УДК 621.396

ВСТРАИВАНИЕ ИНФОРМАЦИОННЫХ ДАННЫХ В НЕПОДВИЖНЫЕ ИЗОБРАЖЕНИЯ С ИСПОЛЬЗОВАНИЕМ ПРЯМОГО РАСШИРЕНИЯ СПЕКТРА

А.А. КУЗНЕЦОВ, А.М. БОТНОВ, П.А. ЛАПТИЙ

Исследуются стеганографические методы встраивания данных в неподвижные изображения для скрытной передачи информации. Рассматривается метод стеганографической защиты, основанный на использовании прямого расширения спектра дискретных сигналов, исследуется его эффективность с точки зрения обеспечиваемой стойкости, пропускной способности и величины вносимых искажений в контейнер-изображение.

Ключевые слова: метод расширения спектра, дискретный сигнал, корреляционный прием, стеганографическая защита информации.

1. ПОСТАНОВКА ПРОБЛЕМЫ В ОБЩЕМ ВИДЕ И АНАЛИЗ ЛИТЕРАТУРЫ

Важным направлением в развитии современных средств защиты информации являются стеганографические системы, которые обеспечивают сокрытие в тайне от противника не только информационного содержания передаваемых данных, но и самого факта передачи сообщений [1, 2]. Наиболее перспективными являются стеганографические методы, построение которых базируется на развитом математическом аппарате теории дискретных сигналов и помехозащищенной передачи данных [3–7].

Целью данной статьи является исследование стеганографического метода встраивания данных в неподвижные изображения [2], основанного на использовании прямого расширения спектра дискретных сигналов, оценка его эффективности с точки зрения обеспечиваемой стойкости, пропускной способности и величины вносимых искажений в контейнер-изображение.

2. ПРЯМОЕ РАСШИРЕНИЕ СПЕКТРА В ТЕОРИИ СВЯЗИ

Для построения современных помехозащищенных систем цифровой связи используются методы теории дискретных сигналов, корреляционного и спектрального анализа [3–7]. С точки зрения эффективного использования частотно-временных и энергетических ресурсов каналов связи наиболее перспективными считаются широкополосные системы с шумоподобными дискретными сигналами и прямым расширением спектра [3, 4].

Под дискретным сигналом будем понимать информационный сигнал, который представляется в виде отдельных значений, взятых по времени. Далее мы будем рассматривать дискретный сигнал как двоичную псевдослучайную последовательность (ПСП) $\Phi_i = \left(\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}}\right)$ длины n из множества $\Phi = \left\{\Phi_0, \Phi_1, ..., \Phi_{M-1}\right\}$ мощности $|\Phi| = M$ [6].

Элементы двоичной ПСП принимают одно из значений:

$$\varphi_{i_z} = \begin{cases} +1 \\ -1 \end{cases}, \; z = 0, ..., n-1 . \qquad (1)$$

Для построения помехозащищенной широкополосной связи используют понятие корреляции дискретных сигналов - статистической взаимосвязи двух или нескольких ПСП. Математической мерой коррелированности (похожести) двух дискретных сигналов $\Phi_i, \Phi_j \in \Phi$ служит коэффициент корреляции $\rho(\Phi_i, \Phi_j)$ [3, 4]:

$$\rho\left(\Phi_i, \Phi_j\right) = \frac{1}{n} \sum_{z=0}^{n-1} \Phi_{iz} \Phi_{j_z} . \qquad (2)$$

Два сигнала Φ_i, Φ_j называют ортогональными, если коэффициент корреляции $\rho(\Phi_i, \Phi_j) = 0$. Если $\rho(\Phi_i, \Phi_j) \approx 0$ будем называть сигналы Φ_i и Φ_j квазиортогональными [5, 7].

В работах [3 – 5, 7] исследованы различные подходы к построению дискретных сигналов с улучшенными ансамблевыми и корреляционными свойствами: производные ортогональные системы сигналов (ПОСС); нелинейные производные кодовые последовательности (НПКП); полные кодовые кольца (ПКК); последовательности Голда. В табл. 1 в качестве примера приведены результаты исследований ансамблевых и корреляционных свойств производных систем сигналов [7].

Таблица 1

Ансамблевые и корреляционные свойства дискретных сигналов

n	M	ρ
64	$\approx 10^3$	$2{,}1/\sqrt{n}$
128	$\approx 10^4$	$2{,}5/\sqrt{n}$
256	$\approx 10^6$	$2{,}9/\sqrt{n}$
512	$\approx 10^7$	$3{,}2/\sqrt{n}$
1024	$\approx 10^8$	$3{,}5/\sqrt{n}$
2048	$\approx 10^9$	$3{,}8/\sqrt{n}$
4096	$\approx 10^{10}$	$3{,}9/\sqrt{n}$

Как следует из приведенных в табл. 1 данных, применение производных ортогональных диск-

ретных сигналов позволяет при сохранении низкой коррелированности дискретных последовательностей ($\rho(\Phi_i,\Phi_j)\approx 0$) существенно повысить мощность M ансамблей дискретных сигналов, с ростом длины последовательностей эта тенденция усиливается.

В современной теории цифровой связи большие ансамбли слабокоррелированных дискретных сигналов используются для построения широкополосных помехозащищенных систем передачи данных. Передаваемые сообщения в таких каналах приобретают вид шумоподобных последовательностей, а за счет большой мощности ансамблей дискретных сигналов и прямого расширения частотного спектра обеспечивается высокая имитостойкость, помехозащищенность и скрытность цифровых каналов связи [3 – 5].

Для передачи данных в широкополосной системе связи информационный сигнал $x(t)=\begin{cases}+1\\-1\end{cases}$ модулируется посредством его умножения на расширяющий кодовый сигнал $g(t)=\Phi_i\in\Phi$ – псевдослучайную последовательность из рассмотренных выше ансамблей дискретных сигналов. Поскольку кодовый сигнал по своим статистическим свойствам подобен шуму, то полученный расширенный сигнал

$$y'(t)=y(t)+e(t) \qquad (3)$$

слабо отличим от шумов в канале связи, что и позволяет осуществить скрытую передачу.

