хранение резервных (архивных) копий конфиденциальных информационных ресурсов; обеспечение энерго- и пожаробезопасности всей ИС в целом и др.; технологический уровень: устранение угроз безопасности информации, связанных с использованием некачественных аппаратно-технических средств обработки и хранения информации и систем передачи данных; контроль качества (в т.ч. целостности) используемого программного обеспечения; организация резервных хранилищ данных, кластеров; периодическое архивирование данных; контроль лицензионной политики; организация защиты от вредоносных и разрушающих программ и т.д.; пользовательский уровень: устранение угроз, связанных с некорректными (случайными, ошибочными и т.д.) действиями персонала или умышленными действиями нелояльных сотрудников компании или третьих лиц (разграничение доступа к информационным ресурсам, защита от НСД, аутентификация пользователей, включая удаленных и мобильных сотрудников компании и т.д.); сетевой уровень: система защиты на этом уровне должна устранить угрозы, исходящие от злоумышленников, находящихся как внутри, так и вне пределов защищаемой КИС на уровне базовой сетевой инфраструктуры (сегментация ЛВС по уровням конфиденциальности обрабатываемой информации, защита информации при ее передаче по внешним и внутренним каналам связи, защита от внешних вторжений и т.д.); уровень управления: организация связи с системой управления ИС; управление, координация и контроль осуществляемых организационных и технических мероприятий на всех нижележащих уровнях СБ; контроль полноты реализации функций защиты на каждом из уровней и неразрывности функционирования СБ при переходе от уровня к уровню; окончательный (а далее периодический) контроль стойкости и комплексности всей СБ в целом (например, путем применения специальных технических средств «дружественного взлома») и т.д.

Следует сказать, что в конкретной автоматизированной системе наличие всех пяти уровней СБ в явном виде не всегда обязательно, хотя стойкость системы защиты напрямую зависит от наличия каждого уровня и его функциональной наполненности. Очевидно также, что стоимость и сложность реализации СБ существенным образом растет от уровня к уровню, причем снизу-вверх. Так, например, значительную часть необходимых функций СБ на физическом уровне можно реализовать простыми и привычными организационными мерами, т.е. практически «бесплатно». А, например, на сетевом уровне для защиты сложных систем необходимо применение уже достаточно дорогостоящих технологий, таких как межсетевое экранирование, VPN, средства обнаружения вторжений и т.д. [3]

Следует отметить, что предлагаемый пятиуровневый «стековый» подход помимо упрощения самого процесса проектирования, позволяет четко формализовать три достаточно сложные задачи, которые неизбежно возникают при создании систем защиты ИС:

– обеспечение целостности (комплексности) системы защиты;

– разграничение требований и функций СБ при защите информации, обладающей различной степенью конфиденциальности;

– обеспечение целостности СБ при защите территориально-распределенных ИС.

При этом при решении указанных задач в абсолютном большинстве случаев удается обеспечить оптимальное соотношение функциональность/стоимость СБ для владельца ИС и должным образом это обосновать.

Проблема обеспечения целостности системы защиты в рамках предложенной модели СБ принимает достаточно понятную и наглядную форму – это обеспечение полноты реализации функций защиты на каждом уровне модели СБ и обеспечение целостности функций защиты при переходе от уровня к уровню. Очевидно, что максимальная степень комплексности СБ достигается в том случае, когда применяемые технические средства, решения и методы обеспечивают защиту каждого уровня в соответствии с самыми жесткими требованиями и при этом все используемые с СБ технические средства проявляют свою функциональность на каждом уровне модели. Очевидно, что построить настолько «комплексную» СБ в принципе возможно только при неограниченных ресурсах проекта. Поэтому на практике необходимо найти разумный и, главное, обоснованный компромисс между «комплексностью» системы, т.е. ее функциональной наполненностью, и совокупной стоимостью ее построения, эксплуатации и восстановления [4]. Под стоимостью эксплуатации подразумевается уровень адаптируемости СБ (т.е. сохранение необходимого уровня защиты) к неизбежным изменениям состава и конфигурации ИС.

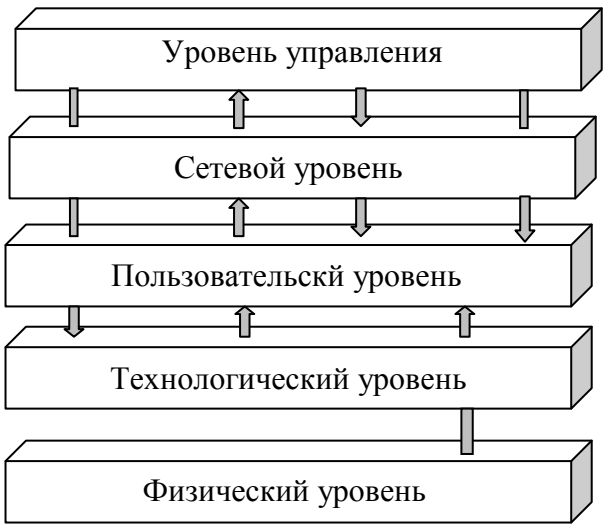


Рис. 2. Пример реализации функциональных связей между уровнями СБ

Практика показывает, что в настоящее время оптимальным подходом для обеспечения необходимой комплексности СБ является построение системы на базе таких продуктов, которые проявляют свои защитные

функции на двух-трех, при этом не обязательно соседних, уровнях СБ (рис. 1). И, очевидно, «проявляемые» на каждом уровне защитные функции должны полностью перекрывать налагаемые на защиту данного уровня требования. Сделать это возможно уже сегодня на основе имеющихся на рынке продуктов по защите информации. Так, например, существующие сегодня развитые продукты по реализации функций межсетевого экранирования позволяют решать не только традиционные для межсетевых экранов задачи по фильтрации трафика на сетевом уровне, но и часть задач пользовательского уровня (аутентификация удаленных пользователей и задание политик безопасности для каждого пользователя при работе в открытой сети), технологического уровня (контроль входящего трафика на предмет наличия разрушающих и вредоносных программ) и уровня управления (целостное управление всем комплексом с единой консоли). Очевидно, что чем больше таких «многоуровневых» средств защиты применяется в СБ, тем легче ее проектирование и полнее и надежнее она выполняет свои функции.

Проблема разграничение системы защиты информации различной степени конфиденциальности заключается в том, что часто на практике в рамках одной ИС приходится «работать» с информацией, требования по защите которой существенно отличаются друг от друга. Так обрабатываемые и хранимые в рамках типовой ИС информационные ресурсы, как правило, разделяются на три группы: открытые информационные ресурсы, конфиденциальные информационные ресурсы, информационные ресурсы ограниченного доступа,

Очевидно, что защита всех разновидностей информационных ресурсов в рамках одной и той же СБ подразумевает, что даже открытые ресурсы будут защищаться по требованиям, предъявляемым к защите секретной информации. Очевидно, это приведет к необоснованно высокой стоимости СБ и большим неудобствам работы для персонала компании. Так же неэффективно будет построение трех различных СБ для каждого из ресурсов, поскольку, во-первых, четко разделить эти ресурсы в рамках одной ИС практически никогда не удается, а во-вторых, это опять приведет к повышению стоимости самой системы.

В рамках предложенной модели указанная проблема может быть решена путем разграничения требований и, соответственно, функциональности для каждого из уровней защиты СБ применительно к каждой группе информационных ресурсов. Сделать это тем более возможно, поскольку в пределах одного уровня требования к защите информации находятся, можно сказать, «в одной системе координат». При этом, если информационные ресурсы в каком либо элементе ИС четко физически не разделены, в рамках СБ необходимо оценить возможность разделения указанных ресурсов на каждом из уровней системы. Если в пределах одного уровня сегментировать информацию не удается, система

требований для данного уровня, очевидно, должна строиться исходя из требований по защите информации максимальной степени конфиденциальности. Если сегментация информации возможна, к уровню может предъявляться двойная (тройная и т.д.) система требований.

В подобных случаях необходимо применение несложных экономических расчетов по оценке эффективности того или другого решения.

Проблема обеспечения целостности защиты территориально-распределенных ИС заключается в том, что, как правило, даже большая корпорация не в состоянии обеспечить одинаковый уровень защиты для ИС Центрального офиса (ЦО) компании и всех ее филиалов (представительств, дочерних компаний и т.д.). На практике чаще всего ЦО защищается в соответствии с самыми жесткими требованиями по ИБ, а для системы защиты филиалов регламентируются только технические параметры взаимодействия с СБ ЦО. В большинстве случаев такой подход является вполне обоснованным, поскольку именно в ИС ЦО сосредоточены основные информационные ресурсы компании. Поэтому проблема, собственно, заключается в том, чтобы обеспечить целостность защиты СБ ЦО при ее информационном взаимодействии с СБ «менее защищенных» филиалов.

В рамках предложенного подхода сохранение целостности защиты корпоративной СБ обеспечивается в том случае, когда при взаимодействии двух систем СБ ЦО дополнительно контролирует те параметры защиты, которые не контролируются в СБ филиала. В случае же «прозрачного» взаимодействия двух систем на одном из уровней СБ (чаще всего пользовательском и сетевом) требования к данному уровню СБ филиала должны соответствовать аналогичным требованиям СБ ЦО.

Для успешного использования современных информационных технологий необходимо эффективно управлять не только сетью, но и системой защиты информации этой сети. Система, реализующая управление составом событий информационной безопасности должна работать автономно, необходимо также разработать модель процесса планирования рационального модульного состава СБ каждого уровня, а также метод формирования рационального комплекса средств защиты на основе общих критериев [5].

В процессе управления в условиях реального времени, *планирование* СБ как функция управления представляет собой процесс последовательного *снятия неопределенности* относительно структуры и *состава* средств защиты в СБ. Процесс планирования $Р_{пл}$ рациональных наборов СрЗ характеризуется с помощью выражения

$$Р_{пл} = Ф \rightarrow S_г, \qquad (9)$$

где Ф – множество функциональных подсистем для контура безопасности;

$S_г$ – выбранный набор средств защиты.

На первом этапе задается множество функциональных подсистем для контура безопасности, результатом планирования является управляющая информация, которая содержит конкретные данные по распределяемым ресурсам, направляемым на достижение целевого состояния СБ.

Процесс принятия решения о выборе рационального варианта набора СрЗ для контура безопасности – это функция преобразования содержания информации о требованиях, предъявляемых к средствам защиты, входящим в набор, о характеристиках средств защиты, в подмножество наилучших вариантов набора $S' \subseteq S$. Множество вариантов набора

$$S = \{S_1,\ldots,S_r,\ldots,S_R\}, \quad (10)$$

где R – число вариантов альтернативных наборов, из которых осуществляется выбор.

Для выбора рационального варианта набора средств защиты используется целевая функция J:

$$S_r = J(S). \quad (11)$$

Совокупность сведений, позволяющих сопоставлять варианты наборов, это характеристики средств защиты функциональных подсистем для рубежа – множество W, включающее в себя два подмножества:

$$W_{зщ_l} \subset W_l \text{ и } W_{и_l} \subset W_l, \quad (12)$$

где $W_{зщ_l}$ – показатель средств защиты «защищенность информации»;

$W_{и_l}$ – показатель средств защиты «издержки» для l-ой функциональной подсистемы.

На основе морфологического подхода модель принятия решений по выбору рационального варианта набора может быть представлена в виде кортежа:

$$ПР: \langle Ц,Ф,П_s,S,W_l,J,S_r(S')\rangle, \quad (13)$$

где Ц – цель принятия решения;

Ф – исходные данные для порождения вариантов набора средств защиты:

$$Ф = \{Ф_1,Ф_2,\ldots,Ф_l,\ldots,Ф_L\};$$

$П_s$ – правило порождения вариантов набора, которое может быть представлено в аналитическом виде как векторное произведение множеств

$$S = Ф_1 \times Ф_2 \times \ldots \times Ф_l \times \ldots \times Ф_L, \quad (14)$$

где $Ф_l$ – множество, состоящее из средств защиты l-ой функциональной подсистемы

$$Ф_l = \{A_{l1},A_{l2},\ldots A_{lm}\ldots,A_{lK_l}\}; \quad (15)$$

S– множество порожденных вариантов набора;

W_l – данные для выбора рациональных вариантов;

J– целевая функция для выбора рационального набора средств защиты (правило выбора);

S_r – рациональный набор средств защиты.

Отмечается, что в условиях автоматизированного управления и при использовании экспертной информации в процессе принятия решения можно говорить (даже в случае формализованного правила выбора) о *рациональном*, а не оптимальном решении.

В соответствии с предлагаемой моделью защиты, основой планирования рационального модульного состава СБ являются функциональные требования к наборам СрЗ для каждого контура безопасности, которые формулируются на основе нормативной документации, в соответствии с уровнем критичности обрабатываемой информации. Альтернативные средства защиты для каждой функциональной подсистемы набора средств защиты выбираются с учетом этих требований. Вариантов наборов, сертифицированных по требуемому классу защищенности, может быть много. Сравнение вариантов наборов средств защиты предлагается производить по количественной мере.

Для решения задачи выбора рациональных вариантов наборов средств защиты для контуров безопасности разрабатывается метод обработки знаний, использующий неформализуемый опыт эксперта в области безопасности, обеспечивающий преобразование сведений о характеристиках средств защиты из базы знаний и вывод решения в аналитической форме – метод формирования рационального комплекса средств защиты для СБ.

1. Разрабатываются варианты набора СрЗ. Множество возможных вариантов решения задачи выбора задается морфологической матрицей.

2. Заполняются вспомогательные матрицы, в которых отмечаются совместимые друг с другом программно-аппаратные средства. Вспомогательная квадратная матрица совместимых решений заполняется следующим образом: для каждой пары средств защиты разных функциональных подсистем определяется, совместимы ли они. Если СрЗ совместимы, то функция совместимости $s(A_{lm}, A_{pr}) = 1$, в противном случае $s(A_{lm}, A_{pr}) = 0$.

