УДК 004.056.52, 004.82 **DOI** 10.15622/sp.53.5

В.Ю. Осипов, В.И. Воробьёв, Д.К. Левоневский

ПРОБЛЕМЫ ЗАЩИТЫ ОТ ЛОЖНОЙ ИНФОРМАЦИИ В КОМПЬЮТЕРНЫХ СЕТЯХ

Осипов В.Ю., Воробьёв В.И., Левоневский Д.К. **Проблемы защиты от ложной информации в компьютерных сетях.**

Аннотация. Анализируется текущее состояние в области защиты от ложной информации в компьютерных сетях и формулируются актуальные проблемы, связанные с этой защитой. Предлагается подход к оценке мероприятий защиты от такой информации на основе использования марковской модели дезинформирования. Раскрывается архитектура перспективной системы анализа информации в компьютерных сетях по требованиям достоверности. В рамках этой архитектуры рассматриваются усовершенствованные методы анализа достоверности текстов. Предлагается комплексный подход к использованию известных и предложенных методов для оперативного выявления ложной информации в компьютерных сетях. Кроме того, метод может применяться в области борьбы с киберпреступностью и терроризмом для поиска сетевых ресурсов и коммуникационных площадок, которые могут быть использованы для организации противоправной деятельности.

Ключевые слова: достоверность, ложная информация, анализ текста, классификация, онтология, информационная безопасность.

1. Введение. В последние годы в условиях различных конфликтов, конкурентной борьбы за новые технологии, рынки сбыта, энергетические, биологические и другие ресурсы резко возросла угроза намеренного распространения ложной информации (дезинформации) в Интернете. Известно, что посредством дезинформации можно ввести людей в заблуждение и навязать неадекватное общественное мнение. Также в некоторой мере возможно манипулирование сознанием и поведением как отдельных личностей, так и групп людей. Злоумышленники могут усиливать или ослаблять взгляды людей на различные события, жизненные ценности, выполняемые работы, на поведение в сложившихся ситуациях и т.п. То, что одним из каналов распространения ложной информации является Интернет, обусловлено особенностями информационной инфраструктуры, к которым относятся простота и дешевизна доступа, размытость государственных границ, широкие возможности манипулирования информацией и ее восприятием, высокий уровень анонимности доступа к сети.

В поисковых системах без труда можно найти многочисленные сайты с объективно ложной информацией. В большом количестве распространяются поддельные новости о якобы произошедших резонансных событиях (теракты, смерть известных людей, финансовые потрясения). Зачастую для привлечения внимания используются приемы, искажающие информацию о произошедших событиях.

Для защиты населения от ложной информации в Интернете необходимо наличие эффективной системы противодействия ей. В общем виде под такой системой понимается совокупность взаимосвязанных, организационных, технических, правовых и других мероприятий, обеспечивающих: своевременную профилактику, выявление, блокирование, удаление ложной информации в Интернете, устранение последствий, ответственность злоумышленников за противоправные действия. Для обеспечения функционирования такой системы требуется соответствующее научно-методическое обеспечение [1].

Проблема защиты от ложной информации существует с давних пор [2], но с появлением и активным развитием Интернета ситуация резко изменилась. Существенно возросли объемы распространяемой информации и возможности по ее использованию с целью ввести в заблуждение людей. Как показывают международные события последних лет, через Интернет осуществимы широкомасштабные информационные воздействия на людей с различными террористическими, политическими, религиозными, экономическими и другими целями. При этом усложнились задачи по оперативному обнаружению, распознаванию информационных воздействий, выявлению источников их порождения и прогнозированию таких угроз. Традиционные методы защиты в этих условиях недостаточно эффективны [3]. Они предусматривают активное участие человека в анализе информации на предмет ее деструктивности и опасности для пользователей Интернета. Это требует существенных затрат людских и материальных ресурсов. В последние годы стали появляться методы и средства, позволяющие автоматизировать этот процесс [4], что значительно повышает его оперативность и снижает трудоемкость [5]. Для некоторых аспектов информационной безопасности решения уже есть. Можно привести в качестве примера известный модуль Avast Online Security [6], работающий с веб-сайтами. В основу работы этого модуля положены методы рейтинговых оценок. Модуль использует базу данных веб-репутации, которая обновляется с учетом информации, поступающей от пользователей. Такой модуль дает возможность пользователю оценивать веб-ресурсы и предоставляет ему информацию об их рейтинге, а также о том, вовлечен ли веб-ресурс в схемы фишинга, мошенничества, содержит ли он вредоносное программное обеспечение. Этот модуль собирает данные о вредоносных страницах и страницах с плохой репутацией и предупреждает пользователя при попытке перехода на такой сайт, блокирует навязчивые всплывающие окна и защищает от отслеживания деятельности пользователя в сети. В целом модуль Avast Online Security позволяет оце-

нивать страницы с точки зрения наличия вредоносного программного обеспечения и мошеннической активности, но не учитывает их достоверность. Для этого в настоящее время применяются различные модели и методы:

– анализа текстов (морфологический, синтаксический и семантический анализ [7], тематическое моделирование [8], оценка достоверности [9], анализ культурологических канонов сетевых сообществ [10]);

– анализа изображений и видео (генерация словесных описаний [11], определение элементарных ситуаций и действий [12], прогнозирования ситуаций [13]);

– анализа звуковых потоков (определение эмоций в речи [14, 15], распознавание истинности речи [16]).

Для обнаружения дезинформирующих ДВ в Интернет прорабатывают применение различных алгоритмов машинного обучения, а для прогнозирования таких воздействий используют методы регрессионного анализа и другие [27].

Однако эти известные методы не в полной мере учитывают всю специфику процессов выявления и защиты от ложной информации в компьютерных сетях и обладают невысоким уровнем их автоматизации. Во многом не совершенны методы выявления смыслового содержания в информационных воздействиях (в виде деструктивной графической, текстовой, числовой, музыкальной и комбинированной информации) и соотнесения его с возможными угрозами для населения. Это существенно затрудняет автоматическую проверку информации по требованиям достоверности.

Требуется разработка новых моделей и методов, позволяющих своевременно выявлять и прогнозировать угрозы от дезинформирующих ДВ, обосновывать целесообразные мероприятия по их блокировке и устранению последствий.

