УДК 613.86;159.913

К.О. Гнидко, А.Г. Ломако

КОНТРОЛЬ ПОТЕНЦИАЛЬНО ОПАСНОГО ИНФОРМАЦИОННО-ПСИХОЛОГИЧЕСКОГО ВОЗДЕЙСТВИЯ НА ИНДИВИДУАЛЬНОЕ И ГРУППОВОЕ СОЗНАНИЕ ПОТРЕБИТЕЛЕЙ МУЛЬТИМЕДИЙНОГО КОНТЕНТА

Гнидко К.О., Ломако А.Г. **Контроль потенциально опасного информационно-психологического воздействия на индивидуальное и групповое сознание потребителей мультимедийного контента.**
Аннотация. В настоящей статье рассмотрены основные принципы построения многоуровневой системы контроля потенциально опасного информационно-психологического воздействия на потребителей мультимедийного контента. Представлены результаты экспериментальных исследований по обнаружению скрытых подпороговых воздействий.
Ключевые слова: психофизиологические воздействия, суггестия, подпороговые сообщения.

Gnidko K.O., Lomako A.G. **Monitoring of Potentially Dangerous Information-Psychological Affect on Individual and Group Consciousness of Multimedia Content Consumers.**
Abstract. This article reviews the basic principles of multi-level system of diagnostics and monitoring of potentially dangerous informational and psychological affect on consumers of multimedia content. The results of experimental studies on the detection of hidden subliminal messages are represented.
Keywords: psychophiziological affections, suggestion, subliminal messages.

1. Введение. Возможности науки и техники в настоящее время позволяют создавать средства и методы для информационного воздействия на индивидуальное, групповое и массовое сознание граждан РФ. Действующая Военная доктрина Российской Федерации подчеркивает тенденцию смещения военных опасностей и военных угроз в информационное пространство и внутреннюю сферу Российской Федерации, а также указывает на одну из главных внутренних военных опасностей: деятельность по информационному воздействию на население, в первую очередь на молодых граждан страны, имеющую целью подрыв исторических, духовных и патриотических традиций в области защиты Отечества [1].

Обычно под информационной безопасностью подразумевается состояние защищенности жизненно важных интересов граждан, общества и государства в информационной сфере [2]. Однако при этом все же недостаточное внимание уделяется личности человека. А именно личность должна стать предметом пристального внимания и защиты от возможных угроз, в том числе информационных, в период бурного развития телекоммуникационных систем. Информационное оружие имеет целью, в том числе, «массовое распространение и внедрение в сознание людей определенных представлений, привычек и поведенче-

ских стереотипов, вызывающих недовольство или панику среди населения, а также провоцирование деструктивных действия различных социальных групп» [3].

Перечень приоритетных проблем научных исследований в области обеспечения информационной безопасности Российской Федерации, утвержденный Советом Безопасности РФ, определяет в качестве одного из наиболее актуальных направлений исследование проблем защиты индивидуального, группового и массового сознания российского общества от деструктивных воздействий различных информационных источников. Наиболее удобной средой для осуществления атак на защищаемые психофизиологические ресурсы в силу ряда особенностей (масштабность, гетерогенность, децентрализованность, отсутствие цензуры, возможность передачи мультимедийных данных любых видов) в настоящее время является глобальная вычислительная сеть Интернет. Таким образом, чрезвычайно актуальной является не только проблема защиты информации, но и проблема защиты от информации, которая в последнее время приобретает международный масштаб и стратегический характер.

Организациями, ведущими активные исследования в области психотехнологий на территории России, являются кафедра психоэкологии Российского университета дружбы народов, Институт психологии Российской академии наук, научная группа кафедры психиатрии Московской медицинской академии им. Сеченова, Всероссийский научно-исследовательский институт телевидения и радиовещания и некоторые другие организации.

В настоящей статье кратко обобщены результаты проведенных авторами исследований в области защиты пользователей автоматизированных систем от потенциально опасного мультимедийного контента. Проект системы защиты несовершеннолетних пользователей от скрытого вредоносного воздействия мультимедийного контента сети Интернет, разработанный в соответствии с изложенными в данной работе принципами, удостоен приза за лучшую инновационную идею на конкурсе инновационных проектов в сфере науки и высшего профессионального образования Санкт-Петербурга в 2012 году [4].

2. Многоуровневая фильтрация потенциально опасных информационно-психологических воздействий в мультимедийных потоках. Эффективная система диагностики и контроля потенциально опасных информационно-психологических воздействий на индивидуальное и групповое сознание должна обеспечивать выполнение следующих функций:

– на внешнем уровне: обнаружение и ликвидация условий, по-

рождающих проявление угроз, а также создание условий, препятствующих их проявлению;

– на граничном уровне: обнаружение проявлений угроз, препятствие доступу реализуемых угроз к защищаемым ресурсам;

– на внутреннем уровне: обнаружение и локализация несанкционированных воздействий на защищаемый ресурс, своевременная ликвидация последствий и причин возникших нарушений информационно-психологической безопасности.

Перечислим подмножества классов негативного контента, подлежащих вскрытию и фильтрации при обработке мультимедийных потоков данных в контурах автоматизированных систем инфраструктуры критически важных объектов.

1. Текстовые массивы. К наиболее значимым параметрам, определяющим воздействие текста на психофизиологическое состояние человека, относятся:

– негативная фоносемантическая окраска;

– фрактальные свойства, определяющие меру суггестивности текста [5, 6].

2. Видеопотоки. Наибольшее значение для деятельности человека-оператора имеет зрительный анализатор, через который поступает около 90% всей обрабатываемой информации. Зрение позволяет воспринимать форму, цвет, яркость и движение предметов. Возможность зрительного восприятия определяется энергетическими, пространственными, временными и информационными свойствами сигналов, поступающих к оператору. Совокупность этих свойств и динамика их изменения во времени (структура видеопотока) определяют информацию, которую визуальный сигнал несет в сферу сознательного и бессознательного психики оператора. В ряде работ [7–10] рассматриваются различные аспекты угрозы скрытого воздействия видеосигнала на сознание, подсознание и психофизиологическое состояние человека. Общеизвестен инцидент, приведший к госпитализации в Японии 16 декабря 2007 года более 700 детей после просмотра мультфильма «Покемоны». Вызвано это было тем, что мелькание цветовых пятен в кадре на протяжении около 10 секунд вызвало эффект «резонанса» с основными частотами функционирования головного мозга.

Экспериментальное подтверждение имеют эффекты воздействия на сознание и подсознание человека скрытых неосознаваемых визуальных вставок, а также световая частотная стимуляция [11, 12]. В связи с этим, фильтрации в видеопотоке подлежат:

– скрытые визуальные вставки;

– диспарантные видеовставки;

– колебания яркости (мерцания) в диапазоне биоэффективных частот.

3. Аудиопотоки. Вредоносный контент, подлежащий фильтрации в аудиопотоках, включает:

– аудиосуггестию (под аудиосуггестией здесь и далее понимается процесс воспроизведения и восприятия особой аудиоинформации, результатом которого является существенное снижение уровня критического восприятия объекта воздействия (суггеренда), изменяется его эмоциональное и психофизиологическое состояние) [13, 14];

– вредоносные бинауральные ритмы в области биоэффективных частот (так называемые «цифровые наркотики»).

