

*В.И. ХАХАНОВ, С.В. ЧУМАЧЕНКО, А.С. МИЩЕНКО, А.В. ЗАЦАРНЫЙ,  
Ю.В. ХАХАНОВА*

## **ИНФРАСТРУКТУРА ДИАГНОСТИРОВАНИЯ ВРЕДНОСНЫХ ПРОГРАММ В ИНДИВИДУАЛЬНОМ КИБЕРНЕТИЧЕСКОМ ПРОСТРАНСТВЕ**

---

Предлагаются методы и архитектуры параллельного анализа информации, представленной в виде аналитических, графовых и табличных форм ассоциативных отношений для поиска, распознавания, диагностирования деструктивных компонентов и принятия решений в  $n$ -мерном векторном кибернетическом индивидуальном пространстве. Рассматриваются векторно-логические процесс-модели актуальных прикладных задач, в том числе – диагностирование спама и восстановление работоспособности программно-аппаратных компонентов компьютерных систем, качество решения которых оценивается неарифметической метрикой взаимодействия булевых векторов. Предлагается концепция саморазвития информационно-компьютерной экосистемы, повторяющая эволюцию функциональностей человека. Представляются оригинальные процесс-модели ассоциативно-логического анализа информации на основе быстродействующего мультипроцессора в  $n$ -мерном векторном дискретном пространстве. Разрабатывается алгебраическая структура, регулирующая векторно-матричные преобразования в дискретном булевом пространстве для анализа информации на основе логических операций над ассоциативными данными. Рассматривается применение алгебры векторной логики для создания процесс-моделей актуальных прикладных задач, качество решения которых оценивается интегральной неарифметической метрикой взаимодействия ассоциативных структур. Состоятельность теории, методов, структур данных и процесс-моделей подтверждается рассмотрением многочисленных примеров параллельного анализа, синтеза, поиска и принятия решений в информационном пространстве.

### **Введение**

Проблема создания эффективной инфраструктуры кибернетического пространства (Cyber Space), а также саморазвивающейся информационно-компьютерной экосистемы (ИКЭС) планеты особенно важна для глобальных компаний, таких как Лаборатория Касперского, Google, Microsoft.

Кибернетическое пространство как объект природы также подвержено влиянию деструктивных компонентов, влияющих на работоспособность субъектов, которыми являются компьютеры, системы и сети. Поэтому сейчас и в будущем важной проблемой остается стандартизация пространства и специализация всех взаимодействующих субъектов, включая негативные, как неотъемлемую часть экосистемы. Данная акция есть постоянно действующая во времени, цель которой – не отставать, но на один шаг опережать появление новых вредоносных компонентов, путем создания инфраструктуры кибернетического пространства, обеспечивающей функционирование компьютерной экосистемы планеты и качество жизни каждого человека.

Среди модулей такой инфраструктуры можно выделить диагностирование отказов и спама путем анализа информации, полученной на стадии тестирования и использования специальных методов встроенного поиска спама на основе стандартов граничного сканирования или ассерционной избыточности, ориентированной на обнаружение спама, что позволит идентифицировать и устранять его без применения внешних средств. Таким образом, можно будет обходиться без сложных внешних программ моделирования, тестирования и диагностирования путем прививки каждого электронного письма тестопригодной интеллектуальной избыточностью кода на стадии его создания. При этом следует использовать предикат узнавания, который оперирует не только булевыми, но регистровыми и матричными переменными, что делает его практически значимым в формальной записи уравнений диагноза или распознавания:

$$x^a \approx x \oplus a = 0 \vee \min_i Q_i \rightarrow x \oplus a \oplus Q = 0;$$

$$x^m \approx x \oplus m = 0 \vee \min_i Q_i \rightarrow x \oplus m \oplus Q = 0;$$

$$T \oplus S = Q \approx \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} \begin{bmatrix} 1 & . & 1 & 1 \\ . & 1 & . & . \\ . & 1 & . & 1 \\ 1 & . & . & 1 \end{bmatrix} \Delta \begin{bmatrix} 1 & . & 1 & 1 \\ . & 1 & 1 & 1 \\ . & . & . & . \\ . & . & . & 1 \end{bmatrix} = \begin{bmatrix} . & . & . & . \\ . & . & 1 & 1 \\ . & 1 & . & 1 \\ 1 & . & . & . \end{bmatrix},$$

$x^a$ ,  $x^m$  – предикатные переменные;  $a$ ,  $m$  – значения переменных;  $Q_i$  – оценка узнавания значения переменной;  $T$  – тест;  $S$  – объект, подлежащий тестированию (программа).

На основе предиката узнавания  $m$ -образа любой сложности, природы и формы можно создавать достаточно компактные уравнения предикатов, формирующие интеллектуальные решения в области распознавания образов, принятий решений, тестирования знаний и технических объектов, диагностирования (узнавания) спама среди электронных писем.

В связи с этим предложенная инфраструктура кибернетического пространства, метрика его измерения и процесс-модели анализа и синтеза субъектов дают возможность создавать эффективные решения компьютерных изделий, ориентированных на быстрый поиск, распознавание, диагностирование не только позитивных, но и негативных субъектов. Конкретно, предложенная инфраструктура может решать задачи: 1) Описание многообразия электронного мусора кибернетического пространства. 2) Формализация процессов взаимодействия триады компонентов <программа, спамность, тесты>. 3) Диагностирование и фильтрация электронных писем. 4) Создание и эффективное использование базы спама. 5) Создание быстродействующих интеллектуальных саморазвивающихся средств сервисного обслуживания и защиты кибернетического пространства.

Сегодня чрезвычайно важно обозначить возможные пути для решения проблемы создания инфраструктуры саморазвивающейся компьютерной экосистемы. Эволюция ИКЭС основывается на использовании трех наиболее важных компонентов: фантазия, математика и технология, где субъектом экосистемы выступает саморазвивающийся компьютер (СРК). Основное отличие СРК от современного компьютера заключается в концепции жизненного цикла. Стратегия настоящего компьютера есть обучение или повторение уже пройденного пути. Принципиальная позиция СРК – постоянный поиск новых путей для самосовершенствования (эволюционирования) на основе мирового опыта, скрытого в информационном пространстве. Согласно запрету Геделя, адаптированному для информационного пространства, нельзя создать компьютер, который способен решать любые задачи, формально представленные спецификацией. Тем не менее, принцип Геделя предоставляет методологическую основу эволюции (саморазвития) ИКЭС, которую можно интерпретировать следующим образом. Для информационного пространства всегда можно придумать такую полезную спецификацию, которая не покрывается существующими у человечества решениями, что обуславливает создание нового функционального или технологического компонента для его последующего включения в планетарную библиотеку. ИКЭС имеет возможность повторить эволюцию человечества, только в тысячи раз более быстрыми темпами. Здесь заложены основные принципы эволюции, явно выраженные уже в современной компьютерной индустрии: 1) Стандартизация – самое главное для эволюции и жизненного цикла СРК – рынок не принимает и не понимает нестандартных по интерфейсу решений. 2) Специализация есть повышение эффективности предоставляемых (персонально ориентированных) сервисов изделия, связанных с быстродействием, качеством, затратами, энергосбережением путем оптимизации структуры и функциональных компонентов, покрывающих спецификацию. 3) Повсеместное использование векторно-логического критерия качества решения в задачах генерирования идей, синтеза и анализа. Генерирование – процесс создания новой функциональности. При этом синтез оперирует существующими в информационном пространстве компонентами для создания структуры. Анализ – оценивание полученного решения. 4) Диаграмма Хассе используется для выработки стратегии