2. Модель процесса дезинформации. В интересах решения поставленной задачи рассмотрим обобщенную модель процесса дезинформации пользователей Интернета с учетом мероприятий противодействия. Структуру такого процесса можно представить в виде рисунка 1.

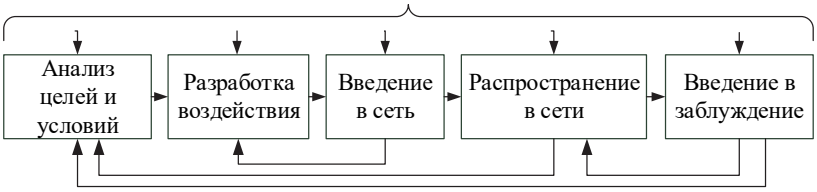

Рис. 1. Обобщенная структура процесса дезинформирования

Процесс начинается с анализа злоумышленником целей дезинформации, объектов воздействия и условий реализации. На основе этого анализа злоумышленником осуществляется разработка дезинформирующего деструктивного воздействия (ДВ). При разработке ДВ злоумышленники стремятся сформировать их так, чтобы они были подобны истинным сигналам по опознавательным статическим и динамическими характеристиками. При этом основной объем таких воздействий по содержанию согласуют с реальными фактами. Ложная информация в таких воздействиях может занимать незначительную, но емкую по содержанию часть. При формировании ДВ принимают во внимание априорную вероятность появления на объектах воздействия истинных сигналов с аналогичными опознавательными параметрами, как у ложных сигналов. После разработки ДВ злоумышленникам требуется получить доступ к интересующим ресурсам Интернета. ДВ могут массированно выставляться на сайтах Интернета или рассылаться по конкретным почтовым адресам. В ряде случаев в ДВ могут быть предусмотрены механизмы саморазмножения в сети. При воздействии ДВ на объекты возможны различные негативные эффекты с далеко идущими последствиями. Цель процесса — порождение следствий, устраивающих заказчика: получение коммерческой выгоды, создание паники в массах или, напротив, ее недопущение, создание «образа врага» или скрытие важных событий за второстепенными фактами.

Для защиты от ДВ проводятся различные мероприятия по противодействию. Они могут быть направлены на снижение предпосылок для разработки ДВ злоумышленниками. Также возможны противодействия попыткам непосредственной разработки ДВ. Для недопущения внесения на значимые сайты Интернета ложной информации и распространения (размножения) ДВ могут предусматриваться свои мероприятия защиты. В интересах снижения рисков от влияния ДВ возможно своевременное выявление, блокирование и удаление таких воздействий.

В интересах прогнозирования угрожающих событий рассмотренному процессу дезинформирования можно поставить в соответствие граф возможных состояний. Опираясь на центральную предельную теорему теории вероятностей для потоков событий, этот процесс можно описать с применением математического аппарата марковских процессов. Влияние мероприятий защиты можно учесть через изменение параметров переходов анализируемых процессов из одного состояния в другое. Используя такую модель, эффективность мероприятий защиты пользователей Интернет от дезинформации можно оценивать по приращениям вероятностей и времени перехода процесса в интересующие состояния за счет проводимых мероприятий противодействия.

Так, в соответствии с рисунком 1 процесс дезинформирования формализуется в виде графа состояний марковского процесса (рисунок 2).

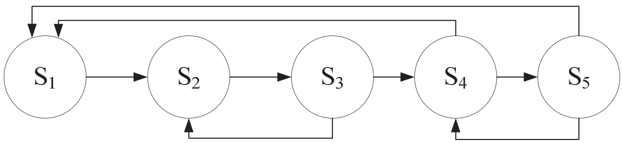

Рис. 2. Граф состояний марковского процесса дезинформирования пользователей Интернета

В качестве состояний на рисунке 2 приводятся: S_1 — анализ злоумышленником целей дезинформирования и условий их реализации, S_2 — разработка ДВ, S_3 — введение ДВ в сеть, S_4 — распространение ДВ в сети, S_5 — введение пользователей сети в заблуждение. Дугам графа ставятся в соответствие значения интенсивностей переходов процесса из одного состояния в другое, которые зависят от реализации мероприятий защиты.

Этому графу соответствует система из пяти дифференциальных уравнений:

$$\frac{dP_1(t)}{dt} = \lambda_{41}P_4(t) + \lambda_{51}P_5(t) - \lambda_{12}P_1(t)$$

$$\text{...} .$$

$$\frac{dP_5(t)}{dt} = \lambda_{45}P_4(t) - (\lambda_{51} + \lambda_{54})P_5(t)$$

В этой системе $P_1(t),...,\ P_5(t)$ — вероятности нахождения процесса в состояниях 1,….,5 на момент времени t; $\lambda_{12},...,\ \lambda_{54}$ — интенсивности переходов процесса из одного состояния в другое. Если эти интенсивности, начальное и интересующее состояния определены, то такая система дифференциальных уравнений легко разрешается известными методами с применением, например пакета прикладных программ MatLab. Умея распознавать выделенные состояния и зная $\lambda_{12},...,\ \lambda_{54}$, можно прогнозировать наступление угрожающих событий. Основные сложности при использовании таких моделей состоят в определении интенсивностей переходов процесса из одних состояний в другие и в распознавании текущих состояний. Для их определения необходимы реальные наблюдения событий. Такие модели, наряду с методами регрессионного анализа могут успешно применяться при планировании и обосновании мероприятий противодействия дезинформирующим ДВ.

В интересах такого противодействия могут создаваться специальные центры. Например, создание центра информационного анализа и противодействия, в обязанности которого входит выявление дезинформации с использованием информационно-коммуникационных технологий, предусмотрено биллем H.R.5181 Конгресса США. Технические меры противодействия заключаются в анализе публикуемой информации и выявлении ее источника, что в дальнейшем может стать основанием для применения юридических мер (расследование, удаление информации, привлечение к ответственности). К техническим мерам относятся действия по обзору сети, извлечению образцов информации из веб-страниц, вычислению характеристик этих образцов информации, классификации образцов, нейтрализации воздействия (блокирование, удаление).