3. Генерируется множество решений по выбору вариантов набора СрЗ с усечением этого множества до подмножества вариантов набора из совместимых между собой программно-аппаратных продуктов.

Множество $S = \{S_1, ... S_r, ..., S_R\}$, состоящее из всех возможных вариантов построения набора СрЗ для рубежа, является декартовым произведением множеств альтернатив.

Элемент множества

$$S_r = \{(A_{1i}, A_{2j}, \ldots A_{lm}, \ldots A_{Ln}) : A_{lm} \in \Phi_l, \ \forall l = \overline{1,L}\},$$

где L – число функциональных подсистем для контура безопасности;

A_{lm} – средство защиты для реализации l-ой функциональной подсистемы.

Генерация множества решений по выбору вариантов набора, состоящих из совместимых между собой СрЗ, осуществляется следующим образом.

Происходит итерационный синтез вариантов набора, состоящих из совместимых СрЗ: на первом шаге перебираются последовательно варианты средств защиты для первой подсистемы, после выбора альтернативы A_{1i} осуществляется переход ко второму шагу. На втором шаге выполняется последовательный перебор вариантов средств защиты второй подсистемы, но выбор осуществляется только для таких альтернатив A_{2j}, для которых функция совместимости $s(A_{1i}, A_{2j}) = 1$ и т.д. При выборе альтернатив из l-ой подсистемы выбор осуществляется только из таких альтернатив A_{lm}, для которых функции совместимости равны единице: $s(A_{l-1,m}, A_{lm}) = 1$, …; $s(A_{2j}, A_{lm}) = 1$, $s(A_{1i}, A_{lm}) = 1$. Таким, образом, выбор СрЗ из каждой строки морфологической матрицы для формирования варианта набора осуществляется только из совместимых между собой программно-аппаратных продуктов.

4. Дальнейшее усечение множества S выполняется методом полного перебора по заданной целевой функции. В качестве целевой функции для выбора варианта набора $S_r = \{A_{1i}, A_{2j}, \ldots A_{lm}, \ldots A_{Ln}\}$, применяется функция

$$J = \max_r \frac{W_{K_{зщ}^1}^{A_{1i}} + \ldots + W_{K_{зщ}^1}^{A_{lm}} + \ldots + W_{K_{зщ}^1}^{A_{Ln}}}{W_{K_{и}^1}^{A_{1i}} + \ldots + W_{K_{и}^1}^{A_{lm}} + \ldots + W_{K_{и}^1}^{A_{Ln}}}, \qquad (16)$$

где $W_{K_{зщ}^1}^{A_{lm}}$ – значения показателя «защищенность»; $W_{K_{и}^1}^{A_{1m}}$ – значения показателя «издержки» средства защиты A_{lm}.

Для оценки средств защиты разных функциональных подсистем наборов разрабатываются иерархические структуры обобщенных критериев качества средств защиты: показатель «защищенность» и показатель «издержки».

Критерии качества средств защиты по иерархии «защищенность» делятся на две группы: показатели обеспечения эффективности оперативных методов защиты и показатели функциональной пригодности. Критерии качества по иерархии «издержки» делятся также на две группы: в первую включена стоимость соответствующего средства защиты, число пользователей по одной лицензии и другие возможные экономические издержки; ко второй группе издержек относятся функциональные издержки, такие, например, как падение производительности информационной системы при использовании данного средства защиты.

Оценка средств защиты и критериев осуществляется попарным сравнением по методу Т. Саати, результаты приводятся в числовом виде. С

использованием иерархических структур критериев качества СрЗ вычисляются нормированные значения собственных векторов средств защиты по всем критериям до показателей «защищенность» $K^1_{зщ}$ и «издержки» $K^1_{и}$ на основании обработки всех матриц попарных сравнений с учетом связей критериев.

После выбора рациональных наборов средств защиты для рубежей защиты получен рациональный модульный состав целостного комплекса средств защиты объекта, удовлетворяющий требованию

$$J \rightarrow \max. \qquad (17)$$

5. Оценивается, удовлетворяет ли сформированный комплекс средств защиты требованию,

$$C_{\Sigma} \le C_{доп}, \qquad (18)$$

где C_{Σ} – суммарные затраты на реализацию комплекса СрЗ;

$C_{доп}$ – выделенные на реализацию комплекса денежные ресурсы.

При этом C_{Σ} вычисляется с помощью следующего выражения:

$$C_{\Sigma} = \sum_{s} \left(\sum_{i_s} C^{Б}_{i_s} + \sum_{j_s} C^{С}_{j_s} + \sum_{k_s} C^{В}_{k_s} + C_{сегм_s} \right) + C_{пр}, \qquad (19)$$

где S – число сетевых сегментов; $C^{Б}_{i_s}$ – стоимость набора средств защиты хоста, на котором обрабатывается информация базового уровня критичности; $C^{С}_{j_s}$ – стоимость набора средств защиты хоста, на котором обрабатывается информация среднего уровня критичности; $C^{В}_{k_s}$ – стоимость набора средств защиты хоста, на котором обрабатывается информация высокого уровня критичности; $C_{сегм_s}$ – стоимость набора средств защиты на границе s-го сетевого сегмента; $C_{пр}$ – стоимость наборов средств защиты периметра.

Выбор комплекса средств защиты для СБ достигается итерационно путем приближения к рациональному составу, удовлетворяющему требованиям к допустимым затратам на его реализацию.

Структура системы интеллектуальной поддержки организационно-технического управления ЗИ, в которой реализуется метод формирования рационального комплекса средств защиты, представлена на рис. 3.

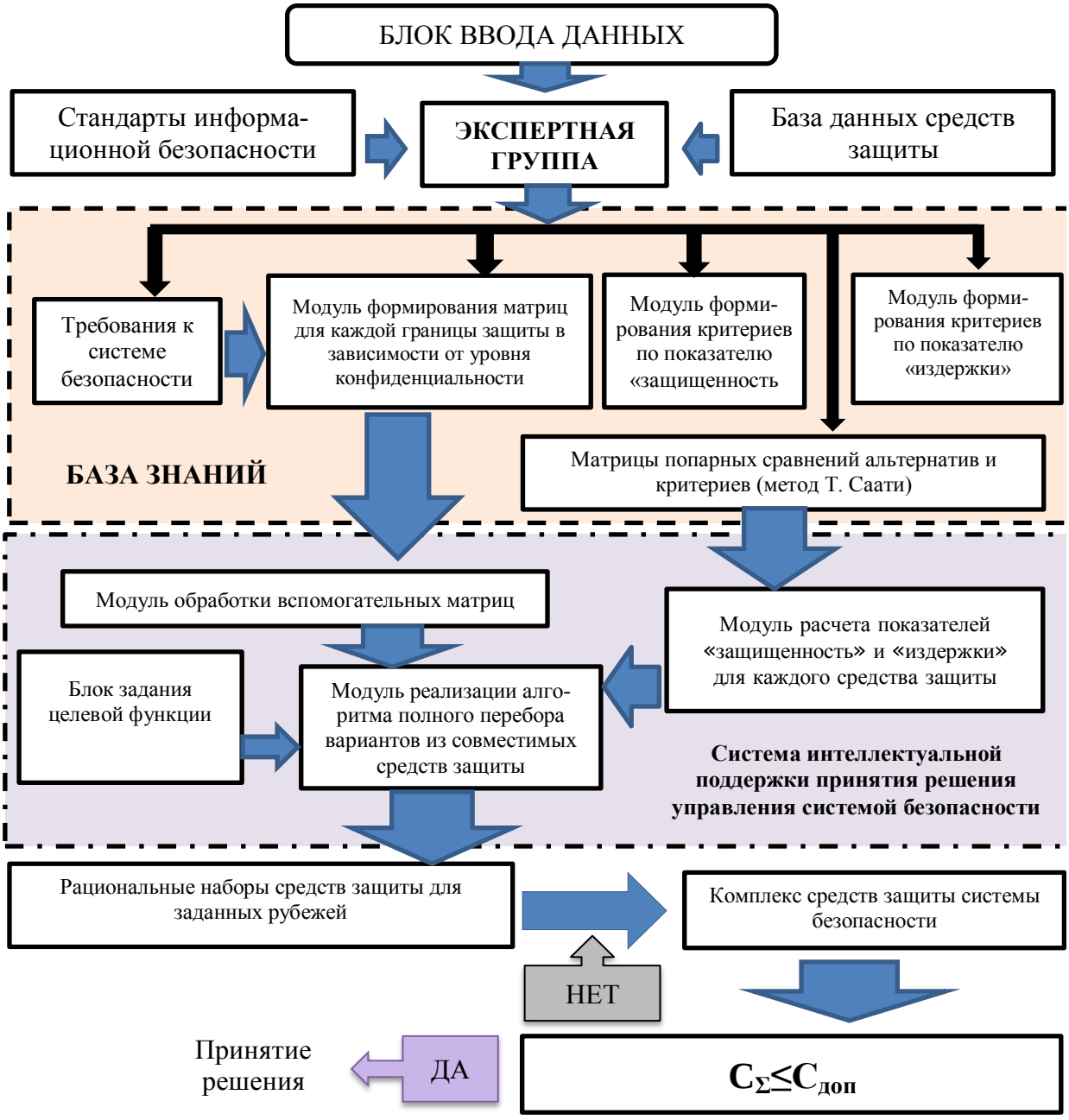


Рис. 3. Структура системы интеллектуальной поддержки принятия решений управления информационной безопасностью

Выводы

В рамках предложенной модели проблема безопасности может быть решена путем разграничения требований и, соответственно, функциональности для каждого из уровней защиты СБ применительно к каждой группе информационных ресурсов и средств защиты. Сделать это тем более возможно, поскольку в пределах одного уровня требования к защите информации находятся, можно сказать, «в одной системе координат». При этом, если информационные ресурсы и средства защиты в

каком либо элементе ИС четко физически не разделены, в рамках СБ необходимо оценить возможность разделения указанных ресурсов на каждом из уровней системы. Если в пределах одного уровня сегментировать информацию не удается, система требований для данного уровня, очевидно, должна строиться исходя из требований по защите информации максимальной степени конфиденциальности.

В системе интеллектуальной поддержки рациональные решения предлагается выбирать на основе использования экспертных знаний; в ней реализуется механизм приобретения знаний в процессе заполнения полей знаний экспертом при взаимодействии его с автоматизированной системой, выполняется совокупность процедур над проблемной областью с использованием многокритериального сравнительного анализа для выявления в заданном экспертом множестве подмножества наилучших по критериям предпочтения вариантов наборов, из которых формируется рациональный комплекс средств защиты.

В заключение необходимо подчеркнуть, что предлагаемый подход к проектированию СБ на базе иерархи пятиуровневой модели носит достаточно общий (методический) характер и оставляет большое поле для творчества компаниям проектировщикам. Предложенный системный подход позволяет существенно сократить сроки разработки СБ и при этом предложить заказчику действительно оптимальное и обоснованное решение.

Литература

1. В.И. Андреев, Ю.Ю. Гончаренко, М.М. Дивизинюк, И.Н. Павлов, В.А. Хорошко. Проектирование систем технической защиты информации. – Севастополь.: Изд. Центр СНУЯЭиП, 2011. – 235 с.

2. Толюпа С.В., Пархоменко І.І. Побудовакомплексних систем захистускладнихінформаційних систем на основі структурного підходу. Науково-технічний журнал “Сучасний захист інформації”.– 2015. - №4. – С. 96-104.

3. Павлов И.Н. Проектирование систем защиты информации. Формальный подход / И.Н. Павлов. – “Правове, нормативне та метрологічнезабезпечення систем захистуінформації в Україні”. – Київ, 2005. – Вып. 11. – С. 54 − 59.

4. Толюпа С.В. Проектирование систем поддержки принятия решений в процессе восстановления и обеспечения комплексной защити информационных системах. // Науково-технічний журнал “Сучасний захист інформації”. – 2012. - №4. – С. 69 - 74.

5. Толюпа С. В., Наконечный В. С., Якименко Ю. М. Оценка защищённости информации в автоматизированных информационных системах с помощью общих критериев. Наукові записки Українського науково-дослідного інституту зв’язку. – 2015. – №6(40). – С. 27 - 31.

6. Толюпа С. В., Наконечный В. С., Якименко Ю. М. Обеспечение безопасности информации в автоматизированных информационных системах. Наукові записки Українського науково-дослідного інституту зв’язку. – 2015. – №5(39). – С. 33 - 37.

МЕТОД ЗАЩИТЫ ИНФОРМАЦИИ НА ОСНОВЕ КОСВЕННОГО СТЕГАНОГРАФИЧЕСКОГО ВСТРАИВАНИЯ В АДАПТИВНОМ ПОЗИЦИОННОМ ПРОСТРАНСТВЕ

Фролов О.В., Баранник В.В., Баранник Н.В.

Введение

Глобализация в сфере информационных технологий, интенсивное развитие технологий информатизации и управления в различных сферах деятельности общества сопровождается повышением значимости информационных ресурсов. С одной стороны, использование информационных систем облегчает сбор, обработку и обмен информацией, с другой стороны наблюдается постоянное увеличение объемов информационных данных, а соответственно и развитие подходов для их защиты.

Наиболее остро стоит вопрос обеспечения защищенности государственных информационных ресурсов. Такие информационные ресурсы является собственностью государства и требует защиты в соответствии с законодательством. Значимость государственного информационного ресурса обусловлена наличием потенциального злоумышленника, действия которого направлены на нарушение конфиденциальности, целостности и доступности государственной информации. Потеря или подмена даже части такой информации могут привести к негативным экономическим и политическим последствие, привести к значительным потерям материальных и человеческих ресурсов.