оптимизации покрытия функциональностей спецификации библиотечными компонентами или их сочетаниями, принадлежащими информационному пространству. Она согласуется с современной Y-технологией, входящей в состав ESL Design, суть которой – использовать библиотечные компоненты на всех уровнях проектирования изделия для покрытия специфицируемой функциональности в процессе синтеза.

На рис. 1 представлен замкнутый цикл эволюции ИКЭС, который фактически изоморфен спирали развития человечества, накрученной на временную ось.

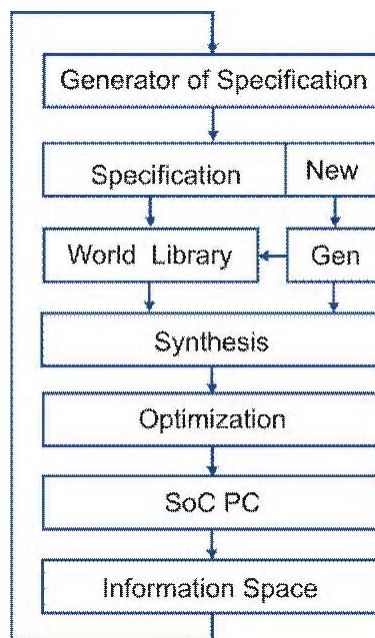


Рис. 1. Цикл ИКЭС

*Целью* данного исследования является существенное повышение качества индивидуального кибернетического пространства (ИКП) пользователя и уменьшение стоимости эксплуатационных расходов за счет вакцинации ИКП путем добавления в пространство инфраструктуры сервисного обслуживания, включающей библиотеки позитивных и негативных сообщений и обеспечивающей тестирование, диагностирование и устранение вредоносных компонентов электронных писем.

*Объект исследования* – индивидуальное кибернетическое пространство, представленное информацией, ее носителями и преобразователями, а также деструктивными компонентами, наносящими вред функциональностям, улучшающим качество жизни человека.

*Предмет исследования* – инфраструктура сервисного обслуживания, включающая библиотеки позитивных и негативных сообщений и встроенную программную избыточность, которая работает в реальном масштабе времени, обеспечивает тестирование, диагностирование и устранение вредоносной и «мусорной» информации электронных писем, описанных в соответствующих библиотеках.

*Мотивация:* 1) Отсутствие на рынке антиспамной защиты встроенных средств тестирования, диагностирования и удаления вредоносных компонентов, составляющих инфраструктуру сервисного обслуживания, подобной тому, как в цифровых системах на кристаллах существуют стандарты граничного сканирования, а в программных продуктах – асерционная избыточность, ориентированные на встроенное тестирование дефектов и ошибок с последующим восстановлением работоспособности аппаратных или программных изделий. 2) Наличие теоретических разработок, связанных с технологией алгебрологического векторного анализа информационных данных, ориентированных на высокое быстродействие решения и оценивания задач распознавания, образов, принятия решений и тестирования объектов. 3) Наличие образцовой производственной и маркетинговой инфраструктуры Лаборатории Касперского, способной поддержать проект создания технологии вакцинации электронных сообщений и авторитетно предложить его рынку информационных технологий.

4) Миниатюризация цифровых и телекоммуникационных систем (телефоны смартфоны, планшеты) требует постоянной защиты от массовых и ненужных электронных писем путем внедрения встроенных антиспамных средств, контролирующих информационный обмен.

*Задачи:* 1) Разработка математического аппарата анализа кибернетического пространства, ориентированного на создание моделей и методов сервисного обслуживания программных продуктов для тестирования, диагностирования и устранения массовых и ненужных электронных писем. 2) Создание типовых процесс-моделей и критериев взаимодействия электронных писем с контентным анализом полезных функциональностей. 3) Разработка технологии анализа структуры программного кода для определения критических точек и установки в них ассерционных операторов наблюдения и управления в процессе его функционирования. 4) Создание инфраструктуры сервисного обслуживания функциональных программ для встроенного тестирования, диагностирования и устранения вредоносных компонентов из программного кода функциональности на основе использования библиотеки спамной информации. 5) Тестирование и верификация встроенной инфраструктуры сервисного обслуживания функциональностей, защищающей программный код от вредоносных компонентов.

### 1. Эволюция Cyber Space и Internet

Для создания схемы, реализующей полезную функциональность, следует генерировать примитивы  $P = \{P_1, P_2, \dots, P_i, \dots, P_n\}$  самого нижнего уровня. Для этого необходимо создавать фильтры  $F = \{F_1, F_2, \dots, F_j, \dots, F_m\}$ , формирующие таблицы примитивных отношений, взятых из информационного пространства планеты (рис. 2). Имея стандартизированные структуры данных для отдельных порталов и браузеров, доставляющих новые сервисы с более высоким быстродействием, следует ожидать постепенного качественного улучшения всех компонентов Cyber Space. Конечная цель такого взаимного и положительного влияния элементов инфраструктуры кибернетического пространства – выработка единых стандартов по интерфейсам и его превращение в саморазвивающуюся интеллектуальную информационно-компьютерную экосистему. Существенное значение будут иметь первичные фильтры или преобразователи для создания новых стандартизированных примитивов, создающих технологичную инфраструктуру для скоростного драйва по Cyber Space с использованием специализированного неарифметического двигателя (I-Computer). Со временем аморфная, или «мусорная» часть Internet будет уменьшаться, а стандартизированная инфраструктура – увеличиваться. К 2020 году информационное пространство планеты должно принять цивилизованные форматы структур данных со стандартизированными интерфейсами, подобно тому, как это происходило с развитием планетарной инфраструктуры транспортных сообщений с терминалами, отелями, заправками, ориентированными на сервисы, удовлетворяющие любые запросы пользователя.