3. Метод выявления ложной информации в компьютерных сетях. В основе предлагаемого метода лежит идея использования нового распределенного программного приложения. Архитектура этого приложения представлена на рисунке 3. Она разделяется на 3 уровня.

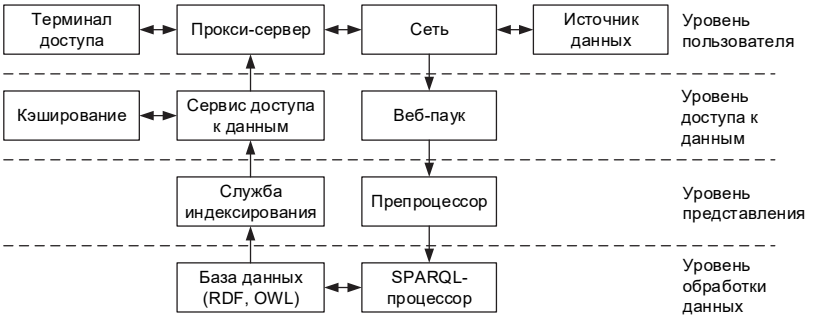

Рис. 3. Схема функционирования приложения

Согласно этой идее на уровне пользователя происходит коммуникация между клиентом (терминалом доступа) и сетевыми ресурсами. Единственным компонентом приложения на этом уровне является прокси-сервер, который выполняет мониторинг трафика (например, HTTP-запросов и ответов), получает сведения о запрашиваемых ресурсах и в зависимости от настроек предоставляет пользователю характеристики этих ресурсов или блокирует их при выявлении нежелательной информации.

На уровне доступа к данным выполняется получение и обработка сведений о сетевых ресурсах. К этому уровню относятся:

– служба поиска — веб-паук, выполняющий обход веб-ресурсов в сети, поиск и загрузку данных из этих ресурсов;

– сервис доступа к данным, который предоставляет прокси-серверу сведения о характеристиках веб-ресурсов;

– сервис кэширования, который хранит часто используемые запросы и ускоряет выдачу результата.

На уровне представления реализуются механизмы поиска ресурсов и информации о них в хранилище данных. К этому уровню относятся:

– сервис индексирования, предназначенный для поиска сведений об образце информации в хранилище данных (например, по хэшу);

– препроцессор, предназначенный для извлечения образцов информации (текста, изображений, аудио, видео) из контейнеров (веб-страниц, документов Word и т.п.), а также извлечение текста и образов из мультимедийных объектов.

На уровне обработки данных выполняется анализ и сопоставление свойств информационных объектов и сохранение информации о них в хранилище данных.

В соответствии с предлагаемым методом выявления ложной информации предусматривается представление и обработка данных на предмет их достоверности. Более детально, согласно ему осуществляется выделение образцов информации из контейнеров (веб-страница, документ), оценка свойств этих образцов и их классификация на основе полученных оценок.

Алгоритм, реализующий этот метод, состоит из ряда шагов:

Шаг 1. Получение контейнера с информацией путем обхода заданного сегмента сети веб-пауком.

Шаг 2. Выделение образцов (объектов) информации из контейнера (текст, изображение, видео, звук) методом парсинга HTML-тегов и подгрузки ссылочных данных.

Шаг 3. Создание для каждого образца информационного объекта вида <ID, информация, метаданные> и сохранение его в базе данных. К метаданным относятся сведения об источнике информации, времени доступа и публикации, ключевые слова.

Шаг 4. Определение и оценка достоверности характеристик (таблица 1) информационного объекта на основе информации, метаданных и связей с другими информационными объектами.

Шаг 5. Определение и оценка достоверности свойств информационного объекта.

Шаг 6. Расчет интегрального показателя достоверности информационного объекта.

В результате мы имеем дело с многоуровневой системой показателей и правил их оценки, позволяющих перейти от оценки достоверности характеристик информационного объекта, вычисляемых непосред-

ственно, к достоверности существенных свойств объекта как носителя информации, а от них — к достоверности информационного объекта в целом. Схему оценки достоверности информационных объектов можно представить в виде рисунка 4, где вершинами обозначены реализуемые функции, а дугам ставятся в соответствие весовые коэффициенты.

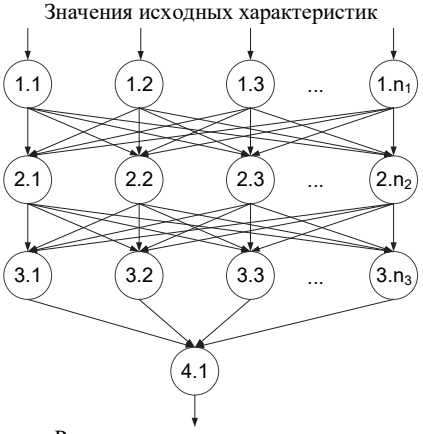

Рис. 4. Схема оценки достоверности информационных объектов

В некотором смысле такую схему оценки можно трактовать как классификатор информационных объектов по уровням достоверности. Поясним такой подход.

Каждому информационному объекту можно поставить в соответствие многоуровневую систему его характеристик (свойств). Отнесение таких объектов к ложной информации в общем случае предусматривает анализ его характеристик на различных уровнях иерархии с учетом их взаимосвязей. Для случаев, когда известна многомерная плотность $f(y_1,...,y_n)$ распределения характеристик $y_1,...,y_n$ информационного объекта, вероятность $P_{\partial u}$ отнесения его к достоверной информации можно рассчитать по формуле:

$$P_{\partial u} = P_a \int_{Y_1} ... \int_{Y_n} f(y_1,...,y_n)dy_1...dy_n,$$

где P_a — априорная вероятность наличия достоверной информации; $Y_1,...,Y_n$ — области значений для истинных характеристик информационного объекта.

Если считать независимыми друг от друга характеристики $y_1,...,y_n$, то:

$$P_{\partial u} = P_a \int_{Y_1} f_1(y_1)dy_1 ... \int_{Y_n} f(y_n)dy_n,$$

где $f_1(y_1), \ldots, f_n(y_n)$ — одномерные плотности распределения характеристик $y_1,...,y_n$. Согласно этим выражениям информационный объект может быть отнесен к ложной информации при низкой априорной вероятности P_a или при низкой вероятности достоверности хотя бы одной из его характеристик.