Наиболее распространенными на практике подходами обеспечения безопасности информации является криптографические методы. Но в тоже время существуют некоторые ограничения при использовании криптографии для обеспечения информационной безопасности, а именно:

- повышенное внимание потенциального злоумышленника к криптографически зашифрованному сообщению;
- развитие современных математических моделей для осуществления криптографического анализа;
- использование стандартизированных и беспроводных каналов передачи государственного информационного ресурса.

Для устранение существующих ограничений при использовании криптографических методов необходимо использовать скрытую передачу данных. На сегодняшний день существует большое количество стеганографических методов, которые позволяют скрытно передавать информацию в контейнере, который не привлекает внимание [1]. Наибольший интерес представляют цифровые стеганографические методы встраивания информации в цифровые изображения. Это обусловлено рядом причин, а именно: широким распространением цифровых

изображений, большим объемом мультимедийного файла, наличием в изображениях областей с визуальной избыточностью [2].

Существующие стеганографические подходы условно можно разделить на методы косвенного и непосредственного встраивания. Непосредственное встраивание реализуется путем замены элемента изображения-контейнера на элемент скрываемого сообщения. Наоборот, в косвенных методах встраивание осуществляется на основе создания зависимости между элементами путем их модификации [3].

В сравнении с методами непосредственного встраивания, косвенные подходы имеют некоторые преимущества при обеспечении безопасности государственного информационного ресурса, а именно:

- повышенной стойкостью встроенных данных к атакам и стеганографическому анализу;
- отсутствием привязки к конкретному формату представления контейнера;
- наличием возможности использования известного на приемной стороне изображения-контейнера.

Анализ существующих методов косвенного стеганографического встраивания выявил следующие недостатки при обеспечении безопасности государственного информационного ресурса [2]:

1. Низкое значение стойкости стеганограммы к визуальным атакам злоумышленника.

2 Низкая устойчивость встроенных данных к активным атакам злоумышленника.

3. Неудовлетворительное значение объема встраиваемых данных.

4. Необходимость наличия на приемной стороне прототипа исходного изображения-контейнера для однозначного изъятия встроенной информации.

Существующие недостатки методов непосредственного встраивания обусловлены тем, что существующие алгоритмы используют для косвенного встраивания психовизуальную избыточность изображения.

Значит, целью ***научно-прикладного исследования*** является проектирование стеганографической системы на основе подхода, который учитывает не только психовизуальные закономерности в изображении, а и структурные зависимости между элементами представления изображения-контейнера.

1. Разработка подхода для проектирования метода косвенного стеганографического встраивания

При проектировании стеганографической системы для обеспечения конфиденциальности служебной информации государственных структур и ведомственных организаций предлагается использовать функционально

преобразование, которое учитывает структурные зависимости между элементами изображения-контейнера.

В качестве функционального преобразования, предлагается использовать функционал для адаптивного позиционного числа, а в качестве элемента изображения-контейнера фрагмент изображения F размерностью m строк и n столбцов.

Предложенное функциональное преобразование позволяет выявить структурные закономерности в изображении. Такие закономерности обусловлены ограничением на динамический диапазон. Величина ψ динамического диапазона представления фрагмента F изображения-контейнера определяется на основе следующего выражения:

$$\psi = \max_{1 \le i \le m} \{c_{i,j}\} + 1, \quad j=\overline{1,n}. \tag{1}$$

Здесь $c_{i,j}$ – j-й элемент в i-ой строке массива F.

В процессе реализации функционального преобразование на основе для адаптивного позиционного числа (АПЧ) фрагмент F исходного изображения рассматривается как множество адаптивных позиционных чисел {C(j)}:

$$C(j) = \{c_{1,j}; ...; c_{i,j}; ...; c_{m,j}\}.$$

Значения кода K (j) будет определяться, как сумма произведений элементов позиционного числа C(j) на их весовые коэффициенты $V_{i,j}$ по формуле:

$$K(j) = \sum_{i=1}^{m} c_{i,j} \; V_i. \tag{2}$$

Здесь $c_{i,j}$ – (i; j)-й элемент адаптивного позиционного числа C(j); V_i - весовой коэффициент элемента $a_{i,j}$ адаптивного позиционного числа C(j) фрагмента F.

Весовой коэффициент $V_{i,j}$ элемента $c_{i,j}$ зависит от его позиции в числе C(j) и вычисляется как сумма оснований всех младших элементов АПЧ. Но учитывая, что основания ψ для всех элементов $c_{i,j}$ позиционного числа C(j) вычисляется адаптивно и принимает одинаковое значение, то весовой коэффициент $V_{i,j}$ будет вычисляется по следующей формуле:

$$V'_{i,j} = V_i = \psi^{m-i}. \tag{3}$$

Второй этап предусматривает формирование кодограммы S(F), которая включает служебную составляющую S(Ψ) и информационную составляющую S(j). Данный этап реализуется при помощи оператора выделения разрядов $\varphi_c(\bullet)$ по формуле:

$$S(F) = \varphi_c(S(j), \Psi),$$

где Ψ - базис, содержащий информацию об основаниях для адаптивных позиционных чисел фрагмента F;

$S(j)$ - кодограмма кодового представления адаптивного позиционного числа $C(j)$.

Кодограмма $S(j)$ имеет следующий вид:

$$S(j) = \{s_1, ..., s_\xi, ..., s_{q(S(j))}\},$$

где $q(S(j))$ – длина двоичной кодограммы $S(j)$;

s_ξ - ξ-й двоичный разряд кодограммы $S(j)$.

Процесс реконструкции элемента $c_{i,j}$ для адаптивного позиционного числа $C(j)$ на основе кода $K(j)$ выполняется по формуле

$$c'_{i,j} = [K(j)/V_i] - [(K(j)/(\psi \ V_i)] \psi \ ,$$

или

$$c'_{i,j} = [\sum_{i=1}^{m} c_{i,j} V_i \ / V_i] - [\sum_{i=1}^{m} c_{i,j} V_i / (\psi \ V_i)] \psi \ .$$

Такое преобразование осуществляется без внесения искажений.

В случае адаптивного позиционного кодирования, значение реконструированного элемента $c_{i,j}$ числа $C(j)$ фрагмента F не меняется в случае кодирования и декодирования с различными основаниями ψ и ψ', т.е.

$$c'_{i,j} = c_{i,j} = [K(j)/V_i] - [(K(j)/(\psi \ V_i)] \psi \ =$$
$$= [K'(j)/V'_i] - [(K'(j)/(\psi' \ V'_i)] \psi' \ = c''_{i,j} \ , \qquad (4)$$

где $c'_{i,j}$ - элемент числа $C(j)$, реконструированный на основе системы оснований Ψ;

$c''_{i,j}$ - элемент числа $C(j)$, реконструированный на основе системы оснований Ψ';

$K(j)$ - кодовое представление числа $C(j)$, сформированное в базисе оснований Ψ;

$K'(j)$ - кодовое представление числа $C(j)$, сформированное в базисе оснований Ψ';

ψ' - значение модифицированного основания элемента $c'_{i,j}$.

Докажем, что значение реконструированного элемента $c_{i,j}$ числа $C(j)$ фрагмента F не меняется в случае кодирования и декодирования с различными основаниями ψ и ψ', при условии что:

$$\psi' \geq \psi .$$

Для этого распишем левую часть выражения (4) с учетом соотношения для кода $K(j)$:

$$K(j) = \sum_{i=1}^{m} c_{i,j} V_i .$$

Тогда получим:

$$c'_{i,j} = [\sum_{i=1}^{m} c_{i,j} V_i) / V_i] - [\sum_{i=1}^{m} c_{i,j} V_i) / (\psi \ V_i)] \cdot \psi \ . \qquad (5)$$

Рассмотрим первое слагаемое выражения (5):

$$[\frac{\sum_{i=1}^{m} c_{i,j} V_i}{V_i}] \quad = \quad [\frac{\sum_{\xi=1}^{i} c_{\xi,j} V_\xi + \sum_{\xi=i+1}^{m} c_{\xi,j} V_\xi}{V_i}].$$

Перепишем данное выражение с учетом следующего соотношения:

$$V_i > \sum_{\xi=i+1}^{m} c_{\xi,j} V_\xi . \qquad (6)$$

В этом случае получим преобразованное значение первого слагаемого выражения (5):

$$[\frac{\sum_{i=1}^{m} c_{i,j} V_i}{V_i}] \quad = \quad [\frac{\sum_{\xi=1}^{i} c_{\xi,j} V_\xi}{V_i}]. \qquad (7)$$

Теперь преобразуем полученное выражение с учетом формулы для весового коэффициента V_i:

$$V_i = \psi^{\,m-i} . \qquad (8)$$

Тогда выражение (8) примет следующий вид:

$$[\frac{\sum_{\xi=1}^{i} c_{\xi,j} V_\xi}{V_i}] \ = [\frac{\sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi^{\,m-\xi}}{\psi^{\,m-i}}] = \sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi^{\,i-\xi} .$$

На следующем этапе рассмотрим второе слагаемое выражения (5):

$$[\frac{\sum_{i=1}^{m} c_{i,j} \ V'_i}{\psi' \ V'_i}] \cdot \psi' .$$

Перепишем данное выражение с учетом следующего неравенства:

$$\psi \ V_{i,j} > \sum_{\xi=i}^{m} c_{i,j} V_i . \qquad (9)$$

В этом случае второе слагаемое выражения (5) примет вид:

$$[\frac{\sum_{i=1}^{m} c_{i,j} V_i}{\psi \quad V_i}] = [\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} V_\xi + \sum_{\xi=i}^{m} c_{\xi,j} V_\xi}{\psi \quad V_i}] = [\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} V_\xi}{\psi \quad V_i}]. \qquad (10)$$

Преобразуем полученное выражение (10) с учетом формулы (9):

$$[\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} V_\xi}{\psi \quad V_i}] = [\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi^{m-\xi}}{\psi \quad \psi_i^{m-i}}] = \sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi^{i-\xi-1}.$$

Перепишем выражение (5) с учетом выполненных преобразований. В этом случае выражение (5) примет следующий вид:

$$c'_{i,j} = \sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi^{i-\xi} - [\sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi^{i-\xi-1}] \cdot \psi =$$

$$= \sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi^{i-\xi} - \sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi^{i-\xi} = c_{i,j}. \qquad (11)$$

Теперь перепишем правую часть выражения (5) с учетом формулы для кода $K'(j)$, сформированного с учетом модифицированного основания ψ':

$$K'(j) = \sum_{i=1}^{m} c_{i,j} V'_i.$$

Тогда получим:

$$c''_{i,j} = [\sum_{i=1}^{m} c_{i,j} V'_i) / V'_i] - [\sum_{i=1}^{m} c_{i,j} V'_i) / (\psi' V'_i)] \cdot \psi'. \qquad (12)$$

Рассмотрим первое слагаемое выражения (12):

$$[\frac{\sum_{i=1}^{m} c_{i,j} V'_i}{V'_i}] = [\frac{\sum_{\xi=1}^{i} c_{\xi,j} V'_\xi + \sum_{\xi=i+1}^{m} c_{\xi,j} V'_\xi}{V'_i}].$$

Перепишем данное выражение с учетом соотношения (6). В этом случае получим первое слагаемое выражения (12) будет иметь вид:

$$[\frac{\sum_{i=1}^{m} c_{i,j} V'_i}{V'_i}] = [\frac{\sum_{\xi=1}^{i} c_{\xi,j} V'_\xi}{V'_i}]. \qquad (13)$$

Теперь преобразуем полученное выражение с учетом формулы (8) для весового коэффициента V'_i:

$$\left[\frac{\sum_{\xi=1}^{i} c_{\xi,j} V'_{\xi}}{V'_i}\right] = \left[\frac{\sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi'^{m-\xi}}{\psi'^{m-i}}\right] = \sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi'^{i-\xi}.$$

Теперь рассмотрим второе слагаемое выражения (12):

$$\left[\frac{\sum_{i=1}^{m} c_{i,j} V'_i}{\psi' V'_i}\right] \cdot \psi' .$$

Перепишем данное выражение с учетом неравенства (9). Тогда получим:

$$\left[\frac{\sum_{i=1}^{m} c_{i,j} V'_i}{\psi' V'_i}\right] = \left[\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} V'_{\xi} + \sum_{\xi=i}^{m} c_{\xi,j} V'_{\xi}}{\psi' V'_i}\right] = \left[\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} V'_{\xi}}{\psi' V'_i}\right]. \quad (14)$$

Преобразуем полученное выражение (14) с учетом формулы (8):

$$\left[\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} V'_{\xi}}{\psi' V'_i}\right] = \left[\frac{\sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi'^{m-\xi}}{\psi' \cdot \psi'^{m-i}}\right] = \sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi'^{i-\xi-1}.$$

Перепишем выражение (12) с учетом выполненных преобразований. В этом случае выражение (12) примет следующий вид:

$$c''_{i,j} = \sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi'^{i-\xi} - \left[\sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi'^{i-\xi-1}\right] \cdot \psi' =$$

$$= \sum_{\xi=1}^{i} c_{\xi,j} \cdot \psi'^{i-\xi} - \sum_{\xi=1}^{i-1} c_{\xi,j} \cdot \psi'^{i-\xi} = c_{i,j}.$$

Теперь перепишем выражение (4) с учетом преобразованной правой и левой части:

$$c'_{i,j} = c''_{i,j} = c_{i,j}.$$

Откуда можно заключить, что значения реконструированных элементов $c_{i,j}$ числа C(j) фрагмента F не меняется в случае кодирования и декодирования с различными основаниями ψ и ψ'. Графически это можно представить, как показано на рис. 1.

