На правах рукописи

ЧЕРНОПРУДОВА Елена Николаевна

ЗАЩИТА ПОЧТОВЫХ СЕРВИСОВ ОТ НЕСАНКЦИОНИРОВАННЫХ РАССЫЛОК НА ОСНОВЕ КОНТЕНТНОЙ ФИЛЬТРАЦИИ ЭЛЕКТРОННЫХ СООБЩЕНИЙ

Специальность: 05.13.19
Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени кандидата технических наук

Уфа 2013

Работа выполнена на кафедре программного обеспечения вычислительной техники и атоматизированных систем ФГБОУ ВПО «Оренбургский государственный университет»

Научный руководитель: доктор технических наук, профессор
СОЛОВЬЕВ Николай Алексеевич,
Оренбургский государственный университет, кафедра программного обеспечения вычислительной техники и автоматизированных систем

Официальные оппоненты: доктор технических наук, доцент
МАШКИНА Ирина Владимировна,
Уфимский государственный авиационный технический университет, кафедра вычислительной техники и защиты информации

кандидат технических наук, доцент
БОРОВСКИЙ Александр Сергеевич,
Оренбургский государственный аграрный университет, институт управления рисками и комплексной безопасности

Ведущая организация: ФГБОУ ВПО «Оренбургский государственный институт менеджмента»

Защита состоится «24» декабря 20_13_ г. в 12 часов на заседании диссертационного совета Д-212.288.07 при Уфимском государственном авиационном техническом университете по адресу: 450000, Уфа-центр, ул. К.Маркса, 12.

С диссертацией можно ознакомиться в библиотеке университета.

Автореферат разослан «__» ноября 2013 г.

Ученый секретарь
диссертационного совета
д-р техн. наук, доцент **И.Л.Виноградова**

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

Почтовые сервисы информационно-телекоммуникационных систем (ИТКС) корпоративных предприятий с территориально-распределенной структурой являются средством документооборота и служебной переписки, важнейшим информационным каналом реализации бизнес-процессов. Одной из проблем использования электронной почты становится массовая рассылка несанкционированных электронных сообщений (НЭС) субъектами коммерческой или иной информации. Отсюда, противодействие НЭС становится актуальной задачей обеспечения информационной безопасности (ИБ) ИТКС.

Степень разработанности темы. Проблемам обеспечения ИБ электронной почты посвящены работы Валеева С.С., Васильева В.В., Зегжда П.В., Машковой И.В., Семеновой М.А., Шварца А.А. и зарубежных исследователей B. Pfahringer, K. Junejo, D. Zhou и других. Обобщая результаты исследований, можно сделать вывод, что в настоящее время сложилась система методов, моделей и средств фильтрации НЭС, позволяющая решать широкий спектр задач ИБ.

Вместе с тем, лавинообразный рост интенсивности НЭС, изменения способов их доставки приводят к ложной классификации контента и, что особенно важно, к частичной потере легитимных сообщений. Кроме того, известные методы фильтрации НЭС идентифицируют спам-рассылки и не учитывают изменяющиеся потребности адресатов служебной корреспонденции. Поэтому развитие методов защиты электронной почты остаётся актуальной задачей научных исследований в области ИБ, **объектом** которых становится защита почтовых сервисов ИТКС от НЭС, **предмет** – методы, модели и средства контентной фильтрации легитимной корреспонденции электронной почты; **границы** исследований – почтовые сервисы ИТКС корпоративных предприятий с территориально-распределенной структурой.

Системный анализ ИБ электронной почты от НЭС выявил ряд противоречий между требованиями практики и состоянием теории спам-фильтрации, основным из которых становится противоречие между существенно возросшей интенсивностью спам-рассылок при наличии ложной классификации и отсутствием методов идентификации легитимной почтовой корреспонденции с учетом изменяющихся потребностей адресатов, работающих в реальном масштабе времени. Отсюда, **целью исследования** становится *повышение достоверности идентификации легитимной почтовой корреспонденции на основе семантической подготовки электронных сообщений к интеллектуальной фильтрации и нейросетевой классификации в условиях изменяющегося контента служебной переписки*.

Задачи исследования:

1. Системный анализ защиты почтовых сервисов ИТКС предприятий с территориально-распределенной структурой.

2. Разработка модели электронного почтового сообщения, учитывающей семантику контента почтовой корреспонденции.

3. Разработка методики и алгоритмов фильтрации легитимных электронных сообщений почтовых сервисов в условиях изменяющихся интересов адресатов служебной корреспонденции.

4. Разработка средств фильтрации легитимной корреспонденции почтовых сервисов ИТКС корпоративных предприятий.

5. Проведение экспериментальной проверки почтовых сервисов со средствами фильтрации служебной корреспонденции и оценка их эффективности.

Научная новизна работы

1. Разработана модель электронного сообщения для средств защиты почтовых сервисов ИТКС, отличающаяся от известных:

– применением меры значимости термов в качестве веса признаков для описания электронных почтовых сообщений (ЭПС), *позволяющей* устранить эффект больших различий в частотах фиксации термов;

– методикой определения меры значимости терма в рамках одного сообщения, *позволяющей* сократить пространство признаков за счет исключения термов с малой информативной нагрузкой;

– методом выделения устойчивых словосочетаний, позволяющей усилить смысловое содержание термов и сократить пространство признаков за счет использования дополнительных мер близости между термами в сообщении и тесноты взаимосвязи между ними.

2. Предложена методика и алгоритмы контентной фильтрации электронной почтовой корреспонденции на основе нейросетевого классификатора ART2a, отличающиеся наличием дополнительного нейрона, обеспечивающего определение меры сходства входящего сообщения с экземплярами обучающей выборки при отнесении сообщения к классу НЭС для исключения ложной классификации легитимной корреспонденции.

Практическая значимость исследования заключается в развитии системного программного обеспечения средств защиты почтовых сервисов ИТКС, обеспечивающего повышение достоверности идентификации легитимных почтовых сообщений в условиях меняющегося контента служебной переписки.

Полученные в ходе исследований **результаты реализованы** в программном обеспечении почтовых сервисов «Интеллектуальная система фильтрации несанкционированных рассылок», **зарегистрированы** в Федеральной службе по интеллектуальной собственности, патентам и товарным знакам, **подтверждены** актами внедрения ООО «ТБинформ» (г. Оренбург) и ФГБОУ ВПО «Оренбургский государственный университет».

Методология и методы исследования. Решения поставленных задач в рамках проводимых исследований опирается на теоретические осно-

вы защиты информации, теорию принятия решений, методы лингвистического анализа текста и контент-анализа, методы искусственного интеллекта, методы теории вероятностей и математической статистики, методы теории эксперимента и оценки эффективности программных систем.