На практике необходимо учитывать неравнозначность вклада каждой характеристики в $P_{\partial u}$. Для определения достоверности множества одноуровневых, но не равнозначных характеристик информационного объекта, отражающих некоторое его свойство, можно использовать выражение:

$$P^*_{\partial u} = \sum_{i \in \Omega} a_i P_{\partial u_i},$$

где $P_{\partial u_i}$ — вероятность достоверности *i*-й характеристики объекта; — относительный вес этой характеристики, $P^*_{\partial u} \le 1$.

С учетом этого интегральный показатель W оценки достоверности информационного объекта можно рассчитать по правилу: $W = W_{zi}$ при $z = Z$ и $i = 1$:

$$W_{zi} = \sum_{j=1}^{n_{z-1}} a_{zij} W_{z-1,j}\ ;\ i = \overline{1, n_z}\ ;\ z = \overline{1, Z}\ ,$$

где z — номер уровня обработки характеристик; Z — число уровней; n_z — число различных характеристик, влияющих на достоверность свойств информационного объекта, на $z+1$ уровне; a_{zij} — относительный вклад показателя $W_{z-1,i}$ в W_{zj}. На уровне $z = 1$ в качестве показателей W_{zi} могут выступать вероятности достоверности исходных характеристик информационного объекта.

Физический смысл такого интегрального показателя — взвешенная сумма частных нормированных показателей.

Базовые исходные характеристики информационных объектов сведены в таблицу 1.

Таблица 1. Набор характеристик образца данных

Характеристика	*Значения*
Метаданные	
Адрес источника	URL
Уровень домена	Целое, больше 1
Протокол доступа	Незащищенный Защищенный без валидного сертификата Защищенный с валидным сертификатом Защищенный с расширенной валидацией
Дата публикации	Дата
Объем	Целое (байты)
Рейтинг	Вещественное, -1…1
Связи	
Количество совпадений	Целое, 0…∞
в т.ч. с более ранней датой публикации	Целое, 0…∞
в т.ч. с более поздней датой публикации	Целое, 0…∞
Лучшая степень совпадения	Целое, 0…∞
в т.ч. с более ранней датой публикации	Целое, 0…∞
в т.ч. с более поздней датой публикации	Целое, 0…∞
Моменты распределения степени совпадения	Вещественный вектор
Количество гиперссылок на объект	Целое, 0…∞
Моменты распределения интегрального показателя и рейтинга ресурсов, ссылающихся на объект	Вещественный вектор
Моменты распределения интегрального показателя и рейтинга ресурсов, на которые объект ссылается	Вещественный вектор
Содержимое	
Стилистические маркеры	Вещественный вектор
Психолингвистические маркеры	Вещественный вектор
Лексические маркеры	Вещественный вектор
Семантические маркеры	Вещественный вектор
Степень соответствия:	
контента и заголовка	Вещественное, 0…1
контента и описания	Вещественное, 0…1
контента и ключевых слов	Вещественное, 0…1
Степень однородности	Вещественное, 0…1

К важным свойствам информационных объектов, прежде всего, следует отнести их структурные, частотные и содержательные особенности отдельных конструкций и объектов в целом.

Целью применения системы правил является классификация информационных объектов и выделение признаков, которые могут характеризовать эти объекты как ложные.

С использованием предложенного метода по значениям интегрального показателя можно относить анализируемые информационные объекты к различным уровням достоверности. Однако при этом необходимо знать области допустимых истинных значений оцениваемых характеристик и свойств информационных объектов.

Рассмотрим далее особенности анализа текстовых данных для определения свойств образцов информации в рамках предложенного метода.

4. Анализ текстовых данных. Ввиду постоянного увеличения количества текстовых документов в электронной форме необходимо наличие эффективных алгоритмов их классификации и анализа.

Рассмотрим n текстовых документов $D = \{D_1, D_2, ..., D_n\}$. Пусть известно, что эти документы могут быть разделены на m тем $T = \{T_1, T_2, ..., T_m\}$. Документы с индексами $I = \{I_1, I_2, .., I_m\}$, $I_j = \{I_{j1}, I_{j2}, ..., I_{jk_j}\}$ принадлежат теме T_j.

Рассмотрим в качестве примера две темы: $T = \{T_1, T_2\}$, $I_1 = \{1, 2\}$, $I_2 = \{10, 11\}$.

В данном примере мы точно знаем, что документы с индексами 1 и 2 принадлежат первой теме, а документы с индексами 10 и 11 принадлежат второй теме. Здесь мы можем определить принадлежность любого документа к определенной теме, то есть считать, что документ D_i принадлежит T_1 или что он не принадлежит T_2. Этот подход может быть расширен с использованием классов слов. Классы объединяют близкие по значению слова. Например, класс «дезинформация» состоит из слов «ложь», «миф», «слух» и так далее. Классы могут также состоять из других классов. Таким образом, класс «большая ложь» содержится в классе «ложь». Если мы знаем, что слово принадлежит классу «большая ложь», мы можем сказать, что оно принадлежит классу «дезинформация».

Естественным образом коллекция всех слов может быть разделена на такие классы, которые образуют древовидную структуру вида тип-подтип. На верхушке этой структуры расположен класс «что-то», который содержит все слова. Решение с использованием

классов отличается тем, что коллекция W определена другим образом. Коллекция W содержит классы слов, не сами слова. Следовательно, она будет меньше.

В [17] рассмотрены три алгоритма классификации текста. Первое решение (алгоритм) основано на использовании ключевых слов. Второе использует классы слов. Третье основано на использовании смешанного алгоритма. Ранее в СПИИРАН в работе [17] для сравнения двух документов была собрана коллекция текстовых документов, которая содержала около 1000 новостей из пяти тем. Размер этих документов был около 1 кб. Исследование проводилось для нескольких тем. Для получения результатов мы изменяли только коллекцию документов D и коллекцию индексов I. В итоге было извлечено более 150 результатов. В 75% случаев метод, основанный на использовании ключевых слов, был точнее. Опираясь на данные результаты предлагается построить систему поиска на основе статистического метода.

