ИНФОРМАЦИЯ И БЕЗОПАСНОСТЬ

Том 20, Выпуск 4 *Ноябрь 2017 г.*

ФГБОУ ВО
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

РЕГИОНАЛЬНЫЙ УЧЕБНО-НАУЧНЫЙ ЦЕНТР
ПО ПРОБЛЕМАМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Решением Президиума ВАК Российской Федерации
журнал «Информация и безопасность» включен в перечень ведущих научных изданий,
в которых публикуются результаты диссертаций на соискание ученой степени доктора наук

АДРЕС РЕДАКЦИИ:
394049, г. Воронеж, ул. Ватутина, д. 1
тел./факс: (473) 252-34-20

e-mail:mnac@comch.ru

УЧРЕДИТЕЛЬ:
ФГБОУ ВО «Воронежский
государственный технический университет»

СОДЕРЖАНИЕ

МЕЖРЕГИОНАЛЬНЫЙ ФОРУМ ЦИФРОВЫХ ТЕХНОЛОГИЙ

МАТЕРИАЛЫ
Научно-практической конференции «Безопасный Интернет»

УДК 004.8

МОДЕЛЬ ОЦЕНКИ БЕЗОПАСНОСТИ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ НА ОСНОВЕ МЕТОДА ВЕЙВЛЕТ-АНАЛИЗА

А.О. Калашников, Е.А. Сакрутина

В работе рассматривается модель оценки безопасности критической информационной инфраструктуры на основе прогнозирования рисков информационной безопасности объектов критической информационной инфраструктуры, находящихся под воздействием компьютерных атак.

Ключевые слова: критическая информационная инфраструктура, оценка безопасности, риски информационной безопасности, вейвлет-анализ, ассоциативный поиск

Введение

Приоритетной целью государственной политики на современном этапе является ускоренный переход к цифровой экономике.

Данный переход характеризуется интенсивным внедрением и использованием перспективных *информационных технологий*[1] в сферах экономики и финансов, промышленности и энергетики, транспорта и связи, государственного и муниципального управления, обороны и безопасности, науки и культуры, образования и здравоохранения и многих других.

Однако, широкое и повсеместное использование информационных технологий немыслимо без повышенного внимания к проблемам их *безопасности*[2].

И в первую очередь это относится к вопросам обеспечения информационной безопасности объектов *критической информационной инфраструктуры Российской Федерации* (далее – КИИ РФ) и КИИ РФ в целом.

О том, что важность указанной проблемы отчетливо осознается, в том числе, на уровне Президента и Правительства Российской Федерации говорит и принятый недавно Федеральный закон от 26.07.2017 № 187-ФЗ «О безопасности критической

Калашников Андрей Олегович – ИПУ РАН, д. т. н., в.н.с., зам. руководителя ВНОЦ, e-mail: aokalash@ipu.ru

Сакрутина Екатерина Алексеевна – ИПУ РАН, н.с., e-mail: consoft@ipu.ru

[1] *Информационная технология* – приемы, способы и методы применения средств вычислительной техники при выполнении функций сбора, хранения, обработки, передачи и использования данных ([1], статья 3.1.9).

[2] *Безопасность информационной технологии* – состояние информационной технологии, определяющее защищенность информации и ресурсов информационной технологии от действия объективных и субъективных, внешних и внутренних, случайных и преднамеренных угроз, а также способность информационной технологии выполнять предписанные функции без нанесения неприемлемого ущерба субъектам информационных отношений ([1], статья 3.1.2).

информационной инфраструктуры Российской Федерации» [2] (далее – ФЗ от 26.07.2017 № 187-ФЗ), который вступает в силу с 01.01.2018.

В соответствии с указанным законом под КИИ РФ понимаются – «объекты критической информационной инфраструктуры, а также сети электросвязи, используемые для организации взаимодействия таких объектов».

В свою очередь, *объектами* КИИ РФ являются – «информационные системы, информационно-телекоммуникационные сети, автоматизированные системы управления субъектов критической информационной инфраструктуры», а *субъектами* – «государственные органы, государственные учреждения, российские юридические лица и (или) индивидуальные предприниматели, которым на праве собственности, аренды или на ином законном основании принадлежат информационные системы, информационно-телекоммуникационные сети, автоматизированные системы управления, функционирующие в сфере здравоохранения, науки, транспорта, связи, энергетики, банковской сфере и иных сферах финансового рынка, топливно-энергетического комплекса, в области атомной энергии, оборонной, ракетно-космической, горнодобывающей, металлургической и химической промышленности, российские юридические лица и (или) индивидуальные предприниматели, которые обеспечивают взаимодействие указанных систем или сетей».

Наконец, под *безопасностью* КИИ РФ в рамках ФЗ от 26.07.2017 № 187-ФЗ понимают – «состояние защищенности критической информационной инфраструктуры, обеспечивающее ее устойчивое функционирование при проведении в отношении ее компьютерных атак».

Среди всех объектов КИИ РФ выделяют подмножество так называемых *значимых объектов* КИИ РФ, где под значимым объектом понимается – «объект критической информационной инфраструктуры, которому присвоена одна из категорий значимости и который включен в реестр значимых объектов критической информационной инфраструктуры».

К значимым объектам КИИ РФ могут относится крупные гидротехнические сооружения, объекты атомной энергетики, вредные химические производства, транспортные узлы, аэропорты и т.п.

Воздействие компьютерных атак на информационно-технологическую инфраструктуру (далее – ИТИ) указанных объектов, приводящее к выходу ее технологических параметров за установленные нормативные пределы, может повлечь за собой реализацию нештатных ситуаций с тяжелыми и даже катастрофическими последствиями.

В соответствии с ФЗ от 26.07.2017 № 187-ФЗ *категорирование* объекта КИИ РФ представляет собой установление соответствия объекта КИИ РФ критериям значимости и показателям их значений, присвоение ему одной из категорий значимости, проверку сведений о результатах ее присвоения.

Категорирование объектов КИИ РФ осуществляется исходя из:

1) социальной значимости, выражающейся в оценке возможного ущерба, причиняемого жизни или здоровью людей, возможности прекращения или нарушения функционирования объектов обеспечения жизнедеятельности населения, транспортной инфраструктуры, сетей связи, а также максимальном времени отсутствия доступа к государственной услуге для получателей такой услуги;

2) политической значимости, выражающейся в оценке возможного причинения ущерба интересам Российской Федерации в вопросах внутренней и внешней политики;

3) экономической значимости, выражающейся в оценке возможного причинения прямого и косвенного ущерба субъектам критической информационной инфраструктуры и (или) бюджетам Российской Федерации;

4) экологической значимости, выражающейся в оценке уровня воздействия на окружающую среду;

5) значимости объекта критической информационной инфраструктуры для обеспечения обороны страны, безопасности государства и правопорядка.

При этом устанавливаются три категории значимости объектов критической информационной инфраструктуры – первая, вторая и третья.

Для успешной реализации мероприятий по обеспечению безопасности объектов КИИ РФ и КИИ РФ в целом необходимо решение целого ряда сложных научно-технических задач, из которых задача оценки текущего (реального) уровня безопасности объектов КИИ РФ и КИИ РФ в целом и прогнозирования его изменения является одной из ключевых.

Именно построению модели оценки безопасности КИИ РФ на основе прогнозирования рисков информационной безопасности объектов КИИ РФ, находящихся под воздействием компьютерных атак будет посвящено дальнейшее содержание настоящей статьи.

Оценка безопасности объектов КИИ РФ

Основные положения.

Объектом нашего исследования является КИИ РФ.

В соответствии с ФЗ от 26.07.2017 № 187-ФЗ будем полагать, что КИИ РФ состоит из объектов двух основных типов: *значимых объектов* (далее – ЗОКИИ) и *сетей электросвязи*, используемых для организации взаимодействия таких объектов.

Состояние ЗОКИИ может быть описано текущим *уровнем информационной безопасности* (далее – УИБ) и связанным с ним *рисковым потенциалом* (далее – РП). УИБ ЗОКИИ определяется текущим уровнем *информационных угроз* (далее – УИУ) и текущим уровнем *защищенности* (далее – УЗ) ЗОКИИ.

В свою очередь, РП представляет собой прогнозную оценку последствий нарушения функционирования ЗОКИИ при реализации компьютерных атак на ЗОКИИ и возникновении в результате этого компьютерных инцидентов.

Оценка РП ЗОКИИ может быть представлена в многокритериальной (векторной) форме, учитывающей РП ЗОКИИ по отдельным направлениям, определенным ФЗ от 26.07.2017 № 187-ФЗ: социальный риск; политический риск; экономический риск; экологический риск; риск для обеспечения обороны, безопасности и правопорядка.

На основании векторной оценки РП может быть построена интегральная оценка РП ЗОКИИ (подробнее см., например, [3]).

На основании оценок РП отдельных ЗОКИИ и с учетом взаимного влияния ЗОКИИ друг на друга осуществляются прогнозные оценки (векторная и интегральная) РП для КИИ РФ в целом.

Основные подходы к разработке моделей ЗОКИИ.

Будем предполагать, что модель ЗОКИИ состоит из трех взаимосвязанных уровней (подробнее см., например, [4-6]). Верхний уровень – уровень «технологических» (бизнес) процессов (далее – ТП) ЗОКИИ.

Каждый ТП может быть описан набором (вектором) *характеристик*, исходя из установленных эксплуатационных пределов и условий функционирования ЗОКИИ.

Тогда *состояние* ТП может быть описано текущим вектором *значений* характеристик.

Состояние ТП ЗОКИИ, в свою очередь, определяет текущий РП ЗОКИИ.

Если значения вектора характеристик не выходят за границы эксплуатационных пределов и установленных условий функционирования ЗОКИИ, то такое состояние ТП ОКИИ может считаться *нормальным.*

В этом случае будем считать, что РП ЗОКИИ соответствует *допустимому* уровню.

Если значения вектора характеристик выходят за границы значений, вытекающих из установленных эксплуатационных пределов и условий функционирования

ЗОКИИ, то такое состояние ТП ОКИИ может считаться *аварийным.*

В этом случае РП ЗОКИИ будет превышать допустимый уровень.

Средний уровень – уровень информационно-технологических процессов (далее – ИТП), которые обеспечивают нормальное функционирование ТП.

Будем считать, что каждый ИТП также может быть описан вектором характеристик (в общем случае отличным от вектора характеристик ТП).

Тогда, для того чтобы обеспечить нахождение ТП в нормальном состоянии, значения вектора характеристик ИТП также должны находиться в определенных диапазонах, которые будем считать диапазонами нормальных значений характеристик ИТП.

Если значения вектора характеристик ИТП находятся в пределах указанных выше диапазонов нормальных значений, то такое состояние ИТП ЗОКИИ может считаться *нормальным.*

Если значения вектора характеристик ИТП выходят за диапазоны нормальных значений, то такое состояние ИТП ЗОКИИ может считаться *аварийным.*

Очевидно, что в этом случае, если ИТП ЗОКИИ находится в аварийном состоянии, то и ТП ЗОКИИ находится в аварийном состоянии.

Нижний уровень – уровень информационно-технологической инфраструктуры (далее – ИТИ) ЗОКИИ, обеспечивающий функционирование ИТП ЗОКИИ.

Будем считать, что состояние ИТИ также может быть описано вектором характеристик (в общем случае отличным от вектора характеристик ТП и ИТП).

Тогда, для того чтобы обеспечить нахождение ИТП в нормальном состоянии, значения вектора характеристик ИТИ также должны находиться в некоторых определенных диапазонах, которые будем считать диапазонами нормальных значений характеристик ИТИ.

Если значения вектора характеристик ИТИ находятся в пределах указанных выше диапазонов нормальных значений, то такое состояние ИТИ ЗОКИИ может считаться *нормальным.*

Если значения вектора характеристик ИТИ выходят за диапазоны нормальных значений, то такое состояние ИТП ЗОКИИ может считаться *аварийным.*

Очевидно, что в этом случае, если ИТИ ЗОКИИ находится в аварийном состоянии, то и ИТП ЗОКИИ, а, следовательно, и ТП ЗОКИИ находятся в аварийном состоянии.

С учетом сказанного выше, параметры вектора характеристик, описывающего состояние ИТИ ЗОКИИ, будем называть *критическими переменными состояния*.

Таким образом, в целом, можно считать, что **существует определенная «функциональная» зависимость между состоянием ИТИ ЗОКИИ, описываемым вектором характеристик ИТИ, и РП ЗОКИИ**, **описываемым вектором РП**.

При этом отклонение состояния ИТИ ЗОКИИ от нормального приводит к соответствующему отклонению РП ЗОКИИ от допустимого уровня.

Параметры указанной выше зависимости могут быть определены экспертно-аналитическими методами в сочетании с методами анализа данных и стохастического имитационного моделирования.

Таким образом, учитывая вышесказанное, при построении модели ЗОКИИ представляется достаточным ограничиться построением **параметрической модели ИТИ ЗОКИИ** и **некоторой регрессионной моделью, описывающей зависимость вектора РП от параметров ИТИ**.

Исходные данные для построения моделей могут быть получены в процессе проведения *категорирования* объектов КИИРФ и *оценки защищенности* ЗОКИИ и КИИ РФ в целом. Далее параметры модели ЗОКИИ могут изменяться при появлении дополнительной информации об изменении структуры ИТИ, появлении новых информационных угроз, воздействий компьютерных атак, результатов расследования компьютерных инцидентов и т.д.

Основные подходы к разработке модели КИИ РФ в целом.

Как было показано выше, КИИ РФ представляет собой совокупность взаимосвязанных и взаимозависимых ЗОКИИ.

Таким образом, изменения параметров одного ЗОКИИ в общем случае должно оказывать определенное влияние на параметры других ЗОКИИ.

Изменения параметров могут носить как *глобальный*, затрагивающий большинство ЗОКИИ или КИИ РФ в целом (появление новых информационных угроз или методов и средств защиты информации, изменение модели нарушителя или защищающегося и т.д.), так и *локальный* характер, затрагивающий один или несколько ЗОКИИ (изменение конфигурации ИТИ, появлении новых компьютерных инцидентов и т.д.).