Значит, значения реконструируемого элемента $c_{i,j}$ в числе C(j) не меняется в случае кодирования и декодирования с различными основаниями ψ и ψ'.

Предлагается использовать свойство однозначности декодирования адаптивных позиционных чисел при создании метода косвенного стеганографического встраивания служебной информации.

При проектировании метода стеганографического встраивания необходимо разработать подход модификации оснований адаптивных позиционных чисел для косвенного встраивания служебной информации.

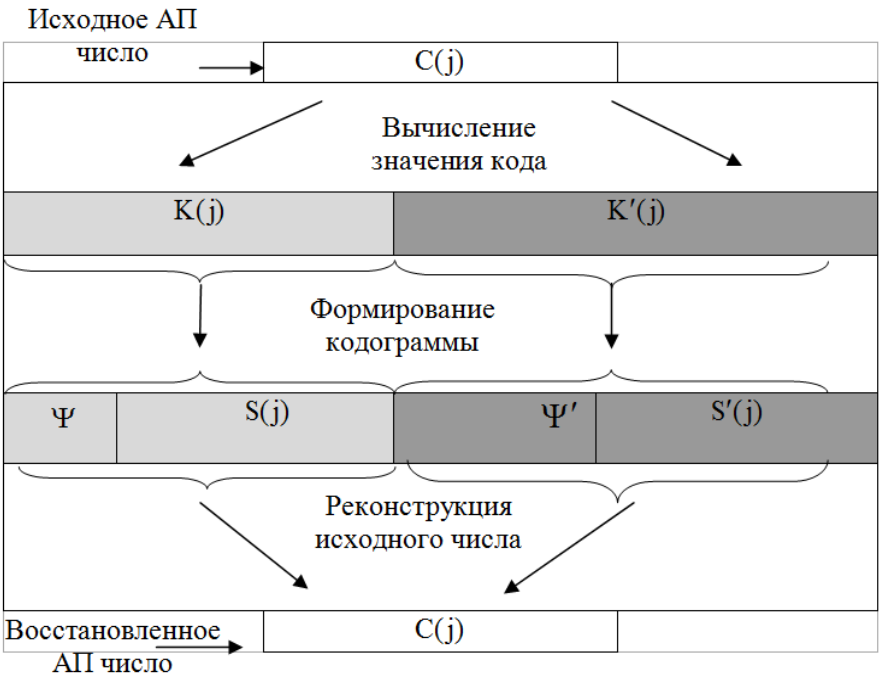


Рис. 1. Графическая интерпретация свойства адаптивного позиционного кодирования

2. Создание метода адаптивного косвенного стеганографического встраивания служебной информации на основе использования функционала для адаптивного позиционного числа

Для обеспечения конфиденциальности служебной информации косвенное встраивание элемента b_ξ скрываемого сообщения $B = \{b_1;...;b_\xi;...;b_v\}$ предлагается проводить в блок изображения-контейнера путем модификации основания ψ_i базиса Ψ на основе следующего правила:

$$\psi' = \psi + k, \text{ где } k = b_\xi.$$

Здесь ψ' - основание, модифицированное в результате косвенного стеганографического встраивания; k - коэффициент модификации.

На следующем этапе вычисляется значение кода $K'(j)$ для числа $C(j)$ с учетом модифицированного основания ψ':

$$K'(j) = \sum_{i=1}^{m} c_{i,j}\ \psi'.$$

Третий этап предусматривает формирование кодограммы $S'(F)$, которая включает служебную составляющую $S(\Psi')$ и информационную составляющую $S'(j)$:

$$S'(F) = \varphi_c(S'(j), \Psi'),$$

где φ_c - оператор выделения разрядов.

Встраивание элементов скрываемого сообщения предлагается осуществлять в двоичном виде $b_\xi \in [0;1]$. В этом случае встраивание будет вносить минимальные искажения. Тогда, коэффициент k модификации предлагается выбирать на основе следующего правила:

$$k = \begin{cases} 0, & b_\xi \to = 0; \\ 1, & b_\xi \to = 1. \end{cases}$$

В этом случае косвенное встраивание бита $b_\xi \in [0;1]$ будет выполняться по формуле:

$$\psi' = \begin{cases} \max\limits_{1 \le j \le n} \{c_{i,j}\} + 1, & b_\xi \to = 0; \\ \max\limits_{1 \le j \le n} \{c_{i,j}\} + 1 + 1, & b_\xi \to = 1. \end{cases} \quad (15)$$

Перепишем выражение (15) с учетом формулы (1). Тогда получим:

$$\psi' = \begin{cases} \psi, & b_\xi \to = 0; \\ \psi + 1, & b_\xi \to = 1. \end{cases} \quad (16)$$

или

$$\psi' = \psi + b_\xi. \quad (17)$$

Таким образом, предложенный подход позволяет осуществить косвенное стеганографическое встраивание сообщения $B = \{b_1; ...; b_\xi; ...; b_\nu\}$, $b_\xi \in [0;1], \xi = \overline{1, \nu}$ в блоки исходного изображения-контейнера.

Теперь рассмотрим этапы функционирования стеганографической системы, построенной на основе разработанного метода косвенного встраивания. Данная система позволяет встроить бит скрываемого сообщения путем модификации основания адаптивного позиционного числа. Обратное стеганографическое преобразование осуществляется по биполярному принципу для авторизированного и неавторизированного пользователя.

Стеганографическая система включает в себя следующие базовые составляющие:

I. Косвенное стеганографическое встраивание.

Косвенное стеганографическое встраивание схематично представлено на рис. 2 и включает следующие действия:

1. Выявление основания в выбранном блоке $F_{\tau,\gamma}$. Данный этап реализуется на основе следующего выражения:

$$\psi = \max_{1 \le i \le m} \{c_{i,j}\} + 1, \quad j = \overline{1, n}.$$

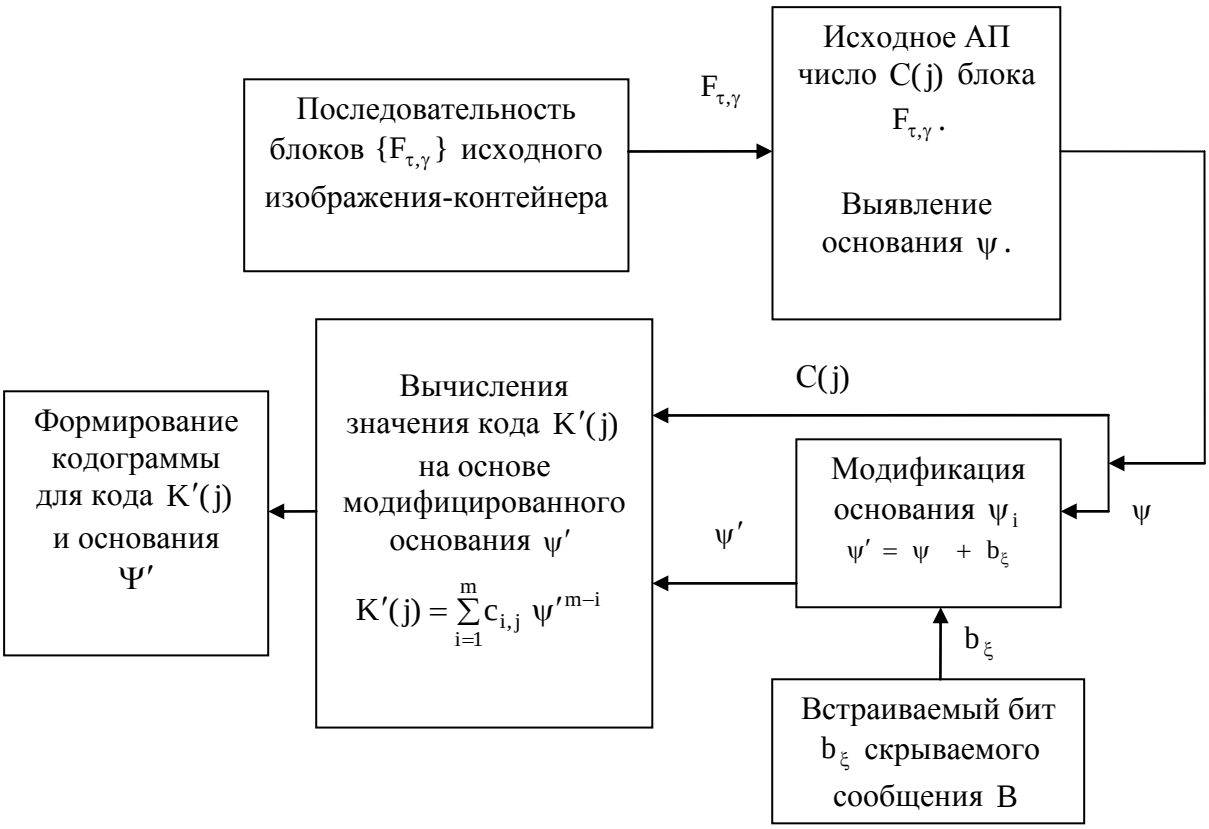


Рис. 2. Схема косвенного стеганографического встраивания

2. Косвенное встраивание бита b_ξ скрываемого сообщения $B = \{b_1;...;b_\xi;...;b_\nu\}, \xi = \overline{1,\nu}$ путем модификации основания ψ по следующему правилу:

$$\psi' = \begin{cases} \psi\ , & b_\xi \to = 0; \\ \psi\ + 1, & b_\xi \to = 1. \end{cases}$$

3. Формирование кода $K'(j)$ для адаптивного позиционного числа $C(j)$ блока $F_{\tau,\gamma}$. Вычисление кода $K'(j)$ выполняется с учетом модифицированного основания ψ' по формуле:

$$K'(j) = \sum_{i=1}^{m} c_{i,j}\ \psi'^{m-1}.$$

4. Формирование кодограммы, которая содержит служебную $S(\Psi')$ (основание ψ') и информационную $S(j)$ составляющую.

II. Теперь рассмотрим процесс извлечения встроенных данных.

Данный этап осуществляется при авторизированном доступе. При этом авторизированному пользователю известно ключевая информация, которая представляет собой ключевое правило встраивания. Обратное стеганографическое преобразование подразумевает проведение оценки области встраивания на основе ключевого правила.

Схема косвенного стеганографического изъятия представлена на рис. 3 и включает следующие этапы:

1. Извлечение из кодограммы кода $K'(j)$ при помощи основания ψ'.

2. Восстановление элементов исходного числа:

$$c'_{i,j} = [K'(j)/V'_i] - [K'(j)/(\psi' V'_i)]\,\psi'.$$

3. Выявление исходного основания ψ по формуле:

$$\psi'' = \max_{1 \le i \le m} \{c'_{i,j}\} + 1, \quad j = \overline{1,n}. \qquad (18)$$

где ψ''_i - i-й основание восстановленного базиса Ψ''.

4. Косвенное изъятие встроенного бита b'_ξ. Данный этап реализуется на основе сравнения модифицированного ψ' и восстановленного ψ'' основания на основе следующего выражения:

$$b'_\xi = \begin{cases} 0, & \to \psi' - \psi'' = 0; \\ 1, & \to \psi' - \psi' = 1. \end{cases}$$

или

$$b'_\xi = \psi' - \psi''. \qquad (19)$$

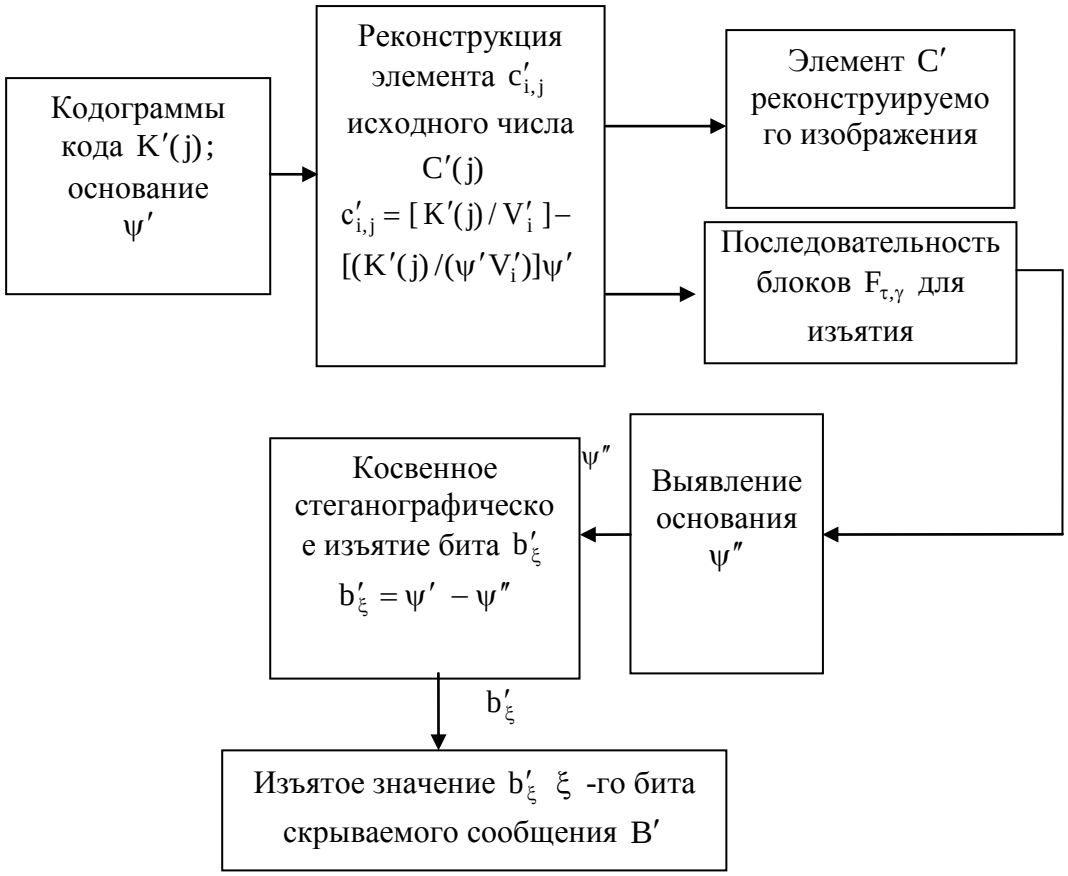


Рис 3. Схема косвенного стеганографического изъятия.