Различные подходы к этой проблеме заявлены рядом патентов [18-22]. Большинство заявленных проектов делают основной акцент на разработку глобальных опознавательно-преобразовательных фильтров, роль которых состоит в схеме кодирования — декодирования многообразия языка и создания надлингвистической базы индексов смысловой нагрузки. Очевидная проблема в подобных подходах, кроме громоздкой лингвистически-смысловой свертки, представляется в процессе декодирования. Можно несколько десятков слов свернуть в один смысловой индекс, но очень трудно дешифровать этот индекс адекватно запросу (эффект обратного перевода). Предлагаемая модель не требует серьезной смысловой интерпретации запроса (ключа), а опознает его по статистическим характеристикам структурных единиц текста. Наиболее близким аналогом из перечисленных работ является [18].

Идеологической основой рассматриваемой нами задачи является попытка преодоления противоречия между литературными (субъективными) законами языка при описании какого-либо предмета (процесса, явления) и формализованными (объективными) алгоритмами распознавания смысла как логического модуля, отображаемого в виде математического образа. Рассмотрим способ извлечения информационной структуры любой статьи, независимо от стиля ее написания. Фактически, информационная характеристика страницы Интернета состоит из четырех составляющих:

– заголовок страницы;
– перечень ключевых слов, составленный автором;
– краткое описание страницы;
– собственное содержание страницы.

Первые три элемента достаточно субъективны. Предлагаемый метод ориентирован на четвертый элемент. Он предусматривает сканирование текста и выделение собственного набора ключевых слов, объективно соответствующих содержанию. В этом случае машина обрабатывает любой безликий объект, ассоциируя его с его порядковым номером. Кроме того, он не требует искусственно заданной аналитической модели (функции распределения), ему безразлична профессиональная направленность информации, то есть не требуется специализированных библиотек, что очень важно в том случае, если содержание статьи находится на стыке различных областей знаний. Следовательно, такой способ идентификации абсолютно объективен, независим от рейтинга, позиции в каталоге и смысловой трактовки информационного модуля. Важна лишь тематическая насыщенность, обеспечивающая устойчивые статистические характеристики, что, кстати, исключит обращения к «мертвым ссылкам».

Конкретная цель данной задачи заключается в моделировании одной из составляющих частей робота-анализатора на основе статистического метода анализа и синтеза показателей при информационном дефиците с применением метода свертки показателей, разработанного в [23]. Суть метода заключается в формировании сводного статистического показателя объекта путем формирования вектора отдельных показателей, выбора синтезирующей функции и определения вектора весовых коэффициентов. В качестве исследуемого предмета мы принимаем текстовой модуль (статью) направленной тематики, например, нечисловой, неточной и неполной информации с целью получить систему статистических показателей.

За исходное предположение принимается тот факт, что заметная повторяемость определенной группы слов (с учетом семантического анализа) в тексте достаточно большого объема должна нести информацию о содержании статьи, что позволяет частоту повторяемости рассматривать как численную характеристику потенциально возможного ключевого слова. Иными словами, мы пытаемся создать числовой образ нечисловой информации, что позволит впоследствии ранжировать числовой ряд, полученный с использованием указанного метода.

Таким образом, считая любое слово как смысловую единицу объектом исследования на начальном этапе, а его повторяемость — численной характеристикой объекта, мы имеем входные данные для расчета статистических показателей составляющих элементов исследуемого текста, что при определенной обработке можно трактовать как математический отпечаток текстового файла.

Другой подход основан на семантической сети (Semantic Web), при котором компьютеры могут использовать обозначения с хорошо определенной и пригодной для машинной интерпретации семантикой, чтобы обмениваться знаниями. Знания, представленные на языках Се-

мантической сети, например на языке структурированного описания ресурсов (RDF), отличаются как от обычно неструктурированного произвольного текста, который можно обнаружить на большинстве Web-страниц, так и от высокоструктурированной информации в базах данных. С другой стороны в рамках одного документа, документы Семантической сети (Semantic Web Document, SWD) могут быть смешением конкретных фактов, определением классов и свойств, логических ограничений и метаданных.

Один из вариантов описания семантики — это построение математической модели языка [24, 25], в которой любое слово русского языка можно рассматривать как имя функции. При этом конкретное значение слово получит только после подстановки аргументов, а его смысл будет вычислен по мере выполнения функции. Предложение в данном случае — это законченная суперпозиция функций, а смысл предложения вычисляется при построении и выполнении этой суперпозиции. Такой подход к семантическому анализу позволяет в том числе построить онтологии, описывающие предложения. Кроме того, онтологии, как и тезаурус, можно использовать еще на этапе семантического анализа. В этом случае наиболее эффективно их совместное использование, где онтология описывает комплекс понятий и отношений предметной области, а тезаурус формирует подобную систему понятий и отношений в рамках лингвистических знаний по предметной области [25]. Описанные выше методики можно обозначить как семантический анализ неструктурированной информации, к которой на сегодняшний день можно отнести и множество документов на языке HTML. Эти исследования в основном делают попытку решения трудных для формализации задач.

Для онтологического моделирования можно использовать формальное представление онтологии в виде:

$$O = (E, D, R, P),$$

где E — множество сущностей; D — множество их определений; R — множество отношений; P — множество правил [26].

Процесс онтологического моделирования начинается с разделения понятий на две категории: классы (концепты) и их свойства (слоты). Классы разрабатываемой онтологии описывают понятия предметной области. Каждый из классов может иметь свой подкласс, который предоставляет более подробное описание, чем его надкласс. Задача слота — описать свойства класса и экземпляра. Свойства дают возможность делать выводы об общих свойствах классов. Свойство — это бинарное отношение. Различают два типа свойств:

1. свойства-значения — это отношения между представителями классов и типами данных;

2. свойства-объекты — это отношения между представителями двух классов.

Описание данных категорий строится на основе языков XML, RDF и OWL, дающие возможность создавать классы, свойства и отдельные экземпляры, то есть создать структуру онтологий. Приведем пример использования этих трех компонентов, определим класс сущностей «Слухи» и некоторые свойства «Слухи», такие как «Слухи-желания», «Слухи-пугала» и «Разобщающие агрессивные слухи».