Учитывая наличие определенной зависимости между состоянием ИТИ ЗОКИИ, описываемым вектором характеристик ИТИ и РП этого ЗОКИИ, описываемым вектором РП, а также влияние изменений параметров одного ЗОКИИ на параметры других ЗОКИИ, можно говорить о наличии **определенной функциональной зависимости, описывающей влияние УИБ и РП одного ЗОКИИ на УИБ и РП других ЗОКИИ**.

Параметры указанной выше зависимости могут быть определены экспертно-аналитическими методами в сочетании с методами анализа данных и стохастического имитационного моделирования.

В этом случае представляется целесообразным использовать в качестве модели КИИ РФ взвешенный ориентированный мультиграф, где вершины графа представляют собой ЗОКИИ, а ребра – возможность влияния изменений параметров одного ОКИИ на параметры других ОКИИ.

Приписываемые вершинам значения факторов представляют собой векторы УИБ и РП данного ЗОКИИ, а характеристики ребер (в качестве этих характеристик могут выступать «веса», «передаточные» функции и т.д.) – адаптированную форму определенной зависимости (вообще говоря, стохастической), описывающей влияние УИБ и РП данного ЗОКИИ на УИБ и РП других ЗОКИИ.

Таким образом, описание КИИ в рамках указанной модели представляет собой многомерную «матрицу» параметров различных ЗОКИИ (УИБ и РП), а описание влияния ЗОКИИ друг на друга – многомерную «матрицу» функциональных зависимостей УИБ и РП одного ЗОКИИ от УИБ и РП других ЗОКИИ.

Основные подходы к оценке УИБ и РП КИИ РФ в целом.

Для разработки методов оценки УИБ и РП КИИ РФ в целом представляется целесообразным использовать обобщенную комплексную оценку УИБ и РП КИИ РФ в целом.

Она может быть получена путем агрегирования локальных оценок УИБ и РП ЗОКИИ с использованием, как было показано выше, методов комплексного оценивания (например, основанных на бинарных матричных свертках [3]).

В свою очередь для решения задачи прогнозирования рисков информационной безопасности объектов КИИ РФ могут быть использованы различные подходы.

Однако, исходя из специфики описанных выше моделей, представляется целесообразным использовать методы прогнозирования параметрических временных рядов на основе вейвлет-анализа, как например [7, 8], о чем подробно будет сказано в следующей части статьи.

Вейвлет-анализ и идентификация критических переменных состояния

При решении задачи идентификации критических переменных состояния можно выделить широкий класс процессов, для управления которыми не достаточно построения линейных моделей.

Кроме того, данные процессы могут иметь некоторые особенности в определенные моменты времени.

Решение задачи построения прогнозирующих моделей для нестационарных процессов является особенно актуальной при оценке безопасности объектов КИИ РФ и КИИ РФ в целом.

Первые работы по вейвлет-анализу временных рядов с выраженной неоднородностью появились в середине 80-х годов [9].

Метод был позиционирован как альтернатива преобразованию Фурье [10], локализующему частоты, но не дающему временного разрешения процесса.

В дальнейшем появилась и развивается теория вейвлетов и ее многочисленные приложения.

На рис. 1 проиллюстрированы подходы совместного использования преобразования Фурье, теории вейвлетов [11] и теории идентификации систем [12].

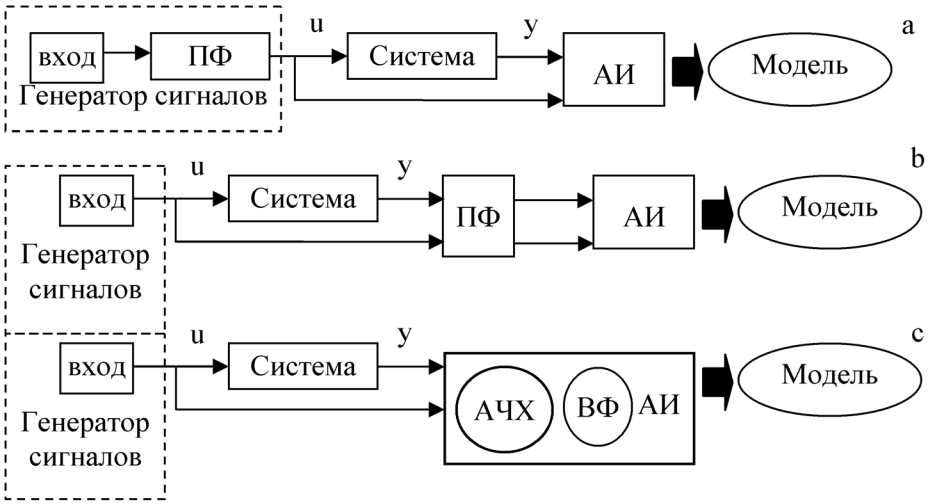


Рис. 1. Основные принципы алгоритмов идентификации: (a, b) показывают классические методы идентификации в частотной области, (c) является модификацией.
ПФ – преобразование Фурье, АИ – алгоритм идентификации, АЧХ – анализ частотных характеристик, ВФ – вейвлет-фильтр

Классический подход к идентификации систем в частотной области состоит в том, что необходимо сначала отфильтровать входные данные, подать отфильтрованные входы в систему, а затем провести идентификацию (см. рис. 1(a)).

Этот подход имеет существенную ограниченность – необходима априорная информация о свойствах частот системы, чтобы выявлять важные частоты при построении фильтра.

Кроме того, если какая-либо резонансная частота не выявлена, то отклик модели не будет ее содержать, и, соответственно, может быть утеряно важное свойство исходной системы.

Проблема может быть решена путем использования фильтрации после подачи входов в систему (см. рис.1(b)).

Другой подход может заключаться в осуществлении вейвлет-фильтрации и частотного анализа неотфильтрованных входных и выходных сигналов.

Однако, чтобы сохранить временную структуру задачи идентификации систем, невозможно разделить вейвлет-преобразование и идентификацию.

Таким образом, и частотный анализ и вейвлет-преобразование должны быть

включены в алгоритм идентификации, как показано на рис. 1(c).

Виды и основные характеристики вейвлетов

С использованием вейвлетов можно приближать сложный *сигнал*[3] с очень высокой точностью. Вследствие выделения локальных особенностей у исследуемого процесса, отсутствующих у рядов Фурье, и достаточного разнообразия, вейвлеты все чаще находят практическое применение для анализа отличительных особенностей сложных сигналов.

Базисные функции данного преобразования занимают промежуточное положение между крайними случаями – гармоническими и импульсными функциями [13].

При формировании вейвлет-преобразований в общем случае применяются две непрерывные, взаимозависимые и интегрируемые по независимой переменной функции:

- вейвлет-функция $\psi(t)$ с нулевым значением интеграла и частотным Фурье-образом $\Psi(\omega)$, характер которой отражает локальные особенности сигнала;
- масштабирующая функция $\varphi(t)$ (так называемая *скейлинг-функция*) с единичным значением интеграла, с помощью которой осуществляется грубое приближение (аппроксимация) сигнала.

Скейлинг-функции определяются только для ортогональных вейвлетов.

Их целесообразно использовать для анализа низкочастотных и высокочастотных составляющих.

Базисные функции $\psi(t)$, формирующие вейвлеты (эти функции выбираются в соответствии со спецификой задач), должны удовлетворять следующим условиям [14]:

- *Локализация*: вейвлет должен быть локализован вблизи нуля аргумента как во временной, так и в частотной области. Например, дельта-функция $\delta(t)$ и гармоническая функция не удовлетворяют этому условию.
- *Нулевое среднее*: $\int_{-\infty}^{\infty}\psi(t)dt = 0$, что означает знакопеременность (осцилляцию) графика исходной функции вокруг нуля на оси времени, и он должен иметь нулевую площадь
- *Равенство нулю n первых моментов*: $\int_{-\infty}^{+\infty} t^n\psi(t)dt = 0$.

Такие вейвлеты называют *вейвлетами n-го порядка*.

Соответствующее вейвлет-преобразование позволяет более детально анализировать высокочастотную составляющую процесса.

- Это может быть полезным для анализа процессов, характеризующихся резкими важными для анализа «пиками».
- *Автомодельность* (*самоподобиевейвлет-преобразования*): все вейвлеты конкретного семейства $\psi_{ab}(t)$ имеют то же число осцилляций, что и материнский вейвлет $\psi(t)$, поскольку получены из него посредством масштабных преобразований (a) и сдвига (b).
- *Ограниченность*: $\int_{-\infty}^{+\infty}|\psi(t)|^2dt < \infty$.
- Вейвлет должен быть достаточно быстро убывающей функцией временной (пространственной) переменной.

Определения и свойства одномерного непрерывного вейвлет-преобразования обобщаются на многомерный и на дискретный случаи.

Примеры наиболее известных вейвлетов представлены в [15]:

- **вещественные непрерывныебазисы**: Гауссовы, DifferenceofGaussians (DOG), Littlewood & Paley (LP);
- **вещественные дискретные**: HAAR-вейвлет (вейвлетХаара), FHAT («Французскаяшляпа» – Frenchhat);
- **комплексные**: Морле (Morlet), Пауля (Paul);

[3] Здесь и далее под сигналом будем понимать временной ряд, описывающий изменениеуровня безопасности объектовКИИ РФ или КИИ РФ в целом.

- **вейвлеты, определяемые итерационными выражениями**: функций Добеши (Daubechies) [16].

За счет выбора определенных вейвлетов можно более полно выявить характерные нестационарные области анализируемого технологического процесса.

Однако, некоторые свойства вейвлет-преобразования сигналов не зависят от выбора материнских вейвлетов.

К таким свойствам относятся: линейность, инвариантность относительно сдвига, инвариантность относительно масштабирования, дифференцируемость [17].

Наиболее часто в задачах анализа используется вейвлет Хаара.

Базис Хаара известен с 1910 г. и широко используется в преобразованиях, например, на его основе разработан стандарт кодирования изображений со сжатием JPEG2000 [18].

Модель оценки безопасности КИИ РФ

Предположим, что данные о оценках безопасности КИИ РФ содержатся в некоторой пополняемой базе знаний (БЗ), в которую заносятся данные о функционировании КИИ РФ. Пусть $x_1(t), \dots, x_S(t)$–характеристики состояния КИИ, $X(t)$ – состояние КИИ РФ характеризующееся вектором характеристик состояния КИИ РФ в момент времениN ($X(t) = \{x_1(t), \dots, x_S(t)\}$), $R(t)$ – оценка безопасности состояния X_t КИИ РФ в момент времени t.

Процесс обработки исторических данных в БЗ сводится к ассоциативному поиску состояний КИИ РФ близких к текущему в БЗ (см. рис. 2).

Критерий близости между состояниями может быть представлен в виде: логической функции – предиката, расстояния в n-мерном пространстве, когда наборы признаков представляют собой векторы в n-мерном пространстве.

Процесс ассоциативного поиска может осуществляться либо как процесс восстановления состояния по частично заданным характеристикам, либо как процесс поиска связанных ассоциативно с данным состоянием других состояний, привязанных к другим моментам времени.

Вместо восстановления состояния КИИ РФ по частично заданным характеристикам может осуществляться восстановление фрагмента состояния в условиях неполной информации.

В работах [17, 19-23] предложен подход к формированию поддержки принятия решения об управлении, основанный на динамическом моделировании процедуры ассоциативного поиска.

Метод прогнозирования выхода объекта (в данном случае оценки безопасности состояния КИИ РФ) на основе ассоциативного поиска, состоит в построении виртуальных прогнозирующих моделей.

Метод предполагает построение новой прогнозирующей модели оценки безопасности в каждый момент времени t, с использованием исторических данных («ассоциаций»), сформированных на этапе обучения и адаптивно корректируемых в соответствии с определенными критериями.

Пусть линейная динамическая модель оценки безопасности КИИ РФ имеет следующий вид:

$$R(t) = \sum_{i=1}^{m} a_i R(t-i) + \sum_{j=1}^{r_s} \sum_{s=1}^{S} b_{j,s} x_s(t-j) \quad (1)$$

где:

$R(t)$ – прогноз оценки безопасности состояния КИИ РФ в момент времени t,

x_s – характеристика КИИ РФ,

m – глубина памяти по оценке безопасности состояния КИИ РФ,

S – размерность вектора характеристик состояния КИИ РФ,

a_i и $b_{j,s}$ – настраиваемые коэффициенты.

Причем, при построении модели (1) из БЗ КИИРФ $x_s(t-j)$ выбираются в каждый момент времени не в хронологическом порядке, а r_s является количеством векторы параметров состояния выбранных по критерию минимума расстояния из БЗ КИИ РФ.

Модель (1) не является классической регрессионной моделью, так как при построении модели из БЗ КИИ РФ выбираются векторы параметров состояния не в хронологической последовательности, а лишь близкие к текущему вектору параметров состояния в смысле определенного критерия.

Для построения виртуальной модели оценки безопасности КИИ РФ в данный момент времени по текущему состоянию КИИ РФ воспользуемся следующим критерием отбора входных векторов из БЗ, приведенным ниже.

Введем в качестве расстояния (нормы в $\Re^S$) между точками S-мерного пространства параметров состояния величину:

$$d_{t,t-j} = \sum_{s=1}^{S} |x_s(t) - x_s(t-j)|, \quad \forall j = \overline{1, t-1}, \tag{2}$$

где $x_s(t)$– компоненты вектора состояний КИИ РФ в момент времени t.