Положения, выносимые на защиту

1. Результаты системного анализа защиты почтовых сервисов информационно-телекоммуникационных систем корпоративных предприятий с территориально-распределенной структурой позволили выявить основные признаки электронных почтовых сообщений, необходимые для классификации электронных рассылок.

2. Модель электронного сообщения на основе устойчивых словосочетаний, отличающаяся использованием дополнительных мер, повышающих семантическую нагрузку термов при сокращении пространства признаков для классификации легитимной почтовой корреспонденции.

3. Методика и алгоритм контентной фильтрации электронной почтовой корреспонденции на основе нейросетевого классификатора ART2a, отличающиеся введением дополнительного нейрона для проверки достоверности отнесения сообщений к классу несанкционированных рассылок на основе меры сходства векторов, позволяющие исключить ложную классификацию легитимных сообщений.

4. Прототип системы защиты почтовых сервисов, основанный на двухуровневой фильтрации почтовых сообщений, отличающийся предварительной подготовкой электронных почтовых сообщений к нейросетевой классификации, обеспечивающий контентную фильтрацию легитимной корреспонденции почтовых сервисов в реальном масштабе времени.

5. Результаты экспериментального исследования эффективности защиты почтовых сервисов от несанкционированных рассылок, основанные на контентной фильтрации электронных сообщений, позволяющей исключить потерю легитимной корреспонденции.

Апробации, публикации.

Научные и практические результаты исследований обсуждались и получили одобрение на VIII, IX, X всероссийских научно-технических конференциях (с международным участием), Оренбург (2009-2013 гг.); международной молодежной конференции, Дрезден-Розендорв, Германия, Уфа, Россия (2010 г.); на Всероссийской научной школа, Воронеж (2011 г.); на конкурсе научно-исследовательских работ студентов, аспирантов и молодых ученых «ЭВРИКА-2011», Новочеркасск; на конкурсе научно-исследовательских работ «IT-Security Conference for the Next Generation» (Москва-Мюнхен, 2011), ЗАО «Лаборатория Касперского» (Диплом II степени); на областной выставка научно-технического творчества молодежи «НТТМ-2010», «НТТМ-2011», г. Оренбург (Сертификат).

Основные результаты исследований опубликованы в 10 печатных работах, две из которых – в издании, определенном ВАК России для опубликования научных результатов диссертаций на соискание ученых степе-

ней доктора и кандидата наук, в одном свидетельстве о государственной регистрации программ.

Структура и объем диссертации. Работа состоит из введения, четырех глав, заключения, изложенных на 126 страницах и 2 приложений, содержит 54 рисунка и 11 таблиц. Список использованных источников включает 107 наименований.

КРАТКОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во **введении** обоснована актуальность работы, определенны объект, предмет и границы исследования, сформулированы цель и задачи исследования, представлено основное содержание научной работы и положения, выносимые на защиту.

В первой главе проведен системный анализ информационных процессов эксплуатации почтовых сервисов корпоративных предприятий с территориально-распределенной структурой, позволивший выявить противоречия меду практикой и теорий обеспечения защиты почтовых сервисов ИТКС, представленные на рисунке 1.

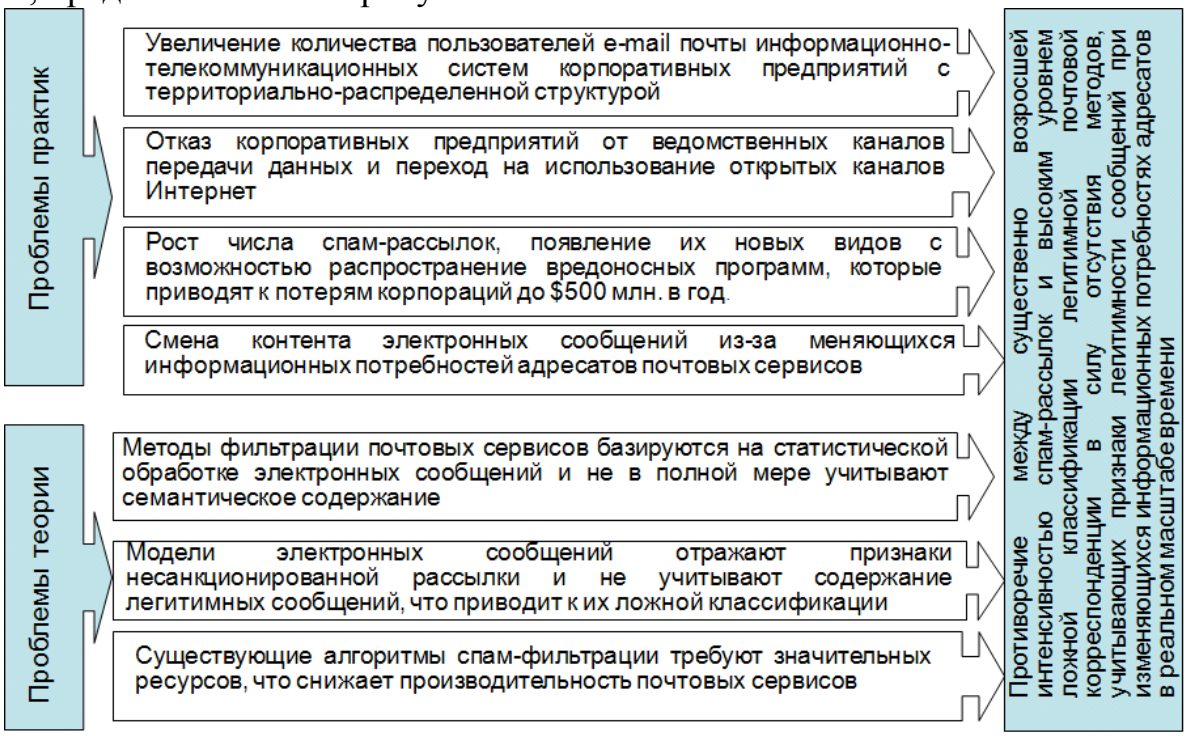

Рисунок 1 – Основные противоречия между практикой и теорией обеспечения защиты почтовых сервисов корпоративных предприятий

На основе анализа результатов эксплуатации почтовых сервисов системы электронного документооборота и управления взаимодействием (СЭДУВ) DIRECTUM определены признаки служебной электронной корреспонденции. Признаки ЭПС разделены на формальные (адреса отправителя, размер и формат сообщения, список получателей, IP-адреса и т.п.) и семантические, отражающие содержание сообщения (термы). Под термом понимаются символы (слова) сообщения, отделенные между собой пробе-

лом. Исследования существующих средств спам-фильтрации электронных сервисов СЭДУВ показали наличие ложной классификации легитимных ЭПС с вероятностью 0,08-0,12, что недопустимо для служебной переписки. При этом спам-рассылки идентифицируются с вероятность 0,7-0,8.