Докажем, что разработанная стеганографическая система позволяет осуществить безошибочное изъятие встроенного бита b'_ξ, т.е.

$$b'_\xi = b_\xi. \quad (20)$$

В этом случае, учитывая выражения (17) и (19) также должно выполняться следующее условие:

$$\psi' - \psi'' = \psi' - \psi \ . \quad (21)$$

Тогда при отсутствии активных и пассивных воздействий на кодовое представление числа, значения исходных и принятых модифицированных оснований равны. Отсюда, безошибочное изъятие будет осуществляться при выполнении следующего условия:

$$\psi'' = \psi \ .$$

Учитывая, что значения оснований ψ' и ψ определяются на основе выражений соответственно (1) и (18), то для доказательства условия (21) необходимо выполнение следующего равенства:

$$c'_{i,j} = c_{i,j}.$$

Доказательство выполнения данного равенства приведено в выражениях (8-14). Отсюда следует, что стеганографическая система косвенного встраивания позволяет изъять 100% встроенной информации.

Теперь рассмотрим обратное стеганографическое преобразование в случае неавторизированного доступа. В этом случае у злоумышленника отсутствует ключевое правило встраивание, а декодирование будет содержать следующие действия:

1. Извлечение из кодограммы кода $K'(j)$ при помощи основания ψ'.

2. Восстановление элементов исходного числа:

$$c''_{i,j} = [K'(j)/V'_i] - [K'(j)/(\psi' V'_i)]\ \psi'.$$

где $c''_{i,j}$- i-й элемент реконструируемого числа $C''(j)$, как составляющей реконструируемого фрагмента F при неавторизированном доступе.

Выводы

1. Разработан подход для осуществления косвенного стеганографического встраивания служебной информации на основе использования прямого и обратного функционального преобразования для адаптивных позиционных чисел. Обосновано свойство однозначного обратного декодирования элементов числа при формировании кодограммы на основе исходных и модифицированных служебных данных.

Научная новизна. Впервые разработан подход для осуществления косвенного стеганографического встраивания скрываемого элемента путем модификации основания для адаптивных позиционных чисел фрагмента изображения-контейнера. В отличии от других стеганографических методов, встраивание осуществляется без модификации самих элементов

изображения, что позволяет устранить внесение визуальных искажений. Это позволяет повысить устойчивость встроенных данных к визуальным атакам злоумышленника.

2. Разработан метод косвенного стеганографического встраивания встроенных данных на основе модификации служебных данных. Для минимизации вносимых искажений предложено встраивать один бит скрываемой информации путем модификации основания адаптивного позиционного числа.

Научная новизна. Впервые разработан метод косвенного стеганографического встраивания бита скрываемого сообщения путем модификации основания фрагмента изображения-контейнера. В отличии от других методов косвенное встраивание осуществляется путем модификации оснований элементов фрагмента изображения с последующим формированием на основе модифицированного основания кодовых конструкций для адаптивных позиционных чисел. Это позволяет повысить конфиденциальность встраиваемых служебных данных в интересах государственных структур и ведомственных организаций.

3. Разработан метод косвенного стеганографического изъятия встроенного бита скрываемого сообщения на основе сравнения исходных и модифицированных оснований. Механизм обратного стеганографического изъятия предусматривает:

1) восстановление элементов исходного фрагмента изображения-контейнера на основе системы модифицированных оснований;

2) выявление системы исходный оснований из фрагмента изображения контейнера;

3) косвенное изъятие бита скрываемого сообщения путем сравнения исходных и модифицированных оснований.

Научная новизна. Впервые разработан метод обратного косвенного стеганографического преобразования на основе сравнения исходных и модифицированных оснований. В отличии от других систем, восстановление исходного фрагмента изображения-контейнера осуществляется для неавторизированного и авторизированного пользователя при наличии ключевой информации. Это позволяет скрытно встраивать бит скрываемой служебной информации в фрагмент изображения-контейнера на основе модификации служебных данных.

Литература

1. Грибунин В.Г., Оков И.Н., Туринцев И.В., Цифровая стеганография. – М.: Солон-Пресс, 2002. – 272 с.

2. Конахович Г.Ф., Пузыренко А.Ю., Компьютерная стеганография. Теория и практика. – К.: « МК-Пресс» 2006. – 288с.

3. Тарасов Д.О., Мельник А.С., Голобородько М.М. Класифікація та аналіз безкоштовних програмних засобів стеганографії // Інформаційні системи та мережі. Вісник НУ “Львівська політехніка” № 673. – Львів 2010. – С. 365 - 374.

ПІДХОДИ ДО ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ЗАХИСТУ ІНФОРМАЦІЇ В ІНФОРМАЦІЙНО-ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМАХ НА СТАДІЇ МОДЕРНІЗАЦІЇ

Юдін О.К., Стрельбіцький М.А.

Вступ

Забезпечення безпеки державних інформаційних ресурсів в умовах стрімкого розвитку інформаційних технологій вимагає наявності високоефективних систем захисту інформації (СЗІ). Зазначені системи являють собою складні організаційно-технологічні структури, створення яких вимагає вирішення комплексу системних задач.

Визначення кількісних та якісних оцінок ефективності СЗІ, як об’єктивного підтвердження якості СЗІ є достатньо складним завданням, що залежить від: призначення ІТС, умов її функціонування, типу інформації яка циркулює в ній, тощо.

Існуючі засоби захисту можна розділити на окремі групи, які забезпечують:

– розмежування доступу;

– захист інформації при передаванні каналами зв’язку;

– захист від витоку можливими каналами витоку інформації;

– захист від шкідливого програмного забезпечення, тощо.

Разом із тим, окремі засоби не в змозі забезпечити захист від значної кількості існуючих загроз, а проста комбінація засобів захисту може навіть знизити ефективність їх застосування з причини їх конфліктності.

Загалом ефективність СЗІ, як складного комплексу програмно-апаратних засобів, технічних, організаційних та інших методів і заходів, це здатність системи протистояти негативним впливам дестабілізуючих факторів, що передбачались як в рамках проектування так і іншим загрозам інформаційній безпеці.

Таким чином, розробка інструменту визначення ефективності СЗІ дозволить обґрунтувати шляхи вдосконалення такого типу систем.

Метою статті є аналіз та визначення раціонального підходу до оцінювання ефективності процесу захисту інформації в інформаційно-телекомунікаційних системах на стадії модернізації.

Результати дослідження. На сучасному етапі розвитку, інформація є одним з найцінніших ресурсів який визначає рівень національної безпеки держави. З розвитком інформаційних технологій і зростанням значущості технічних засобів зв'язку інформація піддається все більшій кількості загроз, які за умови їх реалізації можуть призвести до збитків національного масштабу. У цих умовах ефективність функціонування правоохоронних органів в значній мірі залежить від можливості системи захисту інформації запобігти реалізації загроз. Це особливо актуально для прикордонного відомства, яке в значній мірі пов'язано з особливостями

організації захисту державного кордону України, процесу пропуску через державний кордон осіб і транспортних засобів та вантажів, специфікою функціонування на адміністративній межі та на лінії розмежування в зоні проведення антитерористичної операції. Однак, впровадження та обслуговування таких комплексних систем захисту інформації потребують значних витрат ресурсів. Аналогічні дослідження, проведені в цивільних організаціях показують, що витрати коштів на захист інформації досягають 20-30 % усього бюджету організації на інформаційні технології [1]. У цих умовах актуальною є проблема аналізу ефективності функціонування систем спрямованих на забезпечення інформаційної безпеки.

Над розв’язанням зазначеної проблеми у сфері оцінювання ефективності СЗІ працюють Архипов О.Є., Архипова С.А., Бородавко І.Т., Ворожко В.П., Голубничий О. Г., Конахович Г. Ф., Корченко А.Г., Носок С.А., Потапов В.Г., Пузиренко О. Ю., Риндюк В.А. [2-7]. Разом із тим, на теперішній момент залишаються дискусійними методологічні підходи до оцінювання ефективності систем захисту інформації.

Основний матеріал

Існуючі підходи до оцінювання ефективності базуються на двох категоріях: кількісних та якісних оцінок. Якісні методи оцінювання застосовуються в умовах достатньо значної невизначеності та базуються, як правило, на досвіді експертів в даній галузі. З метою формалізації такого типу оцінок використовуються вітчизняні [8] та міжнародні стандарти [9, 10].

Вітчизняні стандарти оцінювання установлює критерії оцінки захищеності інформації, оброблюваної в комп'ютерних системах, від несанкціонованого доступу та являє собою методологічну базу для визначення вимог з захисту властивостей інформації. В зазначеному стандарті сформовані критерії, які надають порівняльну шкалу для оцінки надійності механізмів захисту інформації від несанкціонованого доступу, реалізованих в комп'ютерних системах.

Міжнародні стандарти оцінювання пропонують деталізований перелік аспектів інформаційної безпеки таких як політика безпеки, організація інформаційної безпеки, керування ресурсами, тощо. Інший стандарт ISO/IEC 15408 «Загальні критерії оцінки безпеки інформаційних технологій» описує інфраструктуру в якій користувачі комп’ютерної системи описують вимоги, розробники заявляють про властивості безпеки, а експерти визначають ступінь відповідності. Базою для зазначеного стандарту є «Критерії оцінки безпеки інформаційних технологій». Стандарт містить два основних види вимог безпеки: функціональні, що висуваються до функцій безпеки і реалізує їх механізмів, і вимоги довіри, які пред'являються до технології та процесу розробки та експлуатації [11].

Більш точнішими являються кількісні методи оцінювання, як результат глибокого опису предметної області дослідження, детального моделювання процесів систем захисту інформації. Використання зазначених методів дає можливість визначити конкретне значення оцінки процесу який досліджується та в подальшому здійснити його оптимізацію.

В роботі [12] проведений аналіз типів критеріїв, які зустрічаються на практиці, а саме:

– критерії типу «ефект-витрати», котрий дозволяє оцінити досягнення мети функціонування СЗІ при заданих витратах, іншими словами економічна ефективність;

– критерії, які дозволяють оцінити якість СЗІ за заданими показниками та виключити варіанти, які не задовольняють заданим обмеженням, при цьому використовуються методи багатокритеріальної оптимізації, методів дискретного програмування;

– штучно сформовані критерії, що дозволяють оцінити інтегральний ефект.

Вищенаведені критерії формують різноманітні методи та методики, які описують процес оцінювання ефективності СЗІ, що свідчить про відсутність єдиного методологічного підходу до вирішення завдання такого класу. Крім того, СЗІ являє собою складну систему, ефективність функціонування якої залежить від її елементів та описується, як правило, множиною критеріїв з антагоністичними характеристиками.

Аналіз підходів [12, 13-18] до оцінювання ефективності СЗІ дозволив сформувати їх класифікацію (рис. 1).

Перший підхід до оцінювання ефективності функціонування СЗІ включає якісні оцінки, які оперують категоріями «краще-гірше». Одним із методів даного підходу є методи експертних оцінок, як способу прогнозування та оцінки майбутніх результатів дій на основі прогнозів фахівців [18]. Ефективність використання зазначеного методу залежить від ступеня володіння експертом проблем в досліджуваній галузі, високий професійний рівень та рівень ерудованості, вміння висловлювати чіткі відповіді на поставлені запитання. Експерти можуть сформулювати найкращу, на їх думку, структуру СЗІ, визначити необхідні засоби захисту інформаціх (ЗЗІ), надати рекомендації з організації захисту інформації.

Методи експертних оцінок поділяються на дві групи: індивідуальні та колективні.

Індивідуальні методи використовують думку експертів без врахування думок інших. До таких методів належать: інтерв'ю та анкетування. Метод інтерв'ю ґрунтується на взаємодії аналітика та експерта, де останній дає відповіді на запитання аналітика. Позитивним даного методу є спілкування аналітика з експертом, де експерт може задати уточнюючі запитання, які з погляду аналітика є не суттєвими, хоча експерт може вважати їх значущими. В результаті співбесіди визначаються

фактори впливу на СЗІ, очікуваний ефект від впровадження запропонованих рішень, напрямів підвищення ефективності функціонування СЗІ, тощо.