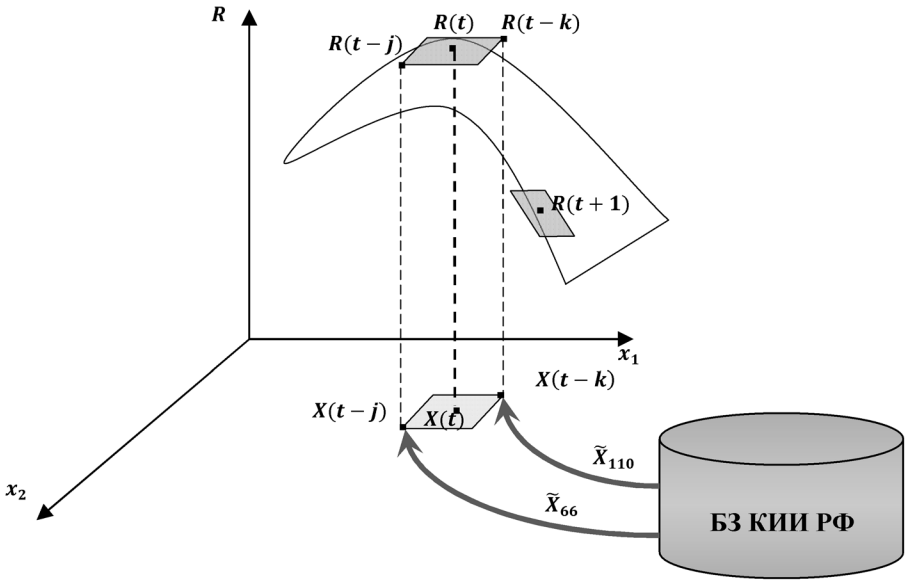


Рис. 2. Ассоциативный поиск в БЗ КИИ РФ

В силу одного из свойств нормы («неравенство треугольника») имеем:

$$d_{t,t-j} \le \sum_{s=1}^{S} |x_s(t)| + \sum_{s=1}^{S} |x_s(t-j)|, \quad \forall j = \overline{1, t-1}. \tag{3}$$

Пусть для текущего вектора состояний КИИ РФ $x_s(t)$:

$$\sum_{s=1}^{S} |x_s(t)| = d_t. \tag{4}$$

Для построения аппроксимирующей гиперповерхности для $X(t)$ отберем из архива входных данных такие векторы $x_s(t-j)$, $j = \overline{1, t-1}$, что для некоторого заданного D_t будет выполнено условие:

$$d_{t,t-j} \le d_t + \sum_{s=1}^{S} |x_s(t-j)| \le D_t, \quad \forall j = \overline{1, t-1}, \tag{5}$$

где D_N может быть выбрано, например, из условия:

$$D_t \geq 2d_t^{max} = 2\max_j \sum_{s=1}^{S} |x_s(t-j)|. \quad (6)$$

Если в выбранной области не наберется достаточного количества параметров состояния для применения МНК, т.е. соответствующая система линейных уравнений окажется неразрешимой, то выбранный критерий отбора точек в пространстве входов можно будет ослабить за счет увеличения порога D_t.

В предположении, что входные воздействия удовлетворяют условиям Гаусса-Маркова, оценки, получаемые по методу наименьших квадратов, являются состоятельными, несмещенными и статистически эффективными.

Условие устойчивости модели прогнозирования

Пусть модель прогнозирования критических параметров состояния на основе ассоциативного поиска имеет вид (1). Для выбранного уровня детализации L текущих векторов состояния получаем кратно-масштабное разложение [13]:

$$\begin{aligned} x_s(t) &= \sum_{k=1}^{N} c_{L,k}^{x_s}(t)\varphi_{L,k}(t) \\ &\quad + \sum_{l=1}^{L}\sum_{k=1}^{N} d_{l,k}^{x_s}(t)\psi_{l,k}(t), \\ R(t) &= \sum_{k=1}^{N} c_{L,k}^{R}(t)\varphi_{L,k}(t) \\ &\quad + \sum_{l=1}^{L}\sum_{k=1}^{N} d_{l,k}^{R}(t)\psi_{l,k}(t), \end{aligned} \quad (7)$$

где L – глубина кратно-масштабного разложения ($1 \leq L \leq L_{max}$, где $L_{max} = \lfloor \log_2 N^* \rfloor$ и N^* – мощность множества состояний системы в БЗ КИИ РФ);

$\varphi_{L,k}(t)$ – масштабирующие функции («скейлинг-функции);

$\psi_{l,k}(t)$ – вейвлет-функции, которые получаются из материнских вейвлетов путем растяжения/сжатия и сдвига:

$$\psi_{l,k}(t) = 2^{l/2}\psi_{\text{материнский}}(2^l t - k),$$

где в качестве материнских вейвлетов рассматриваются вейвлеты Хаара;

l – уровень детализации анализа;

$c_{L,k}$ – масштабирующие коэффициенты;

$d_{l,k}$ - детализирующие коэффициенты. Коэффициенты вычисляются посредством алгоритма Малла [13].

В соответствии с (7) разложим уравнение (1) по вейвлетам:

$$\begin{aligned} &\sum_{k=1}^{N} c_{L,k}^{R}(t)\varphi_{L,k}(t) + \sum_{l=1}^{L}\sum_{k=1}^{N} d_{l,k}^{R}(t)\psi_{l,k}(t) = \\ &+ \sum_{l=1}^{L}\sum_{k=1}^{N}\left(\sum_{i=1}^{m} a_i d_{l,k}^{R}(t-i)\psi_{l,k}(t-i)\right) + \\ &+ \sum_{k=1}^{N}\left(\sum_{s=1}^{S}\sum_{j=1}^{r_s} b_{s,j} c_{L,k}^{x_s}(t-j)\varphi_{L,k}(t-j)\right) + \\ &+ \sum_{l=1}^{L}\sum_{k=1}^{N}\left(\sum_{s=1}^{S}\sum_{j=1}^{r_s} b_{s,j} d_{l,k}^{x_s}(t-j)\psi_{l,k}(t-j)\right). \end{aligned}$$

Рассмотрим отдельно аппроксимирующую (8) и детализирующую (9) части:

$$\begin{aligned} &c_{L,k}^{R}(t)\varphi_{L,k}(t) = \\ &\sum_{i=1}^{m} a_i c_{L,k}^{R}(t-i)\varphi_{L,k}(t-i) + \\ &+ \sum_{s=1}^{S}\sum_{j=1}^{r_s} b_{s,j} c_{L,k}^{x_s}(t-j)\varphi_{L,k}(t-j). \end{aligned} \quad (8)$$

$$\begin{aligned} &d_{l,k}^{R}(t)\psi_{l,k}(t) = \\ &\sum_{i=1}^{m} a_i d_{l,k}^{R}(t-i)\psi_{l,k}(t-i) + \\ &+ \sum_{s=1}^{S}\sum_{j=1}^{r_s} b_{s,j} d_{l,k}^{x_s}(t-j)\psi_{l,k}(t-j) \end{aligned} \quad (9)$$

Объект, описываемый приведенными соотношениями (8) и (9), будет устойчив, если одновременно будут устойчивы N уравнений (соответствующие соотношениям относительно каждого из слагаемых по k $(k = 1, \ldots, N)$ для аппроксимирующей и детализирующей частей.

Тренд и циклическую составляющего можно представить в виде

аппроксимирующей части, а значения флуктуаций на данных интервалах, характеризующих активность (аномальность) сигналов, с учетом случайной шумовой помехи – детализирующей частью[24].

Вид условия устойчивости для прогнозирующей модели оценки безопасности КИИ РФ будет зависеть от вида модели (1). Если $m > r_s$, то условие устойчивости аппроксимирующей части модели (1)имеет вид:

$$\left|\frac{a_1 c_{L,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}{2c_{L,k}^R(t)}\right| < 1,$$

$$\left|-\frac{a_2 c_{L,k}^R(t-2) + \sum_{s=1}^{S} b_{s,2} c_{L,k}^{x_s}(t-2)}{a_1 c_{L,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}\right| < 1,$$

…,

$$\left|-\frac{a_{r_s+1} c_{L,k}^R(t-r_s-1)}{a_{r_s} c_{L,k}^R(t-r_s) + \sum_{s=1}^{S} b_{s,r_s} c_{L,k}^{x_s}(t-r_s)}\right| < 1,$$

$$\left|-\frac{a_{r_s+2} c_{L,k}^R(t-r_s-2)}{a_{r_s+1} c_{L,k}^R(t-r_s-1)}\right| < 1,$$

…,

$$\left|-\frac{2a_m c_{L,k}^R(t-m)}{a_{m-1} c_{L,k}^R(t-m+1)}\right| < 1,$$

а для детализирующей части:

$$\left|\frac{a_1 d_{l,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}{2d_{l,k}^R(t)}\right| < 1,$$

$$\left|-\frac{a_2 d_{l,k}^R(t-2) + \sum_{s=1}^{S} b_{s,2} d_{l,k}^{x_s}(t-2)}{a_1 d_{l,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}\right| < 1,$$

…,

$$\left|-\frac{a_{r_s+1} d_{l,k}^R(t-r_s-1)}{a_{r_s} d_{l,k}^R(t-r_s) + \sum_{s=1}^{S} b_{s,r_s} d_{l,k}^{x_s}(t-r_s)}\right| < 1,$$

$$\left|-\frac{a_{r_s+2} d_{l,k}^R(t-r_s-2)}{a_{r_s+1} d_{l,k}^R(t-r_s-1)}\right| < 1,$$

…,

$$\left|-\frac{2a_m d_{l,k}^R(t-m)}{a_{m-1} d_{l,k}^R(t-m+1)}\right| < 1.$$

Если $m < r_s$, то условие устойчивости аппроксимирующей части модели (1) имеет вид:

$$\left|\frac{a_1 c_{L,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}{2c_{L,k}^R(t)}\right| < 1,$$

$$\left|-\frac{a_2 c_{L,k}^R(t-2) + \sum_{s=1}^{S} b_{s,2} c_{L,k}^{x_s}(t-2)}{a_1 c_{L,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}\right| < 1,$$

…,

$$\left|-\frac{\sum_{s=1}^{S} b_{s,m+1} c_{L,k}^{x_s}(t-m-1)}{a_m c_{L,k}^R(t-m) + \sum_{s=1}^{S} b_{s,m} c_{L,k}^{x_s}(t-m)}\right| < 1,$$

$$\left|-\frac{\sum_{s=1}^{S} b_{s,m+2} c_{L,k}^{x_s}(t-m-2)}{\sum_{s=1}^{S} b_{s,m+1} c_{L,k}^{x_s}(t-m-1)}\right| < 1,$$

…,

$$\left|-\frac{2\sum_{s=1}^{S} b_{s,r_s} c_{L,k}^{x_s}(t-r_s)}{\sum_{s=1}^{S} b_{s,r_s-1} c_{L,k}^{x_s}(t-r_s+1)}\right| < 1,$$

а для детализирующей части:

$$\left|\frac{a_1 d_{l,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}{2d_{lk}^R(t)}\right| < 1,$$

$$\left|-\frac{a_2 d_{l,k}^R(t-2) + \sum_{s=1}^{S} b_{s,2} d_{l,k}^{x_s}(t-2)}{a_1 d_{l,k}^R(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}\right| < 1,$$

…,

$$\left|-\frac{\sum_{s=1}^{S} b_{s,m+1} d_{l,k}^{x_s}(t-m-1)}{a_m d_{l,k}^R(t-m) + \sum_{s=1}^{S} b_{s,m} d_{l,k}^{x_s}(t-m)}\right| < 1,$$

$$\left|-\frac{\sum_{s=1}^{S} b_{s,m+2} d_{l,k}^{x_s}(t-m-2)}{\sum_{s=1}^{S} b_{s,m+1} d_{l,k}^{x_s}(t-m-1)}\right| < 1,$$

…,

$$\left|-\frac{2\sum_{s=1}^{S} b_{s,r_s} d_{l,k}^{x_s}(t-r_s)}{\sum_{s=1}^{S} b_{s,r_s-1} d_{l,k}^{x_s}(t-r_s+1)}\right| < 1.$$

Если $m = r_s \neq 1$,то условие устойчивости аппроксимирующей части модели (1) имеет вид:

$$\left|\frac{a_1 c_{L,k}^{R}(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}{2c_{L,k}^{R}(t)}\right| < 1,$$

$$\left|-\frac{a_2 c_{L,k}^{R}(t-2) + \sum_{s=1}^{S} b_{s,2} c_{L,k}^{x_s}(t-2)}{a_1 c_{L,k}^{R}(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}\right| < 1,$$

…,

$$\left|-\frac{2\left[a_m c_{L,k}^{R}(t-m) + \sum_{s=1}^{S} b_{s,m} c_{L,k}^{x_s}(t-m)\right]}{a_{m-1} c_{L,k}^{R}(t-m+1) + \sum_{s=1}^{S} b_{s,m-1} c_{L,k}^{x_s}(t-m+1)}\right| < 1,$$

а для детализирующей части:

$$\left|\frac{a_1 d_{l,k}^{R}(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}{2d_{l,k}^{R}(t)}\right| < 1,$$

$$\left|-\frac{a_2 d_{l,k}^{R}(t-2) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-2)}{a_1 d_{l,k}^{R}(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}\right| < 1,$$

…,

$$\left|-\frac{2\left[a_m d_{l,k}^{R}(t-m) + \sum_{s=1}^{S} b_{s,m} d_{l,k}^{x_s}(t-m)\right]}{a_{m-1} d_{l,k}^{R}(t-m+1) + \sum_{s=1}^{S} b_{s,m-1} d_{l,k}^{x_s}(t-m+1)}\right| < 1.$$

Если $m = r_s = 1$,то условие устойчивости аппроксимирующей части модели (1) имеет вид:

$$\left|\frac{a_1 c_{L,k}^{R}(t-1) + \sum_{s=1}^{S} b_{s,1} c_{L,k}^{x_s}(t-1)}{c_{L,k}^{R}(t)}\right| < 1,$$

а для детализирующей части:

$$\left|\frac{a_1 d_{l,k}^{R}(t-1) + \sum_{s=1}^{S} b_{s,1} d_{l,k}^{x_s}(t-1)}{d_{l,k}^{R}(t)}\right| < 1.$$

Вывод условий устойчивости приведен в [25].

Не выполнение условий устойчивости прогнозирующей модели указывает на вероятность сохранения тенденции кувеличению или уменьшению оценки безопасности на следующем такте времени.

Таким образом, можно определить опасные события, которые в будущем могут нарушить безопасность КИИ РФ.

Заключение

В работе предложена модель оценки безопасности КИИ РФ, а также условия устойчивости модели на основе кратномасштабного вейвлет-разложения.

На основе исследования динамики коэффициентов вейвлет-разложения можно определять опасные события, которые в будущем могут нарушить безопасность КИИ РФ.

Предложенная модель предназначена для применения в системе мониторинга угроз безопасности объектов КИИ РФ. Литература

1. ГОСТ Р 53114-2008. Защита информации. Обеспечение информационной безопасности в организации. Основные термины и определения. – М.: Стандартинформ, 2009. – 20 с.

2. Федеральный закон от 26.07.2017№ 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации».

3. Управление информационными рисками организационных систем: механизмы комплексного оценивания / А.О. Калашников // Информация и безопасность. – 2016. – № 3. –С. 315 – 322.