Разработанная авторами концепция реализует системный подход к фильтрации потенциально вредоносного контента в мультимедийных потоках данных и учитывает особенности возникновения и реализации угроз на всех этапах их существования.

Достижение поставленной цели предполагает решение следующих задач:

1. Установление подозрения на наличие вредоносных структур в мультимедийных потоках данных. Объемы информации, циркулирующие в современных сетях передачи данных, определяют высокие требования к скорости анализа и выявления потенциально вредоносных конструкций в мультимедийном контенте. При этом сохраняется значительная неопределенность признаков вредоносности. В данных условиях эффективным решением является построение распознавателя верхнего уровня на основе аппаратной реализации нейронной сети прямого распространения, предварительно обученной на репрезентативной выборке эталонных объектов. Входными данными для нейросетевого классификатора являются параметры структуры бинарного потока, полученные в результате его предварительной обработки. Выходной информацией нейросетевого классификатора является решение о принадлежности анализируемого участка бинарного потока к классу нейтральных или потенциально вредоносных (подозрительных) объектов.

2. Выявление скрытых информационно-психологических воздействий в мультимедийных объектах на основе частных алгоритмов распознавания. Входными данными для частных алгоритмов распознавания вредоносных свойств в мультимедийных объектах, являются файлы, содержащие данные объекты. На данном этапе из анализа исключаются файлы, зашифрованные криптографическими средствами, а также мультимедийные файлы, для которых в автоматизированной системе отсутствуют соответствующие декодеры. Выходные данные для каждого из алгоритмов, зависят от типа анализируемого объекта и

класса выявляемой угрозы вредоносного воздействия. В общем случае результатом работы частных алгоритмов является решение о принадлежности анализируемого мультимедийного объекта по признаку наличия вредоносных свойств к одному из двух классов: подозрительные на вредоносность и нейтральные.

3. Подтверждение факта вредоносного воздействия на основе контроля ответных физиологических реакций. Принятие решения о наличии негативного влияния на потребителей мультимедийного контента осуществляется на основе применения нечеткого классификатора, входными данными для которого являются психофизиологические параметры персонала, снимаемые комплексом датчиковой аппаратуры. Базисом для построения классификации на основе теории размытых множеств являются регулярные наблюдения и опыт экспертов в области психофизиологии. Выход нечеткой модели зависит от ее структуры и параметров – функций принадлежностей и весов правил. В общем случае настройка представляет собой нахождение параметров, минимизирующих расстояние между желаемым и действительным поведением нечеткой модели на обучающей выборке.

Вероятность возникновения психофизиологических угроз новых типов предъявляет к разработанной технологии требование по возможности дообучения распознавателей различных уровней. В силу того, что обучаемым элементом системы является вероятностная нейронная сеть, соблюдение данного требования возможно при наличии обратной связи – от подсистемы контроля психофизиологических параметров к нейросетевому распознавателю. При этом выявленные нечетким классификатором вредоносные объекты должны пройти соответствующую первичную обработку.

Обобщенная схема модели системы диагностики и контроля потенциально опасных информационно-психологических воздействий на индивидуальное и групповое сознание потребителей мультимедийного контента представлена на рисунке 1.

Вредоносные конструкции, скрытые в мультимедийных объектах, на бинарном уровне представления обладают, подобно компьютерным вирусам, набором сигнатурных признаков, по которым может быть установлено подозрение на наличие вредоносных свойств в анализируемом участке бинарного потока. Подобные признаки могут быть определены в результате структурного и спектрального анализа бинарных образов эталонных вредоносных объектов.

В соответствии со схемой, представленной на рисунке 1, на вход системы поступает бинарный поток, содержащий мультимедийные данные. Анализируемый бинарный поток является двоичной по-

следовательностью конечной длины и представляет собой конкатенацию некоторого числа L двоичных символов.

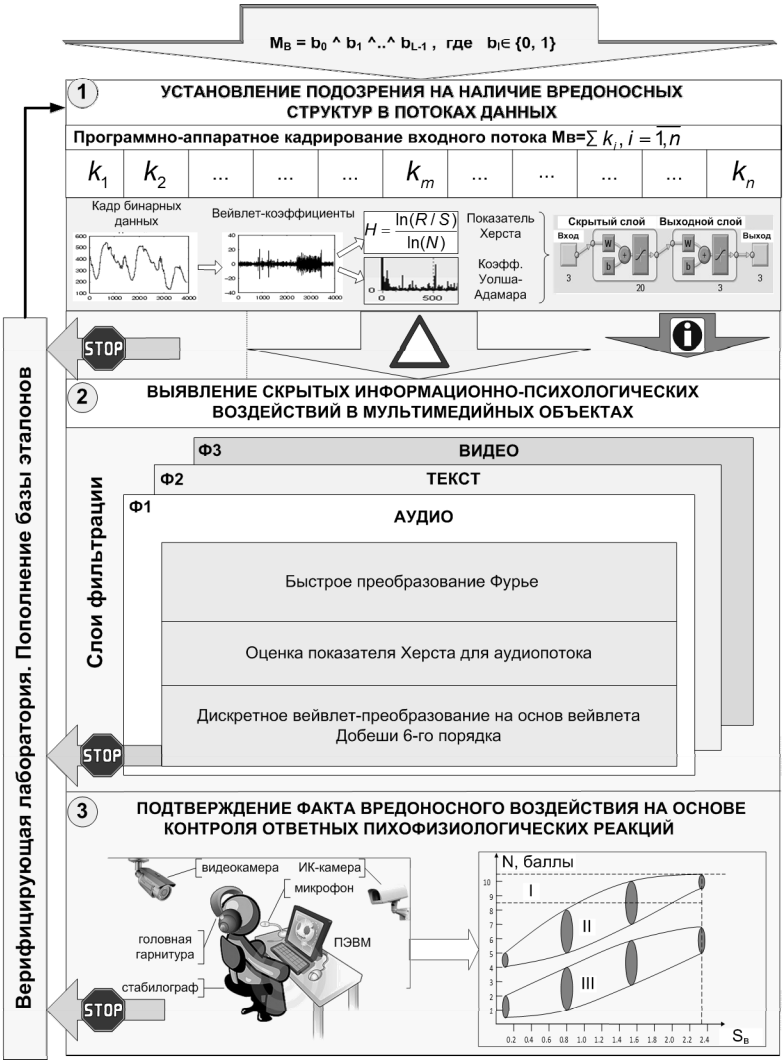


Рис. 1. Модель многоуровневой системы диагностики и контроля потенциально опасных информационно-психологических воздействий на индивидуальное и групповое сознание потребителей мультимедийного контента

Элементарным (неделимым) структурным элементом бинарного

потока является бит (двоичный символ): $M_b=b_0 \wedge b_1 \wedge \ldots \wedge b_{L-1}$, где $b_i \in \{0,1\}$, $i=\overline{1,L-1}$. Как правило, для сокращения объема передаваемой информации и увеличения скорости и надежности передачи информации применяются различные форматы кодирования и сжатия данных. Структура и свойства обработанных таким образом потоков могут не иметь признаков наличия вредоносных объектов. Вместе с тем, для реализации угрозы психофизиологического воздействия негативный контент должен быть предъявлен атакуемому лицу в явном виде, определенном особенностями восприятия и порогами чувствительности органов чувств человека. Данное обстоятельство позволяет осуществлять наиболее результативное выявление и фильтрацию вредоносных информационно-психологических воздействий непосредственно после декодирования бинарного потока на этапе, предшествующем этапу восприятия.