При приеме в демодуляторе полученный сигнал $y'(t)=y(t)+e(t)$ как смесь переданной последовательности $y(t)$ и произошедших в канале связи ошибок $e(t)$ умножается на синхронизированную копию расширяющего сигнала $g(t)$. Другими словами, на приемной стороне осуществляется вычисление коэффициента корреляции (2), значение которого определяет правило принятия решения:

$$\rho\left(y'(t),g(t)\right)=\frac{1}{n}\sum_{z=0}^{n-1}x(t)\Phi_{i_z}\Phi_{i_z}+\frac{1}{n}\sum_{z=0}^{n-1}e(t)\Phi_{i_z}.$$

Учитывая псевдослучайность Φ_i, используемых в качестве $g(t)$, вторым слагаемым в правой части равенства можно пренебречь (количество «+1» примерно равно количеству «-1»), т.е.

$$\begin{aligned}\rho\left(y'(t),g(t)\right)&\approx\rho\left(y(t),g(t)\right)=\\&=x(t)\frac{1}{n}\sum_{z=0}^{n-1}\left(\Phi_{i_z}\right)^2=x(t),\end{aligned} \qquad (4)$$

т.е. значение информационного сигнала на приемной стороне определяется по выражению

$$x(t)=\begin{cases}+1,\text{ при }\rho\left(y'(t),g(t)\right)\approx+1;\\-1,\text{ при }\rho\left(y'(t),g(t)\right)\approx-1;\end{cases} \qquad (5)$$

где знак «≈» предполагает наличие ошибок $e(t)$, вызванных естественными или преднамеренными помехами в канале связи.

Структурная схема приема-передачи информации с использованием прямого расширения спектра приведена на рис. 1.

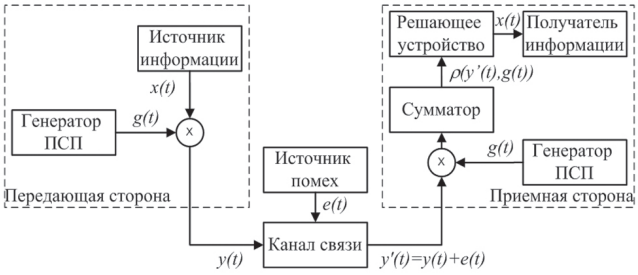

Рис. 1. Структурная схема передачи приема-передачи информации с использованием прямого расширения спектра

Предположим, что временная длительность немодулированного сигнала $x(t)$ равна T, а его частота соответственно равна $F\left(x(t)\right)=\frac{1}{T}$. Передача модулированного сигнала $y(t)$ при той же временной длительности T приведет к расширению частотного спектра передаваемого сигнала, пропорционально числу элементов псевдослучайной последовательности, т.е. пропорционально длине n: $F\left(y(t)\right)=n\frac{1}{T}=nF\left(x(t)\right)$. Тем не менее, использование прямого расширения спектра передаваемого сигнала обеспечивает одновременную передачу многих других информационных сигналов в той же полосе частот. Это следует из взаимной ортогональности (квазиортогональности) применяемых ансамблей дискретных сигналов. Действительно, если на приемной стороне принята аддитивная смесь $\sum_l y_l(t)$ нескольких модулированных сигналов, тогда вычисление коэффициента корреляции даст следующее:

$$\rho\left(\sum_l y_l(t),g(t)\right)=\frac{1}{n}\sum_l\sum_{z=0}^{n-1}x_l(t)\Phi_{l_z}\Phi_{i_z}. \qquad (6)$$

Но все последовательности из множества Φ имеют низкое значение взаимной корреляции, т.е. при $l\neq i$ имеем $\rho(\Phi_l,\Phi_i)=0$ (для ортогональных сигналов имеем равенство $\rho(\Phi_l,\Phi_i)=0$. Следовательно, всеми слагаемыми при $l\neq i$ в правой части равенства (6) можно пренебречь. Отсюда, при наличии в аддитивной смеси $\sum_l y_l(t)$ дискретного сигнала $\Phi_{l=i}$ имеем выражение (4) и соответствующее правило принятия решения (5).

Метод прямого расширения спектра нашел практическое использование в системах цифровой связи с кодовым разделением каналов (CDMA), где для каждого абонента информационного обмена используются уникальные расширяющие кодовые сигналы из ансамбля ортогональных (квазиортогональных) последовательностей. Т.е. для различения кодовых сигналов и разделения соответствующих абонентских каналов используемые ПСП должны быть слабо коррелированны

друг с другом, в идеальном случае – ортогональными.

Так, например, в стандарте CDMA IS-95 для кодового разделения каналов используются ортогональные дискретные сигналы Уолша-Адамара [4]. Они образуются из строк матрицы Адамара H_i, формируемой по рекуррентному правилу:

$$H_i = \begin{bmatrix} H_{i-1} & H_{i-1} \\ H_{i-1} & -H_{i-1} \end{bmatrix}, \; H_0 = [1]. \qquad (7)$$

Многократное повторение правила (7) позволяет сформировать матрицу Адамара любого размера, кратного четырем. Строки сформированных матриц взаимноортогональны, т.е. их скалярное произведение равно нулю. Эти строки и составляют ансамбль $\Phi = \{\Phi_0, \Phi_1, ..., \Phi_{M-1}\}$ дискретных сигналов Уолша-Адамара $\Phi_i = (\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}})$, где n – размерность сформированной матрицы H_i (в IS-95 использованы H_i с $n = 64$).

Для передачи информации одна из строк $\Phi_i \in \Phi$ матрицы Адамара ставится в соответствие абонентскому каналу, например, для связи между базовой станцией и конкретным абонентом. Модуляция осуществляется по правилу (3), т.е. для передачи информационной "1" посылается строка Φ_i, для "0" – посылается последовательность, сформированная путем логического отрицания Φ_i (ее инверсная копия).