4. Управление информационными рисками организационных систем: базовая модель // Системы управления и информационные технологии. – 2008. – № 1.3(31). – С. 366 – 371.

5. Управление информационными рисками автономных организационных

систем / А.О. Калашников // Системы управления и информационные технологии. – 2008. – № 2.2 (32). – С. 262 – 267.

6. Управление информационными рисками взаимодействующих организационных систем / А.О. Калашников // Системы управления и информационные технологии. – 2008. – № 1.3 (31). – С. 375 – 380.

7. Прогнозирование нестационарных временных рядов при несимметричных функциях потерь / В.Ю. Черных, М.М. Стенина // Машинное обучение и анализ данных – 2015. –Т. 1, №14: С. 1893-1909.

8. Алгоритмы построения статистик для анализа и прогнозирования нестационарных временных рядов / К.П. Осминин // Информационные технологии и вычислительные системы – 2009. – №1: С. 3-13.

9. Decomposition of Hardy functions into square integrable wavelets of constant shape / A. Grossman, J. Morlet // SIAM Journals on Mathematical Analysis – 1984. Т. 15. № 4: P. 723–736.

10. Вейвлеты и их использование / И.М. Дремин, О.В. Иванов, В.А. Нечитайло // Успехи физических наук – 2001. Т. 171. № 5: С. 465–501.

11. An Introduction to Wavelets Through Linear Algebra / M.W. Frazier // Springer Verlag, 1999.

12. System Identification: Theory For The User / L. Ljung // Linkoping University, 1999.

13. Малла, С. Вэйвлеты в обработке сигналов / С. Малла – М.: Мир, 2005. – 671 с.

14. Вейвлет – анализ: основы теории и примеры применения / Н.М. Астафьева // Успехи физических наук – 1996. Т. 166. № 11: С. 1145–1170.

15. Яковлев А. Н. Введение в вейвлет-преобразования: учебное пособие / А.Н. Яковлев // НГТУ, 2003. – 104 с.

16. Daubechies I. Ten lectures on wavelets. University of Lowell, Philadelphia: Society for Industrial and Applied Mathematics (SIAM).1992.

17. Идентификация систем на основе вейвлет-анализа / Е.А.Сакрутина, Н.Н.Бахтадзе// Труды XII Всероссийского совещания по проблемам управления (ВСПУ-2014, Москва) – 2014. М.: ИПУРАН, С. 2868–2889.

18. Анализ одномерного сигнала на основе нечетного и четного базисов вейвлетов с компактными носителями / А.Г.Шоберг// Интеллектуальные системы – 2012. Т. 33. №3: С. 150–157.

19. System identification in frequency domain using wavelets: Conceptual remarks / Z. Váňa, H.A. Preisig // Systems & Control Letters – 2012. Т. 61, №. 10. P. 1041–1051.

20. Associative Search Models in Industrial systems / N. Bakhtadze, V.A. Lototsky, E. Maximov, B.V. Pavlov // IFAC Proceedings Volumes – 2007. Т. 40, № 3, P. 105–108.

21.Multi-agent Approach to Design of Multimodal Intelligent Immune System for Smart Grid / N.N. Bakhtadze, I.B. Yadykin, V.A. Lototsky, E.M. Maximov, E.A. Sakrutina // IFAC Proceedings Volumes – 2013. Т. 46, № 9, P. 1164–1169.

22.Development of Intelligent Identification Models and their Applications to Predict the Submarine Dynamics by Use of Computer Simulation Complexes / N.N. Bakhtadze, B.V. Pavlov, E.A. Sakrutina // IFAC Proceedings Volumes – 2013. Т. 46, № 9, P. 1244–1249.

23. Associative Search and Wavelet Analysis Techniques in System Identification / N.N. Bakhtadze, V.A. Lototsky, S.A. Vlasov, E.A. Sakrutina // IFAC Proceedings Volumes – 2012. Т. 45, № 16, P. 1227–1232.

24. Обнаружение вторжений на основе вейвлет-анализа сетевого трафика / Н.А.Тишина, И.Г. Дворовой, Н.А. Соловьев // Вестник УГАТУ - 2010. - №(40): С. 188–194.

25. Some Functions of the “Safety management system” in the Transportation Area Safety Assurance / E. Sakrutina // Proceedings of 2017 International Siberian Conference on Control and Communications (SIBCON) – 2017. P. 1–5.

26. Kauai, HI: IEEE. Freeman, L. C. (1979). Centrality in social networks conceptual clarification. SocialNetwork. 1(3). P. 215–239.

27. Antsupov, A. Ya. Conflictology: the textbook for higher education institutions / A.Ya. Antsupov, A. I. Shilov. – 3rd prod.,

reslave. and additional – SPb.: St. Petersburg, 2007. – 591 p.

28. Voldstad, R. Community Detection on Last.fm Artist Data / R. Voldstad // – 2014.

29. Panagiotis,Karampelas. Techniques and Tools for Designing an Online Social Network Platform. New Hampshire: Hellenic American University / – M.: World, 1992. – 400 p.

30. Paolo, Massa, Martino Salvetti, and DaniloTomasoni. Bowling alone and trust decline in social network sites. In Proc. Int. Conf. Dependable, Autonomic and Secure Computing, pages 658-663, 2016.

31. Tsvetovat, M. Social Network Analysis for network interests: Finding Connections on the Social Web / M. Tsvetovat, A. Kouznetsov // O'Reilly.-2011. — P. 45. — 192 c.

32. Newman, M. E. (2004). Coauthorship networks and patterns of Scientific Collaboration. Proceedings of the National Academy of Sciences, 101(suppl 1): 5200-5205.

33. Ahn, Y. Analysis of topological characteristics of huge onlane social networking services Advogato/ Y. Ahn, S. Han, H. Knak, S. Moon, H. Jeong // 16th International Conference on the World Wide Web. – 2007. – P. 835-844.

34. Mantzaris A. V. Uncovering nodes that spread information between communities in social networks / A. V. Mantzaris // EPJ Data Science. – 2014.. – Vol.286. – P. 509–512.

35. Blondel V. D. Fast unfolding of communities in large networks / Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre // Journal of statistical mechanics, Theory and experiment. – 2008. – Vol. 2008

36. Woo J. Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog / J. Woo, H. Chen // Graduate School of Information Security, Korea University, Anamro, Seoul, Korea. – 2016. – P. 19.

37. Cannarella J. Epidemical modeling of online social network dynamics / J. Cannarella, J.A. Spechler // Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA. – 2014. – P. 66.

Воронежский научно-образовательный центр управления информационными рисками
Voronezh Research and Education Center for Information Risk Management
Институт проблем управления имени В. А. Трапезникова РАН
Institute of Control Sciences named after V.A. Trapeznikov

A MODEL OF THE CRITICAL INFORMATION INFRASTRUCTURE SECURITY ASSESSMENT ON THE WAVELET ANALYSIS BASIS

A.O. Kalashnikov, E.A. Sakrutina

In the paper, a model of the critical information infrastructure security assessment is considered on the basis of predicting risks of the information security of critical infrastructure objects affected by computer attacks

Key words: critical information infrastructure, security assessment, information security risk, wavelet analysis, associative search

УДК 004.932.2

АНАЛИЗ ИЗОБРАЖЕНИЙ МЕТОДОМ НЕЧЁТКОЙ СЕГМЕНТАЦИИ С ИСПОЛЬЗОВАНИЕМ ПРОГРАММИРУЕМОЙ ВЕНТИЛЬНОЙ МАТРИЦЫ

Ю.Ю. Громов, П.И. Карасев, С.К. Стегачев, О.Г. Иванова

Идентификация и сегментация являются сложными этапами процесса поиска изображений. Сложность связана с большим количеством изображений, которые необходимо обработать при поиске данного фрагмента за короткое время, и автоматической идентификацией интересующей области для оптимизации процесса

Ключевые слова: Сегментация изображений, распознавание образов, обнаружение контуров

Введение

Достижения в области цифровой обработки позволили создавать и хранить большие массивы изображений. Необходимость в хороших поисковых системах диктуется потребностями организации эффективного поиска по коллекциям [1]. Первоначально, подход к поиску изображений заключался в поиске текста к изображениям, аннотированным вручную ключевыми словами; это предполагает большое количество ручного труда и противоречивость оценок.

Поиск изображений по содержанию преодолевает эти трудности с помощью автоматического извлечения свойств изображения, таких как, например: цвет, текстура, форма.

Поиск изображений является стадией общего процесса обработки, в который также входят повышение качества изображения, сжатие и интерпретация.

Сегментация используется для разбиения цифрового изображения на множество областей для облегчения анализа. В процессе сегментации определяются объекты и границы [2]. Интерпретация возможна при условии, что изображение разделено на объекты и фон. Первый шаг в процессе анализа изображений - это сегментация изображения по цвету, за которым следует решение таких задач, как обнаружение границ, определение характерных свойств, которые зависят от качества сегментации.

Слишком подробная сегментация выделяет множество лишних деталей объекта, а недостаточно подробная может сгруппировать множество различных объектов в одном регионе, - поэтому требуемый уровень сегментации является решающим фактором для успеха или неудачи анализа. В алгоритме сегментации изображения используются разрывы и сродство серых уровней изображений.

Одними из наиболее часто используемых методов сегментации являются: гистограммы порога, кластеризация, нечеткий метод с использованием векторов [3]. Программируемые матрицы ПВМ – это интегральные схемы, чей внутренний функционал программируется пользователем. Схема ПВМ содержит логические ячейки, которые могут автономно принимать конечное множество конфигураций. В процессе разработки для каждой ячейки указывается логическая функция; она может быть запрограммирована в диапазоне от простого цифрового логического вентиля, до сложных алгоритмов обработки изображений. Хотя ПВМ медленней, чем традиционная интегральная схема, специализированная для решения конкретной задачи (ASIC - Application Specific Integrated Circuit), гибкость ПВМ при разработке является

Громов Юрий Юрьевич – ТГТУ, доктор техн. наук, профессор, действительный член АИН РФ и РАЕН.
Карасев Павел Игоревич – ТГТУ, соискатель.
Стегачев Сергей Константинович – Тамбовский областной онкологический клинический диспансер, канд. мед. наук, врач высшей квалификационной категории по специальности «Рентгенология».
Иванова Ольга Геннадьевна – ТГТУ, канд. техн. наук, доцент.

важным преимуществом. Пользователи могут изменять программу по желанию на любой стадии эксперимента, с экономией времени и затрат. ПВМ широко используются при обработке цифровых сигналов.

Обзор исследований по теме работы

Накано и другие предложили использовать ПВМ для поиска изображений в 2003 году [4]. Ими было разработан код на языке Verilog HDL, который реализует создание списка всех изображений с аналогичными фрагментами, и создана соответствующая ПВМ. Встроенная ИС с конкретным шаблоном операций позволила значительно ускорить процесс поиска. По оценкам, скорость поиска изображений за счет аппаратного решения в 3000 раз лучше скорости программного решения.

Барскар и другие в 2011 году [5] предложили использовать нечеткую логику для определения границ. В предлагаемом методе характерные свойства определяются путем расчета вертикальной и горизонтальной маски для поиска границы. Сходство между введенным изображением и изображением в БД оценивается по извлеченным значениям признаков по формуле расстояний Манхэттена. Для улучшения поиска изображений применяется нечёткая логика, благодаря чему можно использовать естественный язык при формулировке требований. Предлагаемая нечёткая логика для поиска границ при определении величины порога даёт лучшие поисковые результаты в сравнении с другими существующими методами.

Ли и др. в 1994 году [6] предложили новый алгоритм нечеткой энтропии для сегментации изображений. Для расчета нечёткой функции энтропии используются нечёткая ширина области и функция Шеннона. Все максимумы локальной энтропии расположены в изображении для оптимального разбиения на сегменты. Предложенный алгоритм хорошо работает, когда пики и впадины в гистограммах неясны, или вероятностная модель изображения и классы сегментации неизвестны. Эксперименты с использованием различных видов изображений показали хорошую производительность предложенного алгоритма.

Методы исследования

Набор данных для экспериментов по поиску изображений состоит из 44 изображений. Изображения были сегментированы с помощью НАОК-метода, реализованного на ПВМ. Характерные свойства были извлечены из сегментированных изображений с помощью быстрого преобразования Хартли. Наивный классификатор Байеса и индукция деревьев решений были использованы для классификации изображений [7].

В данном разделе рассматриваются

- Предлагаемый НАОК-метод
- Быстрое Преобразование Хартли (БПХ)
- Наивный классификатор Байеса
- Алгоритм J48

НАОК-метод

Граница является базовой характеристикой изображения; обнаружение границы – это процесс идентификации и определения местонахождения края, который является ключевым при сегментации изображений. Искажение, шум, перекрытия, изменения интенсивности являются факторами, затрудняющими поиск границ [8]. В литературе представлено много методов определения границ, подобных нечёткому методу, который обсуждается в данном разделе.

Нечеткая логика выступает эффективным инструментом для расширения и обнаружения границ при моделировании неоднозначности или неопределенности. При использовании нечёткой функции шум эффективно удаляется из изображения, а само изображение может быть преобразовано в нечёткое изображение. При нечёткой сегментации постоянное значение цвета связывается и объединяется для создания областей и представления объектов, которые изображение содержит. Используется метрика близости между компонентами - «соседство», которая определена нечёткой логикой. Мера

соседства зависит как от цветового расстояния, так и топологического соотношения между компонентами. Нечёткая логика в результате даёт более чёткое сегментирование и лучшее время выполнения.

В предлагаемом алгоритме используются два различных метода выделения границ, и нечёткая логика применяется на выходе метода обнаружения двух границ.

В первом методе обнаружения границы в качестве фильтра свертки используется оператор Собеля

$$H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (1)$$

$$H_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2)$$

Фильтры (1) и (2) применяются к изображению для получения вертикального и горизонтального градиента компонент Нх и Ну. Оценка амплитуды градиента дана в уравнении 3

$$|H| = |H_x| + |H_y| \quad (3)$$

Угол ориентации кромки относительно сетки пикселей равен

$$\phi = \arctan\left(\frac{H_x}{H_y}\right) \quad (4)$$

Два фильтра могут быть применены одновременно (рис.1), чтобы получить амплитуду, оценка которой дана в уравнении 5.

$$\begin{matrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{matrix}$$

Рис. 1. Псевдо-фильтр свертки

Во втором способе в качестве фильтра используется оператор Лапласа

$$I = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (6)$$

Амплитуда для рис.2 задается соотношением

$$|I| = |4a_5 - (a_2 + a_4 + a_6 + a_8)| \quad (7)$$

Амплитуды из обоих методов нормированы для получения значения в диапазоне [0,1]. Нормированные величины разделены на три класса: Vl, Vm, Vh. Диапазон значений для каждого класса представлен на рис.2.