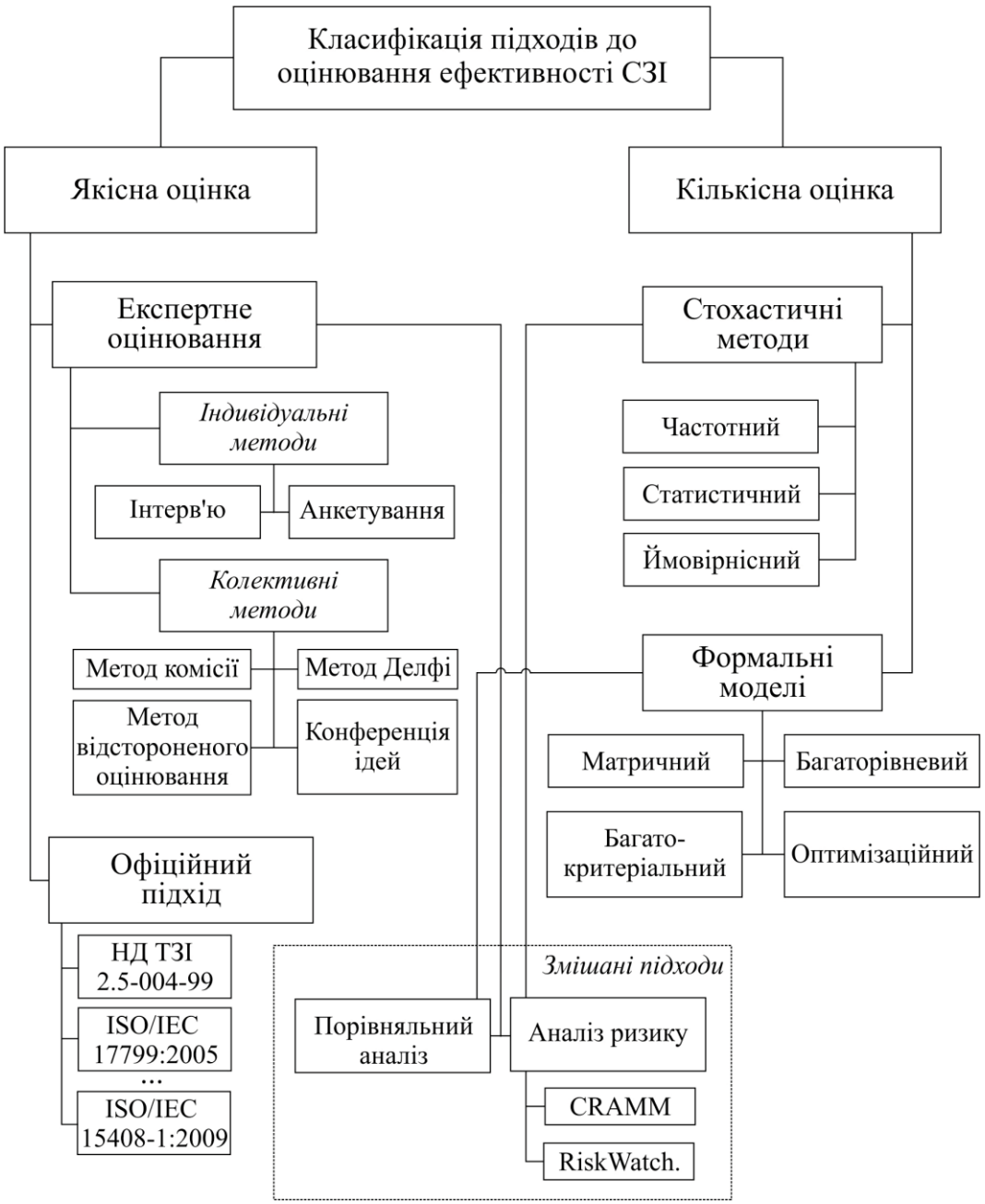


Рис. 1. Класифікація підходів до оцінювання ефективності систем захисту інформації

Метод анкетування полягає у письмовій відповіді експерта на запитання анкети. Разом із тим, даний метод має окремі недоліки, зокрема питання в анкеті можуть бути сформульовані некоректно, то ж і відповіді будуть не точними, тощо.

Наступною групою методів експертної оцінки є колективні експертні методи, які забезпечують формування єдиної спільної думки в результаті

взаємодії залучених фахівців-експертів [18]. Серед зазначених методів виділяють метод комісії, методи Дельфі, відстороненого оцінювання, конференція ідей та інші. Перевагою зазначених методів є можливість залучення фахівців з широким діапазоном знань з теорії та практики. Ступінь забезпечення безпеки системи визначається як [19]:

$$SR = \frac{1}{n}\sum_{i=1}^{n} W_i G_i, \qquad (1)$$

де W_i – суб'єктивний коефіцієнт важливості і-ї характеристики СЗІ;

G_i – призначене експертом значення кожної з характеристик;

n – кількість характеристик.

Ще одної групою якісної оцінки є офіційний підхід. На підставі розроблених раніше стандартів з оцінювання СЗІ, практичного досвіду експлуатації такого типу систем, появи нових загроз інформаційному ресурсу та способів протидії розробляються нормативні документи які регламентують порядок оцінювання СЗІ.

Порядок оцінювання комп'ютерної системи на предмет відповідності критеріям, які наведені в [8] визначається відповідними нормативними документами. Експертна комісія, яка проводить оцінку комп'ютерної системи, визначає, які послуги і на якому рівні реалізовані в даній комп'ютерній системі, і як дотримані вимоги гарантій. Результатом оцінки є рейтинг, що являє собою упорядкований ряд (перелічення) буквено-числових комбінацій, що позначають рівні реалізованих послуг, в поєднанні з рівнем гарантій. Комбінації упорядковуються в порядку опису послуг в критеріях. Для того, щоб до рейтингу комп'ютерної системи міг бути включений певний рівень послуги чи гарантій, повинні бути виконані всі вимоги, перелічені в критеріях для даного рівня послуги або гарантій [8].

Наступним підходом до оцінювання СЗІ є кількісний підхід, який вирішується залученням стохастичних методів та формальних моделей. Використання такого підходу дозволяє об'єктивно визначити значення показників якості системи.

Серед стохастичних методів можна виокремити: частотний, статистичний, ймовірнісний. Частотний метод оцінювання СЗІ базується на аналізі статистичного матеріалу та використовується для визначення збитку від реалізації певної загрози та визначається як [19]:

$$R = 10^{(S+V-4)}, \qquad (2)$$

де S – показник частоти виникнення загрози (обирається в інтервалі 0 – загроза майже не виникає до 7 – висока ймовірність виникнення загрози);

V – показник збитку, обирається в залежності від S.

Недоліком зазначеного методу є потреба у значному об'ємі статистичного матеріалу.

Наступним методом є статистичний, який визначає виникнення певної загрози за визначений період часу, тобто статистична обробка потенційних загроз.

Ще одним із методів стохастичної групи є ймовірнісний. Даний підхід визначає ймовірність відмови системи в результаті дії дестабілізуючих факторів.

Наступною групою якісного оцінювання СЗІ є використання формальних моделей, а саме матричних, багаторівневих, багатокритеріальних, оптимізаційних.

Окремі автори [20] пропонують застосувати матричні або формальні моделі захисту. Відповідно до цього підходу стан системи описується трійкою параметрів: суб'єкти, об'єкти та права доступу. Аналогом такого підходу є модель дискреційного розмежування доступу, де аналогічним чином описується порядок надання доступу користувачам системи, які виступають в рамках даної моделі як суб'єкти, або процеси запущені від їх імені, а також модель Грехема-Денінга [21]. Спосіб визначення показників ефективності агреговано в [12, 18], суть якого полягає у: визначенні параметрів; складання тривимірної матриці відношень; перетворення матриці відношень в двовимірну таблицю; визначення якісних та кількісних значень показників.

Наступним підходом є використання моделей розмежування доступу, в яких вводиться функція рівня безпеки та визначається решітка рівнів безпеки (конфіденційності), які в сукупності визначають допустимі відношення доступу між сутностями системи. Показники ефективності визначаються відповідно до елементів моделі кінцевих станів Белла–Ла Падули,

Ще одним підходом до оцінки ефективності СЗІ є застосування нечітких показників таких як «абсолютно незахищена», «недостатньо захищена», «захищена», достатньо захищена», «абсолютно захищена». Запропонована в [22] методика використовує поняття «рівень безпеки», який нормовано на проміжку [0, 1]. Показники надійності є функцією $\mu^{A}(x_i)$, де x_i – елемент множини вимог безпека, А – множина значень, яка визначає виконання вимог безпеки. Оцінка ефективності за даною методикою здійснюється за попередньо чітко визначеними критеріями безпеки (багатокритеріальний підхід), що є головним недоліком даної методики.

Оптимізаційний або комбінаторний підхід використовує методи дискретного програмування (дискретної оптимізації), де значення функції $\sum_{j=1}^{n} c_j x_j$ необхідно максимізувати (мінімізувати) при умовах: $\sum_{j=1}^{n} a_{ij} x_j \le b_i$, $i = \overline{1,m}$. При даному підході використовуються інструменти лінійного, випуклого програмування, тощо.

Використання тільки одного із запропонованих підходів має недоліки. Наприклад, при використанні тільки методів експертних оцінок є великий ризик впливу певних шаблонів та стереотипів на кінцеве рішення. З іншого боку, використання тільки математичних методів вимагає чітких вихідних даних, які, як правило відсутні. З метою визначення якості функціонування СЗІ в умовах невизначеності вихідних даних запропоновано використати змішані методи. Одним із таких методів є метод порівняльного багатовимірного аналізу, суть якого полягає у визначенні ступеня взаємного впливу загроз та причин їх виникнення [15]. Загальна схема методу являє собою послідовність операцій, а саме:

– складення переліку об'єктів та вибір ознак оцінювання;
– формування матриці ознак;
– нормалізація елементів матриці ознак;
– визначення середнього арифметичного за усіма об'єктами;
– визначення стандартного відхилення;
– визначення матриці відстаней між показниками захищеності.

Отримана матриця відстаней дозволяє попарно зіставити між собою показники захищеності, здійснити їх ранжування, визначити ступінь взаємного впливу. Використання цього методу вимагає формування матриці ознак, для якого застосовується метод експертних оцінок. Наведений метод дозволяє оцінити вплив та взаємодію сформованих експертами загроз та на цій підставі раціоналізувати політику безпеки об'єкту інформаційної діяльності.

Одним із підходів змішаного характеру є управління ризиками, як діяльності, що направлена на прийняття і виконання управлінських рішень з метою зниження ймовірності виникнення негативного результату і мінімізації можливих втрат, які викликані її реалізацією [23]. Одними із найбільш відомих алгоритмів, що сприяє оптимізації сил та засобів при управлінні ризиками є метод CRAMM і RiskWatch.

Метод CRAMM був розроблений Центральним агентством з комп'ютерів та комунікації Великої Британії. Окремі версії методу спрямовані на вирішення завдань Міністерства оборони, цивільних закладів, фінансових структур, приватних організацій. Метою розробки методу було створення формалізованої процедури, яка дозволяє впевнитись, що всі вимоги, які пов'язані з безпекою повністю проаналізовані та задокументовані; уникнути зайвих витрат, які пов'язані із суб'єктивною оцінкою ризику; надавати допомогу при плануванні і здійсненні захисту на всіх стадіях життєвого циклу інформаційних систем; забезпечити проведення робіт у стислі терміни; автоматизувати процес аналізу вимог безпеки; надати обґрунтування для заходів протидії; оцінювати ефективність контрзаходів, порівнювати різні варіанти.

Метод складається із трьох етапів [15]. На першому етапі здійснюється формалізований опис інформаційної системи, її основних

функцій, категорій користувачів, персоналу, який здійснює дослідження системи. Наступним етапом методу є ідентифікація та оцінювання ресурсів. Третім етапом є оцінювання загроз і вразливостей. Оцінка ризику визначається через ймовірність реалізації та величину збитку:

$$Р_{изик} = Р_{реалізації} \cdot З_{биток}, \quad (3)$$

В подальшому здійснюється деталізація ймовірності реалізації:

$$Р_{реалізації} = Р_{загрози} Р_{вразливості}, \quad (4)$$

Таким чином, наведений метод дає підстави для прийняття рішення керівниками організації щодо впровадження нових механізмів безпеки або про модернізацію існуючої СЗІ.

Ще одним із методів аналізу ризиків є метод RiskWatch, перевагами якого є детально опрацьована методика аналізу ризиків, яка добре зарекомендувала себе в багатьох державних організацій США; наявність як кількісної так і якісної оцінки ризиків; достатньо велика база знань щодо загроз, вразливостям та контрзаходам; можливість редагування і вдосконалення бази знань. Структурно метод RiskWatch складається з трьох етапів, на першому з яких здійснюється формалізований опис системи яка досліджується [15]. На другому етапі визначають конкретні характеристики системи. На третьому етапі здійснюється оцінка ризиків, де встановлюються зв’язки між втратами, ресурсами, загрозами та вразливостями та розраховуються математичні очікування втрат за рік:

$$L = PV, \quad (5)$$

де L – сума втрат від загроз за рік;

P – частота виникнення за гроз протягом року;

V – вартість ресурсу.

Наочно, що використання наведених методів вимагає отримання об’єктивних вихідних даних, зокрема статистичних, що може бути проблематичним. Разом із тим, як на наш погляд, характер дії дестабілізуючих факторів (ДФ) на функціонування СЗІ носить ймовірнісний характер. З цієї причини даний підхід є найбільш коректним до опису процесу оцінювання ефективності СЗІ.

Дослідженням такого підходу присвячено ряд робіт [12, 14, 15, 17, 22], але тільки в окремих [24, 25] запропоновано використати підхід теорії ефективності цілеспрямованих процесів. Як на наш погляд, такий підхід найбільш точно описує загалом поняття ефективності системи як ступінь досягнення мети цією системою. Разом із тим, використання такого підходу обмежується наступними причинами: висока ступінь невизначеності вихідних даних, складність формалізації процесів функціонування.

Таким чином, наявність проблем у формуванні і оцінюванні показників ефективності функціонування СЗІ вимагає розробки підходів до оцінювання ефективності захисту інформації на стадії модернізації.

У відповідності до [26] захист інформації в інтегрованій інформаційно-телекомунікаційній системі (ІІТС) це діяльність, спрямована на запобігання несанкціонованим діям щодо інформації в системі. Дане визначення дозволяє сформулювати поняття захисту інформації як цілеспрямований процес з єдиною (що є принциповим) метою – недопущення несанкціонованих дій стосовно інформації під час всього життєвого циклу ІІТС. Закон України "Про захист інформації в інформаційно-телекомунікаційних системах" дає визначення комплексної системи захисту інформації (КСЗІ) як взаємопов'язаної сукупності організаційних та інженерно-технічних заходів, засобів і методів захисту інформації. Варто зазначити, що у наведеному визначенні КСЗІ розглядається ізольовано від зовнішнього середовища, а саме від умов функціонування та умов застосування системи.