Средствами синтаксиса RDF класс и его свойств будут описаны так:

```
<Class ID="hearing"/>
<Property ID="desire"/>
<Property ID="scare"/>
<Property ID="aggressive"/>
```

После выделения классов, можно перейти к описанию отдельных экземпляров. Пусть ложная информация в следующем: «Айболит, президент страны Лимпопо, угрожает России».

Лимпопо будет представляться следующим образом:

```
<Country ID="Россия">
<name>Айболит</name>
<Country>Лимпопо</Country>
<aggressive resource="Россия"/>
</Country>
```

Подобное описание можно обобщить и использовать как метаописание. OWL позволяет так же вводить ограничения на свойства, области распространения и отношения. Известны связанные между собой форматы, которые удовлетворяют описанной структуре семантических данных и предназначены для использования в Web (рисунок 5).

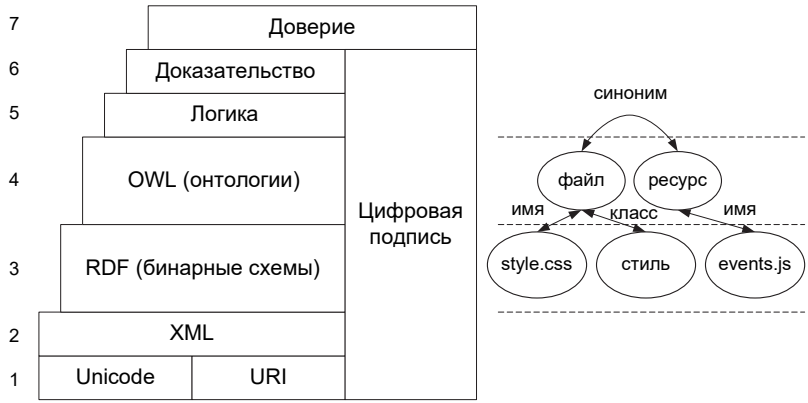

Рис. 5. Форматы семантических данных

На рисунке 5 изображена схема обработки данных, где данные описываются на языках RDF и OWL, поисковые запросы — на формальном языке запросов SPARQL, который наследует конструкции SQL — SELECT, FROM, WHERE, процедура поиска осуществляется при помощи процессора SPARQL. Таким образом, совокупность собранных вместе элементов позволяет искать информацию в предварительно подготовленных, формализованных на языках RDF и OWL семантических данных.

5. Заключение. Рассмотренные проблемы и методы защиты от ложной информации в компьютерных сетях уточняют взгляды на теорию и практику такой защиты. Противодействие процессу дезинформации можно свести к решению ряда частных задач. Для прогнозирования процессов дезинформации в компьютерных сетях и обоснования мероприятий противодействия предлагается использовать предложенную модель этого процесса. В интересах оперативного выявления такой информации разработан усовершенствованный метод ее анализа по требованиям достоверности. Оценку достоверности информации рекомендуется осуществлять по интегральному показателю — взвешенной сумме частных нормированных показателей достоверности ее свойств. Применение разработанного метода совместно с известными решениями позволяет сформировать комплексный подход, ориентированный на широкий круг возможных ситуаций, проявляющихся в компьютерных сетях. Предложенный метод применим при разработке приложений-ассистентов с более широкой функциональностью, позволяющей пользователю сети отфильтровать информацию, не отвечающую требованиям достоверности. Такой метод также может использоваться при борьбе с киберпреступностью и терроризмом для поиска сетевых ресурсов и коммуникационных площадок, которые могут быть использованы для организации противоправной деятельности.

Литература

1. *Осипов В.Ю., Юсупов Р.М.* Информационный вандализм, криминал и терроризм как современные угрозы обществу // Труды СПИИРАН. 2009. №8. С. 34–45.
2. *Осипов В.Ю., Ильин А.П., Фролов В.П., Кондратюк А.П.* Радиоэлектронная борьба. Теоретические основы // Петродворец: ВМИРЭ. 2006. 302 с.
3. *Алексеева И.Ю. и др.* Информационные вызовы национальной и международной безопасности // М.: ПИР-Центр. 2001. 328 с.
4. *Bartlett J., Reynolds L.* The State of the Art 2015: a literature review of social media intelligence capabilities for counter-terrorism // Demos. 2015. 98 p.
5. *Котенко И.В., Чечулин А.А., Комашинский Д.В.* Автоматизированное категорирование веб-сайтов для блокировки веб-страниц с неприемлемым содержимым // Проблемы информационной безопасности. Компьютерные системы. 2015. №2. С. 69–79.
6. Avast Online Security browser extension: Overview. URL: www.avast.ru/faq.php?article=AVKB18 (дата обращения: 14.04.2017).