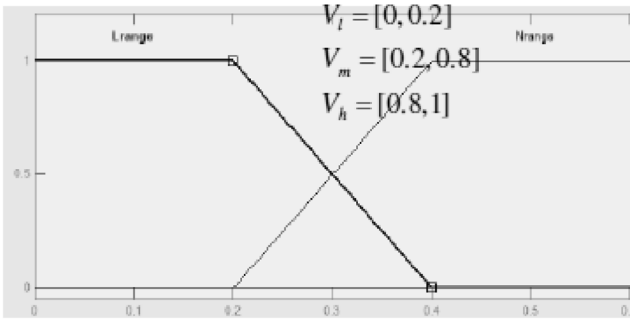


Рис. 2. Функция принадлежности на входе с ранжированием на три класса

Выход нечёткой системы на основе заданных нечётких правил определяет, принадлежит ли пиксель границе или нет. Функция принадлежности на выходе показана на рис.3.

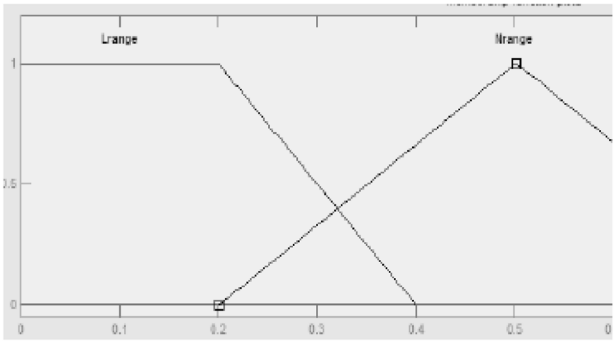


Рис. 3. Функция принадлежности на выходе

Три класса значений на выходе, El, Em, Eh определяют низкую, среднюю или

высокую вероятность того, что пиксель принадлежит границе.

Если p1 и p2 соответствующие значения вероятностей из двух методов обнаружения границы, тогда задаются следующие нечеткие правила:

- Если p1 в Vl и p2 в Vh, то вероятность принадлежности границе классифицируется как Em
- Если p1 в Vh и p2 в Vh, то вероятность принадлежности границе классифицируется как Eh, и т.д.

В общей сложности возможны 9 комбинаций с генерацией 9 правил. Значения внешней границы используются для создания маски, а характерные особенности извлекаются с помощью быстрого преобразования Хартли.

Быстрое преобразование Хартли

Преобразования Хартли переводят вещественные временные или пространственные функции в вещественные частотные функции с двумя независимыми наборами синусоидальных компонент [9]. Для расчетов дискретная адаптация непрерывного преобразования Хартли определяется как

$$H\left(k\Omega_v\right)=\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}h\left(nT\right)cas\left(k\Omega_v nT\right) \quad (8)$$

А обратное преобразование

$$h\left(nT\right)=\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}H\left(k\Omega_v\right)cas\left(k\Omega_v nT\right) \quad (9)$$

Энергетические характеристики извлекаются с помощью быстрого преобразования Хартли на определенных сегментах. Полученные энергии используются для обучения наивной классификации Байеса и алгоритма J48.

Наивная классификация Байеса

Наивная классификация Байеса – то вероятностный метод, используемый для предсказаний. Он основан на условной вероятности, определяемой как

$$P\left(D/a\right)=\frac{P\left(a/D\right).P\left(D\right)}{P\left(a\right)} \quad (10)$$

для данного значения a и класса D. Также, вероятность a можно определить как

$$P\left(D/a\right)=P\left(D\right).\prod P\left(a_j/D\right) \quad (11)$$

Во время обучения условные вероятности каждого атрибута в предсказанном классе оцениваются по множеству данных обучения. Как правило, наивная классификация Байеса обеспечивает хорошие результаты и элементарную вероятностную интерпретацию. Недостаток классификатора – это предположение, что возникновение признаков является независимым событием, т.е., корреляция между атрибутами игнорируется.

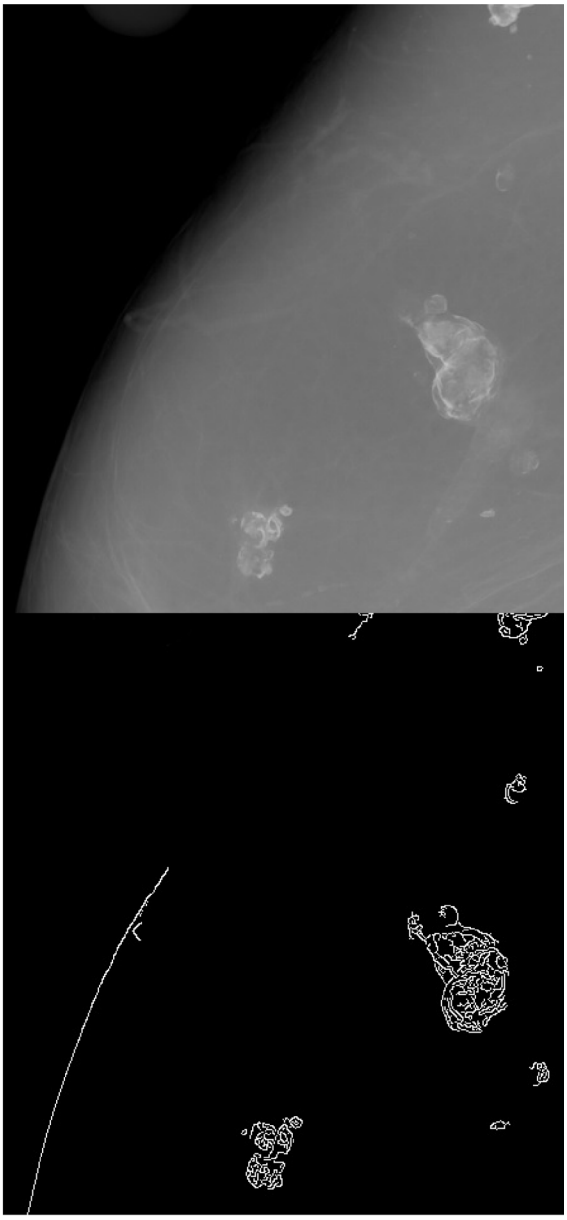


Рис. 4. Исходное изображение и граница, полученная с помощью предложенного НАОК-метода

Результаты

Предложенный НАОК-метод был реализован с использованием Matlab и ModelSim. Форма сигнала, полученного из ModelSim, представлена на рис.5. Для данного исходного изображения (рис.4) показана полученная граница, по которой можно проводить сегментацию.

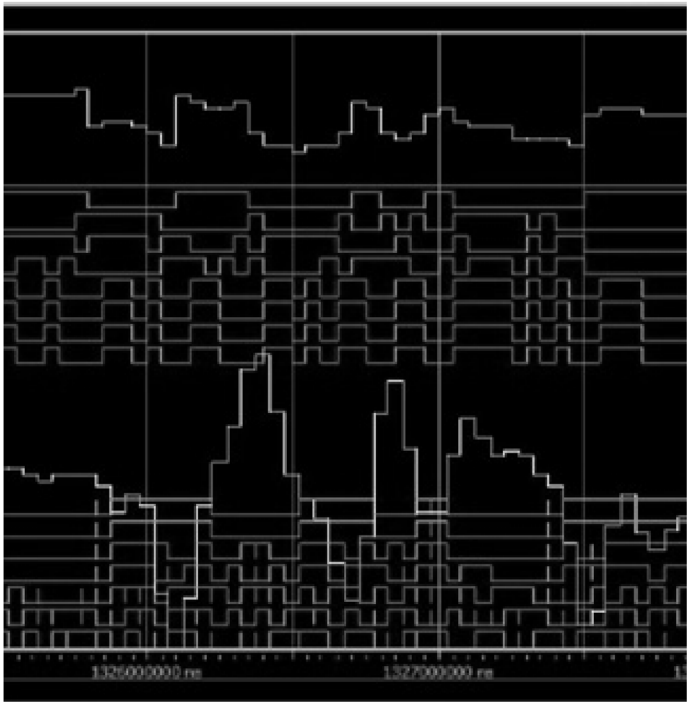


Рис. 5. Моделирование выходного сигнала

Точность классификации использованных изображений с помощью предложенного метода на основе наивной классификации Байеса и алгоритма J48 представлена на рис. 6.

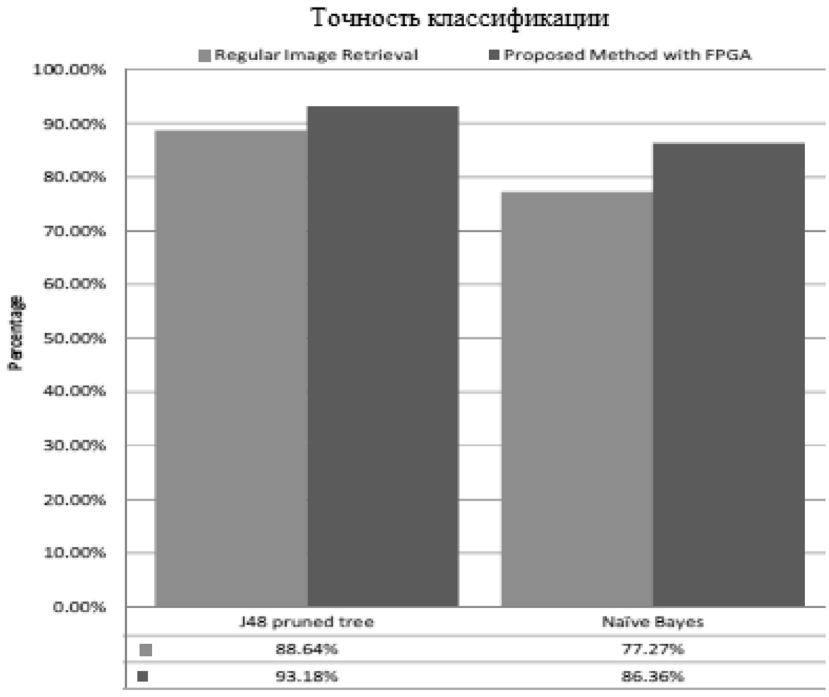


Рис. 6. Сравнение точности классификации с использованием предложенного метода с ПВМ и обычного поиска изображений

Выводы

В данной работе предложен новый алгоритм сегментации с ПВМ для улучшения точности поиска изображений. Характерные особенности извлекались из сегментированного изображения с помощью быстрого преобразования Хартли (БПХ) и использовались для обучения наивной классификации Байеса и алгоритма J48. Точность классификации по предложенному методу лучше результатов обычной процедуры на 4,54% и 9,09% соответственно. Задачи предложенной архитектуры были достигнуты. Дальнейшую работу необходимо направить на выделение характерных свойств на уровне ПВМ, что обеспечит увеличение скорости поиска изображений.

Литература

1. Потапов А.А. Новейшие методы обработки изображений / А.А. Потапов [и др.]; под ред. А.А. Потапова. – М.: ФИЗМАТЛИТ, 2008. – 496 с. /

2. Li, X.Q, Zhao .Z.W, Cheng H.D, Huang C.M, Harris R.W, “A fuzzy logic approach to image segmentation”, Pattern Recognition, 1994. Vol. 1 – Conference A: Computer Vision & Image Processing., Oct 1994, pp 337-341.

3. *Edge-Detection Method for Image Processing Based on Generalized Type-2 Fuzzy Logic* / *P. Melin* [*et al*.] // *IEEE Transactions on Fuzzy Systems*. – 2014. – *No.* 22 (6), *art*. *no*. 6698367. – *P*. 1515-1525.

4. *Yüksel*, *M.E. Edge detection in Noisy images by neuro-fuzzy processing* / *M.E. Yüksel* // *AEU - International Journal of Electronics and Communications*. – 2007. – *No*. 61 (2). – *P*. 82-89.

5. *Nie*, *Y. The fuzzy transformation and its applications in image processing* / *Y. Nie*, *K.E. Barner* // *IEEE Transactions on Image Processing*. – 2006. – *No*. 15 (4). – *P*. 910-927.

6. Бейтс, Р. Восстановление и реконструкция изображений: Пер. с англ. / Р. Бейтс, М. Мак-Доннелл. – М.: Мир, 1989. – 336 с.: ил.

7. Федотов, Н.Г. Теория признаков распознавания на основе стохастической геометрии и функционального анализа / Н.Г. Федотов. - М.: ФИЗМАТЛИТ, 2010. – 304 с.

8. Местецкий, Л.М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры / Л.М. Местецкий. – М.: ФИЗМАТЛИТ, 2009. – 288 с.

9. Акинин, М.В. Нейросетевые системы искусственного интеллекта в задачах обработки изображений / М.В. Акинин, М.Б. Никифоров, А.И. Таганов. – М.: Горячая линия, 2016. – 152 с.: ил.

.

ФГБОУ ВО «Тамбовский государственный технический университет»
Tambov State Technical University

IMAGES ANALYSIS BY METHOD OF FUZZY SEGMENTATION USING THE PROGRAMMABLE VENTILATED MATRIX

Y.Y. Gromov, P.I. Karasev, S.K. Stegachev, O.G. Ivanova

Image recognition and segmentation pose a challenge to Image Retrieval process. Challenges posed include many number of images to be processed for the image retrieval problem at a very fast time and identifying the location of interest automatically to optimize the search

Key words: Image segmentation, pattern recognition, edge detection

УДК 004.056.57

МОДЕЛИРОВАНИЕ И АНАЛИЗ ЭПИДЕМИЧЕСКИХ ПРОЦЕССОВ В КОРПОРАТИВНОЙ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ С ЯРКО ВЫРАЖЕННОЙ НЕОДНОРОДНОСТЬЮ

В.А. Волков, В.В. Исламгулова, О.Н. Чопоров, Е. Ружицкий, В.М. Питолин

В данной работе рассматриваются модели эпидемических процессов заражения корпоративных информационно-телекоммуникационных систем (ИТКС) в результате деструктивного контента различного вида

Ключевые слова: риск, сети, эпидемия, деструктивный контент

Объектом исследования является гетерогенная информационно-телекоммуникационная сеть с ярко выраженной неоднородностью, в которой имеет место распространение деструктивного контента. Для моделирования этого процесса необходимо исходные данные сети преобразовать в формат трехместного предиката, а затем загрузить эти данные непосредственно в созданную программу. В табл. 1 представлены метрики анализируемой сети. В табл. 2 представлены параметры микрофрактала, определяющего вероятности переходов пользователей из одного состояния в другое. Получившиеся вероятности являются усредненными значениями [1]. В табл. 3 представлены характеристики, описывающие состояния узла сети (вершина, конкретный пользователь) в эпидемическом процессе.