На основе анализа методов, применяемых для контроля потока ЭПС, выявлены причины ложной классификации – используемые методы фильтрации ЭПС ориентированы на идентификацию спам-рассылок и не учитывают смысловое содержание легитимных ЭПС, которое, в свою очередь, изменяется в процессе служебного переписки.

Анализ эксплуатации СЭДУВ DIRECTUM показал, что экономические потери от ложной классификации ЭС становятся соизмеримы с ценой сопровождения почтовых сервисов.

Таким образом, системный анализ проблем обеспечения ИБ почтовых сервисов ИТКС позволил выявить основное противоречие между существенно возросшей интенсивностью спам-рассылок и высоким уровнем ложной классификации легитимной почтовой корреспонденции в силу отсутствия методов, учитывающих признаки легитимности при меняющемся контенте сообщений адресатов служебной корреспонденции.

Снизить процент ложных срабатываний предлагается и за счет разработки двухуровневой системы фильтрации, представленной на рисунке 2, и состоящей из формального и интеллектуального фильтров.

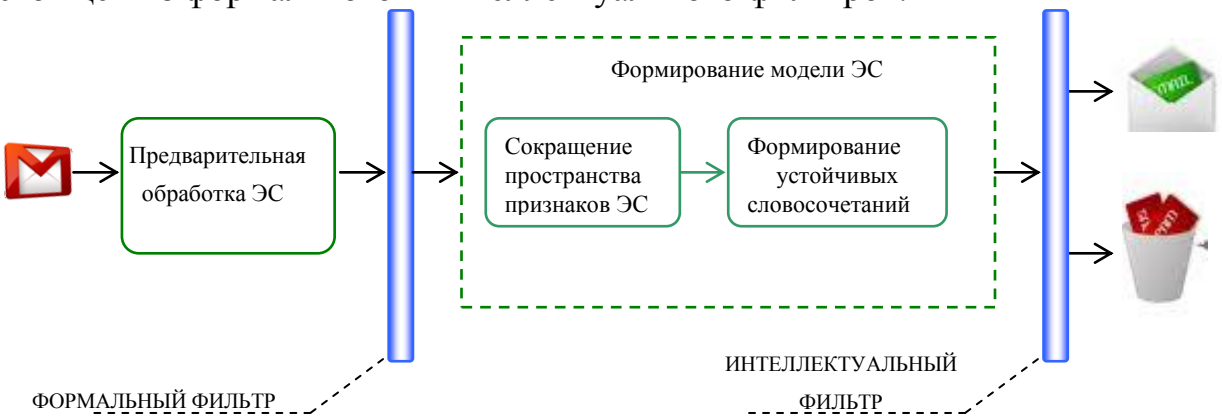

Рисунок 2 – Технология фильтрации ЭПС

Предварительная обработка ЭПС заключается в приведении к стандартному типу кодировки, удалению стоп-слов, гиперссылок и знаков пунктуации. Формальный фильтр использует адреса (IP, e-mail), разделяющие ЭПС на разрешенные и запрещенные, и представляет собой базу данных признаков, формируемую администратором.

Интеллектуальный фильтр осуществляет семантическую классификацию ЭПС конкретного адресата почтовой корреспонденции, что требует предварительного обучения классификатора.

Определена и формализована задача фильтрации ЭПС в средствах защиты почтовых сервисов ИТКС корпоративного предприятия.

Пусть $L \in \{Let_i\}$ множество ЭПС, предназначенных для обучения классификатора. Модель L характеризуется пространством признаков

$P=(p_1,p_2,p_3,....p_l)$, где p_l – значение l-го признака ЭПС; А – алгоритм классификации, относящий L к одному из классов $K \in \{k_1,k_2\}$ (spam/legitim).

Задача фильтрации заключается в построении такого решающего правила, при котором классификация проводится с минимальным числом ошибок R в реальном масштабе времени.

Тогда процедура автоматической фильтрации P_f L ЭПС на множестве классов K примет вид целевой функции

$$R(L(p_l),A(k_j)) \xrightarrow{P_f} \min$$

Таким образом, результаты системного анализа защиты почтовых сервисов информационно-телекоммуникационных систем корпоративных предприятий с территориально-распределенной структурой свидетельствуют о необходимости развития существующих методов фильтрации ЭПС на основе моделей, отражающих семантику почтовой корреспонденции с изменяющимся контентом для исключения ложной классификации легитимных ЭПС.

Во второй главе предложена модель ЭПС, основанная на матричном представлении контента сообщений, отличающаяся от известных векторных моделей текста сокращенным пространством признаков, обеспечивающих семантическую классификацию ЭПС в реальном масштабе времени.

На основе исследований способов описания текстов принято решение об использовании модели ЭПС в векторном пространстве признаков, которое отражает содержание сообщения L_i с помощью термов t, множество которых образует тезаурус $T_k \subseteq \{t_1,....t_j\}$, $j=\overline{1,n}$ определенного класса k *(spam/legetim)*. Каждому терму $t \in T$ в L_i ставится в соответствии вес w_{tj}, т.е. $L_i = (w_{1j},...,w_{tj})$.

Существующие меры взвешивания термов текстовых сообщений формируются из весовых коэффициентов w_{ij}, равных частоте термов f_{ij} в сообщении. Для устранения эффекта больших различий в частотах термов (высокочастотные, среднечастотные и низкочастотные) предложено в качестве меры взвешивания термов использовать величины частот в логарифмическом масштабе по зависимости вида:

$$w_{ij} = \frac{\log_2(f_{ij}+1)\log_2\left(\frac{M}{M_j}\right)}{\sqrt{\sum_{ij=1}^{N}\left[\log_2(f_{ij}+1)\log_2\left(\frac{M}{M_j}\right)\right]^2}} \qquad (1)$$

где M – общее число сообщений в выборке;

N – число термов в выборке после удаления стоп-слов;

M_j – общее число сообщений, содержащих терм t_j.

Отсюда, модель ЭПС можно описать матрицей признаков, столбцами которой будут признаки сообщений служебной корреспонденции, а строками признаки их термов:

$$L_k = \begin{bmatrix} w_{11} & w_{21} & \cdots & w_{M1} \\ w_{12} & w_{22} & \cdots & w_{M2} \\ \vdots & \vdots & \vdots & \vdots \\ w_{1j} & w_{2j} & \cdots & w_{Mj} \\ \vdots & \vdots & \vdots & \vdots \\ w_{1N} & w_{2N} & \cdots & w_{MN} \end{bmatrix} \quad (2)$$

где $i=1,\ldots,M,\ j=1,\ldots N,\ k=2$

Однако матрица признаков легитимных ЭПС служебной корреспонденции имеет размерность, обработка которой требует значительных вычислительных ресурсов и временных затрат.