Таким образом, входной бинарный поток данных должен быть декодирован с применением декодеров соответствующего типа и разделен по типам мультимедийных объектов (текст, видео-, аудиообъект). Поэтому необходимым условием для корректного выполнения последующих процедур фильтрации является наличие в системе установленного пакета кодеков (кодеров-декодеров) для всех типов обрабатываемых мультимедийных объектов.

Основным объектом анализа на первом этапе разработанной технологии являются особенности структуры, в которых могут быть выявлены признаки, характерные для вредоносных информационно-психофизиологических объектов. Для выявления параметров структуры бинарный поток разбивается на отдельные кадры k_i, представляющие собой битовые последовательности длиной d. Определение битовой длины d кадров, на которые разбивается входной бинарный поток, осуществляется с учетом доступных вычислительных ресурсов и зависит от мощности и конфигурации привлекаемой для анализа автоматизированной системы. Уменьшение длины кадра снижает нагрузку на вычислительные ресурсы, однако увеличивает время обработки данных.

Для каждого из кадров осуществляется дискретное вейвлет-преобразование. Преимущество вейвлет-анализа перед другими видами спектрального и структурного анализа заключаются в том, что с его помощью могут быть подвергнуты анализу как общие свойства набора данных, так и локальные особенности, например, резкие изменения, скачки, диссонирующие с общим фоном. На основе полученного набора аппроксимирующих вейвлет-коэффициентов для каждого из кадров бинарного потока выполняется быстрое преобразование Уол-

ша-Адамара и рассчитывается фрактальная размерность на основе вычисления показателя Херста, что позволяет при общем сокращении признакового пространства выделить наиболее значимые и информативные признаки исследуемого бинарного потока.

Данные, полученные в результате указанных преобразований, поступают на вход нейронной сети, предварительно обученной на основе репрезентативной обучающей выборки. Обучающая выборка должна включать бинарные образы объектов двух классов: безопасных относительно оказываемого психофизиологического воздействия и априорно рассматриваемых как вредоносные. С учетом этого нейросетевой классификатор осуществляет разделение кадров бинарного потока по признаку вредоносности на следующие классы:

1) Имеющие признаки вредоносных конструкций. Кадрам данного класса по результатам распознавания присваивается идентификационный маркер «красный».

2) Не имеющие признаков наличия вредоносных конструкций (нейтральные). Кадрам данного класса присваивается классифицирующий маркер «зеленый».

2.1. Применение нейронной сети прямого распространения для обнаружения подозрительных на вредоносные конструкции в бинарном потоке данных. Цель обучения нейронной сети состоит в такой подстройке ее параметров, чтобы заданному входному вектору X сеть ставила в соответствие целевой выходной вектор Z**.** Вместе эти векторы составляют обучающую пару, а группа обучающих пар составляет обучающее множество. Алгоритм обучения искусственной нейронной сети в общем виде включает в себя следующие шаги (рисунок 2):

1) Выбор обучающей пары векторов из обучающего множества и подача входного вектора на входы сети.

2) Вычисление выходного вектора сети.

3) Вычисление разности между выходным и целевым векторами.

4) Коррекция весов сети с целью минимизации ошибки.

5) Повтор шагов с 1 по 4 для каждой пары обучающего множества до тех пор, пока ошибка на всем множестве не достигнет заданного уровня.

После завершения обучения нейросеть готова к решению задачи классификации конструкций, подозрительных на вредоносные, в бинарном потоке данных. В рабочем режиме сети предъявляется входной вектор, состоящий из тройки параметров (максимальное значение коэффициента Уолша-Адамара, порядковый номер коэффициента, при котором достигается его максимальное значение, и показатель Херста,

рассчитанный для анализируемого бинарного кадра). Входной вектор соответствующим образом активирует нейроны слоя образцов. Каждый нейрон слоя образцов выдает на своем выходе некоторый уровень активности $y_i(x)$.

Шаг 1. Вычисление дискретных значений амплитудной вейвлет-функции:

$$W_A(a_i, b_j) = \frac{1}{n(a_i, b_j)} \sum_{k=0}^{N-1} x_k \psi\left(\frac{t_k - b_j}{a_i}\right)$$

Шаг 2. Вычисление скейлограмы:

$$S(a_i, b_j) = \left|W_A(a_i, b_j)\right|^2$$

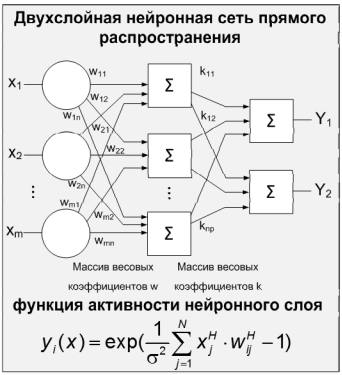
Шаг 3. Масштабирование вейвлета:

$$d_a = 2\Delta_t a$$

Шаг 4. Расчет параметров для обучения нейронной сети:

$$NL(S(a_i, b_j)) = \min d(f; f_{вфф}); \qquad W = \max(NL); I = index(W); \qquad H = \frac{\ln(R/S)}{\ln(N)}$$

Шаг 5. Передача параметров H, W, I ***на вход нейросети прямого распространения:***



Шаг 6. Нормализация входного вектора значений

$$x_j^H = \frac{x_j}{\sqrt{x_1^2 + x_2^2 + \ldots + x_N^2}} = \frac{x_j}{\sqrt{\sum_{j=1}^{N} x_j^2}}$$

Шаг 7. Выбор обучающей пары векторов из обучающего множества и подача входного вектора на входы сети.

Исходные обучающие векторы		Нормлизованные обучающие векторы	
x_1	x_2	x_1^H	x_2^H
3,0	5,0	0,5145	0,8575

Шаг 8. Вычисление выходного вектора сети.

$$Y^k = \sum_{l-1}^{L_k} \exp\left(\frac{1}{\sigma^2} \sum_{j=1}^{N} x_j^H w_{ij}^H - 1\right), k = \overline{1,m}$$

Класс	№ нейрона	w_1	w_2	$y_i(X)$	$Y^k(X)$
A	1	0,5145	0,8575	0,0008	0,5459
	2	0,7071	0,7071	0,3950	
	...	...	...	...	
B	7	0,9615	0,2747	0,0011	0,1389
	8	0,9191	3939	0,0516	
	...	...	...	...	

Шаг 9. Вычисление разности между выходным и целевым векторами.

$$d_q = \left|Y_q^k(X_l) - S_q(a_i, b_j)\right|$$

$$W = \begin{Vmatrix} w_{11} & w_{12} & \cdots & w_{1L} \\ w_{21} & w_{22} & \cdots & w_{2L} \\ \cdots & \cdots & \cdots & \cdots \\ w_{N1} & w_{N1} & \cdots & w_{NL} \end{Vmatrix} = \begin{Vmatrix} x_{11} & x_{12} & \cdots & x_{1L} \\ x_{21} & x_{22} & \cdots & x_{2L} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N1} & \cdots & x_{NL} \end{Vmatrix}$$

Шаг 11. Повтор шагов с 6 по 9 для каждой пары обучающего множества до тех пор, пока ошибка на всем множестве не достигнет приемлемого уровня.