7. *Смирнов И.В., Шелманов А.О., Кузнецова Е.С., Храмоин И.В.* Семантико-синтаксический анализ естественных языков. Часть II. Метод семантико-синтаксического анализа текстов // Искусственный интеллект и принятие решений. 2014. № 1. С. 11–24.
8. *Karpovich S., Smirnov A., Teslya N., Grigorev A.* Topic Model Visualization With IPython // Proceedings of the 20th Conference of FRUCT association. 2017. pp. 131–137.
9. *Dong X.L. et al.* Knowledge-Base Trust: Estimating the Truthworthiness of Web Sources. URL: arxiv.org/pdf/1502.03519v1.pdf (дата обращения: 11.04.2017).
10. *Александров В.В., Зайцева А.А., Кулешов С.В.* Построение глоссариев культурологических канонов кибер-социальных групп в социальных сетях // Международный научный журнал «Инновационная наука». 2016. №12-2. С. 13–17.
11. *Karpathy A., Fei-Fei L.* Deep Visual-Semantic Alignments for Generating Image Descriptions // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015 (CVPR 2015). 2015. pp. 3128–3137.
12. *Lillo I., Niebles J.C., Soto A.* Sparse composition of body poses and atomic actions for human activity recognition in RGB-D videos // Image and Vision Computing. 2017. vol. 59. pp. 63–75.
13. *Batchuluun G., Kim J.H., Hong H.G.* Fuzzy system based human behavior recognition by combining behavior prediction and recognition // Expert Systems with Applications. 2017. vol. 81. pp. 108–133.
14. *Yogesh C.K. et al.* Hybrid BBO_PSO and higher order spectral features for emotion and stress recognition from natural speech // Applied Soft Computing. 2017. vol. 56. pp. 217–232.
15. *Kaya H., Karpov A., Salah A.* Robust Acoustic Emotion Recognition based on Cascaded Normalization and Extreme Learning Machines // Proceedings of the 13th International Symposium on Neural Networks. 2016. LNCS 9719. pp. 115–123.
16. *Budkov V., Vatamaniuk I., Basov V., Volf D.* Investigation of Speech Signal Parameters Reflecting the Truth of Transmitted Information // Proceedings of the 18th International Conference on Speech and Computer (SPECOM 2016). 2016. LNAI 9811. pp. 419–426.
17. *Воробьев В.И и др.* Исследование и выбор криптографических стандартов на основе интеллектуального анализа документов // Труды СПИИРАН. 2016. №5(48). С. 69–87.
18. *Харламов А.А.* Способ формирования смыслового портрета текста и устройство для его осуществления // Патент РФ. № 2000127135. 2003.
19. *Poncet J. et al.* Access by content based computer system. US Patent no. WO 2001033419 A2. 2003.
20. *Sheth A., Avant D., Bertram C.* System and method for creating a semantic web and its applications in browsing, searching, profiling, personalization and advertising. US Patent. no. WO 2001069428 A1. 2001.
21. *Omoigui N.* System and method for knowledge retrieval, management, delivery and presentation. US Patent. no. 20100070448 A1. 2003.
22. *Gardner S.* Ontology-based information management system and method. US Patent. no. 7225183 B2. 2007.
23. *Хованов Н.В.* Оценка сложных объектов в условиях дефицита информации. К столетию метода сводных показателей А.Н. Крылова // Труды 8-й международной научной школы «Моделирование и анализ безопасности и риска в сложных системах». 2008. СПб.: ИПМАШ РАН. С. 18-28.
24. *Перминов С. В., Афанасьев С. В.* Семантический способ поиска информационных аномалий через Web // Труды СПИИРАН. 2006. №3. Т. 1. С. 279–287.
25. *Перминов С.В.* Система семантического поиска // Информационно-измерительные и управляющие системы. 2008. №4. Т. 6. С. 45–50.

26. *Нариньяни А. С.* Кентавр по имени ТЕОН: Тезаурус + Онтология // Труды международного семинара Диалог'2001 по компьютерной лингвистике и ее приложениям. 2001. Т. 1. С. 184–188.
27. *Ferrara E. et al.* Predicting online extremism, content adopters, and interaction reciprocity // Proceedings of the 8th International Conference on Social Informatics. 2016. Part II. pp. 22–39.

Осипов Василий Юрьевич — д-р техн. наук, профессор, заведующий лабораторией информационно-вычислительных систем и технологий программирования, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: интеллектуальные системы, моделирование, информационная безопасность. Число научных публикаций — 100. osipov_vasiliy@mail.ru; 14-я линия В.О., 39, Санкт-Петербург, 199178; р.т.: +7(812)328-08-87, Факс: +7(812)328-44-50.

Воробьев Владимир Иванович — д-р техн. наук, профессор, главный научный сотрудник лаборатории информационно-вычислительных систем и технологий программирования, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: информационная безопасность, облачные и параллельные вычисления. Число научных публикаций — 110. vvi@iias.spb.su; 14-я линия В.О., 39, Санкт-Петербург, 199178; р.т.: +7(812)3284369, Факс: +7(812)3284450.

Левоневский Дмитрий Константинович — научный сотрудник лаборатории информационно-вычислительных систем и технологии программирования, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: информационная безопасность, защита информации, компьютерные сети, моделирование компьютерных процессов, технологии программирования. Число научных публикаций — 20. dl@iias.spb.su; 14-я линия В.О., 39, Санкт-Петербург, 199178; р.т.: +7(812)328-43-69, Факс: +7(812)328-44-50.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проект №16-29-09482).

DOI 10.15622/sp.53.5

V.Yu. OSIPOV, V.I. VOROBIEV, D.K. LEVONEVSKIY

PROBLEMS OF PROTECTION AGAINST FALSE INFORMATION IN COMPUTER NETWORKS

Osipov V.Yu, Vorobiev V.I., Levonevskiy D.K. **Problems of Protection against False Information in Computer Networks.**

Abstract. This paper provides an analysis of the present state in the field of protection against false information in computer networks and formulates current problems related to this protection. An approach to assessing protection activities on the basis of the Markov chain of the disinformation process is proposed. The architecture of a future system of data analysis is described. It implies enhanced methods of text trustworthiness analysis. The proposed complex approach, based on the known and suggested methods, enables detecting false information in computer networks promptly. Furthermore, the proposed method can be used for countering terrorist activities and cybercrimes in order to search for network resources which may be involved in unlawful activities.

Keywords: trustworthiness, false information, text analysis, classification, ontology, information security.

Osipov Vasiliy Yurievich — Ph.D., Dr. Sci., professor, head of computing & information systems and programming technologies laboratory, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: intelligent systems, modeling, information security. The number of publications — 100. osipov_vasiliy@mail.ru; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)328-08-87, Fax: +7(812)328-44-50.

Vorobiev Vladimir Ivanovich — Ph.D., Dr. Sci., professor, chief researcher of computing & information systems and programming technologies laboratory, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: information security, distributed and cloud computations. The number of publications — 110. vvi@iias.spb.su; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)3284369, Fax: +7(812)3284450.

Levonevskiy Dmitriy Konstantinovich — researcher of computing & information systems and programming technologies laboratory, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: information security, computer security, computer networks, modeling of information processes, programming technology. The number of publications — 20. dl@iias.spb.su; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)328-43-69, Fax: +7(812)328-44-50.

Acknowledgements. This research is supported by RFBR (grant 16-29-09482).