Для сокращения (2) использован закон Ципфа, согласно которому слова, встречающиеся в тексте чаще других, являются малоинформативными, не имеющими решающего смыслового значения. Это послужило основанием снижения размерности матрицы признаков за счет избавления от малоинформативных термов. Предложено для каждого терма в L определенного класса использовать величину $RF_{t_j}^k$, характеризующую значимость терма для класса k логарифм отношения числа ЭПС, содержащих t_j-ый терм и относящихся к классу k к числу ЭПС, содержащих t_j-ый терм и не относящихся к классу k. Установлено, что термы, у которых $RF_{t_q}^k \le 1{,}5$, без потери смыслового содержания, можно исключить из соответствующего класса k. В результате использования $RF_{t_j}^k$ пространство анализируемых термов сокращается на 14%.

Размерность матрицы (2) определенного класса уменьшится, если помимо последовательности термов и их значимости учесть связи между термами, т.е. выделить ключевые термы, которые отражают смысловую специфику ЭПС.

Пусть $D_i=\{d_{jq}\}$, $_{j=1,\ldots,N}$ характеристика связи между термами в i-ом сообщении, а d_{jq} – мера смысловой близости t_j-го и t_q-го термов.

В качестве меры близости между термами d_{jq} в сообщении предложено использовать расстояние Дайса. Данная статистическая мера позволяет объединить термы в устойчивые (ключевые) словосочетания, характеризующие семантическое содержание сообщений.

Мера d_{jq} рассчитывается по зависимости вида:

$$d(t_j,t_q) = \log_2\left(\frac{2*(f(t_j,t_q))}{f(t_j)+f(t_q)}\right) \quad (3)$$

где $f(t_j)$ и $f(t_q)$ – частота встречаемости термов t_j и t_q в сообщении,

$f(t_j,t_q)$ – частота совместной встречаемости термов t_j и t_q.

Предложена методика формирования устойчивых словосочетаний:

1) выделение значимых термов с учетом $RF_{t_j}^k$ для соответствующего класса k (spam/legitim);

2) расчет мер близости термов и принятие решения о формировании устойчивого словосочетания;

3) подтверждение смысловой значимости словосочетания.

Решение о формировании устойчивого словосочетания для каждой пары термов принимается, если значение d_{jq} равное или выше, чем в соседних парах термов (левой и правой).

Для подтверждения смысловой значимости полученных устойчивых словосочетаний оценивается теснота взаимосвязи между термами в словосочетании, метрикой которой могут выступать меры ассоциации K_a или контингенции K_k, которые определяются по зависимостям:

$$K_a = \frac{ad - bc}{ad + bc}\ ,\ K_k = \frac{ad - bc}{\sqrt{(a+b)(b+d)(a+c)(c+d)}} \qquad (4)$$

где a – число ЭПС, имеющих терм t_j, который встречается в классе k;

b – число ЭПС , в которых терм t_j встречается с другим классом;

c – число ЭПС, имеющих терм t_q, который встречается в классе k;

d – число ЭПС, в которых терм t_q встречается с другим классом.

Экспериментально установлено, что связь между элементами словосочетания считается подтвержденной, если $K_a \geq 0{,}5$ или $K_к \geq 0{,}3$.

В результате использования устойчивых словосочетаний пространство анализируемых термов сокращается на 10-15%.

Обобщенно модель ЭПС предлагается представлять в виде:

$$L(p_i) = <T^k, w^*(t_j)>,$$

где T^k – терм устойчивых словосочетаний в сообщении;

$w^*(t_j)$ – вес терма после сокращения матрицы признаков (2).

Таким образом, предложена модель ЭПС, основанная на матричном представлении контента сообщений, отличающаяся от известных векторных моделей текста сокращенным пространством признаков без потери их смыслового содержания, обеспечивающих контентную классификацию ЭПС в реальном масштабе времени.

В третьей главе предложена методика и разработан алгоритм контентной фильтрации ЭПС, обеспечивающий исключение ложной классификации легитимной корреспонденции.

Методика контентной фильтрации ЭПС, основанная на адаптивной нейронной сети ART2a, отличается введением дополнительного нейрона для проверки достоверности отнесения сообщений к классу несанкционированных рассылок за счет определения меры сходства по коэффициенту Жаккара, которая определяется отношением числа общих термов, встре-

чаемых во входном сообщении и сообщении хранящимся в обучающей выборке сообщении, к разности между суммой числа термов входного сообщения с числом термов эталона, хранящегося в базе, и числом общих термов.

Входной слой нейронной сети содержит столько нейронов, сколько термов в ЭПС, элементами которого являются значения весов термов $w^*(t_j)$. Слой распознавания представляет собой набор нейронов (тезаурус Т), каждый из которых отвечает за один экземпляр класса.

Алгоритм функционирования нейросетевого классификатора представлен на рисунке 3.