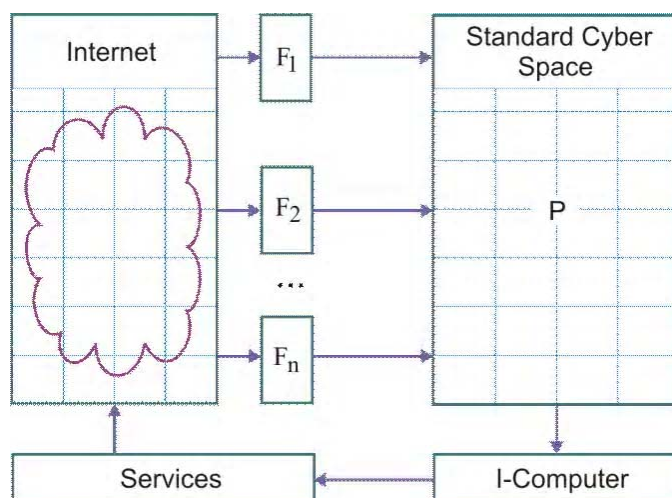


Рис. 2. Эволюция Cyber Space и Internet

В настоящее время отсутствуют стандарты формирования и хранения структур одно-типных данных на сайтах и порталах Internet. Это затрудняет понимание информации поисковыми системами для распознавания и принятия правильных решений. Трудно ожидать, что информационное пространство в одночасье перейдет на рельсы принятых всеми форматов и интерфейсов. Единственным выходом может служить эволюционирование структур данных. Для этого необходимо разрабатывать преобразователи (фильтры) форматов данных. Наличие первичных фильтров позволяет автоматизировать времязатратные процессы создания библиотек базовых примитивов. Например, нетрудно создать генераторы логических отношений (фильтры, см. рис. 2) для каждой части речи русского языка, ориентированные на анализ существующих электронных словарей или признанных эталонных текстов в целях формирования первоначальных таблиц отношений в виде <часть речи> <словообразующие компоненты>. Имея такого рода базовые примитивы, можно автоматически создавать более сложные конструкции и процесс-модели анализа, синтеза предложений, коррекции ошибок, что всегда будет актуально и востребовано на рынке информационно-компьютерных сервисов.

Имея спецификацию, представленную после обработки вербального описания в форме вектора входных и выходных переменных, нетрудно записать стратегию создания новой функциональности как задачу поиска покрытия библиотечными элементами обобщенного вектора  $\langle X, Y \rangle$ . Общее решение задачи похоже на синтез автоматной модели, определяющей взаимодействие компонентов во времени и в пространстве. Однако многообразие примитивов, заранее не заданных, исключает такую возможность, что означает – необходим переход из строгого детерминизма цифровых автоматов в область эволюционных и квазиоптимальных решений.

Условие задачи: имеется спецификация в виде вектора существенных переменных, которые необходимо покрыть минимальным множеством примитивов из библиотеки и сформировать выходной вектор. Красивое решение задачи синтеза функциональной структуры по спецификации есть ключ к самостоятельному генерированию компьютером новых решений. После этого остается решить только две задачи на пути к созданию интеллекта компьютера – самогенерирование оригинальной функциональности, необходимой для решения задачи покрытия, и спецификаций новых полезных для человека или компьютера сервисов.

## 2. Интегральная метрика оценивания диагноза

Инфраструктура мозгоподобных алгоритмов для обнаружения спама включает модели, методы и ассоциативно-логические структуры данных, ориентированные на поддержку процессов поиска, распознавания и принятия решений на основе векторных неарифметических операций. Оценка решения задачи определяется векторно-логическим критерием качества взаимодействия запроса (вектора  $m$ ) с системой ассоциативных векторов (ассоциаторов), в результате которого сгенерируется конструктивный ответ в виде одного или нескольких ассоциаторов, а также, пока еще, численной характеристики степени принадлежности (функции качества) входного вектора  $m$  к найденному решению:  $\mu(m \in A)$ . Входной вектор  $m = (m_1, m_2, \dots, m_i, \dots, m_q)$ ,  $m_i \in \{0, 1, x\}$  и матрица  $A_i$  ассоциаторов  $A_{ijr} (\in A_{ij} \in A_i \in A) = \{0, 1, x\}$  имеют одинаковую размерность, равную  $q$ . Далее степень принадлежности  $m$ -вектора к вектору  $A$  будет обозначаться как  $\mu(m \in A)$ .

Существует 5 типов теоретико-множественного (логического)  $\Delta$ -взаимодействия двух векторов  $m \cap A$ , определенных на рис. 3. Они формируют все примитивные варианты реакции обобщенной ПРП-системы (Поиска, Распознавания и Принятия решения) на входной вектор-запрос. В технологической отрасли знаний – технической диагностике (Design & Test) – указанная последовательность действий изоморфна маршруту: поиск дефектов, их распознавание, принятие решения на восстановление работоспособности. Все три стадии технологического маршрута нуждаются в метрике оценивания решений для выбора оптимального варианта.

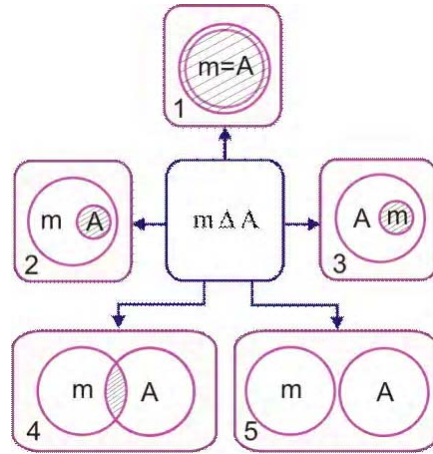


Рис. 3. Результаты пересечения двух векторов

**Определение.** Интегральная теоретико-множественная метрика для оценивания качества запроса есть функция взаимодействия многозначных векторов  $m \cap A$ , которая определяется средней суммой трех нормированных параметров: кодовое расстояние  $d(m, A)$ , функция принадлежности  $\mu(m \in A)$  и функция принадлежности  $\mu(A \in m)$ :

$$Q = \frac{1}{3}[d(m, A) + \mu(m \in A) + \mu(A \in m)],$$

$$d(m, A) = \frac{1}{n}[n - \text{card}(m_i \cap A_i = \emptyset)];$$

$$\mu(m \in A) = 2^{\text{card}(m \cap A) - \text{card}(A)} \leftarrow \text{card}(m \cap A) =$$

$$= \text{card}(m_i \cap A_i = x) \& \text{card}(A) = \text{card}(\bigcup_{i=1}^n A_i = x);$$

$$\mu(A \in m) = 2^{\text{card}(m \cap A) - \text{card}(m)} \leftarrow \text{card}(m \cap A) =$$

$$= \text{card}(m_i \cap A_i = x) \& \text{card}(m) = \text{card}(\bigcup_{i=1}^n m_i = x).$$
(1)

Пояснения. Нормирование параметров позволяет оценить уровень взаимодействия векторов в интервале  $[0, 1]$ . Если зафиксировано предельное максимальное значение каждого параметра, равное 1, то векторы равны между собой. Минимальная оценка,  $Q = 0$ , фиксируется в случае полного несовпадения векторов по всем  $n$  координатам. Если мощность пересечения  $m \cap A = m$  равна половине пространства вектора  $A$ , то функции принадлежности и качества соответственно равны:

$$\mu(m \in A) = \frac{1}{2}; \mu(A \in m) = 1; d(m, A) = 1; Q(m, A) = \frac{5}{2 \times 3} = \frac{5}{6}.$$

Аналогичное значение будет иметь параметр, если мощность пересечения равна половине пространства вектора  $m$ . Если мощность пересечения равна половине мощностей пространств векторов  $m$  и  $A$ , то функции принадлежности равны:

$$\mu(m \in A) = \frac{1}{2}; \mu(A \in m) = \frac{1}{2}; d(m, A) = 1; Q(m, A) = \frac{4}{2 \times 3} = \frac{4}{6} = \frac{2}{3}.$$

Следует заметить, если пересечение двух векторов равно пустому множеству, то степень двойки от символа «пусто» принимается равным нулю:  $2^{\text{card}(m \cap A) = \emptyset} = 2^{\emptyset} = 0$ . Это действительно означает, что количество общих точек при пересечении двух пространств равно нулю.