Табл. 1

Статистические данные сети

Метрика	Количество вершин	Количество рёбер	Вес сети	Диаметр сети	Плотность графа
Значение	121	714	9364944.87	1	0.049

Табл. 2

Параметры микрофрактала для анализируемой сети [1]

Параметр	P_I	P_E	P_R	P_M
Значение	0.245	0.151	0.299	0.073

Табл. 3

Определение вероятностей перехода для сети

Параметр	Содержание параметра
P_I	Вероятность перехода узла в инфицированное состояние
P_E	Вероятность перехода узла в латентную стадию
P_R	Вероятность смерти (удаления) узла
P_M	Вероятность получения иммунитета

Волков Владимир Андреевич – ВГТУ, студент,
e-mail: mnac@comch.ru
Исламгулова Виктория Викторовна – ВГТУ, соискатель
каф. СИБ, e-mail: mnac@comch.ru
Чопоров Олег Николаевич – ВГТУ, профессор,
e-mail: mnac@comch.ru
Ружицкий Евгений – Пан-Европейский Университет (Словакия), к.т.н., декан, доцент,
e-mail: eugen.ruzicky@paneurouni.com
Питолин Владимир Михайлович – ВГТУ, профессор,
к.т.н.,e-mail: mnac@comch.ru

Далее будут рассмотрены несколько вариантов моделирования эпидемического процесса в зависимости от структурно-функциональных особенностей сети, вершин, которые первоначально необходимо заразить, и их степеней. Механизм распространения и результаты эпидемического процесса будут различны, исходя из того, какой элемент (вершина) сети будет заражен.

Специфика исследуемой корпоративной среды (неоднородная сеть), заключается в том, что разброс степени ее вершин сильно разнится, и наблюдается большое количество элементов с высокой степенью центральности, т.е. у большей части вершин количество связей выше среднего значения по сети. Следовательно, наиболее показательными для исследования будут случаи, когда заражаемая вершина является наиболее высокостепенной, и когда заражена вершина с наименьшей степенью, и когда вершина входит в состав кластера.

В первую очередь выбор заражаемой вершины зависит от количества ресурсов у атакующей стороны, которые она может затратить на достижение поставленной цели. В этом конкретном случае целью может являться самая центральная вершина с наибольшей степенью, с которой остальные пользователи так или иначе контактируют. В исследуемой сети таких вершин две, и, чем больше их заразить, тем скорее в сети распространится контент.

Обычно самые важные узлы сети хорошо защищены от различного рода вторжений и атак, поэтому во втором случае для снижения затрат на заражение сервера атакующая сторона может заразить одного или нескольких единичных пользователей, удаленных от высокостепенной вершины, которые, в свою очередь, распространят деструктивный контент глубже в сеть.

В третьем случае можно заразить одного пользователя. Отличие от второго случая состоит в том, что зараженный пользователь должен являться участником определенного кластера сети, то есть работник одной из служб или отдела в организации.

Чтобы проанализировать больше исходов, необходимо запускать моделирование эпидемического процесса несколько раз для каждого из трех описанных случаев. Из полученных результатов необходимо выбрать эталонный результат и в дальнейшем проводить его анализ.

Из рис. 1 и 2 видно, что в сети присутствует большое количество высокостепенных вершин.

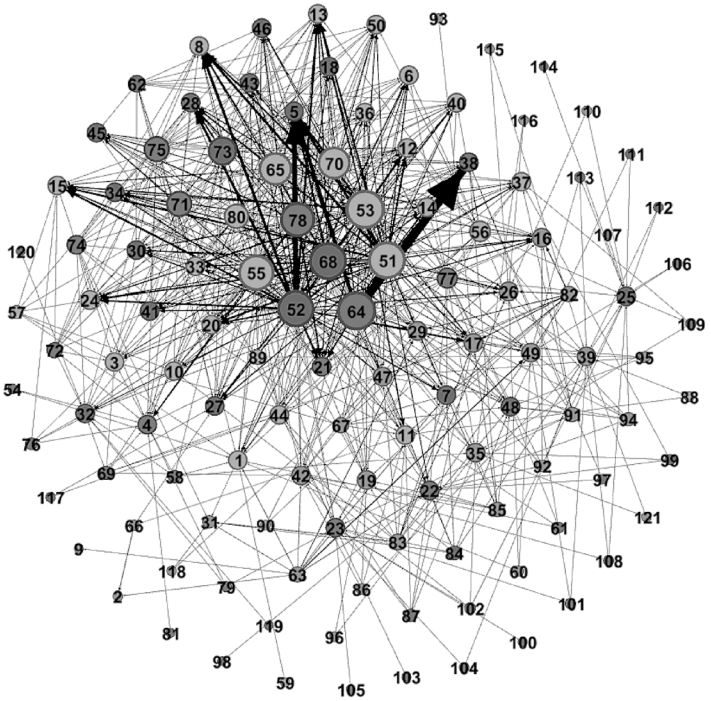


Рис. 1. Графическое представление сети

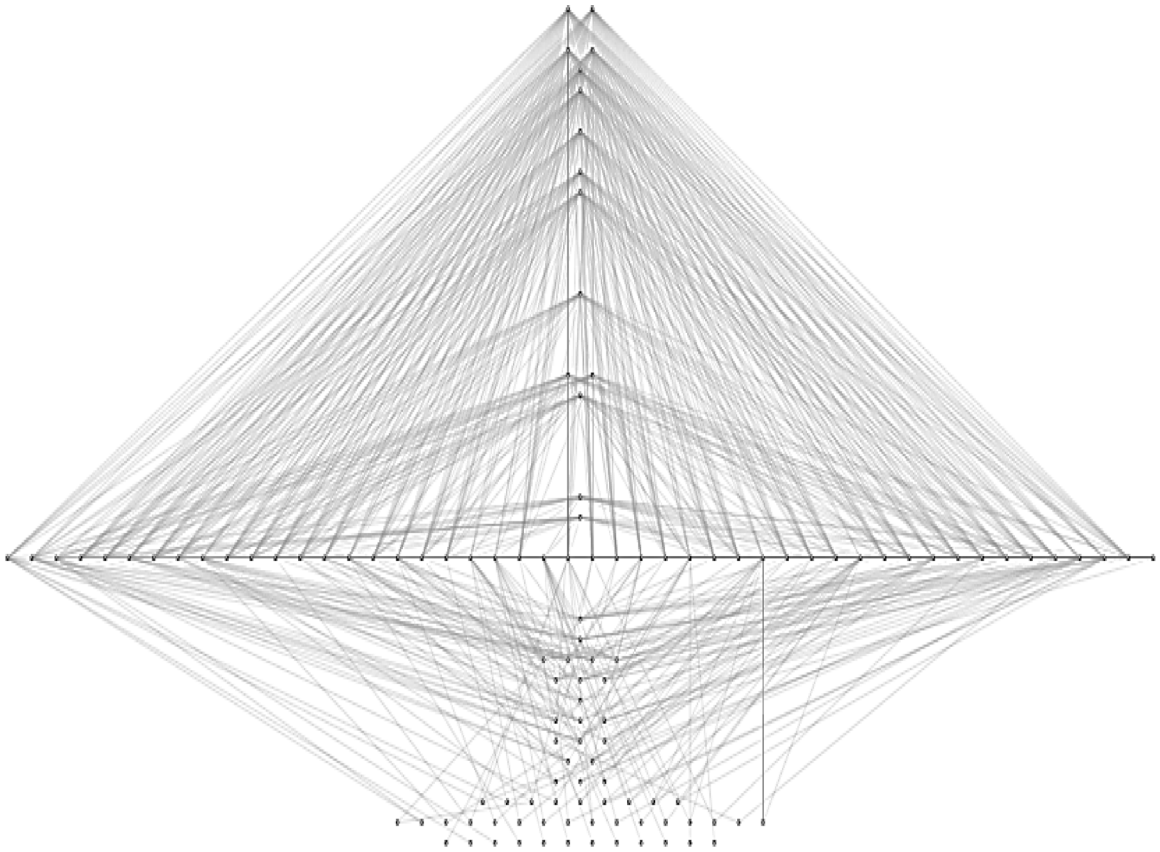


Рис. 2. Послойная модель сети

Первый случай. Для первого случая необходимо выбрать 1 критическую вершину с наибольшей степенью. Количество шагов моделирования напрямую зависит от реального времени, за которое осуществляется один шаг.

В подобных сетях эпидемический процесс стабилизируется на 30-м шаге, поэтому для чистоты эксперимента увеличим этот показатель до 50. Для более удобного представления сети в используемом программном обеспечении можно применить Fruchterman-Reingold укладку графа.

В данном графе (рис. 2) вершины имеют порядковые номера, сформированные на основе IP-адресов и имен компьютеров из предоставленных данных, чтобы не раскрывать структурную специфику исследуемого предприятия.

Данные пяти испытаний мало чем отличаются, а отклонения укладываются в допустимую погрешность.

Эпидемический процесс протекает в соответствии с заданным микрофракталом. В большинстве случаев зараженная вершина удалялась администратором сети на вторых шагах, но к этому моменту эпидемия уже распространялась на другие узлы сети, что не способствовало прекращению эпидемии.

На рис. 3 представлен график, показывающий состояние вершин после эпидемии, где состоянию S соответствуют восприимчивые вершины, E – латентно зараженная вершина, не передающая заболевания соседям, I – зараженная вершина, заражающая соседа, AM – состояние, в котором вершина приобрела иммунитет самостоятельно, либо подвергшаяся вмешательству со стороны администрации сети.

Вершины, находящиеся в состоянии R являются удаленными самостоятельно, или под действием администрирования.

На оси ординат будет откладываться количество вершин, а на оси абсцисс – дискретное время (шаги эпидемии).

В табл. 4 представлен фрагмент матрицы эпидемии для первого случая, откуда можно сделать следующие выводы:

- согласно заданным данным на нулевом шаге моделирования, т.е. до запуска эпидемии количество инфицированных узлов равнялось 1;

- на 2-м шаге количество инфицированных узлов выросло до 5 и продолжило постепенно увеличиваться;

- на 19-м шаге эпидемии количество инфицированных вершин равнялось 16, и эпидемия не прекратила свое действие, однако, вплоть до предпоследнего шага это количество оставалось неизменным, что говорит о стабилизации эпидемии внутри рассматриваемой сети;

- неожиданным всплеском эпидемии было появление 27 инфицированных узлов на 50-м шаге, однако после этого их количество стало снижаться, благодаря их удалению из информационного процесса, после чего эпидемия стабилизировалась окончательно.

Показательным результатом данного испытания стало то, что эпидемия вывела из строя две самые высокостепенные вершины уже на 2-м шаге, то есть взаимодействие пользователей стало затрудненным. Следовательно, атака прошла успешно.

Вывод по первому случаю: теоретически, если атакующая сторона имеет неограниченные ресурсы для проведения различного рода атак, то самая центральная вершина выступает наиболее приоритетной целью для этого, так как вывод ее из строя за короткий промежуток времени затруднит работу всей сети в целом, что будет видно в дальнейшем из графика риска (рис. 6).

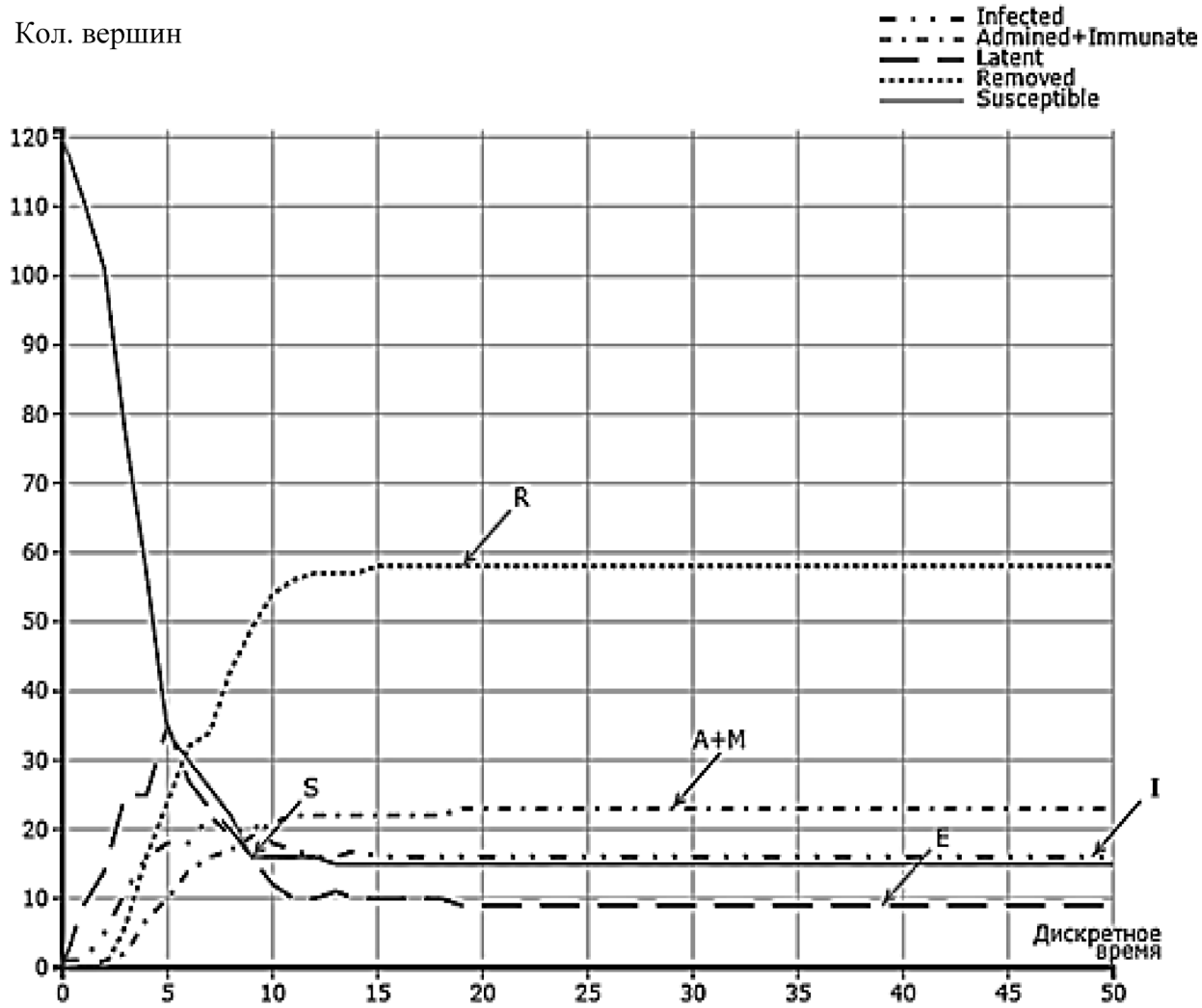


Рис. 3. Состояние вершин после эпидемического процесса

Табл. 4

Состояние вершин для первого случая

Шаг	Инфицированные	Защищённые	Латентные	Удалённые	Восприимчивые
0	1	0	0	0	120
1	1	0	9	0	111
2	5	1	14	0	101
3	11	2	25	6	77
4	16	7	25	16	57
5	18	10	35	24	34
6	18	14	27	32	30
7	22	16	23	34	26
…	…	…	…	…	…
19	16	23	9	58	15
…	…	…	…	…	…
50	27	44	217	124	317