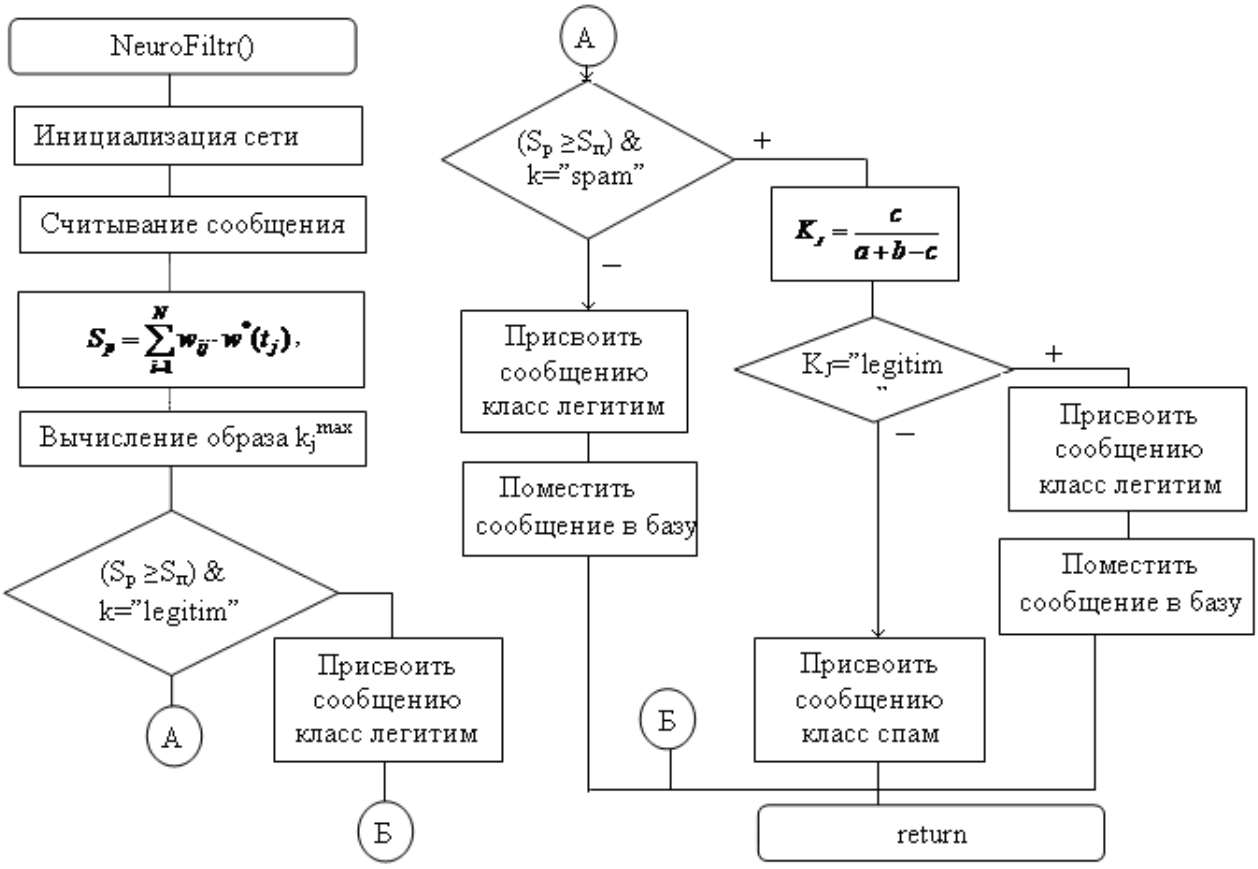

Рисунок 3 – Укрупненный алгоритм нейросетевого классификатора

Отнесение ЭПС к соответствующему классу реализовано блоком сравнения скалярного произведения векторов ***Sp*** с пороговым значением ***Sn,*** определяемым экспериментально. Установлено, что при дополнении к скалярному произведению векторов ***Sp*** меры сходства $K_J \geq 0{,}8$ решение об отнесении сообщения к классу спам отвергается.

Функциональная модель программного проекта прототипа средств фильтрации легитимных ЭПС почтовых сервисов ИТКС представлена на рисунке 4.

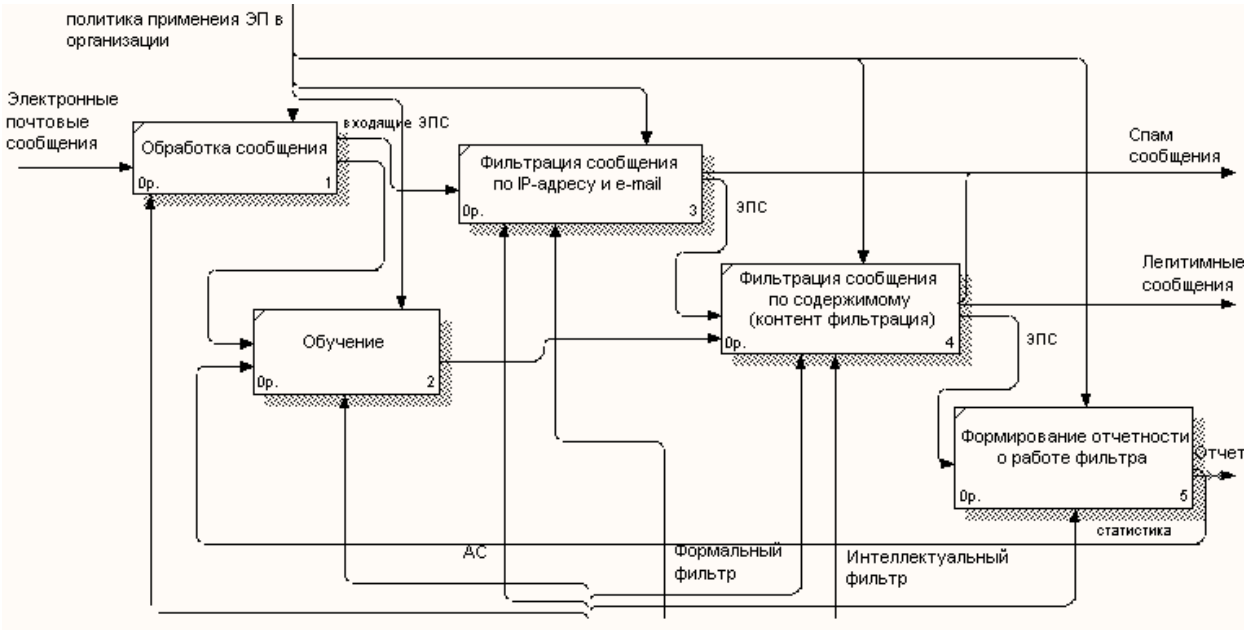

Рисунок 4 - Функциональная модель прототипа системы фильтрации

Архитектура прототипа системы защиты почтовых сервисов ИТКС на основе спам-фильтра представлена на рисунке 5.