Шаг 12. Завершение обучения. Применение нейрораспознавателя

Рис. 2. Алгоритм обучения искусственной нейронной сети

Каждый *k*-нейрон слоя суммирования суммирует уровни активности $y_i(X)$ всех нейронов слоя образцов своего *k*-класса и выдает на своем выходе общий уровень активности данного *k*-класса $Y^k(X)$. Выходной нейрон на основании вычисленных сетью уровней активности по каждому классу $Y^k(X)$ определяет, какой нейрон слоя суммирования имеет максимальный выходной сигнал. Тем самым (по номеру *k* нейрона), определяется номер класса *k*, к которому с большей вероятностью принадлежит предъявленный входной образ X.

Таким образом, в результате работы нейросетевого классификатора делается вывод о принадлежности анализируемого участка бинарного кода к классу подозрительных на вредоносные.

В рамках экспериментальных исследований на стенде применялся нейросетевой классификатор на основе вероятностной нейронной сети прямого распространения [15]. Количество выходных состояний сети – два. Для первого состояния характерно выявление бинарных кадров, «подозрительных» на наличие вредоносных закладок, второе состояние характеризуется выбором бинарных кадров, не имеющих признаков наличия негативного информационного контента. Выборка разбивается на 3 части (обучающую, проверочную и тестовую). Рекомендуемым является следующее соотношение: обучающая часть составляет 75% от общего объема выборки, проверочная и тестовые – по 15% от общего объема. Обучающее, проверочное и тестовое множества бинарных кадров являются непересекающимися, что позволяет оценить способность обученной нейронной сети к обобщению. Результаты оценивания качества разработанного распознавателя (на этапах обучения, коррекции весов нейронов и тестирования) при обнаружении в сетевом трафике признаков аудиосуггестии приведены в виде матриц неточностей на рисунке 3.

Рис. 3. Результаты оценивания качества разработанного нейросетевого распознавателя при обнаружении в бинарном потоке данных аудиосуггестии

2.2. Подтверждение факта вредоносного воздействия на основе нечеткой классификации физиологических показателей пользователя. Подтверждение факта воздействия негативного контента данных может быть осуществлено на основе регистрации в процессе деятельности психофизиологических параметров потребителей мультимедийного контента. Сложность определения функционального состояния человека по совокупности полученных параметров заключается в нечетком характере ответных реакций организма на входные информационные стимулы и размытости границ их диапазонов. Учитывать данные особенности позволяет применение нечеткого классификатора на основе размытых множеств.

Достоверность определения психофизиологического состояния персонала повышается пропорционально увеличению количества регистрируемых параметров и частоте съема данных параметров. Вместе с тем, комплект датчиковой аппаратуры должен причинять минимальные неудобства пользователю системы. Исходя из данного противоречия и доступных к настоящему времени аппаратных средств контроля физиологических параметров, определен рациональный состав комплекта регистрирующей аппаратуры. Он включает в себя: видеокамеру, инфракрасную камеру, головную гарнитуру с интегрированным датчиком частоты сердечных сокращений, микрофон с интегрированным датчиком частоты дыхания и стабилограф. Указанный набор датчиков, с одной стороны позволяет с требуемой достоверностью определять функциональное состояние человека-оператора, а с другой стороны не отвлекает его от выполнения основных обязанностей по предназначению. Вариант расположения датчиковой аппаратуры показан на рисунке 4.

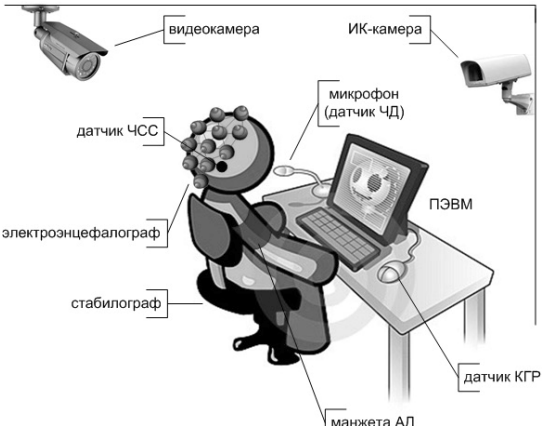


Рис. 4. Автоматизированное рабочее место, оснащенное аппаратурой контроля текущего функционального состояния пользователя

Алгоритм нечеткой классификации во множестве описаний образов функциональных состояний операторов в деятельности может быть представлен следующим образом.

1) Выбрать из состава обучающей выборки b-й образ функционального состояния (ФС) и принять вектор ее характеристик за начало отсчета многомерной поверхности координат ФС.

2) По известным методикам оценки расстояний между образами параметров объектов конкретного типологического содержания определяются векторы расстояний между образами ФС из состава обучающей выборки (ОВ) и образа ФС, принятого за начало отсчета в данной ОВ. Полученные векторы расстояний рассматриваются как изоморфные соответствующим значениям параметров ФС при условии начала их координат в точке, определяемой характеристиками объекта.

3) В соответствии с алгоритмом многомерной размытой классификации вычисляются параметры шкалы оценок функций принадлежности ФС выделенным классам и вектора функций принадлежности остальных образов ФС из состава ОВ, для всех X.

4) Вычисляется степень близости полученных параметров и исходного вектора (характеристического вектора), заданного в информационном векторе ОВ, по величине дисперсионного отношения.

5) Если все образы ФС из состава ОВ просчитаны, то переход к п. 6. Если нет, то переход к п.1.

6) Определение оптимального с точки зрения безошибочности классификации начала отсчета параметров ФС из состава ОВ.

Относительно этой точки в дальнейшем приводится измерение координат классифицируемых функциональных состояний операторов. Данной точке отсчета будут соответствовать рациональные параметры шкалы классификации, а именно центры классов и весовые коэффициенты типологически различных групп понятий в описании ФС.

Обобщенная структурная схема многомерного размытого классификатора, в котором реализуются такие режимы работы как расчет параметров классификационной шкалы, внутреннее структурирование ОВ, контроль непротиворечивости ОВ, коррекция ОВ и собственно классификация новых объектов приведена на рисунке 5.

В результате работы классификатора, основанного на нечеткой логике и аппарате размытых множеств, система генерирует одно из 3-х возможных сообщений:

– пользователь подвергся вредоносному психофизиологическому воздействию – высокая «физиологическая цена» деятельности (интегральный показатель принял значение «красный»);

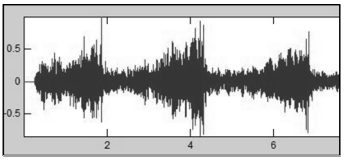
– существует вероятность оказания вредоносного психофизиологического воздействия – нормальная «физиологическая цена» деятельности (интегральный показатель принял значение «желтый»);

– пользователь не подвергался вредоносному психофизиологическому воздействию – низкая «физиологическая цена» деятельности (интегральный показатель принял значение «зеленый»).