References

1. Osipov V.Yu., Yusupov R.M. *Informacionnyj vandalizm, kriminal i terrorizm kak sovremennye ugrozy obshhestvu* [Information vandalism, crime and terrorism as modern threats to the society]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2009. vol. 8. pp. 34–45. (In Russ.).

2. Osipov V.Yu., Ilyin A.P., Frolov V.P., Kondratyuk A.P. *Radioelektronnaya borba. Teoreticheskiye osnovy* [Electronic warfare. Theoretical basis]. Petrodvorets: VMIRE. 2006. 302 p. (In Russ.).
3. Alekseeva I.Yu. et al. *Informatsionnye vyzovy natsionalnoy i mezhdunarodnoy bezopasnosti* [Information challenges againt national and international security]. M.: PIR-Centr. 2001. 328 p. (In Russ.).
4. Bartlett J., Reynolds L. The State of the Art 2015: a literature review of social media intelligence capabilities for counter-terrorism. Demos. 2015. 98 p.
5. Kotenko I.V., Chechulin A.A., Komashinsky D.V. [Automated categorization of websites for inappropriate content blocking]. *Problemy informatsionnoy bezopasnosti. Kompyuternye sistemy – Information security problems. Computer systems*. 2015. vol. 2. pp. 69–79. (In Russ.).
6. Avast Online Security browser extension: Overview. Available at: https://www.avast.ru/faq.php?article=AVKB18 (accessed: 14.04.2017).
7. Smirnov I.V., Shelmanov A.O., Kuznetsova E.S., Khramoin I.V. [Semantic-syntactic analysis of natural languages. Part II. Method for semantic-syntactic analysis of texts]. *Iskusstvenny intellekt i prinyatie resheniy – Artificial intelligence and decision making*. 2014. vol. 1. pp. 11–24. (In Russ.).
8. Karpovich S., Smirnov A., Teslya N., Grigorev A. Topic Model Visualization With IPython. Proceedings of the 20th Conference of FRUCT association. 2017. pp. 131–137.
9. Dong X.L. et al. Knowledge-Base Trust: Estimating the Truthworthiness of Web Sources. Available at: arxiv.org/pdf/1502.03519v1.pdf (accessed: 11.04.2017).
10. Alexandrov V.V., Zaytseva A.A., Kuleshov S.V. [Building glossaries of culturological canons of cyber-social groups in social networks]. *Mezhdunarodny nauchny zhurnal "Innovatsionnaya nauka" – International scientific journal "Innovative science"*. 2016. vol. 12-2. pp. 13–17. (In Russ.).
11. Karpathy A., Fei-Fei L. Deep Visual-Semantic Alignments for Generating Image Descriptions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015 (CVPR 2015). 2015. pp. 3128–3137.
12. Lillo I., Niebles J.C., Soto A. Sparse composition of body poses and atomic actions for human activity recognition in RGB-D videos. Image and Vision Computing. 2017. vol. 59. pp. 63–75.
13. Batchuluun G., Kim J.H., Hong H.G. Fuzzy system based human behavior recognition by combining behavior prediction and recognition. Expert Systems with Applications. 2017. vol. 81. pp. 108–133.
14. Yogesh C.K. et al. Hybrid BBO_PSO and higher order spectral features for emotion and stress recognition from natural speech. Applied Soft Computing. 2017. vol. 56. pp. 217–232.
15. Kaya H., Karpov A., Salah A. Robust Acoustic Emotion Recognition based on Cascaded Normalization and Extreme Learning Machines. Proceedings of the 13th International Symposium on Neural Networks. 2016. LNCS 9719. pp. 115–123.
16. Budkov V., Vatamaniuk I., Basov V., Volf D. Investigation of Speech Signal Parameters Reflecting the Truth of Transmitted Information. Proceedings of the 18th International Conference on Speech and Computer (SPECOM 2016). 2016. LNAI 9811. pp. 419–426.
17. Vorobiev V.I. et al. [Cryptographic Standards Research and Selection on the Basis of Document Intelligent Analysis]. *Trudy SPIIRAN – SPIIRAS Proceedings*. vol 5(48). 2016. pp. 69–87. (In Russ.).
18. Kharlamov A.A. *Sposob formirovaniya smyslovogo portreta teksta i ustroystvo dlya ego osushchestvleniya* [The semantic text portrait building method and the device implementing it]. Patent RF. no. 2000127135. (In Russ.).

19. Poncet J. et al. Access by content based computer system. US Patent. no. WO 2001033419 A2. 2003.
20. Sheth A., Avant D., Bertram C. System and method for creating a semantic web and its applications in browsing, searching, profiling, personalization and advertising. US Patent. no. WO 2001069428 A1. 2001.
21. Omoigui N. System and method for knowledge retrieval, management, delivery and presentation. US Patent no. 20100070448 A1. 2003.
22. Gardner S. Ontology-based information management system and method. US Patent. no. 7225183 B2. 2007.
23. Hovanov N.V. [Complex objects estimation in conditions of lack of information]. *Trudyi 8-y mezhdunarodnoy nauchnoy shkolyi "Modelirovanie i analiz bezopasnosti i riska v slozhnyih sistemah"* [Proceedings of the 8th international scientific school "Security and risk modeling and analysis in complex systems"]. SPb.: IPMAShHRAN. 2008. pp. 18–28. (In Russ.).
24. Perminov S.V., Afanasyev S.V [Semantic Method of Information Anomaly Search via Web]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2006. vol. 3. Issue 1. pp. 279–287. (In Russ.).
25. Perminov S.V. [Semantic search system]. *Informatsionno-izmeritelnye i upravlyayushchie sistemy* – Information-measuring and Control Systems. 2008. vol. 4. Issue 6. pp. 45–50. (In Russ.).
26. Narinyani A.S. [Centaurus named TEON: Thesaurus + Ontology]. *Trudy mezhdunarodnogo seminara Dialog'2001 po kompyuternoy lingvistike i ee prilozheniyam* [Proceedings of the international workshop Dialog'2001 on computer linguistics and its applications]. 2001. vol. 1. pp. 184–188. (In Russ.).
27. Ferrara E. et al. Predicting online extremism, content adopters, and interaction reciprocity. Proceedings of the 8th International Conference on Social Informatics. 2016. Part II. pp. 22–39.