Цель введения векторно-логического критерия качества решения заключается в существенном повышении быстродействия при подсчете качества  $Q$  взаимодействия компонентов  $m$  и  $A$  при анализе ассоциативных структур данных путем использования только векторных логических операций. Арифметический критерий (1) без усреднения функций принадлежности и кодового расстояния можно трансформировать к виду:

$$Q = d[m, A_{i(j)}] + \mu[m \in A_{i(j)}] + \mu[A_{i(j)} \in m],$$

$$d(m, A_{i(j)}) = \text{card}[m \oplus_{i(j)=1}^{n(m)} A_{i(j)} = 1];$$

$$\mu(m \in A_{i(j)}) = \text{card}[A_{i(j)} = 1] - \text{card}[m \wedge_{i(j)=1}^{n(m)} A_{i(j)} = 1]; \quad (2)$$

$$\mu(A_{i(j)} \in m) = \text{card}[m = 1] - \text{card}[m \wedge_{i(j)=1}^{n(m)} A_{i(j)} = 1].$$

Первый компонент, составляющий критерий, формирует степень несовпадения  $n$ -мерных векторов – кодовое расстояние, путем выполнения операции хог, второй и третий определяют степень непринадлежности результата конъюнкции к числу единиц каждого из двух взаимодействующих векторов. Понятия принадлежности и непринадлежности являются взаимодополняющими, но в данном случае технологичнее вычислять непринадлежность. Таким образом, идеальный критерий качества равен нулю, когда два вектора равны между собой. Оценка качества взаимодействия двух двоичных векторов убывает по мере роста критерия от 0 к 1. Чтобы окончательно уйти от арифметических операций при подсчете уже векторного критерия качества, необходимо выражения (2) преобразовать к виду:

$$Q = d(m, A) \vee \mu(m \in A) \vee \mu(A \in m),$$

$$d(m, A) = m \oplus A;$$

$$\mu(m \in A) = A \wedge \overline{m \wedge A}; \quad (3)$$

$$\mu(A \in m) = m \wedge \overline{m \wedge A}.$$

Здесь критерии представлены уже не числами, а векторами, которые оценивают взаимодействие компонентов  $m, A$ . При этом увеличение числа нулей в трех векторах качества повышает критерий, а наличие единиц индицирует ухудшение качества взаимодействия. Для сравнения оценок необходимо определять мощность единиц в каждом векторе без выполнения операций суммирования. Это можно сделать с помощью регистра (рис. 4), который позволяет за один такт выполнить сдвиг влево и уплотнить все единичные координаты  $n$ -разрядного двоичного вектора.

После процедуры сжатия номер правого единичного бита уплотненной серии единиц формирует индекс качества взаимодействия векторов. Для двоичных наборов  $m = (110011001100), A = (000011110101)$  определение качества их взаимодействия по формулам (3) представлено в следующем виде (нулевые координаты отмечены точками):

$m$	1 1 . . 1 1 . . 1 1 . .
$A$	. . . . 1 1 1 1 . 1 . 1
$\overline{m \wedge A}$	. . . . 1 1 . . . 1 . .
$m \wedge A$	1 1 1 1 . . 1 1 1 . 1 1
$d(m, A) = m \oplus A$	1 1 . . . . 1 1 1 . . 1
$\mu(A \in m) = m \wedge \overline{m \wedge A}$	1 1 . . . . . . 1 . . .
$\mu(m \in A) = A \wedge \overline{m \wedge A}$	. . . . . . 1 1 . . . 1
$Q = d(m, A) \vee \mu(m \in A) \vee \mu(A \in m)$	1 1 . . . . 1 1 1 . . 1
$Q(m, A) = (6/12)$	1 1 1 1 1 1 . . . . . .

Рис. 4. Регистр сдвига и уплотнения единиц



Здесь сформирована не только оценка взаимодействия векторов, равная  $Q(m, A) = (6/12)$ , но, что самое главное, единичные координаты строки  $Q = d(m, A) \vee \mu(m \in A) \vee \mu(A \in m)$  идентифицируют все существенные переменные, по которым имеется некачественное взаимодействие векторов. Для сравнения двух решений, полученных в результате логического анализа, используются сжатые векторы качества  $Q$ , над которыми выполняется процедура, включающая следующие векторные операции:

$$Q(m, A) = \begin{cases} Q_1(m, A) \leftarrow \text{or}[Q_1(m, A) \wedge Q_2(m, A) \oplus Q_1(m, A)] = 0; \\ Q_2(m, A) \leftarrow \text{or}[Q_1(m, A) \wedge Q_2(m, A) \oplus Q_1(m, A)] = 1. \end{cases} \quad (4)$$

Вектор-бит  $\text{or}$ -оператор девекторизации формирует двоичное битовое решение на основе применения логической операции  $\text{or}$  к  $n$  разрядам вектора существенных переменных критерия качества. Схемотехническое решение процедуры выбора и аналитическая процесс-модель имеют три операции, которые представлены на рис. 5.

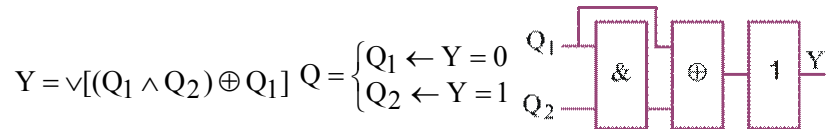


Рис. 5. Процесс-модель выбора решения

Для двоичных векторов, представляющих собой критерии качества, выполнена процедура выбора лучшего из них на основании выражения (рис. 6):

$Q_1(m, A) = (6,12)$	1 1 1 1 1 1 . . . . .
$Q_2(m, A) = (8,12)$	1 1 1 1 1 1 1 1 . . . . .
$Q_1(m, A) \wedge Q_2(m, A)$	1 1 1 1 1 1 . . . . .
$Q_1(m, A) \oplus Q_1(m, A) \wedge Q_2(m, A)$	. . . . .
$Q(m, A) = Q_1(m, A)$	1 1 1 1 1 1 . . . . .