Рис. 5. Структурная схема многомерного размытого классификатора

Таким образом, применение нечеткого классификатора позволяет осуществить подтверждение факта оказания вредоносного информационного воздействия на потребителей мультимедийного контента. Преимуществом применения нечеткого классификатора является возможность распознавания ранее неизвестных типов информационных закладок по вызываемым ими ответным физиологическим реакциям оператора. Бинарный образ скомпрометированного мультимедийного потока передается для дальнейшего исследования в специализированную лабораторию и, при обнаружении устойчивой связи между воздействием содержащегося в нем объекта на органы чувств испытуемого и ухудшением его психофизиологического состояния, принимается решение об обнаружении вредоносного воздействия неизвестного типа. Бинарный образ вредоносного объекта подается на вход блока вейвлет-анализа. Полученные параметры пополняют базу эталонов и передаются на вход нейросетевого классификатора для его дообучения.

2.3. Результаты экспериментальных исследований по обнаружению вредоносных конструкций в аудиопотоках. Рассмотрим процедуру обнаружения потенциально вредоносного контента в аудиопотоке на примере анализа фрагмента аудиозаписи суггестивного содержания (т.н. «громкий зикр») с параметрами: тип кодирования – PCM (*.wav), 16 бит, частота дискретизации – 44100 Гц. Его амплитудно-временное представление и бинарный образ изображены на рисунке 6.



Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0000D210	F2	EE	2F	EE	44	ED	3F	EC	3B	EB	5C	EA	C4	E9	8E	E9
0000D220	CE	E9	83	EA	9F	EB	0C	ED	A6	EE	50	F0	F5	F1	87	F3
0000D230	0A	F5	8D	F6	24	F8	E2	F9	D5	FB	FD	FD	50	00	B8	02
0000D240	13	05	45	07	37	09	D9	0A	2B	0C	39	0D	13	0E	CC	0E
0000D250	72	0F	0A	10	94	10	01	11	47	11	57	11	2A	11	C5	10
0000D260	34	10	89	0F	D9	0E	34	0E	A0	0D	18	0D	90	0C	F1	0B
0000D270	23	0B	1A	0A	CB	08	3D	07	81	05	AC	03	D7	01	13	00
0000D280	69	FE	D6	FC	50	FB	C6	F9	28	F8	73	F6	A9	F4	DC	F2
0000D290	25	F1	9F	EF	64	EE	7E	ED	EB	EC	9D	EC	77	EC	5A	EC

а) б)

Рис. 6. Фрагмент аудиозаписи суггестивного содержания: а) амплитудно-временное представление; б) бинарный образ

Результат кадрирования бинарного потока, применения к нему вейвлет-преобразования на основе вейвлета Добеши 6-го порядка и получения аппроксимирующих коэффициентов (cA) представлен на рисунке 7.

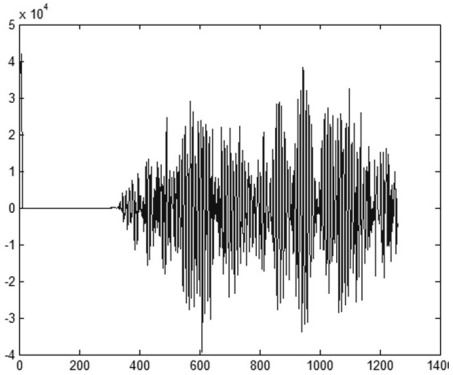


Рис. 7. Аппроксимирующие вейвлет-коэффициенты

Результат применения к полученным коэффициентам быстрого преобразования Уолша-Адмара представлен на рисунке 8. Расчетное значение показателя Херста для анализируемого аудиофрагмента: H=0,2.

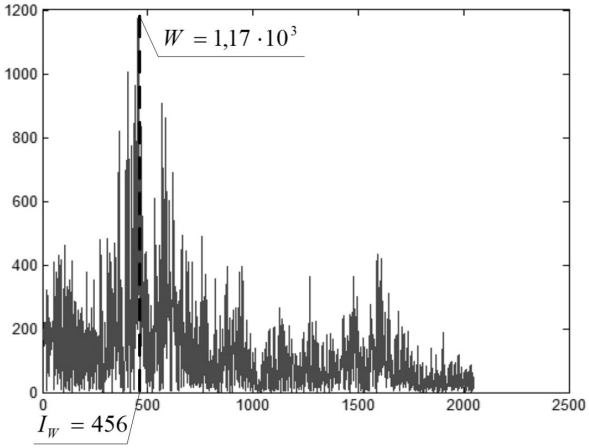


Рис. 8. КоэффициентыУолша-Адамара

Предварительно обученный нейросетевой классификатор с одним скрытым слоем нейронов характеризуется параметрами, представленными на рисунке 9.

Номер нейрона	Матрица весовых коэффициентов входов			Матрица весовых коэффициентов скрытого слоя нейронов	
1	-1,303	-3,417	-1,029	0,135	-0,731
2	0,370	3,821	0,148	-0,956	0,114
3	-2,649	3,642	-0,025	1,917	-1,655
4	0,567	2,547	-3,088	-0,196	-0,622
5	-1,655	-2,891	-1,822	0,127	-0,770
6	1,390	-0,384	-3,804	0,871	-0,439
7	1,415	-1,462	-2,346	-0,186	0,139
8	2,161	2,759	-1,663	0,351	-0,759
9	-1,178	-1,662	-3,690	1,238	-1,818
10	-1,419	2,375	-3,222	0,717	-1,012
11	-2,051	-2,361	2,866	-0,591	0,722
12	2,041	-1,788	2,896	-0,792	0,844
13	-4,571	-3,197	-0,625	-2,620	2,903
14	2,589	-2,685	-0,227	0,983	-0,247
15	2,316	4,122	-0,946	0,663	-1,161
16	2,865	0,232	2,928	-1,864	1,389
17	-3,911	-0,783	-1,499	0,141	-0,528
18	-1,324	-2,157	-4,723	-1,701	1,344
19	-2,510	2,215	1,807	-0,092	0,980
20	3,689	0,897	-1,279	0,765	-1,084

Рис. 9. Параметры нейросетевого классификатора

Отображение нейросетью полученных входных значений в признаковое пространство и решение задачи классификации продемонстрировано на рисунке 10.

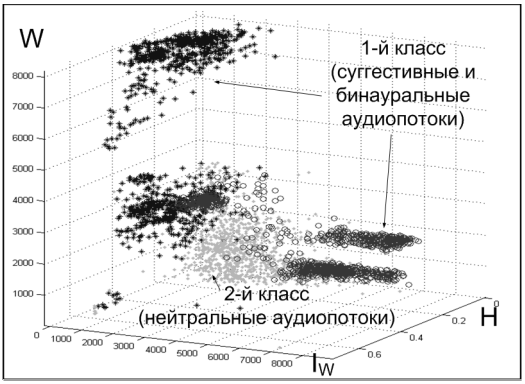


Рис. 10. Отображение нейросетью полученных входных значений в признаковое пространство

В приведенном примере веса выходных нейронов, соответствующих классам мультимедийного контента, приняли следующие значения: Y_1 (1-й класс, суггестивные мультимедийные объекты)=0,996;Y_2 (2-й класс, нейтральные мультимедийные объекты)=0,002, что полностью совпадает с априорной экспертной оценкой.

На рисунке 11 представлены результаты тестирования частных алгоритмов распознавания негативного контента в аудиопотоках на выборке общим объемом более 10000 фрагментов, содержавшей аудиообъекты суггестивного характера (зикры, мантры и т.д.), бинауральные воздействия, скрытые аудиовставки, а также аудиозаписи нейтрального содержания.