Для выделения сигнала на приемной стороне используется корреляционный приемник, т.е. вычисляется коэффициент корреляции (6). При точном совпадении начала пришедшей последовательности и имеющейся копии Φ_i наблюдаются пики корреляционной функции положительной и отрицательной полярностей – в зависимости от передаваемого бита. То есть, детектирование сигнала происходит следующим образом:

$$x(t) = \begin{cases} \text{"1", при } polarity > 0; \\ \text{"0", при } polarity < 0; \\ \text{сторонний сигнал, при } polarity = 0, \end{cases} \qquad (8)$$

где *polarity* – полярность пика корреляционной функции.

Таким образом, применение ортогональных систем дискретных сигналов Уолша-Адамара позволяет обеспечить высокоэффективную широкополосную цифровую связь. В тоже время число образуемых абонентских каналов связи не может превышать мощности M ансамбля сигналов, в данном случае оно не превышает размерности матрицы H_i, $M = n$. Другими словами, максимальное число возможных ортогональных кодов ограничено их длиной. Для рассмотренного примера имеем $M = 64$ (по спецификации IS-95 образуются 61 абонентский и 3 служебных канала). В этом смысле квазиортогональные дискретные сигналы (с $M > n$) имеют неоспоримое преимущество (см. таблицу 1), их применение потенциально позволит существенно повысить абонентскую мощность системы связи. Кроме того, для рассмотренных сигналов функция взаимной корреляции равна нулю лишь при отсутствии временного сдвига между последовательностями. Как следствие такие сигналы используются лишь в синхронных системах и преимущественно в прямых каналах (от базовой станции к абоненту).

Таким образом, перспективным направлением в развитии современных систем широкополосной связи с прямым расширением спектра является разработка и исследование методов синтеза больших ансамблей квазиортогональных дискретных сигналов с улучшенными ансамблевыми, структурными и корреляционными свойствами.

Рассмотренный подход к организации цифровых помехозащищенных каналов связи нашел применение при построении стеганографических методов защиты информации. Так, в работе [2] расширение спектра прямой последовательностью использовано для создания стеганографического метода встраивания данных в неподвижные изображения. Рассмотрим один из вариантов реализации этого метода, авторами которого являются Смит (J.R. Smith) и Комиски (B.O. Comiskey) [2], проведем исследования его эффективности с точки зрения обеспечиваемой пропускной способности стеганографического канала связи и достигаемой стойкости к несанкционированному извлечению информационных сообщений.

3. ПРЯМОЕ РАСШИРЕНИЕ СПЕКТРА В СТЕГАНОГРАФИИ

В методе Смита-Комиски [2], как и в рассмотренных выше системах связи с прямым расширением спектра, информационное сообщение побитно модулируется путем умножения на ансамбль ортогональных сигналов. Затем промодулированное сообщение встраивается в контейнер – неподвижное изображение.

Введем некоторые условные обозначения и математические соотношения, которые, по аналогии с рассмотренными выше системами широкополосной цифровой связи позволят исследовать особенности построения и информационного обмена данных в стеганосистеме.

Представим информационное сообщение m, подлежащее встраиванию в цифровой контейнер-изображение, в виде блоков m_i равной длины, т.е. $m = (m_0, m_1, ..., m_{N-1})$, где каждый блок m_i – последовательность (вектор) из n бит: $m_i = (m_{i_0}, m_{i_1}, ..., m_{i_{n-1}})$.

Контейнер-изображение будем рассматривать как массив данных C размерностью $K \cdot L$, разбитый на подблоки размером $k \cdot l = n$. В качестве элементов массива C могут выступать, например, растровые данные используемого изображения.

Секретными ключевыми данными является набор базисных функций

$$Key = \Phi = \{\Phi_0, \Phi_1, ..., \Phi_{M-1}\},$$

где все базисные функции $\Phi_i = (\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}})$ – взаимно ортогональные дискретные сигналы с

длиной, равной размеру n блока сообщения m_i, т.е. для любых $i, j \in [0,...,M-1]$ выполняется равенство

$$\rho\left(\Phi_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1} \Phi_{i_z} \Phi_{jz} = \begin{cases} +1, \text{ при } i = j; \\ -1, \text{ при } i \neq j. \end{cases}$$

Формальное графическое представление информационного сообщения, контейнера-изображения и ключевых данных приведено на рис. 2.

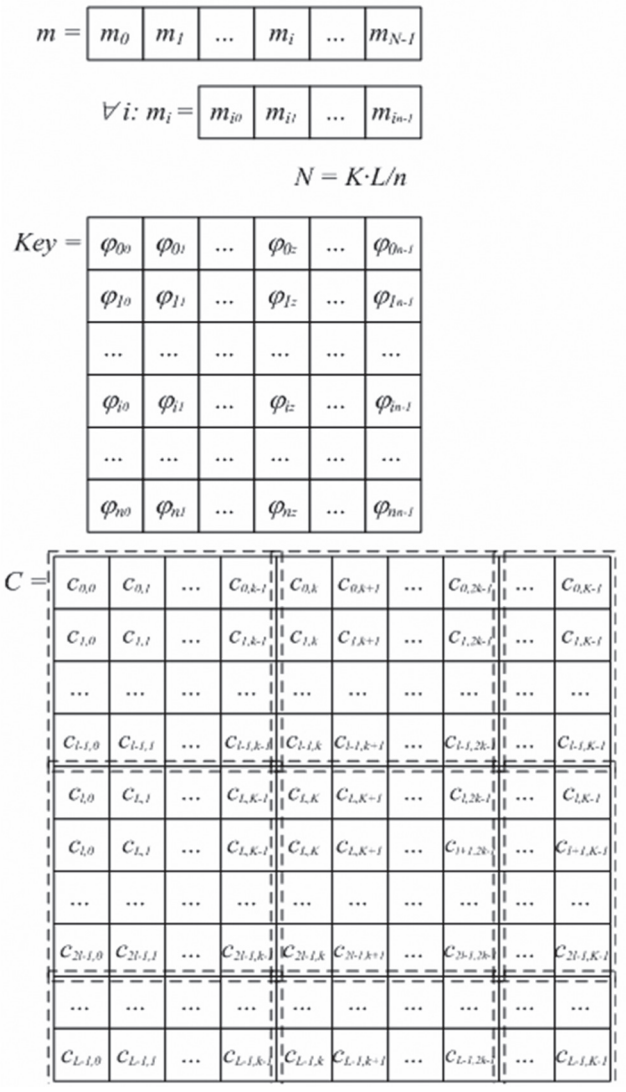

Рис. 2. Формальное представление информационного сообщения, контейнера-изображения и ключевых данных

Целью стеганографического преобразования информации является встраивание каждого отдельного блока сообщения m_i в соответствующий блок контейнера-изображения. В блок данных цифрового изображения размерностью $K \cdot L$ элементов может быть встроено $K \cdot \frac{L}{n}$ блоков информационного сообщения, т.е. до $K \cdot L$ битов.