Второй случай. Для второго случая необходимо выбрать несколько вершин, с наименьшей степенью, которые могут находиться на разном удалении от сервера. Теоретически, при неограниченном количестве ресурсов атакующей стороне необходимо заразить как можно больше низкостепенных вершин, чтобы добиться наилучшего результата во время распространения деструктивного контента. Однако на практике количество ресурсов ограничено, поэтому предположим, что заражены были 3 вершины с наименьшей степенью. Количество шагов можно оставить прежним. Данные испытаний разнятся, поэтому требуют отдельного внимания. Результаты испытаний чередуются между собой. В трех испытаниях из пяти результат похож на тот, что был в первом случае, центральная вершина выходила из строя уже на 6-м шаге, то есть атака проходила успешно, и деструктивный контент распространялся глубоко в сеть. В остальных случаях уже на 3-м шаге первоначальные вершины удаляются администратором сети и не разносят деструктивный контент, то есть эпидемический процесс останавливается. Эталонным испытанием будем считать первый вариант, так как при увеличении количества начальных заражений увеличится вероятность заражения как центральных вершин, так и всей сети. Далее представлен график, который показывает состояние вершин после эпидемического процесса (рис. 4).

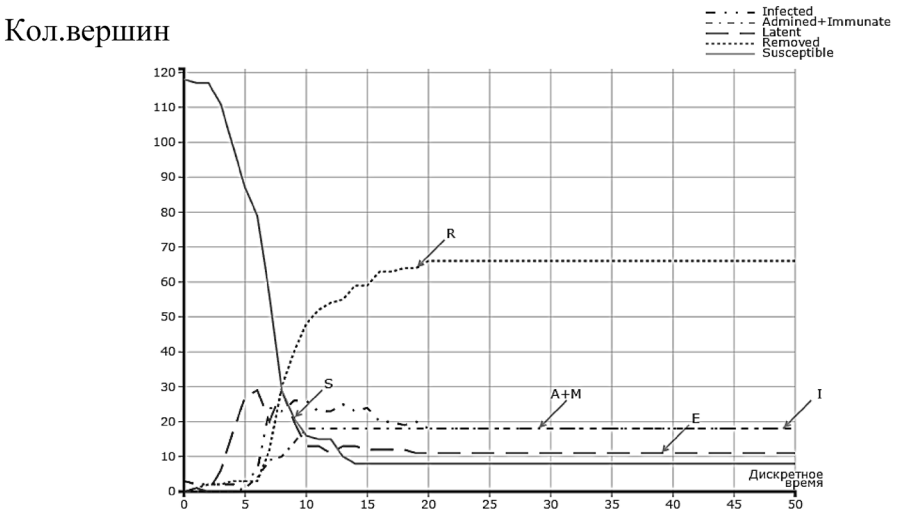


Рис. 4. Состояние вершин после эпидемического процесса

Согласно табл. 5, где представлены состояния вершин для эталонного испытания для второго случая, можно сделать следующие выводы:

- до запуска эпидемии количество инфицированных узлов равнялось 3, и их количество сначала снизилось до 2-х, а затем увеличилось до 6 на 6-м шаге развития эпидемии;
- на 7-м шаге количество инфицированных узлов резко возросло до 24 и продолжило увеличиваться до максимального значения в 26 вершин, после чего оно стало постепенно уменьшаться;
- на 17-м шаге эпидемии количество вершин, участвующих в инфицировании, остановилось на отметке в 20 штук и не изменилось только в самом конце рассматриваемого периода, что говорит о постепенной стабилизации эпидемии внутри сети.

Показательным результатом данного эксперимента является то, что даже при наличии малого количества зараженных изначально вершин есть достаточно большая вероятность вывести из строя всю сеть за небольшой промежуток времени с начала распространения деструктивного контента. Представленную эпидемию можно сравнить с реальным случаем заражения на исследуемом объекте, когда после подключения старого носителя с вредоносным ПО типа «сетевой червь» к одному компьютеру за один час рабочего времени заразилось 3 компьютера, после чего его выявили и обезвредили.

Табл. 5

Состояние вершин для второго случая

Шаг	Инфицированные	Защищённые	Латентные	Удалённые	Восприимчивые
0	3	0	0	0	118
1	2	0	1	1	117
2	2	0	0	2	117
3	2	0	6	2	111
4	2	0	17	3	99
5	3	1	27	3	87
6	6	4	29	3	79
7	24	9	20	12	56
8	23	10	29	30	29
9	26	14	20	40	21
10	26	18	13	48	16
11	23	18	13	52	15
12	23	18	11	54	15
13	25	18	13	55	10
14	23	18	13	59	8
15	24	18	12	59	8
16	20	18	12	63	8
17	20	18	12	63	8
…	…	…	…	…	…
50	18	18	11	66	8

Вывод по второму случаю: как уже было сказано выше, у атакующей стороны вероятность вывода из строя всей сети прямо пропорциональна количеству изначально зараженных узлов, то есть она растет с увеличением количества изначальных критических вершин сети.

Третий случай. Для третьего случая необходимо выбрать хотя бы одну вершину с небольшой степенью, которая обязательно

должна быть непосредственным участником определенного кластера сети - это может быть сотрудник отдела или службы предприятия. Процесс заражения схож со вторым случаем, здесь так же лучше заразить как можно больше вершин одного кластера, однако это все снова упирается в ресурсы атакующей стороны. Количество шагов можно оставить прежним.

В четырех испытаниях из пяти центральные вершины выходили из строя уже на 3-4 шагах, однако, деструктивный контент успевал распространиться по сети, и большая часть пользователей так же выходила из информационного обмена. Во время четвертого испытания уже на 2 шаге первоначальная вершина была удалена администратором сети и перестала разносить деструктивный контент, то есть эпидемический процесс остановился. Как и во втором случае, по тем же причинам, эталонным испытанием будет считаться первый вариант.

На графике (рис. 5) и в табл. 6 изображены состояния вершин для эталонного испытания третьего случая (табл. 6). Из них можно сделать следующие выводы:

- до запуска эпидемии количество инфицированных узлов равнялось 1, и их количество увеличилось до 2 на 2-м шаге, и уже к 3-му шагу развития эпидемии их количество возросло до 7;

- на 4-м шаге количество инфицированных узлов выросло до 17, что является максимальным количеством за рассматриваемый период, после чего оно стало постепенно уменьшаться;

- на 26-м шаге эпидемии количество вершин, участвующих в инфицировании, остановилось на отметке в 10 штук и не изменилось до самого конца рассматриваемого периода, что говорит о стабилизации эпидемии внутри сети.

Показательным результатом данного теста является то, что в третьем случае наибольший ущерб от атаки в начальный момент получит определенный кластер исследуемой сети, то есть отдел или служба предприятия, сотрудника которого использовали для изначального распространения деструктивного контента.

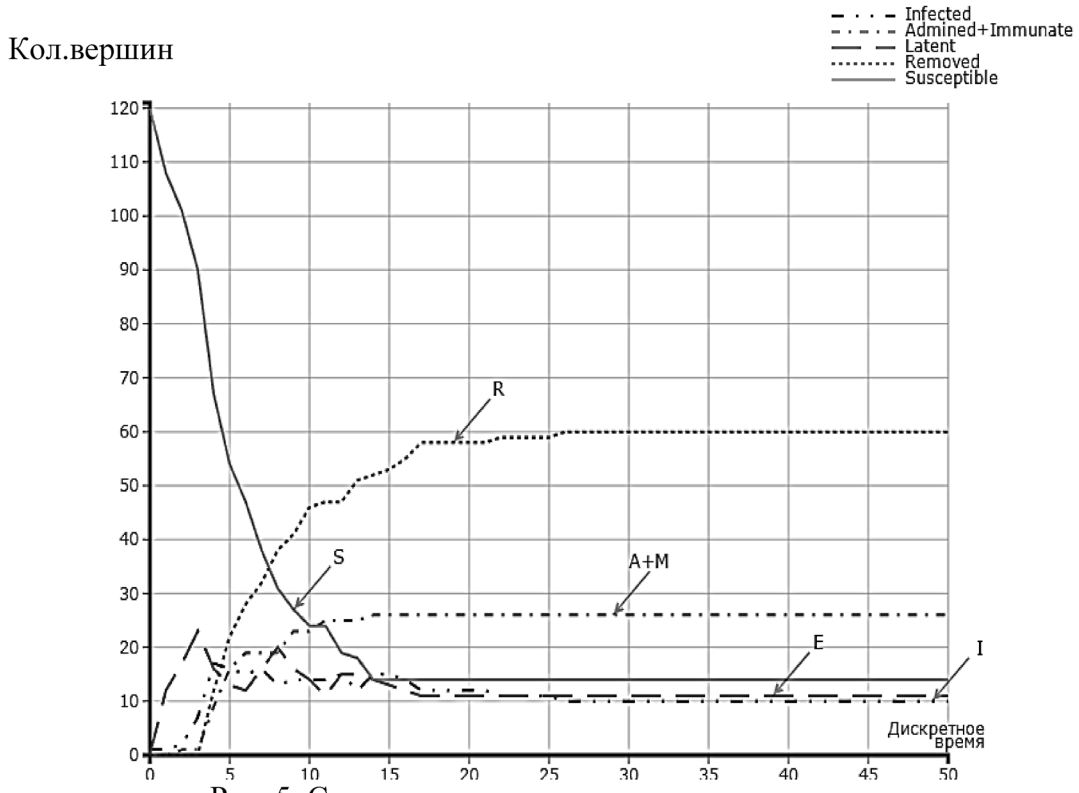


Рис. 5. Состояние вершин после эпидемического процесса

Табл. 6

Состояние вершин для третьего случая

	Инфицированные	Защищённые	Латентные	Удалённые	Восприимчивые
0	1	0	0	0	120
1	1	0	12	0	108
2	2	1	17	0	101
3	7	1	23	0	90
4	17	9	16	12	67
5	16	16	13	22	54
6	15	19	12	28	47
7	16	19	16	32	38
8	13	19	20	38	31
9	14	23	16	41	27
…	…	…	…	…	…
15	15	26	13	53	14
16	14	26	12	55	14
17	12	26	11	58	14
…	…	…	…	…	…
22	11	26	11	59	14
…	…	…	…	…	…
26	10	26	11	60	14
…	…	…	…	…	…
50	10	26	11	60	14

Представленную в третьем случае эпидемию также можно сравнить с реальным случаем заражения на исследуемом объекте.

В начальный момент времени сетевой червь «Blaster» был занесен всего на один компьютер, сообщающийся со многими другими, однако его выявили только через 8 часов, в конце рабочего дня.

За это время он заразил 100 компьютеров (при общем числе около 120).

Вывод по третьему случаю: исходя из вышесказанного, при заражении деструктивным контентом всего одного узла в сети, входящего в какой-либо кластер, велика вероятность, что будет выведена из строя сразу вся сеть предприятия, того что, очевидно, негативно скажется на процессе информационного обмена на столь важном объекте.

График риска по эталонным вариантам испытаний для каждого из трех рассматриваемых случаев представлен на рис. 6.

На графике под темой подразумевается исследуемый вариант развития событий.

Данному графику соответствует матрица рисков, представленная в табл. 7.