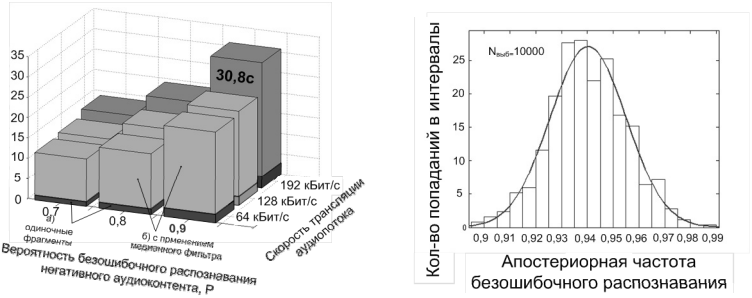


Рис. 11. Результаты тестирования частных алгоритмов распознавания негативного контента в аудиопотоках

При фиксированной апостериорной вероятности безошибочного распознавания P≥0,9 наибольшее время сбора и обработки информа-

ции ($\Delta t^{сб\&обр}$) относительно режима реального времени составило 30,8 с для аудиопотока в формате WAV качеством 192 кбит/с.

Практическая реализация разработанных алгоритмов была выполнена в виде аппаратно-программного комплекса, позволяющего осуществлять обнаружение потенциально вредоносных аудиоданных в различных форматах из Интернет-радиотрансляций, файлов с аудиоданными, звуковых дорожек файлов с видеоданными, звуковых дорожек Интернет-телевидения. Результаты фильтрации аудиопотока в режиме реального времени выводятся на экранную форму модуля обнаружения аудиовоздействий (рисунок 12).

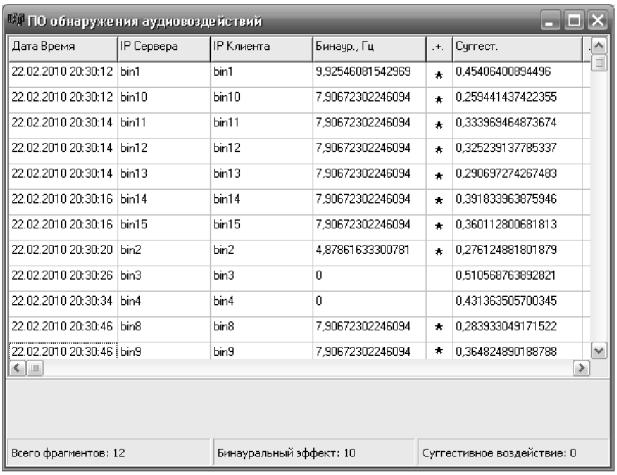


Рис. 12. Сигнализация об обнаруженных вредоносных свойствах аудиопотока

Столбцы экранной формы отражают информацию о времени и дате создания анализируемого файла, IP-адресе сервера и IP-адресе клиента между которыми осуществляется информационное взаимодействие, значение биоэффективной частоты (в Гц), на которой выявлено бинауральное воздействие, а также значение индекса суггестивности. В том случае, если распознаватель принимает решение о наличии в анализируемом аудиопотоке вредоносных вставок, то в столбце «.+.» таблицы, будет отображен символ «*».

2.4. Фильтрация текстов суггестивного содержания. Внушающее воздействие текста, помимо прочего, определяется особой упорядоченностью его элементов, сходной в некотором отношении со структурой фрактала. Подобная организация обеспечивает «естественность» восприятия и повышает эффективность внушения. Рассмотрим текст как линейно развертываемую во времени последовательность

фонетических единиц (звукобукв), каждая из которых может быть закодирована целым положительным числом. Для преобразования текста в численный ряд и его последующего компьютерного анализа могут быть использованы коды символов, определенные стандартом ASCII или Unicode.

Пусть текст представлен в виде целочисленного ряда длиной N. преобразуем его во временной ряд длиной N-1, исходя из логарифмических соотношений:

$$n=\ln\left(\frac{N_{i+1}}{N_i}\right),\ i=1,2,3,\ldots,N\text{-}1.$$

Среднее арифметическое для указанного ряда вычисляется по формуле:

$$M_k=\frac{1}{k}\cdot\sum_{i=1}^{k} n_i\ ,$$

а накопленные отклонения как:

$$D_{k,n}=\sum_{i=1}^{k}(n_j-M_k\ ),\ k=1,2,\ldots,i.$$

Тогда величина размаха определяется следующим образом:

$$R_k=\max\left(D_{k,n}\right)-\min\left(D_{k,n}\right),k\le n,$$

а среднеквадратическое отклонение:

$$S_k=\sqrt{\frac{1}{n}}\cdot\sum_{j=1}^{k}\left(n_j-m_k\right)^2\ .$$

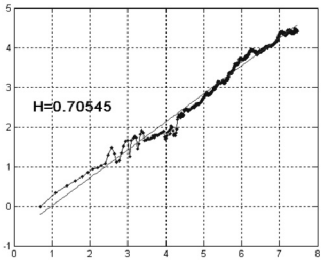
После этого каждый диапазон R_k нормируется путем деления на соответствующее значение S_k. Показатель Херста представляет собой тангенс угла наклона на графике зависимости $\ln(R_k/S_k)$ от $\ln(n)$.

В соответствии с проведенными экспериментальными исследованиями, для текстов, направленных на оказание внушения, значение оценки показателя Херста находится в области $0\le H\le 0{,}5$, что соответствует антиперсистентным рядам. Для нейтральных текстов информационного содержания значение оценки показателя Херста принадлежит области $0{,}5\ge H\ge 1$.

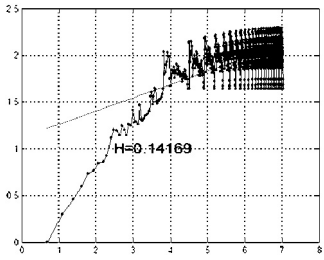
Помимо этого, графики, полученные в результате анализа суггестивных текстов методом нормированного размаха, обладают характерным общим свойством – наличием самоподобных периодически повторяющихся участков спад-подъем. Данный факт, а также низкие значения показателя Херста, являются отражением особой ритмической структуры текстов данного класса. Значения показателя Херста,

полученные в результате анализа внушающих текстов, тем ниже, чем короче формула внушения и чем чаще она повторяется в тексте.

Проведем сравнительный анализ текстов информационно-прагматического содержания: отрывка из романа «Отцы и дети» и лечебного заговора. На рисунке 13 представлены результаты оценки показателя Херста для указанных текстов.



а)



б)

Рис.13. Оценка показателя Херста для текстов нейтрального и суггестивного содержания: а) для отрывка из романа «Отцы и Дети» И.С.Тургенева. H=0,71; б) для лечебного заговора. H=0,14

Полученные значения показателя H для большинства исследованных текстов, априорно отнесенных к классу суггестивных, не превышают 0,5, что соответствует антиперсистентным рядам. Помимо этого, графики, полученные в результате анализа суггестивных текстов методом нормированного размаха, обладают характерным общим свойством – наличием самоподобных периодически повторяющихся участков спад-подъем. Данный факт, а также низкие значения показателя Херста, являются отражением особой ритмической структуры текстов данного класса. Значения показателя Херста, полученные в результате анализа подобных текстов, тем ниже, чем короче формула внушения и чем чаще она повторяется в тексте. Существенным является тот факт, что указанные свойства проявляются при анализе текстов внушающего воздействия независимо от языка, на котором они составлены.