Рис. 6. Выбор лучшего вектора качества

Векторные логические критерии качества взаимодействия ассоциативных наборов позволяют получать оценку поиска, распознавания и принятия решения с высоким быстродействием логических параллельных операций, что особенно существенно для критических систем реального времени.

### 3. Процесс-модель диагностирования СПАМА

Метрика качества, представленная в (3), дает возможность оценивать близость пространственных объектов друг к другу, а также взаимодействие векторных пространств. Практическим примером полезности интегрального критерия качества может служить стрельба по цели, которая иллюстрируется ранее приведенными диаграммами (см. рис. 3) взаимодействия векторов: 1) Снаряд попал точно в цель и поразил ее полностью. 2) Мишень поражена необоснованно большим калибром снаряда. 3) Калибр снаряда недостаточен для поражения крупной цели. 4) Неэффективный и неточный выстрел снарядом большого калибра. 5) Снаряд пролетел мимо мишени.

Процесс-модель взаимодействия  $P(m, A)$  сопровождается интегральным критерием качества, который оценивает не только попадание или промах, но и эффективность использования калибра оружия. Аналитическая запись процесс-модели для выбора лучшего взаимодействия входного запроса  $m$  с системой логических ассоциативных отношений представлена в виде:

$$P(m, A) = \min_{i=1}^n Q_i(m \Delta A_i) = \vee[(Q_i \bigwedge_{j=1, n}^{j \neq i} Q_j) \oplus Q_i] = 0;$$

$$Q(m, A) = (Q_1, Q_2, \dots, Q_i, \dots, Q_n);$$

$$A = (A_1, A_2, \dots, A_i, \dots, A_n);$$

$$\Delta = \{\text{and, or, xor, not, slc, nop}\};$$



$$\begin{aligned}
A_i &= (A_{i1}, A_{i2}, \dots, A_{ij}, \dots, A_{is}); \\
A_{ij} &= (A_{ij1}, A_{ij2}, \dots, A_{ijr}, \dots, A_{msq}); \\
m &= (m_1, m_2, \dots, m_r, \dots, m_q). \\
Q_i &= d(m, A_i) \vee \mu(m \in A_i) \vee \mu(A_i \in m), \\
d(m, A_i) &= m \oplus A_i; \\
\mu(m \in A_i) &= A_i \wedge \overline{m} \wedge A_i; \\
\mu(A_i \in m) &= m \wedge \overline{A_i} \wedge A_i.
\end{aligned} \tag{5}$$

*Комментарии:* 1) Функциональность  $P(m, A)$  задает аналитическую модель вычислительного процесса в виде высказывания, минимизирующего интегральный критерий качества. 2) Структуры данных представлены в виде вершин-таблиц графа  $A = (A_1, A_2, \dots, A_i, \dots, A_m)$ , логически взаимодействующих между собой. 3) Вершина графа задается упорядоченной совокупностью вектор-строк ассоциативной таблицы  $A_i = (A_{i1}, A_{i2}, \dots, A_{ij}, \dots, A_{is})$  явных решений, где строка  $A_{ij} = (A_{ij1}, A_{ij2}, \dots, A_{ijr}, \dots, A_{msq})$  представляет собой истинное высказывание.

Поскольку функционал, представленный в виде таблицы, не имеет постоянных во времени входных и выходных переменных, то данная структура отличается от последовательной машины фон Неймана, задаваемой конечными автоматами Мили и Мура. Равнозначность всех переменных в векторе  $A_{ij} = (A_{ij1}, A_{ij2}, \dots, A_{ijr}, \dots, A_{msq})$  создает одинаковые условия их существования, что означает инвариантность решения задач прямой и обратной импликации в пространстве  $A_i \in A$ . Ассоциативный вектор  $A_{ij}$  определяет собой явное решение, где каждая переменная задается в конечном, многозначном и дискретном алфавите  $A_{ijr} \in \{\alpha_1, \alpha_2, \dots, \alpha_i, \dots, \alpha_k\} = \beta$ . Взаимодействие  $P(m, A)$  входного вектора-запроса  $m = (m_1, m_2, \dots, m_r, \dots, m_q)$  с графом  $A = (A_1, A_2, \dots, A_i, \dots, A_m)$  формирует множество решений с выбором лучшего из них по минимальному критерию качества:

$$P(m, A) = \min Q_i [m \wedge (A_1 \vee A_2 \vee \dots \vee A_i \vee \dots \vee A_m)].$$

Конкретное взаимодействие вершин графа между собой создает функциональность  $A = (A_1, A_2, \dots, A_i, \dots, A_m)$ , которая может быть оформлена в следующие структуры: 1) Единственная ассоциативная таблица, содержащая все решения логической задачи в явном виде. Преимущество – максимальное быстродействие параллельного ассоциативного поиска решения по таблице. Недостаток – максимально высокая аппаратная сложность размещения таблицы большой размерности. 2) Древоподобная (графовая) структура бинарных отношений между функциональными примитивами, каждый из которых формирует таблицу истинности для незначительного количества переменных. Преимущество – максимально низкая аппаратная сложность решения задачи. Недостаток – минимальное быстродействие последовательного ассоциативного поиска решения по дереву. 3) Компромиссная графовая структура логически понятных для пользователя отношений между примитивами, каждый из которых формирует таблицу истинности для логически сильно взаимосвязанных переменных.

В целях детализации структуры векторных вычислений ниже представлены аналитические и структурные процесс-модели, которые приводятся к анализу  $A$ -матрицы по столбцам или строкам. Первая из них изображена на рис. 7 и предназначена для определения множества допустимых решений относительно входного запроса  $m_b$ .

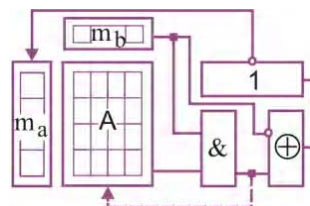


Рис. 7. Поиск всех допустимых решений

Вторая структура (рис. 8) осуществляет поиск оптимального решения на множестве найденных в первой процессной модели путем анализа строк. Кроме того, вторая модель имеет и самостоятельное применение, ориентированное на определение однозначного и многозначного решения, например, при поиске дефектов в цифровой системе на кристалле.

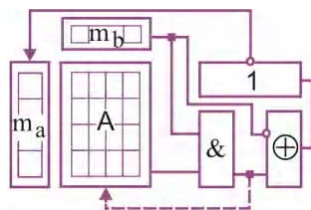


Рис. 8. Структура выбора оптимального решения

Предложенные процесс-модели анализа (графа) ассоциативных таблиц, идентифицирующих компоненты СПАМа, а также введенные критерии качества логических решений позволяют решать задачи квазиоптимального покрытия, диагностирования многообразия спамовых сообщений в индивидуальном кибернетическом пространстве (ИКП) пользователя. Модель векторных вычислений послужила основой для разработки специализированной мультимикропроцессорной архитектуры, ориентированной на поиск, распознавание и принятие решений на использование структур ассоциативных таблиц (рис. 9).