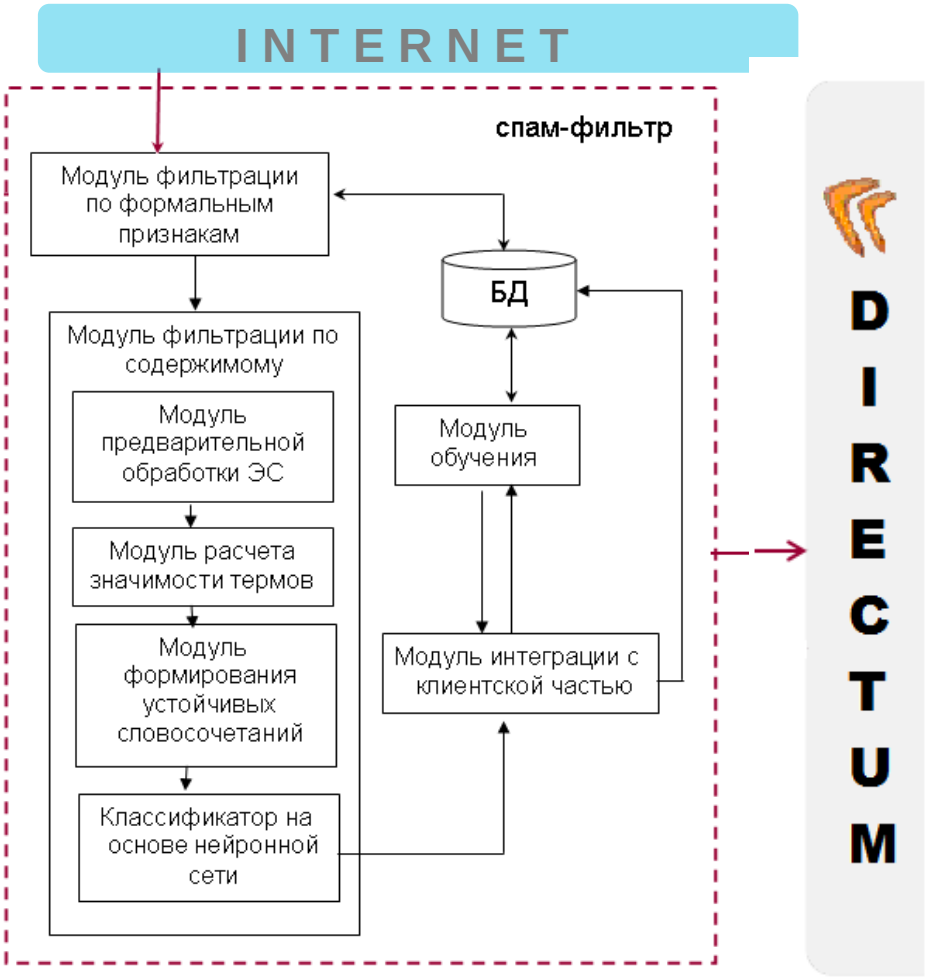

Рисунок 5 – Архитектура прототипа системы защиты почтовых сервисов ИТКС на основе спам-фильтра

Таким образом, предложена методика и разработан алгоритм контентной фильтрации электронных почтовых сообщений, основанные на адаптивной нейронной сети ART2a, отличающиеся использованием дополнительного нейрона внутреннего слоя для проверки достоверности отнесения сообщений к классу несанкционированных рассылок на основе расчета меры сходства векторов, позволяющей исключить ложную классификацию легитимных сообщений. Кроме того, предложен прототип системы защиты почтовых сервисов, основанный на двухуровневой фильтрации почтовых сообщений, отличающийся предварительной подготовкой электронных сообщений к нейросетевой классификации, обеспечивающий контентную фильтрацию легитимной корреспонденции почтовых сервисов в реальном масштабе времени.

В четвертой главе предложена методика оценки эффективности системы фильтрации легитимных ЭПС и проведен имитационный эксперимент, позволивший обосновать порог соответствия входящего ЭПС классу legitim.

В основу обоснования порога соответствия входящего ЭПС определенному классу положен метод теории статистических решений в форме проверки двухальтернативной гипотезы $\boldsymbol{H_0}$ и $\boldsymbol{H_1}$, отражающих предположения о легитимности ЭПС или наличии НЭС. Для того чтобы данная задача обрела математическую содержательность введены показатели эффективности средств фильтрации ЭПС – ошибки классификации. Ошибка первого рода $\boldsymbol{\alpha}$ определяется вероятностью принять решение о легитимности сообщения, когда оно ложно. Ошибка второго рода $\boldsymbol{\beta}$ – вероятность отвергнуть решение о легитимности сообщения, когда оно легитимно.

Показателями качества фильтрации приняты мера полноты (*precision*), оценивающая долю правильной классификации относительно всех объектов определенного класса, мера точности (*recall*), оценивающая долю верной классификации относительно всех объектов, и сводная оценка качества классифицирования (*F-мера*).

Оценка предложенных средств фильтрации ЭПС проводилась на основе имитационного эксперимента, схема которого представлена в правой части рисунка 6, по методу k–подмножеств, алгоритм которого представлен в левой части рисунка 6. Сущность подхода заключается в разбиении экспериментальной выборки на k равных частей. В результате каждого запуска средств фильтрации фиксировались: средние значения двух вероятностных характеристик – α и β, сводная оценка качества классификации (F-мера), меры полноты и точности.

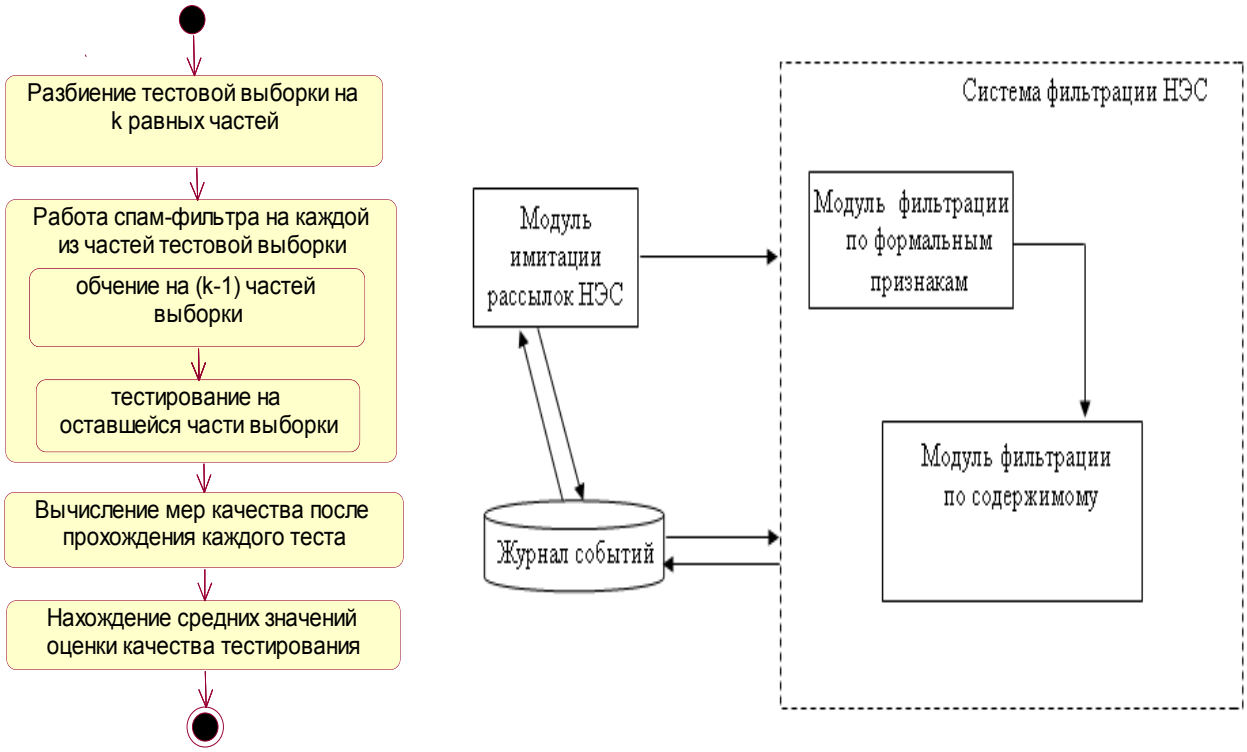

Рисунок 6 – Схема имитационного эксперимента и алгоритм оценки результатов методом k–подмножеств (k-foldes)