Разбиение контейнера на блоки может быть произвольным, однако, как показывает практика, наиболее целесообразным (меньший, в отличие от одномерного представления, численный разброс значений в блоке) является двумерное разбиение, приведенное на рис. 2. В качестве ключевых данных (массива базисных функций $Key = \Phi$) будем использовать рассмотренные выше ансамбли ортогональных дискретных сигналов Уолша-Адамара.

Встраивание информационного сообщения осуществляется следующим образом. Каждый блок сообщения $m_{i_j}, j = 0,...,n-1$ сопоставляется с отдельным блоком контейнера-изображения. Каждый информационный бит блока $m_{i_j}, j = 0,...,n-1$ представляется в виде информационного сигнала

$$m_{i_j}(t) = \begin{cases} +1, m_{i_j} = 1; \\ -1, m_{i_j} = 0 \end{cases}$$

и по аналогии с (3) модулируется расширяющим кодовым сигналом (базисными функциями), т.е. ПСП .

В результате, для каждого информационного блока m_i формируется модулированный информационный сигнал

$$E_i(t) = \sum_{j=0}^{n-1}\sum_{z=0}^{n-1} m_{i_j}(t)\Phi_{j_z} . \qquad (9)$$

Полученный блок сообщения E_i попиксельно суммируется с подблоком контейнера.

Обозначим блоки контейнера следующим образом (см. рис. 3):

$$C_0 = \begin{pmatrix} c_{0,0} & c_{0,1} & \dots & c_{0,k-1} \\ c_{1,0} & c_{1,1} & \dots & c_{1,k-1} \\ \dots & \dots & \dots & \dots \\ c_{l-1,0} & c_{l-1,1} & \dots & c_{l-1,k-1} \end{pmatrix},$$

$$C_1 = \begin{pmatrix} c_{0,k} & c_{0,k+1} & \dots & c_{0,2k-1} \\ c_{1,k} & c_{1,k+1} & \dots & c_{1,2k-1} \\ \dots & \dots & \dots & \dots \\ c_{l-1,k} & c_{l-1,k+1} & \dots & c_{l-1,2k-1} \end{pmatrix}, \dots,$$

$$C_{N-1} = \begin{pmatrix} c_{L-l-1,K-k-1} & c_{L-l-1,K-k} & \dots & c_{L-l-1,K-1} \\ c_{L-l,K-k-1} & c_{L-l,K-k} & \dots & c_{L-l,K-1} \\ \dots & \dots & \dots & \dots \\ c_{L-1,K-k-1} & c_{l-1,k+1} & \dots & c_{L-1,K-1} \end{pmatrix}.$$

Соответствующие модулированные информационные сигналы $E_i(t)$ представим в виде двумерного массива данных:

$$E_i = \begin{pmatrix} E_{i_0} & E_{i_1} & \dots & E_{i_{k-1}} \\ E_{i_k} & E_{i_{k+1}} & \dots & E_{i_{2k-1}} \\ \dots & \dots & \dots & \dots \\ E_{i_{(l-1)(k-1)-k+1=n-k+1}} & E_{i_{(l-1)(k-1)-k+2=n-k+2}} & \dots & E_{i_{(l-1)(k-1)=n-1}} \end{pmatrix},$$

$$i = 0,..., N-1 .$$

Тогда стеганограмма (заполненный контейнер) формируется посредством объединения массивов данных S_i, $i = 0,..., N-1$:

$$S_i = C_i + E_i \cdot G, \quad (10)$$

где $G > 0$ – коэффициент усиления расширяющего сигнала, задающий «энергию» встраиваемых бит информационной последовательности.

Таким образом, заполненный контейнер S образуется из сформированных блоков S_i, $i = 0,...,N-1$ посредством их объединения как это показано на рис. 2 для исходного (пустого) контейнера C.

На этапе извлечения данных нет необходимости владеть информацией о первичном контейнере C. Операция декодирования заключается в восстановлении скрытого сообщения путем проецирования каждого блока S_i, полученного стеганоизображения S на все базисные функции $\Phi_j \in \Phi$, $i = 0,...,N-1$. Для этого каждый блок S_i представляется в форме вектора $S_i = \left(S_{i_0}, S_{i_1}, ..., S_{i_{n-1}}\right)$, $i = 0,...,N-1$.

Чтобы извлечь j-ый бит сообщения из i-го блока стеганоизображения, необходимо вычислить коэффициент корреляции между Φ_j и принятым блоком S_i (представленного в виде вектора):

$$\rho\left(S_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1} S_{i_z} \Phi_{j_z} = \\ = G \cdot \frac{1}{n}\sum_{z=0}^{n-1} E_{i_z} \Phi_{j_z} + \frac{1}{n}\sum_{z=0}^{n-1} C_{i_z} \Phi_{j_z}, \quad (11)$$

где под C_i понимается одномерный массив, т.е. соответствующий блок контейнера, представленный в форме вектора.