Проведенные экспериментальные исследования на больших выборках текстов позволили сформулировать решающее правило для определения принадлежности анализируемого текста к классу потенциально опасной информации в следующем виде:

$$U=\begin{cases}1, \text{ если } 0\le H\le 0{,}49\\ 0, \text{ если } 0{,}49<H\le 1'\end{cases}$$

где H– оценка показателя Херста.

К достоинствам описанного подхода следует отнести возможность перехода от интуитивной качественной оценки эмоционального и суггестивного оценивания текста к количественному представлению, что позволяет выявлять в автоматизированном режиме потенциально вредоносные текстовые документы.

3. Заключение. Таким образом, многообразие форм представления данных в современных информационных средах, способных содержать негативный контент, исключает возможность поиска эффективного распознающего механизма в рамках одной отдельно взятой математической модели. Представленная система реализует многомодельный подход к обнаружению потенциально вредоносного воздействия в информационных потоках.

Четкая классификация с последующим дообучением в рамках разработанной технологии осуществляется на основе нейронных сетей, обладающих такими преимуществами, как высокая скорость обработки входной информации, устойчивая работа при наличии ошибочных данных, высокая результативность решения задач классификации даже при малых объемах обучающих выборок.

Частные смысловые распознаватели наличия негативного контента в мультимедийных потоках данных построены на основе адаптированных алгоритмов, а именно: для текстовых данных – фрактального и фоносемантического анализа, для видеофайлов – быстрого преобразования Фурье (БПФ) и анализа дифференциальной яркости кадров, для аудиоданных – БПФ, вейвлетного преобразования на основе вейвлета Добеши и метода нормированного размаха Херста.

Для подтверждения вредоносного информационно-психофизиологического воздействия на потребителей мультимедийного контента используется нечеткий классификатор на основе аппарата размытых множеств. Наибольшая скорость выявления потенциально вредоносных информационных объектов достигается на уровне применения дообучаемого нейросетевого классификатора, а возможность обнаружения неизвестных ранее типов вредоносных воздействий обеспечивается применением нечеткого классификатора на основе аппарата размытых множеств.

Разработанное на основе изложенных методов и алгоритмов программное обеспечение целесообразно применять в местах логического подключения локальных вычислительных сетей к каналам сети Интернет, что позволит своевременно обнаруживать и при необходимости блокировать потенциально опасные мультимедийные объекты, не допуская их воздействия на конечного пользователя.

Литература

1. Военная доктрина Российской Федерации. Утверждена 26.12.2014 г.
2. Юсупов Р.М. Наука и национальная безопасность // СПб: Наука. 2011. 376 с.
3. Пояснительная записка "К проекту Федерального закона "Об информационно-психологической безопасности". URL: http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=PRJ;n=42885 (дата обращения: 20.01.2015).
4. Победители Конкурса лучших инновационных проектов в сфере науки и высшего профессионального образования Санкт-Петербурга в 2012 году. URL: http://knvsh.gov.spb.ru/closedcontests/view/15/ (дата обращения: 25.01.2015).
5. *Гнидко К.О., Горемыкин Д.В.* Исследование фрактальных свойств внушающих текстов в интересах защиты оператора автоматизированной системы военного назначения от вредоносного информационно-психофизиологического воздействия // Труды Военно-космической академии имени А.Ф.Можайского. 2009. С. 91–92.
6. *Гнидко К.О.* Воздействие динамических фрактальных структур в мультимедийных объектах на функциональное состояние человека // Труды II Всероссийской научно-практической конференции молодых ученых и специалистов «Инновационные подходы к развитию вооружения, военной и специальной техники». М.: ВА ГШ. 2012.
7. *Зелинский С.А.* Информационно-психологическое воздействие на массовое сознание // СПб.: Издательско-Торговый Дом "Скифия". 2008. 280 с.
8. *Ткачева Л.О.* Воздействие фрактальных динамических изображений на функциональное состояние человека // Вестник СПбГУ. 2010. Т. 12. № 2. С. 378–387.
9. *Касперски К.* Тайные рычаги подсознания. Методы психовизуальной атаки // Хакер. 2007. № 9(105). С. 134–137.
10. *Крапивенко А.В.* Технологии мультимедиа и восприятие ощущений // М.: Бином. Лаборатория знаний. 2009. 271 с.
11. *Гнидко К.О., Ломако А.Г., Пономарев Ю.А.* Особенности структурной организации суггестивных аудиопотоков // Материалы XI Международной научно-практической конференции «Информационная безопасность». Ч1. Таганрог: ТИ ЮФУ. 2010. С. 203–205.
12. Рекламу Citroen запретили после эпилептического припадка у телезрителя. URL: http://medportal.ru/mednovosti/news/2012/01/19/epilepsy/ (дата обращения: 22.01.2015).
13. *Гнидко К.О. и др.* Обнаружение психоакустических воздействий на человека с учетом его индивидуальных физиологических особенностей // Материалы Всероссийской научно-практической конференции: «Теоретические и прикладные проблемы клинической психологии». СПб.: ЛГУ им.А.С.Пушкина. 2011. С. 421–426.
14. *Ростовцев Ю.Г., Гнидко К.О., Пилькевич С.В.* Генерация фрактальных сигналов заданной структуры для воздействия на функциональное состояние человека // Материалы IX межведомственной конференции «Научно-техническое и информационное обеспечение деятельности спецслужб». 2012. Т. 8. С. 58–63.
15. *Замарин А.И., Зайцев И.Е., Карелов И.Н.* Синтез нейросетевого алгоритма идентификации линейных корректирующих кодов // Известия ВУЗов – Приборостроение. 1998. Т. 41. № 8.