В процессе эксперимента исследованы несколько версий средств фильтрации, представленные в таблице 1. Для каждой из версии изменялся порог на соответствие классу S_n. Экспериментальная выборка ЭПС для оценки эффективности прототипа системы фильтрации состояла из легитимных сообщений документооборота и спам-рассылок. Тематика сообщений экспериментальной выборки представлена в таблице 2. Всего исследовано 908 ЭПС (424 легитимных сообщений и 484 спам-сообщений) и осуществлено 13 запусков прототипа предложенной системы фильтрации ЭПС. Порог соответствия S_n изменялся в диапазоне от 0,4 до 0,9.

Таблица 1 – Варианты построения классификатора

Название	Модель ЭС	Вес	Метод сокращения признакового пространства	Выделение устойчивых словосочетаний	Алгоритм классификации
Met1	векторная	Tf-idf	RF	+	нейрон. сеть Art
Met2	векторная	Ltc	RF	+	нейрон. сеть Art
Met3	векторная	Ltc	RF	-	нейрон. сеть Art
Met4	векторная	Ltc	IG	+	нейрон. сеть Art
Met5	векторная	Tf-idf	IG	+	нейрон. сеть Art

Таблица 2 – Тематика сообщений

№	Вид сообщения	Тематика сообщений
1	Спам сообщения	“Пустые” сообщения, содержащие только ссылки или вложения.
2	Спам сообщения	Реклама товаров
3	Спам сообщения	Реклама услуг (юридических, бухгалтерских, строительных, образовательных, туристических, медицинских и проч.)
4	Спам сообщения	Приглашения на курсы, предложения схем «отмывания» денег («нигерийские» письма)
5	Легитимные сообщения	Деловая переписка (свободная форма)
6	Легитимные сообщения	Деловая переписка (приказы, распоряжения, отчеты и т.п.)
7	Легитимные сообщения	Приглашения на участия в грантах, конференциях, выставках и т.п.

На рисунке 7 представлены показатели эффективности версий и сравнительные результаты оценки предложенного фильтра легитимных ЭПС при различных значениях порога.

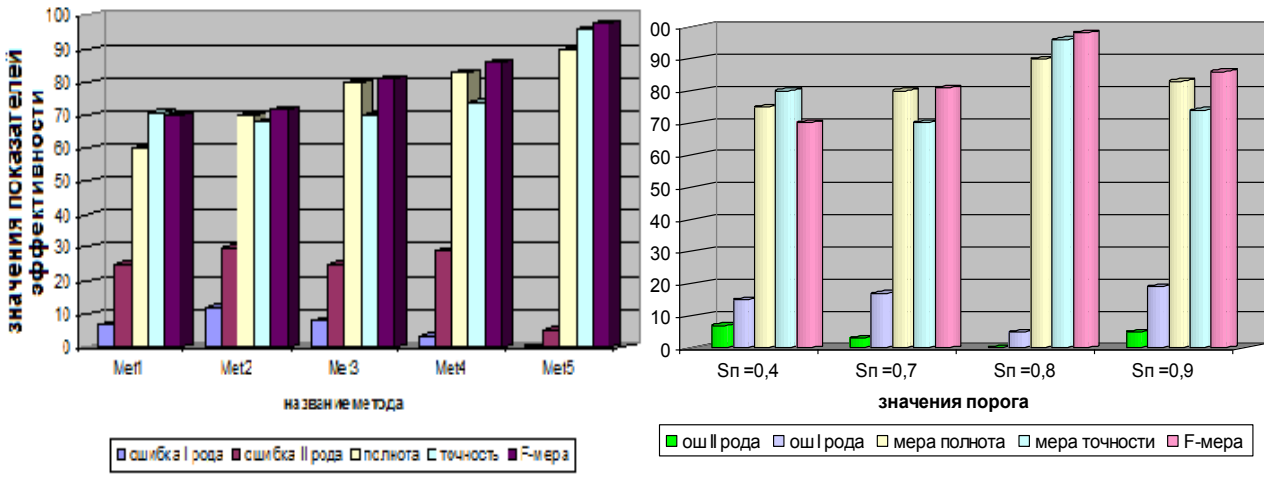

Рисунок 7 – Диаграммы результатов имитационного эксперимента

Как видно из результатов имитационного эксперимента наиболее эффективна версия met5. Анализ результатов исследований met5 показал, что при изменении порога соответствия S_n изменяются показатели качества фильтрации ЭПС. При установке порога $S_п$ = 0,4 число легитимных сообщений принятых за спам составляет 7% , а число спам-сообщений принятых за легитимные составляет 15%. При увеличении порога $S_п$ до 0,7 снижается уровень ошибки 2 рода до 3%, однако уровень ошибки 1 рода составляет 17% и при дальнейшем увеличении порога S_n продолжает расти, что свидетельствует о высокой требовательности нейронной сети (при установленном пороге, близком к единице, нейронная сеть требует почти полного соответствия входного сообщения и прототипа хранящегося в базе). Установка порога $S_п$=0,8 показывает лучшие результаты: ошибка 2 рода стремиться к 0 и составляет 0,001, ошибка 1 рода – 0,07.

Доля НЭС, выявленная предложенной системой фильтрации, выше, чем у байесовского фильтра, при вероятности ложного срабатывания не более 0,05.

Таким образом, результаты экспериментальных исследований прототипа системы защиты почтовых сервисов на основе двухуровневой контентной фильтрации входящих сообщений подтверждают достижение поставленной цели и свидетельствуют о повышении достоверности идентификации легитимной почтовой корреспонденции по ошибке классификации легитимных сообщений до 0,1% , а по ошибке классификации спам-рассылок до 7%.