Предположим, что массив C_i имеет случайную статистическую структуру, т.е. положим, что второе слагаемое в правой части выражения (11) близко к нулю и им можно пренебречь. Тогда имеем:

$$\rho\left(S_i, \Phi_j\right) \approx G \cdot E_i \cdot \Phi_j = G \cdot \sum_{l=0}^{n-1}\sum_{z=0}^{n-1} m_{i_x}(t) \cdot \Phi_{l_z} \Phi_{j_z}. \quad (12)$$

По аналогии с (6) отметим, что все последовательности из множества Φ взаимноортогональны, т.е. при $l \neq j$ имеем $\rho\left(\Phi_l, \Phi_j\right) = 0$. Следовательно, всеми слагаемыми в правой части равенства (12) при $l \neq j$ можно пренебречь. Отсюда имеем:

$$\rho\left(S_i, \Phi_j\right) \approx G \cdot m_{i_j}(t) \cdot \frac{1}{n}\sum_{z=0}^{n-1}\left(\Phi_{j_z}\right)^2 = G \cdot m_{i_j}(t). \quad (13)$$

По аналогии с выделением полезного сигнала (8) значения $m_{i_j}(t)$ могут быть легко восстановлены с помощью знаковой функции.

Поскольку $G > 0$ и $n > 0$ знак $\rho\left(S_i, \Phi_j\right)$ в (13) зависит только от $m_{i_j}(t)$, откуда имеем:

$$m_{i_j}(t) = \operatorname{sign}\left(\rho\left(S_i, \Phi_j\right)\right) = \begin{cases} -1, \text{ при } \rho\left(S_i, \Phi_j\right) < 0; \\ +1, \text{ при } \rho\left(S_i, \Phi_j\right) > 0; \\ ?, \text{ при } \rho\left(S_i, \Phi_j\right) = 0. \end{cases} \quad (14)$$

Если $\rho\left(S_i, \Phi_j\right) = 0$ в (14) будем полагать, что встроенная информация была утрачена.

Структурная схема встраивания информации в контейнер-изображение с использованием прямого расширения спектра для скрытной передачи сообщений представлена на рис. 3.

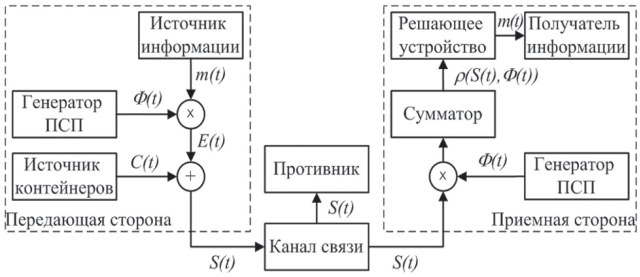

Рис. 3. Структурная схема встраивания информации в контейнер-изображение для скрытной передачи сообщений

Из рисунка следует, что процесс встраивания информационных сообщений для скрытной передачи очень похож на процесс расширения спектра дискретных сигналов в системах связи (см. рис. 1). Поэлементное сложение модулированного сообщения $E(t)$ с контейнером-изображением $C(t)$ (см. выражение (10)) следует интерпретировать как наложение ошибок $e(t)$ на полезный сигнал в канале связи $y(t)$. Задача извлечения сообщения $m(t)$ из $S(t)$ на приемной стороне стеганосистемы эквивалентна задаче детектирования $x(t)$ из смеси полезного сигнала и помехи $y'(t) = y(t) + e(t)$ в широкополосной системе связи. Другими словами, рассмотренная стеганосистема наследует все преимущества широкополосных систем связи: устойчивость к несанкционированному извлечению встроенных сообщений (аналог скрытности в системе связи), устойчивость к разрушению или модификации встроенных сообщений (аналог помехозащищенности), устойчивость к навязыванию ложных сообщений (аналог имитостойкости в системе связи).

Таким образом, использование прямого расширения спектра дискретных сигналов позволяет осуществить встраивание информационных данных в неподвижные изображения для скрытной передачи и реализовать, таким образом, стеганографическую защиту информации.

4. ОЦЕНКА ЭФФЕКТИВНОСТИ СТЕГАНОСИСТЕМЫ

Под эффективностью технической системы в широком смысле понимают соответствие результата выполнения некоторой операции требуемому. При этом техническая система выступает в роли средства реализации исследуемой операции.

Применительно к рассматриваемому процессу стеганографическая система выступает в роли технического средства реализации операции, целью которой является сокрытие от противника факта осуществления скрытной передачи информации. Таким образом, с учетом функционально-

го назначения стеганосистемы, введем следующие показатели эффективности.

1. *Пропускная способность* – отношение объема V встраиваемой в контейнер информации к общему объему D контейнера

$$Q = \frac{V}{D}. \qquad (15)$$

2. *Объем ключевых данных* (в битах)

$$l_{Key} = \log_2\left(|Key|\right), \qquad (16)$$

где $|Key|$ – мощность множества ключевых данных.

3. *Стойкость стеганографического метода* будем оценивать как величину, обратную мощности множества секретных ключевых данных. Ее можно трактовать как вероятностный показатель подбора секретного ключа:

$$W = \frac{1}{|Key|} = 2^{-l_{Key}}. \qquad (17)$$

4. *Величина вносимых искажений* как процентное отношение среднеарифметического всех абсолютных значений Δ-изменений данных контейнера к максимально возможному значению $\Delta_{\max}$:

$$I = \frac{\Delta_{cp}}{\Delta_{\max}} \cdot 100 = \frac{100}{\Delta_{\max} \cdot D} \cdot \sum_{i=1}^{D} |\Delta_i|, \qquad (18)$$

где Δ_i – Δ-изменения i-го элемента контейнера.

5. *Вероятность ошибочного извлечения* информационных данных сообщения

$$P_{ош} = \lim_{D\to\infty} \frac{V_{ош}}{D} = 1 - \lim_{D\to\infty} \frac{V - V_{ош}}{D}, \qquad (19)$$

где $V_{ош}$ – объем ошибочно извлеченных данных.

Используя показатели (15) – (19), оценим эффективность рассмотренного стеганографического метода защиты информации.

1. *Пропускная способность.* На каждый n-элементный блок S_i заполненного контейнера (стеганограммы) приходится n-битный вектор встроенного сообщения m_i (см. выражения (9), (10)). Следовательно, $Q = \frac{1}{B}$, где B – объем данных, приходящийся на один элемент контейнера. Для случая встраивания в растровые данные изображения (цветовая модель R,G,B) с 8-битным кодированием каждого цвета имеем $B = 8$ и $Q = \frac{1}{8}$.