В роботах [25, 27-28] використовується поняття зовнішнього середовища, як основної умови забезпечення функціонування ІТС, якість якої залежить не тільки від ступеня захисту інформації яка циркулює в ній, а й від здатності запобігання негативного впливу зовнішнього середовища та шкідливого програмного забезпечення. В основному всі автори, які розглядають в своїх роботах поняття «зовнішнє середовище» оцінюють тільки його негативний вплив на якість функціонування системи. Зовнішнє середовище розглядається як джерело загроз інформації: діяльність організацій (окремих осіб), вплив стихійного лиха, тощо. Такий підхід до розгляду зазначеного поняття має раціональне підґрунтя, а саме є сенс розглядати тільки негативний вплив, так як позитивний не призводить до зниження якості функціонування ІТС в цілому. Разом із тим, вплив зовнішнього середовища може здійснювати і позитивний вплив на функціонування СЗІ або компенсувати негативні дії на її якість. В рамках такого підходу зовнішнє середовище ІТС розглядається не тільки з негативної сторони, а як неконтрольований вплив на функціонування системи в цілому.

Дамо визначення поняття "навколишнє середовище" в рамках терміну "захист інформації", а саме як сукупність об’єктів, які не входять в СЗІ (КСЗІ) та безпосередньо не приймають участь в процесі захисту інформації, але здійснюють вплив на досягнення мети захисту інформації. В подальшому поняття "навколишнє середовище" будемо розуміти як сукупність умов функціонування та умов застосування системи захисту інформації.

В загальному, функціонування системи захисту інформації можна представити як складну людино-машинну (ергатичну) систему з множиною можливих станів яка взаємодіє із зовнішнім середовищем та оперує власними ресурсами із завданням досягнення мети (рис. 2).

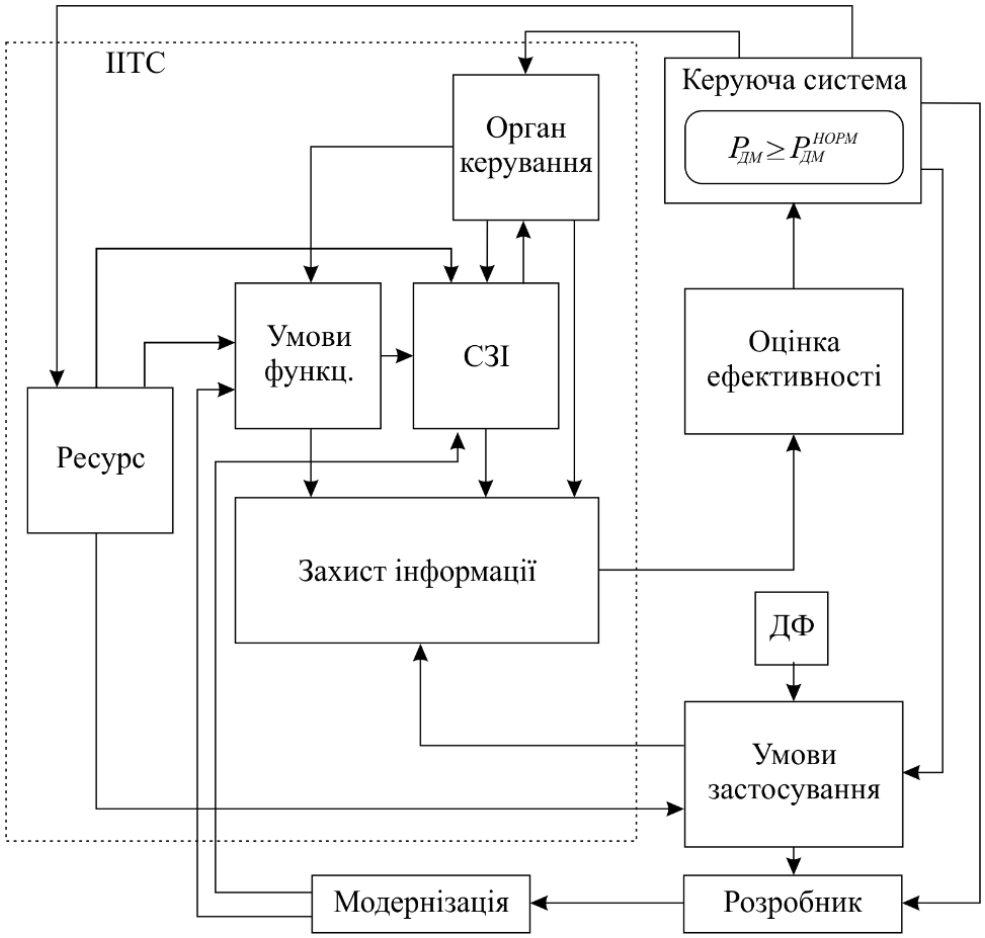


Рис. 2. Узагальнена структура функціонування СЗІ

По відношенню до ІТС «зовнішнім середовищем» вважаються умови застосування системи на які діють різного типу дестабілізуючі фактори об'єктивного характеру, а також керуюча система та процес модернізації розробником.

Керуюча система (розпорядник ІІТС) безпосередньо не входить до складу СЗІ як підсистеми ІТС, але здійснює безпосередній вплив на наявність ресурсів системи, визначає умови застосування та, через орган керування, умови функціонування ІІТС. Наприклад, у прикордонному відомстві розпорядником ІТС є Адміністрація ДПС України. При створенні ІТС визначаються технічні умови розгортання зазначеної системи, розпорядником виділяються ресурси на її створення. Реалізуючи розпорядчу функцію Адміністрація через відповідні керівні документи формує порядок застосування СЗІ в органах охорони державного кордону. Наявність зворотного зв'язку, через оцінку ефективності СЗІ дозволяє Адміністрації здійснювати корегування умов функціонування та застосування системи в цілому. При потребі модернізації існуючої системи розпорядник системи надсилає розробнику технічне завдання на модернізацію системи. Розробник, в свою чергу, отримавши завдання на модернізацію системи враховує умови застосування та здійснює зміни в

системі, в тому числі стосовно СЗІ. Таким чином, модернізація складових ІІТС приводить до зміни показника ефективності процесу захисту інформації, який після її оцінки порівнюється з нормативним значенням та на цій підставі приймається рішення керуючою системою щодо запровадження внесених змін, зміни умов функціонування системи чи визначення допустимих умов застосування ІІТС або регулювання ресурсів, які виділяються на процес захисту інформації з метою дотримання відповідності показника ефективності захисту заданому критерію. Такий підхід передбачає вирішення зворотної задачі оцінювання ефективності СЗІ, а саме задачу синтезу СЗІ при заданих умовах застосування та функціонування. Вищенаведене вимагає використання інших критеріїв: критеріїв переваги та оптимальності, що виходить за межі дослідження.

На систему захисту інформації в ІТС впливають: орган керування, який визначає умови функціонування ІТС та СЗІ, зокрема; наявний ресурс, що забезпечує функціонування СЗІ та визначає умови функціонування ІТС в цілому. Безпосередньо на сам процес захисту впливають умови функціонування (внутрішній фактор) та умови застосування (зовнішній фактор), структура та організація СЗІ, орган керування.

Аналіз узагальненої структури функціонування показав, що вплив дестабілізуючих факторів на процес захисту інформації здійснюється опосередковано через умови застосування системи. Керуюча система з метою дотримання заданого (нормативного) рівня захисту має можливість визначати допустимі умови застосування.

Під умовами функціонування СЗІ будемо розуміти сукупність факторів, які впливають на характеристики СЗІ (стабільність, надійність, відновлюваність, керованість, тощо). До умов функціонування віднесемо також природні та техногенні умови в яких функціонує система, способи її застосування (постійний, періодичний), структуру та організацію СЗІ, кількість та якість ресурсів. У якості прикладу умов функціонування СЗІ наведемо мобільний програмно-технічний комплекс автоматизації прикордонного контролю «Гарт 1/П» який базується на легковому автомобілі та застосовується відповідно до рішення начальника органу охорони кордону на різних ділянках відповідальності та в різний період. Таким чином, умови функціонування такого комплексу є різними (погодні умови, спосіб електроживлення, наявність каналів зв’язку, тощо), спосіб застосування – періодичний.

Умови застосування СЗІ – сукупність факторів організаційно-ситуаційного характеру, які впливають на ситуацію в якій СЗІ виконує свої завдання та визначає допустимі результати виконання завдань функціонального характеру. До умов застосування ІТС ДПСУ відноситься оперативно-стратегічна обстановка, як некерований фактор, що визначається дійсним розвитком прикордонного відомства так і впливом ризиків та загроз його розвитку. Крім того, стратегія майбутнього

використання ІТС визначає основний характер їх цільового застосування. Зазначені фактори за своєю суттю є випадковими, тобто до моменту запуску в експлуатацію ІТС (або після модернізації) їх значення невідомі. Це призводить до неможливості отримання розрахункового значення показника ефективності. Тому, з метою усунення цієї невизначеності необхідно визначити ймовірнісні характеристики всіх випадкових факторів

Відповідно до визначення, поняття "захист інформації" можна розглядати як сукупність (послідовність) узгоджених дій протягом певного часу які спрямовані на досягнення мети цього процесу. При оцінці ефективності необхідно звернути увагу на те, що це властивість процесу, а не самої системи. Тому в подальшому під поняттям ефективності захисту інформації будемо розуміти комплексну властивість цілеспрямованого процесу, який характеризується ступенем досягнення мети, а саме захисту інформації.

При оцінюванні якості СЗІ, яка описується n-вимірним векторним показником $Y_{\langle n \rangle}$ необхідно визначити сукупність критеріїв, які належать класу критеріїв придатності $\{G\}$, математичне формулювання якого має вигляд [29]:

$$G:\left(Y_{\langle n \rangle} \in \left\{Y^{A}_{\langle n \rangle}\right\}\right), \tag{6}$$

де $Y_{\langle n \rangle}$ - показник якості СЗІ;

$\left\{Y^{A}_{\langle n \rangle}\right\}$ - множина допустимих значень показника якості СЗІ.

Таким чином, СЗІ для якої виконується умова (6) придатна до використання за призначенням та виконує свої функції.

Серед множини властивостей системи захисту інформації істотними є ті, які визначають якість процесу захисту інформації. Структурою критеріїв захищеності інформації [8] визначені функціональні критерії які описують вимоги до послуг, що забезпечують захист від загроз одного із чотирьох основних типів: конфіденційності, цілісності, доступності, спостереженості, що визначає множину типів показників якості СЗІ (властивостей інформації):

$$p = \{i, c, a, u\}, \tag{7}$$

де i – цілісність (**i**ntegrity);

c – конфіденційність (**c**onfidentiality);

a – доступність (**a**vailability);

u – спостереженість (acco**u**ntability).

Разом із тим, в процесі захисту інформації витрачаються ресурси задля підтримання функціонування СЗІ на заданому рівні ефективності. Таким чином, СЗІ в будь-який момент часу можна охарактеризувати трійкою властивостей:

– результативністю – властивістю системи забезпечити захист інформації;

– ресурсоємністю, яка характеризується витратою ресурсів системи (матеріально-технічних, часових, енергетичних, фінансових, людських тощо);

– оперативністю – властивістю системи здійснювати захист інформацій протягом зазначеного терміну часу.

Із зазначеного вище можна зробити висновок, що якість захисту інформації не може бути охарактеризована окремими властивостями, а визначається тільки їх сукупністю, тобто трійкою властивостей.

Введемо позначення зазначених властивостей:

$V_{\langle n_1 \rangle}$ – показник результативності захисту інформації;

$R_{\langle n_2 \rangle}$ – показник витрат ресурсів;

$T_{\langle n_3 \rangle}$ – показник часу.

Тоді, показником якості СЗІ буде n-вимірний вектор, котрий містить три групи властивостей:

$$Y_{\langle n \rangle} = \left\langle V_{\langle n_1 \rangle}, R_{\langle n_2 \rangle}, T_{\langle n_3 \rangle} \right\rangle, \tag{8}$$

де $n = n_1 + n_2 + n_3$

Всередині груп можливо згортання часткових показників шляхом введення узагальнених показників. Так, в більшості випадків витрати ресурсів можна привести до витрат коштів, тоді (8) прийме вигляд:

$$Y = \langle v_i, v_c, v_a, v_u; r; \tau \rangle, \tag{9}$$

де v_i – показник цілісності;

v_c – показник конфіденційності;

v_a – показник доступності;

v_u – показник спостереженості;

$r = \sum_{r_i \in R} r_i$ - показник витрат ресурсів;

τ – показник часу.

Разом із тим необхідно врахувати, при згортанні різнорідних показників узагальнений показник губить фізичний сенс, тому при багатокритеріальному аналізі коректним є згортання показників тільки всередині груп показників результатів. Згортання показників якості функціонування системи із різних груп є недопустимим.

Фізичний сенс показників результативності захисту інформації полягають у визначенні часу, протягом якого властивість інформації не буде порушена.

Фізичний сенс показника часу полягає у визначенні часу роботи всіх засобів забезпечення захисту при якому забезпечується нормативний

рівень їх функціонування. З точки зору надійності це напрацювання до відмови та описується відомими функціональними залежностями теорії надійності. Даний показник являється складовим у формуванні показників результативності захисту інформації та може бути згорнутий в них.