В заключении отражены основные результаты диссертационных исследований

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ ДИССЕРТАЦИИ

1. Системный анализ защиты почтовых сервисов ИТКС корпоративных предприятий с территориально-распределенной структурой свидетельствует о необходимости развития существующих методов фильтрации ЭПС на основе моделей, отражающих семантику легитимной почтовой корреспонденции и учитывающих изменяющиеся информационные потребности адресатов, для исключения ложной классификации легитимных ЭПС. Выявлены основные признаки электронных почтовых сообщений, необходимые для классификации электронных рассылок.
2. Разработана модель электронного сообщения в форме устойчивых словосочетаний, которая позволяет без потери смыслового содержания обеспечить классификацию легитимной электронной корреспонденции в реальном масштабе времени. Эффект достигается применением меры значимости термов для устранения больших различий в частотах фиксации термов, исключением термов с малой информативной нагрузкой, выделением устойчивых словосочетаний, позволяющих усилить смысловое содержание термов и сократить пространство признаков на 25% за счет использования дополнительных мер близости между термами в сообщении и тесноты взаимосвязи между ними.
3. Предложена методика и разработаны алгоритмы контентной фильтрации электронной корреспонденции почтовых сервисов на основе нейросетевого классификатора АРТ2а, отличающиеся использованием дополнительного нейрона для проверки сообщений, идентифицируемых как несанкционированные сообщения, мерой сходства векторов Жаккара.
4. Предложен прототип системы защиты почтовых сервисов, основанный на двухуровневой фильтрации электронных почтовых сообщений, отличающийся предварительной подготовкой сообщений к нейросетевой классификации и обеспечивающий контентную фильтрацию легитимной корреспонденции в реальном масштабе времени.
5. Результаты экспериментальных исследований предложенного прототипа системы защиты почтовых сервисов свидетельствуют о повышение достоверности идентификации почтовой корреспонденции по ошибке классификации легитимных сообщений до 0,1% , а по ошибке классификации спам-рассылок до 7%

Перспективы дальнейшей разработки темы. Перспективными направлениями для дальнейших исследований являются разработка методов и средств обнаружения спам-рассылок, анализ рисков возникновения спам-атак, их последствий и определение фактической степени необходимой защиты.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

В рецензируемом журнале списка ВАК:

1. Чернопрудова, Е. Н. Интеллектуальная фильтрация несанкционированных рассылок на основе нейронной сети / Е. Н. Чернопрудова, Н. А. Соловьев // Интеллект. Инновации. Инвестиции. – Спец. вып. - 2011. - С. 106-107.

2. Чернопрудова, Е. Н. Формирование устойчивых словосочетаний в задаче контентной фильтрации электронных сообщений / Е. Н. Чернопрудова, Н. А. Соловьев // Вестник Оренбургского государственного университета. – 2013. – № 11 (160). – С. 98-1106.

Другие публикации:

3. Чернопрудова, Е. Н. Развитие концепции обнаружения вторжений / Е. Н. Чернопрудова, Н. А. Соловьев // Современные информационные технологии в науке, образовании и практике : материалы VIII Всерос. науч.-техн. конф. (с международным участием) / Оренбург. гос. ун-т. – Оренбург, 2009. - С. 66-67.

4. Чернопрудова, Е. Н. Методика оценки критичности и вероятности реализации уязвимостей систем электронного документооборота / Е. Н. Чернопрудова, И. Г Дворовой // Опыт использования и проблемы внедрения инноваций в науке, промышленности, энергетике и строительстве : сб. материалов науч.-практ. конф. / Оренбург. гос. ин-т менеджмента. – Оренбург, 2009. - С. 213-216.

5. Чернопрудова, Е. Н. Нейросетевая модель интеллектуальной фильтрации несанкционированных рассылок / Е. Н. Чернопрудова // Современные информационные технологии в науке, образовании и практике : материалы IX Всерос. науч.-техн. конф. (с международным участием). – Оренбург, 2010. - С. 44 - 47.

6. Чернопрудова, Е. Н. Развитие методов интеллектуальной spam-фильтрации электронных сообщений / Е. Н. Чернопрудова // Сборник работ победителей отборочного тура Всероссийского конкурса научно-исследовательских работ студентов, аспирантов и молодых ученых по нескольким междисциплинарным направлениям / М-во образования и науки РФ, Юж.-Рос. гос. техн. ун-т (НПИ). – Новочеркасск, 2011. - С. 37 - 39.

7. Чернопрудова, Е. Н. Алгоритм spam-фильтрации электронных сообщений на основе нейронной сети / Е. Н. Чернопрудова // Информационно-телекоммуникационные системы и управление : материалы Всерос. науч. школы. – Воронеж, 2011. – С. 439-444.

8. Чернопрудова, Е. Н. Проект системы фильтрации нежелательной корреспонденции / Е. Н. Чернопрудова, Н. А. Соловьев // Современные информационные технологии в науке, образовании и практике : материалы X Всерос. науч.-практ. конф. / Оренбург. гос. ун-т. – Оренбург, 2012. - С. 107-112.

9. Чернопрудова, Е. Н. Методика экспериментальной оценки эффектив-

ности распознавания спам - рассылок электронной почты / Е. Н. Чернопрудова, А. Н. Фазылова // Современные информационные технологии в науке, образовании и практике : материалы X Всерос. науч.-практ. конф. (с международным участием). – Оренбург, 2012.– С. 117-119.

10. Чернопрудова, Е. Н. Экспериментальная оценка интеллектуальной системы фильтрации спама / Е. Н. Чернопрудова // Сборник трудов международной молодежной конференции «Интеллектуальные технологии обработки информации и управления», 17-20 июля 2012 г., Уфа / Уфим. гос. авиац. техн. ун-т. – Уфа, 2012. - С. 91-93.

11. Свидетельство о государственной регистрации программы для ЭВМ № 2013617655 «Программа фильтрации спама на основе нейронной сети» / Е. Н. Чернопрудова, Н. А.Соловьев, В. А. Пучков. Заявка № 2013617655. Дата поступления 21 мая 2013 г. Зарегистрировано в Реестре программ для ЭВМ 21 августа 2013 г.

Диссертант Е.Н. Чернопрудова

ЧЕРНОПРУДОВА Елена Николаевна

ЗАЩИТА ПОЧТОВЫХ СЕРВИСОВ ОТ НЕСАНКЦИОНИРОВАННЫХ РАССЫЛОК НА ОСНОВЕ КОНТЕНТНОЙ ФИЛЬТРАЦИИ ЭЛЕКТРОННЫХ СООБЩЕНИЙ

Специальность: 05.13.19
Методы и системы защиты информации,
информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 14.11.2013. Формат 60x84 1/16.
Бумага офсетная. Печать плоская. Гарнитура Таймс.
Усл. печ. л. 1,0. Усл. кр. – отт. 1,0. Уч. –изд. л. 0,9.
Тираж 100 экз. Заказ № 594

ООО «Издательство ЦДУМ»
450000 г.Уфа, ул. Тукаева, 50