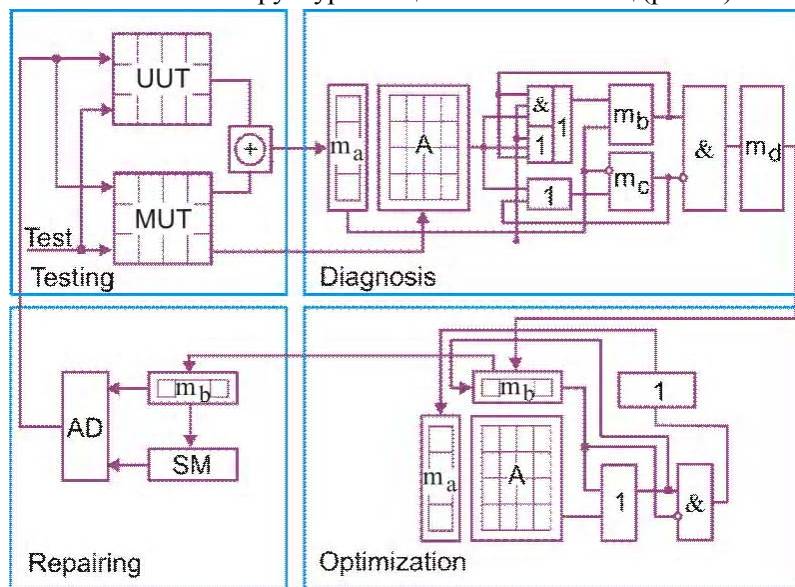


Рис. 9. Модель встроенного тестирования и восстановления ИКП

Оценка эффективности (рис. 10) проектного решения под эгидой специализации и стандартизации  $S_p \cup S_t$  основывается на совместном использовании трех взаимно противоречивых параметров: качество  $Y$ , быстродействие  $T$ , программные затраты  $H$ :

$$E = F(Y, T, H),$$

$$Y = (1 - P)^{n(1-Q)};$$

$$T = \frac{1}{f} \times S \times d;$$

$$H = 2(H^S \times n).$$

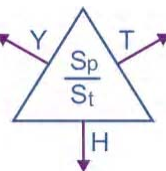


Рис. 10. Оценка эффективности процесс-модели

#### 4. Практические результаты внедрения инфраструктуры

В качестве объекта исследования был выбран SquirrelMail — клиент электронной почты с веб-интерфейсом, написанным на PHP. Приложение может быть проинсталлировано практически на любом веб-сервере, на котором установлен PHP и имеется связь с

почтовым сервером по протоколам IMAP и SMTP. Интерфейсное окно представлено на рис. 11.

Данный клиент достаточно легко расширяем различными плагинами. Для проведения исследования был написан плагин, который реализовывал анализ определения полезности информации на основании предпочтений пользователей.

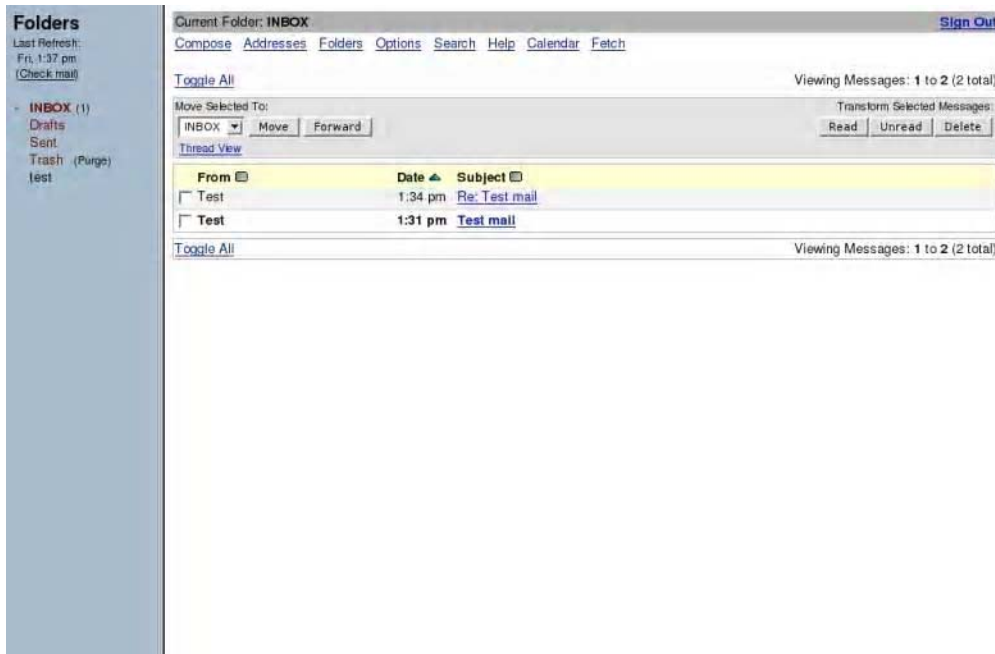


Рис. 11. Интерфейс SquirrelMail

Процесс-модель функционирования плагина представлена на рис. 12.

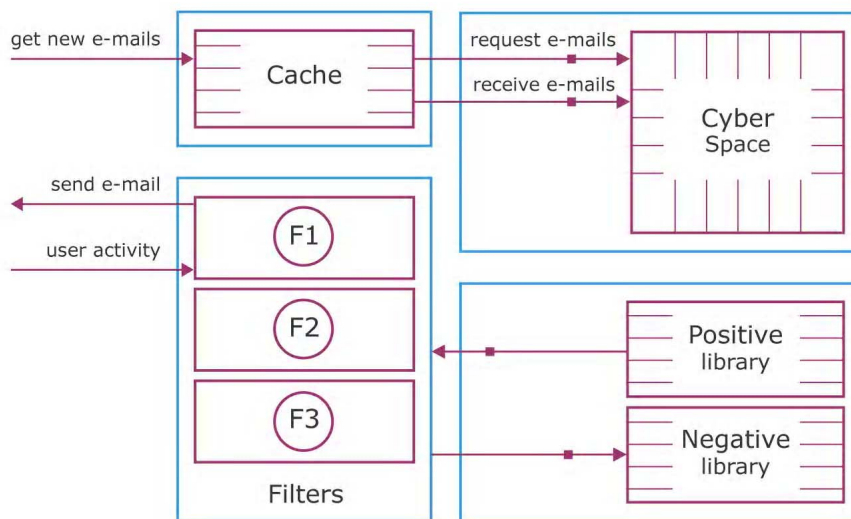


Рис. 12. Процесс-модель функционирования плагина под SquirrelMail

На основании активности пользователя и атрибутов, контентного анализа письма была выбрана и обучена система фильтров. При загрузке новых писем из индивидуального киберпространства (в данном случае киберпространство было представлено подмножеством электронных писем) информация фильтровалась не на основании «спам» или «не спам», а на основании персональных предпочтений пользователя (рис. 13).