References

1. Voennaya doktrina Rossijskoj Federacii. [Military doctrine of Russian Federation]. 2014. (In Russ.).

2. Yusupov R.M. *Nauka i nacional'naja bezopasnost'* [Science and national security]. SPb: Nauka. 2011. 376 p.
3. Poyasnitel'naya zapiska «K proektu Federal'nogo zakona «Ob informacionno-psixologicheskoj bezopasnosti» [Explanatory note «To the draft of Federal Law on information-psychological security»]. Available at: http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=PRJ;n=42885 (accessed: 20.01.2015). (In Russ.).
4. Pobediteli Konkursa luchshix innovacionnyx proektov v sfere nauki i vysshego professional'nogo obrazovaniya Sankt-Peterburga v 2012 godu [Winners of the innovative projects competition in the field of science and higher professional education in Saint-Petersburg in 2012]. Available at: http://knvsh.gov.spb.ru/closedcontests/view/15/ (accessed: 25.01.2015). (In Russ.).
5. Gnidko K.O., Goremykin D.V. [Study of fraclal properties of suggestive texts in order to protect an operator of a military automated system from harmful information-psychological affects]. *Trudy Voenno-kosmicheskoj akademii imeni A.F.Mozhajskogo – Proceedings of Mozhaisky Military Space Academy*. Saint-Petersburg. 2009. pp. 91–92. (In Russ.).
6. Gnidko K.O. [Impacts of dynamic fractal structures in multimedia objects on the functional state of a person]. *Trudy II Vserossijskoj nauchno-prakticheskoj konferencii molodyh uchenyh i specialistov «Innovacionnye podhody k razvitiju vooruzhenija, voennoj i special'noj tehniki»* [Proceedings of II All-Russian scientific-practical conference of young scientists and specialists «Innovative approaches to the development of weapons, military and special equipment»]. M.: VA GSh. 2012 (In Russ.).
7. Zelinskij S.A. *Informacionno-psixologicheskoe vozdejstvie na massovoe soznanie* [Informational and psychological impact on the public consciousness]. SPb.: Izdatel'sko-Torgovyj Dom "Skifija". 2008. 208 p. (In Russ.).
8. Tkacheva L.O. [Impact of fractal dynamic images on the functional state of the person]. *Vestnik Sankt-peterburgskogo gosudarstvennogo universiteta – Herald of the St. Petersburg State University*. 2010. vol. 12. no. 2. pp 378–387. (In Russ.).
9. Kasperski K. [Secret levers of subconscious. Methods of psychovisual attack]. *Haker – Xaker*. 2007. vol. 9(105). pp. 134–137. (In Russ.).
10. Krapivenko A.V. *Texnologii mul'timedia i vospriyatie oshhushhenij* [Multimedia technology and the perception of sensations]. Moscow: Binom. Laboratoriya znanij, 2009. 271 p. (In Russ.).
11. Gnidko K.O., Lomako A.G., Ponomarev Yu.A. [Special features of the structural organization of suggestive audio streams]. *Materialy XI Mezhdunarodnoj nauchno-prakticheskoj konferencii «Informacionnaja bezopasnost'». Ch. 1* [Proceedings of the XI International scientific-practical conference. «Information Security». Part 1]. Taganrog: TI SFU. 2010. pp. 203–205. (In Russ.).
12. Reklamu Citroen zapretili posle e'pilepticheskogo pripadka u telezritelya [Citroen commercial banned after an epileptic seizure of the viewer]. Available at: http://medportal.ru/mednovosti/news/2012/01/19/epilepsy/ (accessed: 22.01.2015). (In Russ.).
13. Gnidko K.O. et al. [Detection of psychoacoustic effects on humans, taking into account their individual physiological characteristics]. *Materialy Vserossijskoj nauchno-prakticheskoj konferencii: «Teoreticheskie i prikladnye problemy klinicheskoj psihologii»* [Proceedings of the All-Russian Scientific-Practical Conference «Theoretical and applied problems of clinical psychology SPb.: LGU im.A.S.Pushkina. 2011. pp. 421–426. (In Russ.).
14. Rostovcev Yu.G., Gnidko K.O., Pil'kevich S.V. [Generation of fractal signals of given structure to influence the functional state of the person]. *Materialy IX*

mezhvedomstvennoj konferencii «Nauchno-tehnicheskoe i informacionnoe obespechenie dejatel'nosti specsluzhb» [Proceedings of the IX interdepartmental conference «Scientific and technical and information support of special services»]. 2012. vol. 8. pp. 58–63. (In Russ.).

15. Zamarin A.I., Zajcev I.E., Karelov I.N. [Synthesis of neural network algorithm for identification of linear error-correcting codes]. *Izvestiya VUZov Priborostroenie – Proceedings of the universities instrument engineering*. 1998. vol. 41. no. 8. (In Russ.).

Гнидко Константин Олегович — к-т техн. наук, докторант, Военно-космическая академия имени А.Ф. Можайского. Область научных интересов: информационно-психологическая безопасность, распознавание образов, извлечение знаний из неструктурированных массивов данных. Число научных публикаций — 27. greeny598@gmail.com; ул. Ждановская, 13, 197198, г. Санкт-Петербург; р.т.: +7(812) 237-19-60.

Gnidko Konstantin Olegovich — Ph.D., doctoral student, Mozhaisky Military Space Academy. Research interests: information-psychological security, image recognition, data mining. The number of publications — 27. greeny598@gmail.com; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812) 237-19-60.

Ломако Александр Григорьевич — д-р техн. наук, профессор кафедры систем сбора и обработки информации, Военно-космическая академия имени А.Ф. Можайского. Область научных интересов: информационная безопасность, теоретическое и системное программирование, синтез и верификация корректности моделей программ. Число научных публикаций — 250. lomako_ag@mail.ru; ул. Ждановская 13, 197198, Санкт-Петербург; р.т.: +7(812) 237-19-60

Lomako Aleksandr Grigor'evich — Ph.D., Dr. Sci., professor of system for collecting and processing information department, Mozhaisky Military Space Academy. Research interests: information security, theoretical and system programming, synthesis and verification of program models. The number of publications — 250. lomako_ag@mail.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812) 237-19-60.

РЕФЕРАТ

Гнидко К.О., Ломако А.Г. **Контроль потенциально опасного информационно-психологического воздействия на индивидуальное и групповое сознание потребителей мультимедийного контента.**

В современных подходах к обеспечению системной безопасности наблюдается существенный перекос в сторону технических элементов защищаемых систем. Вместе с тем, международный масштаб и стратегический характер приобретает проблема защиты человеческой психики от потенциально вредоносного мультимедийного контента.

В настоящей работе кратко изложены результаты теоретических и практических исследований в области построения автоматизированной системы мониторинга и фильтрации мультимедийных данных от контента, способного нанести вред психофизиологическому состоянию пользователей.

Представленная система реализует многомодельный подход к обнаружению потенциально опасных объектов в аудио-, видеопотоках и текстовых массивах. Для оперативного обнаружения сигнатур вредоносных объектов применяется дообучаемый нейросетевой классификатор прямого распространения. Частные смысловые распознаватели построены на основе алгоритмов фрактального и фоносемантического анализа, быстрого преобразования Фурье и анализа дифференциальной яркости кадров, вейвлетного преобразования и метода вычисления нормированного размаха. Для подтверждения факта вредоносного информационно-психофизиологического воздействия на потребителей используется нечеткий классификатор на основе аппарата размытых множеств.

Разработанное на основе изложенных методов и алгоритмов программное обеспечение может применяться в местах логического подключения локальных вычислительных сетей к каналам сети Интернет, что позволит своевременно обнаруживать и при необходимости блокировать потенциально опасные мультимедийные объекты, не допуская их воздействия на конечного пользователя.

SUMMARY

Gnidko K.O., Lomako A.G. **Monitoring of Potentially Dangerous Information-Psychological Affect on Individual and Group Consciousness of Multimedia Content Consumers.**

In the modern approach to system safety ensuring there is a significant bias towards technical elements of the systems being protected. At the same time, the problem of protection of the human psychics from potentially harmful media content currently has the international scope and strategic nature.

This paper summarizes the results of theoretical and practical study in the field of developing of automated systems for monitoring and filtering of multimedia data from the content that can harm psychophysiological states of its consumers.

This system implements a multi-model approach to the detection of potentially dangerous objects in the audio, video streams and texts. The direct distribution neural network classifier is used for prompt detection of malicious objects' signatures. Particular content-focused detectors use algorithms based on fractal methods, phonosemantic analysis, fast Fourier transform, differential frame brightness, wavelet transform and the method of computation of the normalized amplitude. A classifier based on fuzzy sets does confirmation of the fact of harmful effect.

The software developed on the basis of the methods and algorithms described can be used in points of logical connection of local area networks to the Internet network channels. It will enable timely detection and, if necessary, blocking of potentially dangerous multimedia objects, avoiding their impact on the end user.