2. *Объем ключевых данных.* Ключевыми данными является ансамбль дискретных сигналов, образованный строками матрицы Адамара порядка n. Следовательно, под множеством ключевых данных следует понимать множество различных (неизоморфных) матриц Адамара, каждая из матриц задает ансамбль дискретных сигналов. В [7] получены некоторые оценки мощности M_A этого множества, которые приведены в табл. 2.

Таблица 2

Число ансамблей дискретных сигналов Уолша-Адамара [7]

n	M_A
64	19
100	1
256	54
512	102
1024	162
2000	9
4000	16
10000	10

Приведенные оценки мощности M_A дают оценку числа ансамблей дискретных сигналов Уолша-Адамара, т.е. оценку мощности неэквивалентных ключей стеганосистемы. Следовательно, объем ключевых данных оценивается как $l_{key} = \log_2(M_A)$.

3. *Вероятность подбора* секретного ключа $W = (M_A)^{-1}$.

4. Для оценки *величины вносимых искажений* воспользуемся выражением (10). Второе слагаемое в правой части (10) определяет величину Δ-изменений элементов данных контейнера. Сомножитель E_i формируется в результате суммирования n дискретных сигналов (принимающих значения ± 1) с соответствующими полярностями (задаваемыми $m_{i_j}(t)$). Следовательно, все элементы E_i будут принимать значения из диапазона $[-n,...,+n]$, а соответствующие Δ-изменения элементов контейнера не будут превышать $|\Delta_i| \le n \cdot G$. Откуда имеем верхнюю оценку величины вносимых искажений:

$$I = \frac{\Delta_{cp}}{\Delta_{\max}} \cdot 100 \le \frac{n \cdot G}{\Delta_{\max}} \cdot 100. \qquad (20)$$

Для случая встраивания в растровые данные изображения (цветовая модель R,G,B) с 8 битным кодированием каждого цвета и использования дискретных сигналов с $n = 256$ даже при $G = 1$ вносимые искажения могут достигать 100%. Снизить вносимые искажения можно за счет сокращения числа встраиваемых бит данных m_{i_j} (уменьшив число слагаемых в (9)), что неизбежно приведет к снижению пропускной способности стеганографического канала связи.

5. *Вероятность ошибочного извлечения.* Извлечение информационного сообщения, также как и при организации помехозащищенной связи (см. (3) - (5)), осуществляется корреляционным способом (см. (11) – (14)). Следовательно, ошибка извлечения произойдет при изменении знака коэффициента корреляции $\rho(S_i, \Phi_j)$ в выражении (14).

Представим коэффициент $\rho(S_i, \Phi_j)$ в виде

$$\rho(S_i, \Phi_j) = \rho(C_i + E_i \cdot G, \Phi_j) = \rho(C_i, \Phi_j) + \rho(E_i \cdot G, \Phi_j).$$

Последнее слагаемое не изменяет знак $\rho(S_i,\Phi_j)$, событие

$$\rho(S_i,\Phi_j)=\rho(E_i\cdot G,\Phi_j)$$

соответствует безошибочному извлечению сообщения (см. (12), (13)).

Следовательно, ошибка извлечения информационного бита m_{i_j} сообщения произойдет при наступлении события

$$\left|\rho(C_i,\Phi_j)\right|>\rho\left|E_i\cdot G,\Phi_j\right|=\left|G\cdot m_{i_j}\right|=G\,, \qquad (21)$$

т.е. в том случае, когда абсолютное значение коэффициента корреляции используемого для встраивания бита m_{i_j} дискретного сигнала Φ_j с блоком контейнера C_i, в который этот бит встраивается, превзойдет коэффициент усиления G.

Таким образом, запишем

$$P_{ош}=P\left(\left|\rho(C_i,\Phi_j)\right|>G\right),$$

где $P(x)$ – вероятность наступления случайного события x.

Другими словами, правильное извлечение встроенного сообщения является случайным событием, вероятность $P_{б.ош}$ которого непосредственно связана со статистическими свойствами используемого контейнера-изображения. Для безошибочного извлечения сообщения

$$P_{ош}=0, P_{б.ош}=1-P_{ош}=1, \qquad (22)$$

следует стремиться к взаимной ортогональности отдельных фрагментов изображения C_i и используемых в качестве секретных ключей дискретных сигналов Φ_j. В этом случае событие

$$\left|\rho(C_i,\Phi_j)\right|=0<G\,,$$

для всех $i=0,...,N-1$ является достоверным и выполняется (22).

В тоже время, как показали экспериментальные исследования, коэффициент корреляции, как правило, значительно больше нуля $\left|\rho(C_i,\Phi_j)\right|>>0$ и очень часто возникает событие (21). Дело в том, что элементы дискретных сигналов $\Phi_j\in\Phi$ принимают значения $\begin{cases}+1\\-1\end{cases}$, а соответствующий нормированный коэффициент корреляции $\rho\left(\Phi_i,\Phi_j\right)$ по абсолютному значению не превосходит длины n последовательности (см. (2)) и лежит в диапазоне $[0,...,1]$, откуда собственно и следует условие (21).

Однако элементы контейнера C_i принимают значение из числового поля $[0,...,Y]$, размерность которого задается способом кодирования данных изображения. Например, при встраивании информации в растровые данные изображения (цветовая модель R,G,B) с 8 битным кодированием каждого цвета соответствующие C_i принимают значения из диапазона целых чисел $[0,...,255]$. Другими словами, абсолютное значение нормированного относительно n коэффициента корреляции $\left|\rho(C_i,\Phi_j)\right|$ будет лежать в диапазоне $[0,...,Y]$ и для безошибочного извлечения всех битов сообщения (21) необходимо выполнить условие $G>Y$.

В тоже время повышение G ведет к неизбежному росту величины вносимых искажений (20), которые при $I>2...3\%$ (порог зрительной чувствительности человека) становятся заметны стороннему наблюдателю [1, 2], что компрометирует стеганоканал и делают невозможным использование рассмотренной стеганосистемы.