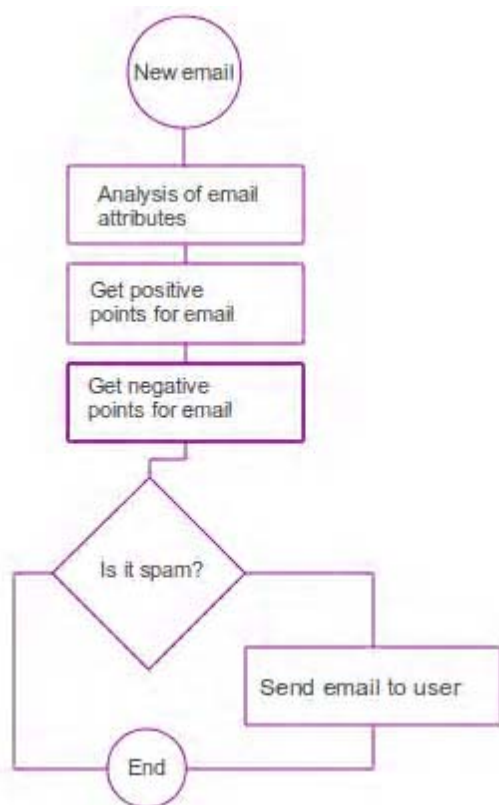


Рис. 13. Граф-схема анализа письма

Из всего множества писем, приходящих на почтовый ящик, «спам» (анонимные массовые рассылки) был частично идентифицирован уже не как спам, а как рассылка, которая может нести интересную информацию для пользователя.

На рис. 14 показана эффективность внедрения инфраструктуры сервисного обслуживания ИКП для одного пользователя, где TL – Total Letters, S – Spam, SAI – Spam after Infrastructure, UL – Useful Letters. Если предположить, что рыночная привлекательность инфраструктуры порядка 1 млрд пользователей, то экономия времени в целом на рынке пользователей кибернетического пространства составит ( $T_{\Sigma}$  – суммарная экономия времени за год;  $k$  – коэффициент уменьшения спама при внедрении инфраструктуры;  $L$  – количество писем в месяц;  $N$  – потенциальное количество пользователей в Украине;  $T$  – время анализа одного письма;  $M$  – количество месяцев в году;  $H_{\Sigma}$  – годовая экономия финансовых средств от внедрения инфраструктуры;  $C_h$  – стоимость одного часа рабочего времени пользователя в Украине):

$$T_{\Sigma} = k \times L \times N \times T \times M = 0,9 \times 800 \times 10000000 \times 1 \times 12 = 86400000000 \text{c} \approx 2740 \text{ years} \approx 24002400 \text{ hours};$$

$$H_{\Sigma} = T_{\Sigma} \times C_h = 24002400 \times \$5 = \$120012000.$$

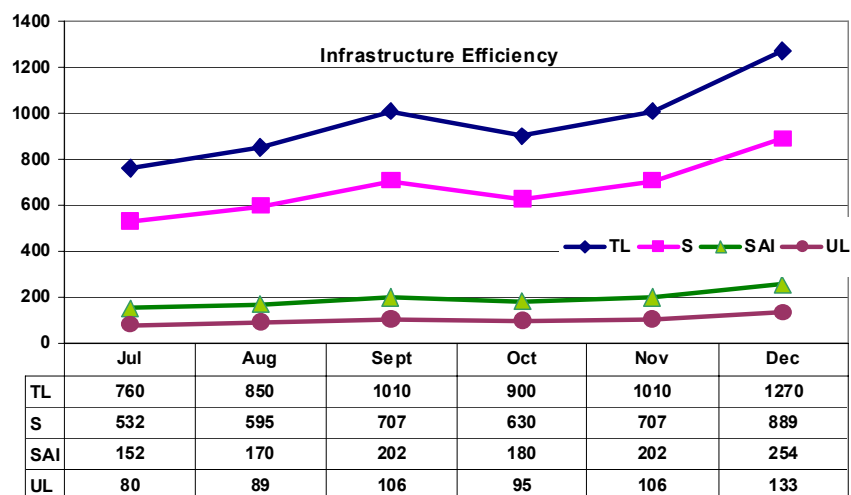


Рис. 14. Эффективность внедрения инфраструктуры ИКП

### Заключение

1. *Научная новизна* результатов исследования заключается в том, что впервые предложена инфраструктура сервисного обслуживания индивидуального кибернетического пространства, которая характеризуется наличием встроенных средств тестирования, диагностирования и восстановления ИКП и двух пополняемых библиотек позитивных и негативных сообщений, что дает возможность существенно (в несколько раз) уменьшить время анализа получаемой информации.

2. *Практическая значимость результатов исследования:* инфраструктура сервисного обслуживания ИКП ориентирована на повышение качества жизни всех субъектов планеты, использующих почтовые сервисы для коммуникации с внешним миром. При этом ИКП – модель будущего общения человека с внешним миром, которая инвариантна по отношению к техническим средствам доступа в киберпространство планеты. Годовой экономический эффект от внедрения инфраструктуры ИКП для украинских пользователей может составить более 120 млн долларов.

3. *Направление будущих научных исследований.* Актуальной представляется проблема создания теории, методов и архитектуры параллельного анализа информации, представленной в виде аналитических, графовых и табличных форм ассоциативных отношений для поиска, распознавания, диагностирования деструктивных компонентов и принятия решений в  $n$ -мерном векторном дискретном пространстве. Здесь целесообразно использовать векторно-логические процесс-модели актуальных прикладных задач, в том числе – диагностирование вирусов и восстановление работоспособности программно-аппаратных компонентов компьютерных систем, качество решения которых оценивается неарифметической метрикой взаимодействия булевых векторов. Решение проблемы ориентировано на поиск, распознавание, диагностирование деструктивных компонентов аппаратно-программными методами в дискретном кибернетическом пространстве. Общность представленной теории синтеза и анализа кибернетического пространства основана на использовании равенства нулю триады равноценных компонентов, соединенных операцией  $\text{xor } m \oplus A \oplus Q = 0$ , формулирующей условия решения проблемы. Здесь первый компонент  $m$  – входной код программы, второй  $A$  – эталонные модели деструктивов, третий  $Q$  – результат взаимодействия первых двух, который может вырождаться в критерий качества отношения или принятия решения, оценку распознавания объектов или образов. Цель – существенное повышение качества программных продуктов и уменьшение стоимости эксплуатационных расходов за счет их вакцинации путем введения в код встроенной программной избыточности в виде инфраструктуры сервисного обслуживания, обеспечивающей тестирование, диагностирование и устранение вредоносных компонентов, классифицируемых в библиотеках. Объект исследования – кибернетическое пространство, представленное информацией, ее носителями и преобразователями, а также деструктивными компонентами, наносящими вред функциональностям, улучшающим качество жизни человека. Субъект исследования – инфраструктура сервисного обслуживания в виде встроенной программной избыточности, работающей в реальном масштабе времени, обеспечивающей тестирование, диагностирование и устранение вредоносных компонентов, описанных в соответствующих библиотеках.