Аналогічно, показник витрат ресурсів також залежить від часу функціонування системи. Разом із тим, розгортання об'єкта інформаційної діяльності передбачає витрату певних ресурсів та при їх відсутності або недостатній кількості дозвіл на функціонування об'єкту інформаційної діяльності не надається. Таким чином, даний показник не потребує дослідження і може бути виключений із вектору показників якості СЗІ та враховуватись окремо при порівнянні двох систем з однаковими значеннями показника якості.

Таким чином, вектор показників якості функціонування СЗІ (9) прийме вигляд:

$$Y = \langle v_i, v_c, v_a, v_u \rangle . \qquad (10)$$

Варто зазначити, що компоненти вектору Y являються кількісними характеристиками кількісних результатів самого процесу захисту інформації. Будемо вважати, що їх якісна характеристика завчасно забезпечується ще до початку експлуатації СЗІ. Аналогічне зауваження застосуємо до якісної характеристики ресурсного забезпечення.

Система захисту інформації, як взаємопов'язана сукупність організаційних та інженерно-технічних заходів, засобів і методів захисту [25] повинна мати властивості системи, а не бути просто сукупністю певних засобів. Крім того, системний підхід повинен застосовуватись на всіх етапах життєвого циклу – від підготовки технічного завдання до експлуатації СЗІ. Зазначимо, що система такого типу повинна мати чітке призначення, причому чим конкретніше сформульована мета системи, тим адекватніше буде її описувати показник ефективності. Складність формулювання мети СЗІ ІІТС полягає у її значній розосередженості та багатофункціональності, причому кожна із підсистем постійно підлягає модернізації. При такому підході значущість властивостей окремих елементів СЗІ зменшується, а загальносистемні завдання, такі як визначення раціональної структури і режимів функціонування, організація взаємодії між складовими системи, вплив умов застосування та функціонування системи. Системне об'єднання складових СЗІ створюють ефект емержентності, тобто появу властивостей які не притаманні жодному елементу окремо.

Наочно, що аналіз ефективності функціонування СЗІ вимагає формування та вирішення завдання кількісного оцінювання характеристик системи. Зазначені дані, які отримані або математичним моделюванням або експериментальним шляхом повинні описувати властивості системи, основним з яких є ефективність функціонування СЗІ. Дана властивість

системи агрегує в собі інші системні властивості, такі як надійність, керованість, оперативність відновлення після збоїв, тощо. Разом із тим, кількісна оцінка ефективності дозволяє здійснити порівняння системи при експлуатації в різних умовах та визначити допустимі межі експлуатаційних умов, при яких дотримується нормативне значення показника захисту інформації.

В більшості випадків при проектуванні СЗІ застосовують емпірично-інтуїтивний підхід, що наглядно показано в оцінюванні рівня безпеки інформаційних систем, коли оперують нечіткими поняттями, наприклад «достатньо захищений» [27]. Нечітке визначення у відношенні до інформаційної безпеки пов'язане з нечіткою постановкою завдання, вимог до захисту в умовах стохастичних впливів дестабілізуючих факторів. Як правило, це пояснюється тим, що СЗІ проектується після розробки основного функціоналу ІТС та не є складовою системи на стадії проектування.

Для оцінювання ефективності функціонування системи захисту інформації необхідно розробити показник ефективності процесу захисту інформації, який повинен відповідати основним вимогам [29]: показовість (адекватність), критичність (чутливість), комплексність (повнота), стохастичність, простота.

На показники результативності СЗІ впливають зовнішні та внутрішні фактори, які визначаються середовищем її функціонування.

Кожна з компонент вектору Y залежить від характеристик СЗІ та її організації, умов функціонування та умов застосування системи.

$$\mathrm{Y} = \mathrm{Y}\left(\mathrm{A}_1, \mathrm{A}_2, \mathrm{B}_1, \mathrm{B}_2\right), \tag{11}$$

де A_1 - характеристики СЗІ;

A_2 - характеристики організації процесу ЗІ;

B_1 - характеристики умов функціонування ЗСІ

B_2 - характеристики умов застосування ЗСІ.

В свою чергу, компоненти вектору Y^A допустимих значень теж залежать від умов застосування системи і визначаються керуючою системою.

$$\mathrm{Y}^{\mathrm{A}} = \mathrm{Y}^{\mathrm{A}}\left(\mathrm{B}_2\right). \tag{12}$$

В загальному випадку на характеристики СЗІ, її організації, умови функціонування та застосування СЗІ діє множина випадкових факторів, що визначає зазначені величини як випадковими. Разом із тим, апріорі випадковими є і допустимі значення вектору Y^A, який залежить від умов застосування системи, так як завчасно невідомо, які повинні бути результати роботи СЗІ, щоб забезпечити необхідний рівень захисту. Окремі дослідження при визначенні умов застосування та функціонування системи приймають припущення про найгірший їх варіант (з точки зору

захисту інформації), тобто величини B_1 та B_2 є не випадковими. Зазначене припущення призводить до неправомірно великих витрат ресурсів.

Таким чином, всі складові вектору показників якості функціонування СЗІ носять ймовірнісний характер, тому:

$$\hat{Y} = Y\left(\hat{A}_1, \hat{A}_2, \hat{B}_1, \hat{B}_2\right), \quad (13)$$

$$\hat{Y}^A = Y^A\left(\hat{B}_2\right).$$

В результаті реальних умов експлуатації СЗІ критерій придатності (6) прийме вигляд:

$$G : \left(\hat{Y} \in \left\{\hat{Y}^A\right\}\right). \quad (14)$$

З виразу (13) можна зробити висновок, що придатність процесу захисту інформації – випадкова подія, яка безпосередньо не може відображати якість процесу. Тому, характеристикою якості СЗІ є ймовірність випадкової події:

$$P_{ДМ} = P\left(\hat{Y} \in \left\{\hat{Y}^A\right\}\right). \quad (15)$$

Таким чином, ймовірність $P_{ДМ}$ - це показник ефективності СЗІ, який визначає ступінь виконання СЗІ своїх функціональних завдань. На її основі формується критерій придатності системи, тобто $P_{ДМ} \geq P_{ДМ}^{НОРМ}$.

Висновки

На підставі проведеного аналізу існуючих підходів до оцінювання ефективності систем захисту інформації запропоновано використати підхід теорії ефективності цілеспрямованих процесів, як такий, що найбільш точно описує поняття ефективності системи як ступінь досягнення мети цією системою. Разом із тим, його використання обмежується наступними причинами: висока ступінь невизначеності вихідних даних, складність формалізації процесів функціонування. Аналіз узагальненої структури функціонування СЗІ показав, що вплив дестабілізуючих факторів на процес захисту інформації здійснюється опосередковано через умови застосування системи. Керуюча система з метою дотримання заданого (нормативного) рівня захисту має можливість визначати допустимі умови застосування.

Вищенаведене дозволило розробити показник ефективності процесу захисту інформації, який відповідає основним вимогам: показовості, критичності, комплексності, стохастичності, простоті. На підставі розробленого показника сформовано критерій придатності системи.

Література

1. Петренко С. А. Информационная безопасность: экономические аспекты / С. Петренко, С. Симонов, Р. Кислов // Jet Info Online. - 2003. - №10. [Электронный ресурс]. - Режим доступа: http://citforum.ru/security/articles/sec/index.shtml.

2. Архипов А. Е. Технологии експертного оценивания в задачах защиты информации / А. Е. Архипов, С. А. Архипова, С. А. Носок // Інформаційні технології та комп'ютерна інженерія : міжнар. наук.-техн. журн. – 2005. – № 1. –С. 89 - 94.

3. Архипов О. Є. Оцінювання ефективності системи охорони державної таємниці : монографія / О. Є. Архипов, І. Т. Бородавко, В. П. Ворожко. – К., 2007. – 63 с.

4. Архипов О. Є. Системні аспекти оцінювання рівня важливості секретної інформації / О. Є. Архипов, В. П. Ворожко // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні : наук.-техн. зб. – К., 2007. – Вип. 2 (15). – С. 10 - 12.

5. Корченко А. Г. Построение систем защиты информации на нечетких множествах. Теория и практические решения / А. Г. Корченко. – К.: "МК-Прес". – 2006. – 320 с.

6. Корченко А. Г. Экспертиза в системе ТЗИ на основе нечетких множеств / А. Г. Корченко, В. Г. Потапов, В. А. Рындюк // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – К., 2003. – Вип. 7. – С. 118 - 127.

7. Конахович, Г. Ф., О. Г. Голубничий, О. Ю. Пузиренко. Оцінка ефективності систем захисту інформації в телекомунікаційних системах. // Проблеми інформатизації та управління. – 2007. – С. 75 - 83.

6. Критерії оцінки захищеності інформації в комп'ютерних системах від несанкціонованого доступу НД ТЗІ 2.5-004-99. Затверджено Наказом департаменту спеціальних телекомунікаційних систем та захисту інформації служби безпеки України від "28" квітня 1999 р. № 22. Із змінами згідно наказу адміністрації Держспецзв'язку від 28.12.2012 № 8069. ISO/IEC 17799:2005 Information technology -- Security techniques -- Code of practice for information security management [Электронный ресурс]. - Режим доступа: https://www.iso.org/standard/39612.html

10. ISO/IEC 15408-1:2009 Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 1: Introduction and general model [Электронный ресурс]. - Режим доступа: https://www.iso.org/standard/50341.html

11. Міжнародні стандарти забезпечення інформаційної безпеки підприємста (СР№1) [Электронный ресурс]. - Режим доступа: https://nikitenko11.wordpress.com/2012/11/23/міжнародні-стандарти-забезпечення-і/

12. Маслова Н. А. Методы оценки эффективности систем защиты информационных систем / Н. А. Маслова // Искусственный интеллект. – 2008. – № 4. – С. 253–264.

13. Павлов І.М. Неформальний підхід в методиці оцінки ефективності ескізного проектування комплексних систем захисту інформації. // Захист інформації, 12 (1 (46)).

14. Пігур Н. В., В. Д. Погребенник. Оцінювання ефективності комплексних систем захисту інформації."(2013). [Электронный ресурс]. - Режим доступа: http://ena.lp.edu.ua:8080/bitstream/ntb/23063/1/45-61-61.pdf

15. Гарасимчук О. І., Костів Ю. М. Оцінка ефективності систем захисту інформації. Вісник КНУ ім. М. Остроградського. – Кременчук: КНУ ім. М. Остроградського. – 2011 (66). Частина 1. – Випуск 1.

16. Власов, О. М., С. В. Толюпа. "Комплексний підхід оцінки ефективності систем захисту інформації в інфокомунікаційних мережах нового покоління." Наукові записки Наукові записки УНДІЗ. Науково-вироб. зб. 3. – 2011. – С. 19.

17. Янчук, В. О. Методика оцінювання стану захисту інформації локальних об’єктів системи електронного врядування. [Электронный ресурс]. - Режим доступа: http://academy.gov.ua/ej/ej11/txts/10ivoseu.pdf

18. С. В. Толюпа, Ю. Я. Самохвалов, Н. В. Цьопа. Комплексні системи захисту інформації спеціальних об’єктів та методика їх оцінки." Сучасний захист інформації. – 2014.

18. Теорія економічного аналізу: Навч. посіб. / Купалова Г.І. – К., 2008. – 639 с.

19. Домарев В.В. Безопасность информационных технологий. Методология создания систем защиты – К.: ООО "ТИД "ДС", 2002. – 688 с.

20. А. В. Артемов Информационная безопасность. Курс лекций [Электронный ресурс]. - Режим доступа: http://fictionbook.ru/static/trials/09/06/63/09066361.a4.pdf

21. Щеглов А.Ю. Защита компьютерной информации от несанкционированного доступа. – М: Наука и техника, 2003. – 384 с.

22. Чипига А.Ф. Оценка эффективности защищённости автоматизированных систем от несанкционированного доступа / Чипига А.Ф., Пелешенко В.С. // Вестник СевКавГТУ / Серия «Физико- химическая» –2004. – № 1(8).

23. Алексеев, Антон. Управление рисками. Метод CRAMM. IT Expert.– Электрон. дан.–М.: ЗАО “ИТ Эксперт (2010). [Электронный ресурс]. - Режим доступа: http://www.itexpert.ru/rus/ITEMS/ITEMS_CRAMM.pdf

24. Грибунин В. Г. Комплексная система защиты информации на предприятии: учеб. пособие для студ. высш. учеб. заведений / В. Г. Грибунин, В.В.Чудовский. – М.: Издательский центр Академия. – 2009. – 416 с.

25. Гончарова Л.Л., Возненко А.Д., Стасюк О.І., Коваль Ю.О. Основи захисту інформації в телекомунікаційних та комп’ютерних мережах. – К., 2013. – 435 с.

25. Закон України Про захист інформації в інформаційно-телекомунікаційних системах.

27.Основы информационной безопасности. Учебное пособие для вузов / Е. Б. Белов, В. П. Лось, Р. В. Мещеряков, А. А. Шелупанов. – М.: Горячая линия – Телеком, 2006. – 544 с.

28. Кавун С.В. Інформаційна безпека. Навчальний посібник. Ч.1 / С.В. Кавун, В.В. Носов, О.В. Мажай. – Харків: Вид. ХНЕУ, 2008. – 352 с.

Научное издание

НАУКОЕМКИЕ ТЕХНОЛОГИИ В ИНФОКОММУНИКАЦИЯХ: ОБРАБОТКА ИНФОРМАЦИИ, КИБЕРБЕЗОПАСНОСТЬ, ИНФОРМАЦИОННАЯ БОРЬБА

Коллективная монография

Под общей редакцией В.М. Безрука, В.В. Баранника

(В авторской редакции)

Корректура и компьютерная верстка *Н.А. Харченко*
Дизайн обложки *В.В. Твердохлеб*