Таким образом, в ходе исследований выявлены следующие противоречия, лежащие в основе разработки и использования стеганографических систем с расширением спектра дискретных сигналов:

– вероятность правильного извлечения встроенных данных $P_{б.ош}$ лежит в прямой зависимости от величины вносимых искажений I;

– величина вносимых искажений I лежит в прямой зависимости от объема встраиваемых бит данных, т.е. от пропускной способности стеганоканала Q;

– вероятность правильного извлечения встроенных данных $P_{б.ош}$ непосредственно зависит от статистических свойств используемого контейнера-изображения.

Для экспериментального исследования эффективности рассмотренного метода встраивания сообщений в неподвижные изображения разработана программная реализация, получены следующие эмпирические оценки:

– зависимости величины вносимых искажений I от пропускной способности Q стеганоканала;

– зависимости величины вносимых искажений I и частоты ошибок извлечения $P^*_{ош}\approx P_{ош}$ от коэффициента усиления G;

– зависимости величины вносимых искажений I от частоты ошибок извлечения $P^*_{ош}\approx P_{ош}$.

Исследования проводились при встраивании информационных данных в растровые данные изображения (цветовая модель R,G,B) с 8 битным кодированием каждого цвета. Полученные эмпирические зависимости приведены на рис. 4–7.

Анализ экспериментально полученных зависимостей подтверждает сделанные ранее выводы, сходимость результатов эксперимента с теоретическими рассуждениями свидетельствует о достоверности полученных результатов.

Из приведенной на рис. 4. зависимости следует, что повышение пропускной способности стеганоканала ведет к резкому увеличению вносимых искажений в контейнер-изображение. Незаметные для стороннего наблюдателя искажения (лежащие ниже порога чувствительности зрительной системы человека) вносятся лишь при $Q\le 0{,}005$. Это соответствует встраиванию не более 10 битов в один блок изображения, т.е. модулированию до десяти информационных сигналов $m_{i_j}(t)$, $j=0,...,9$ в выражении (9).

Зависимости, приведенные на рис. 5, 6, свидетельствуют, что коэффициент усиления, используемый в выражениях (10) – (13), позволяет существенно снизить вероятность ошибочного извлечения информационных данных. К сожалению, это достигается за счет резкого повышения вносимых искажений в используемый контейнер-изображение. Зависимости получены при $Q = 0{,}005$. Очевидно, что для такой величины пропускной способности коэффициент усиления не может превосходить 1 .. 1,5 (см. рис. 5). Однако даже для таких значений вероятность ошибочного извлечения велика и лежит в диапазоне 0,1 .. 0,5.