Мотивация: 1) Отсутствие на рынке антивирусной защиты встроенных средств тестирования, диагностирования и удаления вредоносных компонентов, составляющих инфраструктуру сервисного обслуживания, подобной тому, как в цифровых системах на кристаллах существуют стандарты граничного сканирования, а в программных продуктах – ассерционная избыточность, ориентированные на встроенное тестирование дефектов и ошибок с последующим восстановлением работоспособности аппаратных или программных изделий. 2) Наличие теоретических разработок, связанных с технологией алгебрологического векторного анализа информационных данных, ориентированных на высокое быстродействие решения и оценивания задач распознавания образов, принятия решений и тестирования объектов. 3) Наличие образцовой производственной и маркетинговой инфраструктуры (Лаборатория Касперского), способной поддержать проект создания технологии вакцинации программных продуктов и авторитетно предложить его рынку информационных технологий. 4) Миниатюризация цифровых и телекоммуникационных систем (телефоны смартфоны, IP-фоны, планшеты) требует постоянной защиты от несанкционированного доступа путем внедрения встроенных антивирусных средств, контролирующих информационный обмен. Задачи: 1) Разработка математического аппарата анализа кибернетического пространства, ориентированного на создание моделей и методов сервисного обслуживания программных продуктов для тестирования, диагностирования и устранения вредоносных компонентов. 2) Создание типовых процесс-моделей и критериев взаимодействия вредоносных компонентов с программными кодами полезных функциональностей. 3) Разработка технологии

анализа структуры программного кода для определения критических точек и установки в них ассерционных операторов наблюдения и управления в процессе его функционирования. 4) Создание инфраструктуры сервисного обслуживания функциональных программ для встроенного тестирования, диагностирования и устранения вредоносных компонентов из программного кода функциональности на основе использования библиотеки деструктивных элементов. 5) Тестирование и верификация встроенной инфраструктуры сервисного обслуживания функциональностей, защищающей программный код от вредоносных компонентов.

4. *Ожидаемые результаты и их рыночная привлекательность*: 1) Инфраструктура встроенной защиты программного кода от несанкционированной модификации, приводящей к изменению функциональности. 2) Избыточность инфраструктуры программного кода, которая автоматически синтезируется на стадии проектирования и верификации, составляет не более 5% от специфицированной функциональности. 3) Рыночная привлекательность инфраструктуры, определяемая многообразием программных продуктов, умноженной на уровень продаж каждого изделия, составляет в год порядка одного миллиарда экземпляров. 4) Стоимость создания инфраструктуры для программного продукта составляет 20% затрат от разработки функционального кода. Если уровень продаж – не менее 500 копий, то затраты на создание встроенного антивируса вполне окупаемы в течение года. 5) Внедрение запатентованной технологии вакцинации программных продуктов при их рождении может принести компании порядка 2-х миллиардов долларов в первые 3 года ее эксплуатации. 6) Маркетинговая проблема глобальной компании (Лаборатория Касперского) заключается в убеждении разработчиков программных продуктов имплементировать существующие внешние антивирусы вовнутрь кода полезной функциональности.

**Список литературы:** 1. *Инфраструктура мозгоподобных вычислительных процессов* / М.Ф. Бондаренко, О.А. Гузь, В.И. Хаханов, Ю.П. Шабанов-Кушнаренко. Харьков: Новое слово, 2010. 160 с. 2. *Проектирование и тестирование цифровых систем на кристаллах. Verilog & System Verilog* / В.И. Хаханов, Е.И. Литвинова, О.А. Гузь. Харьков: ХНУРЭ. 2009. 484с. 3. *Проектирование и верификация цифровых систем на кристаллах* / В.И. Хаханов, И.В. Хаханова, Е.И. Литвинова, О.А. Гузь. Харьков: Новое слово, 2010. 528с. 4. *Семенец В.В., Хаханова И.В., Хаханов В.И.* Проектирование цифровых систем с использованием языка VHDL. Харьков: ХНУРЭ. 2003. 492 с. 5. *Хаханов В.И., Хаханова И.В.* VHDL+Verilog = синтез за минуты. Харьков: ХНУРЭ. 2006. 264с. 6. *Хаханов В.И., Чумаченко С.В.* Модели пространств в научных исследованиях // Радиоэлектроника и информатика. 2002. №1. С. 124-132. 7. *Zorian Yervant.* Guest Editor's Introduction: Advances in Infrastructure IP // IEEE Design and Test of Computers. 2003. P.49-55. 8. *Bulent I. Dervisoglu.* A Unified DFT Architecture for Use with IEEE 1149.1 and VSIA/IEEE P1500 Compliant Test Access Controllers. Proceedings of the Design Automation Conference. 2001. P. 53-58. 9. *Bergeron J.* Writing Testbenches using SystemVerilog. Springer US. 2006. 414 p.

*Поступила в редколлегию 11.11.2010*

**Хаханов Владимир Иванович**, декан факультета КИУ ХНУРЭ, д-р техн. наук, профессор кафедры АПВТ ХНУРЭ. Научные интересы: техническая диагностика цифровых систем, сетей и программных продуктов. Увлечения: баскетбол, футбол, горные лыжи. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326. E-mail: hahanov@kture.kharkov.ua.

**Чумаченко Светлана Викторовна**, д-р техн. наук, профессор кафедры АПВТ ХНУРЭ. Научные интересы: математическое моделирование, методы дискретной оптимизации. Увлечения: спорт, путешествия, любительское фото. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326, e-mail: ri@kture.kharkov.ua.

**Мищенко Александр Сергеевич**, аспирант кафедры АПВТ ХНУРЭ. Научные интересы: техническая диагностика цифровых систем, сетей, антивирусная защита, web-технологии. Увлечения: английский язык, web-дизайн. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326, e-mails: alex@simplesolutions.com.ua, alexander.mischenko@gmail.com, santific@gmail.com.

**Зацарный Артем Вячеславович**, студент гр. КИ-08-6 факультета КИУ ХНУРЭ. Научные интересы: техническая диагностика цифровых систем, сетей, антивирусная защита, web-технологии. Увлечения: футбол. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326.

**Хаханова Юлия Владимировна**, студентка гр. КИ-08-6 факультета КИУ ХНУРЭ. Научные интересы: техническая диагностика цифровых систем, сетей, антивирусная защита, web-технологии. Увлечения: футбол, горные лыжи. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326.