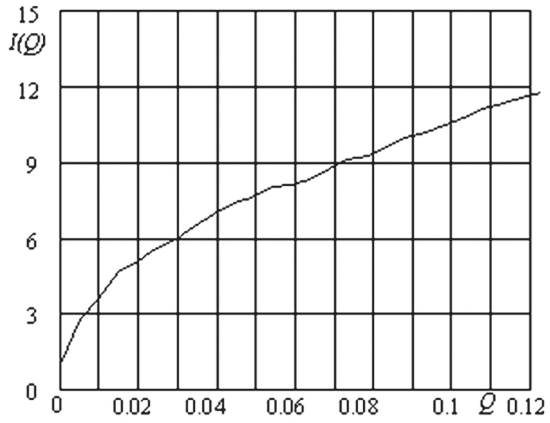

Рис. 4. Зависимость $I(Q)$

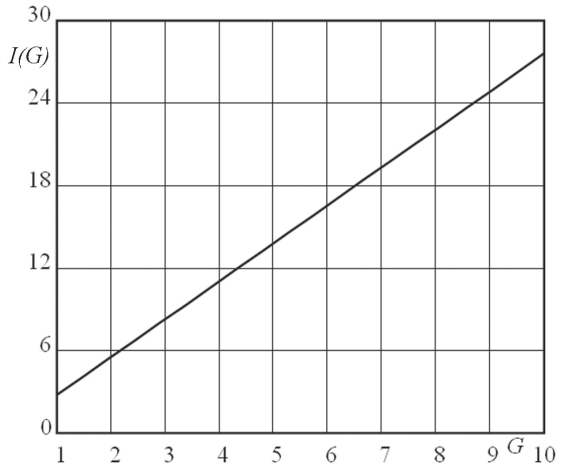

Рис. 5. Зависимость $I(G)$ при $Q = 0{,}005$

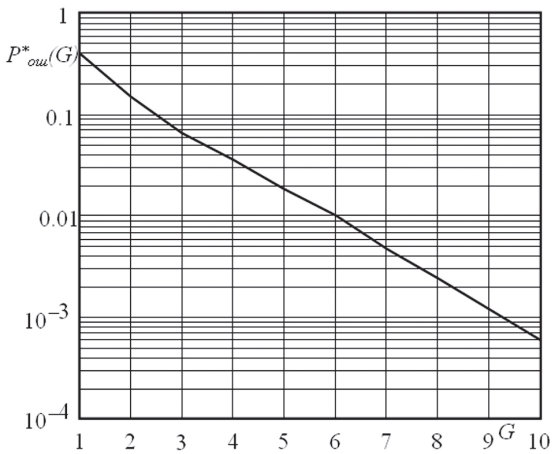

Рис. 6. Зависимость $P^*_{ош}(G)$ при $Q = 0{,}005$

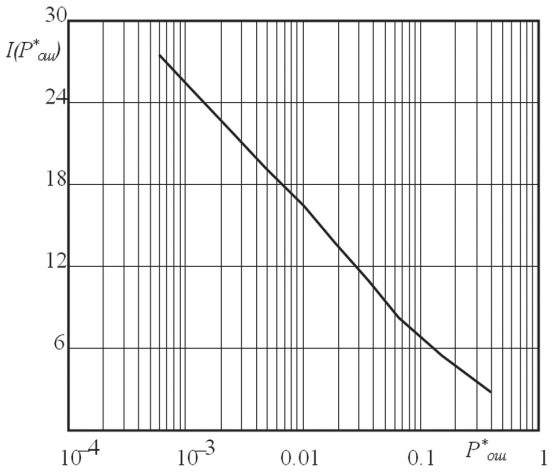

Рис. 7. Зависимость $I(P^*_{ош})$ при $Q = 0{,}005$

Интегральная зависимость $I(P^*_{ош})$, приведенная на рис. 6, обобщает приведенные на рис. 5, 6 данные. Для фиксированной пропускной способности $Q = 0{,}005$ получена эмпирическая кривая, характеризующая зависимость величины вносимых искажений в контейнер-изображение и вероятности ошибочного извлечения информационных данных. Для $Q = 0{,}005$ добиться низких искажений, лежащих ниже порога зрительной чувствительности человека ($I \le 2...3\%$), можно только при очень высокой вероятности ошибочного извлечения информационных данных ($P_{ош} \ge 0{,}1$). Очевидно, что практическое применение подобных стеганосистем необходимо сочетать с помехоустойчивым кодированием информационных данных, что позволит существенно снизить $P_{ош}$.

ВЫВОДЫ

В результате проведенных исследований показано, что использование в стеганографических целях прямого расширения спектра дискретных сигналов позволяет осуществить скрытное встраивание информационных сообщений в неподвижные изображения. Задача извлечения сообщения на приемной стороне стеганосистемы эквивалентна задаче обнаружения информации из смеси полезного сигнала и помехи в широкополосной системе связи.

В ходе исследований выявлены следующие недостатки стеганографических систем с расширением спектра дискретных сигналов: вероятность правильного извлечения встроенных данных зависит от величины вносимых искажений, которая в свою очередь зависит от обеспечиваемой пропускной способности стеганоканала. Иначе говоря, практическое построение стеганосистемы сопряжено с поиском компромисса между величиной вносимых искажений, вероятностью правильного извлечения сообщения на приемной стороне и обеспечиваемой пропускной способностью. Кроме того, в ходе исследований установлено, что вероятность правильного извлечения встроенных данных непосредственно

зависит от статистических свойств используемого контейнера-изображения.

Перспективным направлением дальнейших исследований, по мнению авторов, является использование больших ансамблей слабокоррелированных (квазиортогональных) дискретных сигналов для построения стеганосистем с прямым расширением спектра. Это позволит, с одной стороны, без значительного повышения вносимых искажений в контейнер-изображение существенно повысить пропускную способность стеганоканала. С другой стороны, за счет адаптивного формирования (выбора) дискретных сигналов по критерию минимизации коэффициента корреляции с контейнером изображением это позволит существенно снизить вероятность ошибочного извлечения встроенных данных.

Литература.

[1] *Конахович Г.Ф., Пузыренко А.Ю.* Компьютерная стеганография. Теория и практика. – К.: «МК-Пресс», 2006. – 288 с., ил.

[2] *J. Smith, B. Comiskey,* Modulation and Information hiding in Image. // Information hiding: First Int. Workshop "InfoHiding'96", Springer as Lecture Notes in Computing Science, vol 1174. 1996. – pp. 207-227.

[3] Цифровые методы в космической связи. /Под ред. С. Голомба.- М.: Связь, 1969. – 272 с.

[4] *Скляр Б.* Цифровая связь. Теоретические основы и практическое применение. – М.: Вильямс, 2003. – 1104 с.

[5] *Горбенко И.Д., Стасев Ю.В.* Анализ производных ортогональных систем сигналов // Радиотехника. – 1989. – № 9. – С. 16 – 18.

[6] *Стасев Ю.В., Кузнецов А.А., Носик А.М., Качур Л.Н.* Формирование больших ансамблей дискретных сигналов с использованием избыточных кодов // Збірник наукових праць ХУПС. – Харків: ХУПС. – 2008. – Вип. 2 (17). – С. 102-109.

[7] *Стасєв Ю.В.* Основи теорії побудови сигналів. – Х.: ХВУ, 1999. – 87с.

Поступила в редколлегию 5.07.2010.

Кузнецов Александр Александрович, доктор технических наук, профессор, профессор кафедры БИТ ХНУРЭ. Область научных интересов: криптография, теория обработки и передачи данных, стеганографические методы защиты информации.

Ботнов Антон Михайлович, студент 5 курса факультета компьютерной инженерии и управления ХНУРЭ. Область научных интересов: методы обработки и передачи данных, стеганографические методы защиты информации.

Лаптий Павел Александрович, студент 5 курса факультета компьютерной инженерии и управления ХНУРЭ. Область научных интересов: методы обработки и передачи данных, стеганографические методы защиты информации

УДК 519:616-079.4:616.5

Вбудовування інформаційних даних в нерухомі зображення з використанням прямого розширення спектру / О.О. Кузнєцов, д.т.н., проф., А.М. Ботнов, П.О. Лаптій / / Прикладна радіоелектроніка: наук.-техн. журнал. – 2010. Том 9. № 3. – С. 470-478.

У статті досліджуються стеганографічні методи захисту інформації на основі прямого розширення спектра дискретних сигналів. Пропонується стеганографічний метод приховування даних в нерухоме зображення з використанням квазіортогональних дискретних сигналів. Показано, що запропонований метод дозволяє підвищити пропускну здатність стеганографічного каналу зв'язку та зменшити частку спотворень, що вносяться в контейнер-зображення.

Ключові слова: метод розширення спектру, дискретний сигнал, кореляційний прийом, стеганографічний захист інформації.

Табл. 02. Іл. 06. Бібліогр.: 7 найм.

UDC 519:616-079.4:616.5

Data embedding in stationary images by using direct spectrum expansion / A.A Kuznetsov, A.M Botnov, P.A. Laptii // Applied Radio Electronics: Sci. Mag. – 2010. Vol. 9. № 3. – P. 470-478.

Stenographic methods of embedding data into stationary images for secretive information transmission are considered. The paper considers the method of stenographic protection based upon the use of the direct expansion of a spectrum of discrete signals, investigates, its efficiency from the point of view of provided strength, capacity and magnitude of distortions entered in a container-image.

Key words: spectrum expansion method, discrete signal, correlation reception, steganographic data protection.

Tab. 02. Fig. 06. Ref.: 7 items.