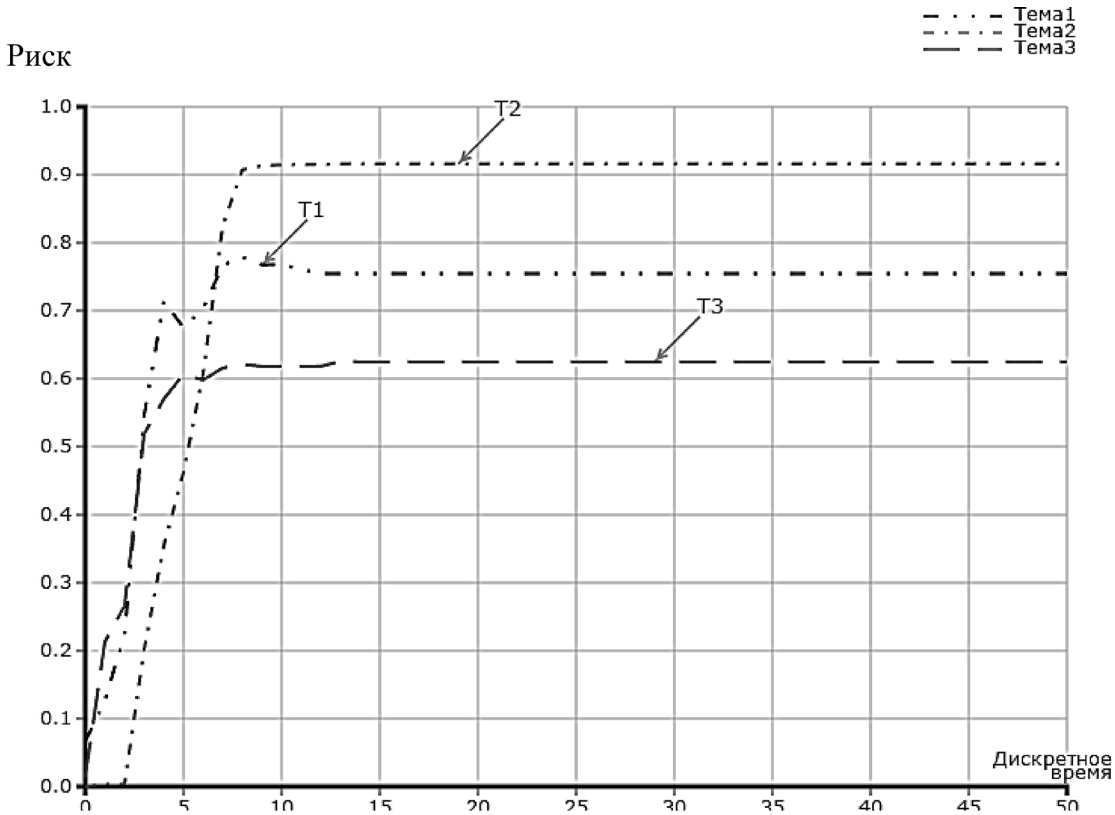


Рис. 6. График риска для всех трех случаев

Табл. 7

Матрица рисков для эпидемии по темам

Шаг	Случай 1	Случай 2	Случай 3
0	0.065965298	3.09143E-05	0.012858587
1	0.126383482	0.003637903	0.213133602
2	0.224101131	0.003637903	0.265883063
3	0.545702923	0.202021325	0.518289093
4	0.713794068	0.354372927	0.569931512
…	…	…	…
10	0.768246859	0.914643052	0.618061234
…	…	…	…
20	0.754511032	0.916159295	0.624742461
…	…	…	…
30	0.754511032	0.916159295	0.624742461
…	…	…	…
40	0.754511032	0.916159295	0.624742461
…	…	…	…
50	0.754511032	0.916159295	0.624742461

В итоге, проанализировав все результаты исследования рассматриваемой сети, осуществляющей обмен данными между сотрудниками на предприятии, можно прийти к выводу, что структурно-функциональные особенности сети при массовой эпидемии во всех случаях на первых шагах распространения нежелательной информации допускают вывод из строя множества узлов как единичных сотрудников, так и критически важных точек.

Зачастую происходит так, что администратор сети блокирует доступ к узлу с наибольшей центральностью, чтобы исключить дальнейшее распространение заражения в сети, однако, за большой промежуток времени большая часть сети,

которая существует независимо от высокостепенной вершины, будет заражена, поэтому необходимо в кратчайшие сроки с начала эпидемии выявлять и блокировать зараженные узлы. По этой причине остро встает вопрос о том, какими способами возможно эффективно минимизировать риск и повысить структурную живучесть для корпоративных информационно-телекоммуникационных сетей с ярко выраженной неоднородностью их элементов.

Литература

1. Woo J. Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog /Woo J., Chen H.// SpringerPlus. – 2016. 5: 66. doi:10.1186/s40064-016-1675

2. Newman, M. E. (2004). Coauthorship networks and patterns of Scientific Collaboration. Proceedings of the National Academy of Sciences, 101(suppl 1): 5200-5205.

3. Ahn, Y. Analysis of topological characteristics of huge onlane social networking services Advogato/ Y. Ahn, S. Han, H. Knak, S. Moon, H. Jeong // 16th International Conference on the World Wide Web. – 2007. – P. 835-844.

4. Alba, R.A. graph-theoretic definition of a sociometric clique / Richard D. Alba / Journal of Mathematical Sociology. – 1973. – P. 113–126.

5. Абрамов, К.Г. Распространение нежелательной информации в социальных сетях Интернета / К.Г. Абрамов, Ю.М. – С.45-48.

6. Johnson, S. Entropic origin of disassortativity in complex networks / S. Johnson, J.J. Torres, J. Marro, M.A. Muñoz / Physical Review Letters. – 2010. – 4 p.

7. Das, S. Anonymizing Edge-Weighted Social Network Graphs / S. Das, O. Egecioglu, A. El Abbadi // Computer Science, UC Santa Barbara, Tech. Rep. – 2009.

8. Majumdar, A. Music Recommendations based on Implicit Feedback and Social Circles: The Last FM Data Set / A. Majumdar, A. Kumar, S. Manohar// – 2009. – 12 p.

9. Maxwell, A. Pretzlav Last.fm Explorer: An Interactive Visualization of Hierarchical Time-Series Data / A. Maxwell // – 2008. – 11p.

10. Konstas, I. On Social Networks and Collaborative Recommendation / I. Konstas // – 2012. – 5p.

11. Byrd K. War with many unknowns / K. Byrd//Computerra. – M.: 2009. – No.

12. Grinyaev, S. Russia in global information society: threats, risks and possible ways of their neutralization / S. Grinyaev, – Electron. it is given. – Access mode:http://www.noravank.am/upload/pdf/419_ru.pdf.

13. Valerio,Arnaboldi, Andrea Passarella, Marco Conti, Robin I.M. Dunbar. Online Social Networks: https://www.flickr.com/account/prefs/privacy

14. Panagiotis,Karampelas. Techniques and Tools for Designing an Online Social Network Platform. New Hampshire: Hellenic American University, 2013. – 172 p.

Воронежский научно-образовательный центр управления информационными рисками
Voronezh Research and Education Center management of information risks
Пан-Европейский Университет
Pan-European University

MODELING AND ANALYSIS OF EPIDEMIC PROCESSES IN A CORPORATE INFORMATION AND TELECOMMUNICATION NETWORK WITH BRIGHTLY INNOVATED INHOMOGENEITY

V.A. Volkov, V.V. Islamgulova, O.N. Choporov, E. Ruzicky, V.M. Pitolin

In this paper, we consider models of epidemic processes of infecting corporate information and telecommunication systems (ITKS) as a result of destructive content of various kinds

Key words: risk, networks, epidemic, destructive content

УДК 004.056

СОЦИАЛЬНАЯ СЕТЬ TWITTER: СТРУКТУРНО – ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ ПРОЦЕССОВ РАСПРОСТРАНЕНИЯ КОНТЕНТА

А.Н. Разгоняев, Е.С. Соколова, С.С. Куликов, Д.Н. Рахманин, Ю. Штефанович

Рассматривается анализ социальной сетизакладок в контексте распространения деструктивного контента, включая реализацию конкретных процедур необходимых для регулирования риска/шанса и управления ими

Ключевые слова: риск, шанс, социальная сеть закладок, контент

Социальная сеть Twitter – веб-сервис для публичного обмена сообщениями при помощи веб-интерфейса, SMS, средств мгновенного обмена сообщениями или сторонних программ-клиентов для пользователей сети Интернет любого возраста. Публикация коротких заметок в формате блога получила название «микроблогинг», сообщения состоят из 140-символов, называемых «твитами». Данная социальная сеть относится к категории социальных сетей закладок [1, 2,3, 4].

Для изучения трафика в социальной сети, необходимо построить структурно-функциональную схему социальной сети Twitter, которая предоставляет пользователям различный контент: текстовая информация, изображения, видеозаписи. Также существует гибридный вид контента, включающий в себя: текст и изображения, текст и видеозаписи т.д. [5].

На рис. 1 изображена схема классификации контента социальной сети Twitter.

Рис. 1. Классификация контента социальной сети Twitter

Разгоняев Антон Николаевич – ВГТУ, студент,
e-mail:mnac@comch.ru
Соколова Елена Сергеевна – ВГТУ, ассистент,
e-mail:mnac@comch.ru
КуликовСергейСергеевич – ВГТУ, доцент
email: nmac@comch.ru
Рахманин Дмитрий Николаевич – ВГТУ, доцент
email: nmac@comch.ru
Штефанович Юрий - Пан-Европейский Университет (Словакия), к.т.н., зам. декана,
e-mail: juraj.stefanovic@paneurouni.com

Для детального рассмотрения контента сети было проанализировано порядка 2000 твитов, среди которых были твиты как пользователей РФ, так и США и Европы. В результате исследования, все написанные посты в сети можно разделить на 6 категорий: светская беседа, разговоры, ретвиты, самореклама, спам, новости[6, 14].

- Светская беседа - 41 %;
- Разговоры - 38 %;
- Ретвиты - 9 %;
- Самореклама - 6 %;
- Спам - 4 %;
- Новости - 4 %;

Для наглядности данные исследования представим в виде гистограммы:

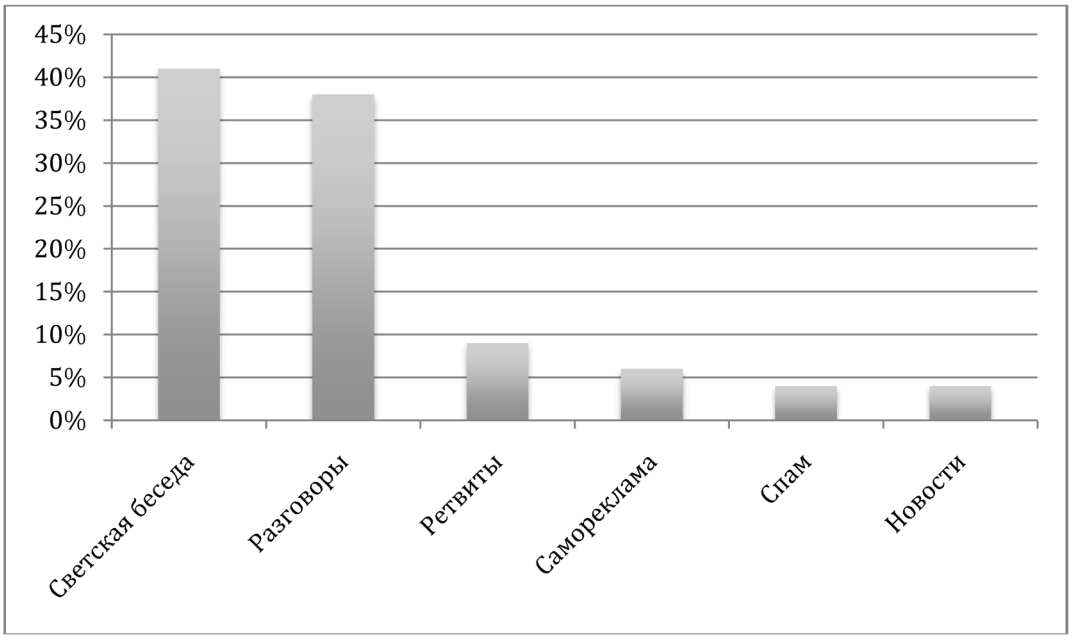


Рис. 2. Гистограмма постов социальной сети Twitter

Под светской беседой принимается социальная забота или осведомленность об окружении – когда человек хочет знать, что другие люди думают, делают и чувствуют, даже если они совершенно не знакомы друг с другом.

Из проведенного исследования, можно сделать вывод, что пользователи сети больше предпочитают общаться друг с другом [5].

Но полученное процентное соотношение категорий сообщений сети остается таковым не всегда. В разгар любых мировых событий, будь то спортивные мероприятия и тд, преобладают «разговоры», «ретвиты» и «новости». Так например во время Чемпионата мира по футболу или вручения премии Оскар, количество ретвитов постов различных людей достигало нескольких миллионов. Самый цитируемый твит собрал 3,3 млн ретвитов.

Каждое такое событие создает поток твитов на определенную тематику. В связи с этим, большая часть сообщений сопровождается определенными хэштегами, которые впоследствии формируют тренды в каждой стране и мире в целом [5].

Посты по конкретной теме (тренду), редко задерживаются дольше недели, обычно не более 3-4 дней. Большинство тем возникает единожды, а затем умирает, как правило, никогда не возвращаясь. 85% таких трендов связано с новостями. Возможно, причиной этого является то, что для описания тренда используются особые слова и выражения, часто весьма специфические.

При анализе аудитории (рис. 2) данной социальной сети целесообразно выделить два основных типа субъектов, задействованных в публикации, просмотре, комментировании и обсуждении той или иной информации, циркулирующей в

социальной сети в соответствующих разделах [6].

К первому типу можно отнести активных субъектов. Они подразделяются на администраторов сетевых ресурсов и активных пользователей. Администраторы имеют повышенные права, следят за порядком в социальной сети и занимаются проверкой жалоб на определенные твиты или любой другой контент, который циркулирует в данной социальной сети. Также есть модераторы, обладающие меньшим набором прав, чем администраторы

Активными считаются пользователи, создающие твиты (авторы твитов) и пользователи, принимающие непосредственное участие в обсуждении данных твитов (ретвиты, оценки, комментирование).

В свою очередь, авторы подразделяются на подтвержденных и неподтвержденных. Возможность пройти верификацию и получить статус подтвержденного профиля доступна на всем, а только лишь очень популярным страницам, пользователи которых ведут активную жизнь в сети[7].

Рис. 2. Классификация субъектов социальной сети Twitter

Второй тип субъектов – пассивный.

Он представлен авторизованными неактивными пользователями и неавторизованными пользователями. Авторизованные подразделяются на неактивных и бездействующих пользователей.

Под авторизованными неактивными пользователями понимаются учетные записи, которые зарегистрированы в социальной сети, но в определенный период времени не принимающие участия в обсуждении твитов, которые не создают твиты и не оценивают другие твиты. Пассивные пользователи ознакомляются с новыми твитами на главной странице социальной сети или с помощью новостной ленты. Бездействующие – забытые

страницы, на которых прекращена активность. Оставшаяся масса людей – неавторизованные пользователи, которые имеют ограниченные возможности в данной социальной сети. Им доступен только просмотр информации на главной странице сайта, включая возможность чтения комментариев [7].

Представим данную классификацию на следующем рисунке (рис. 2).

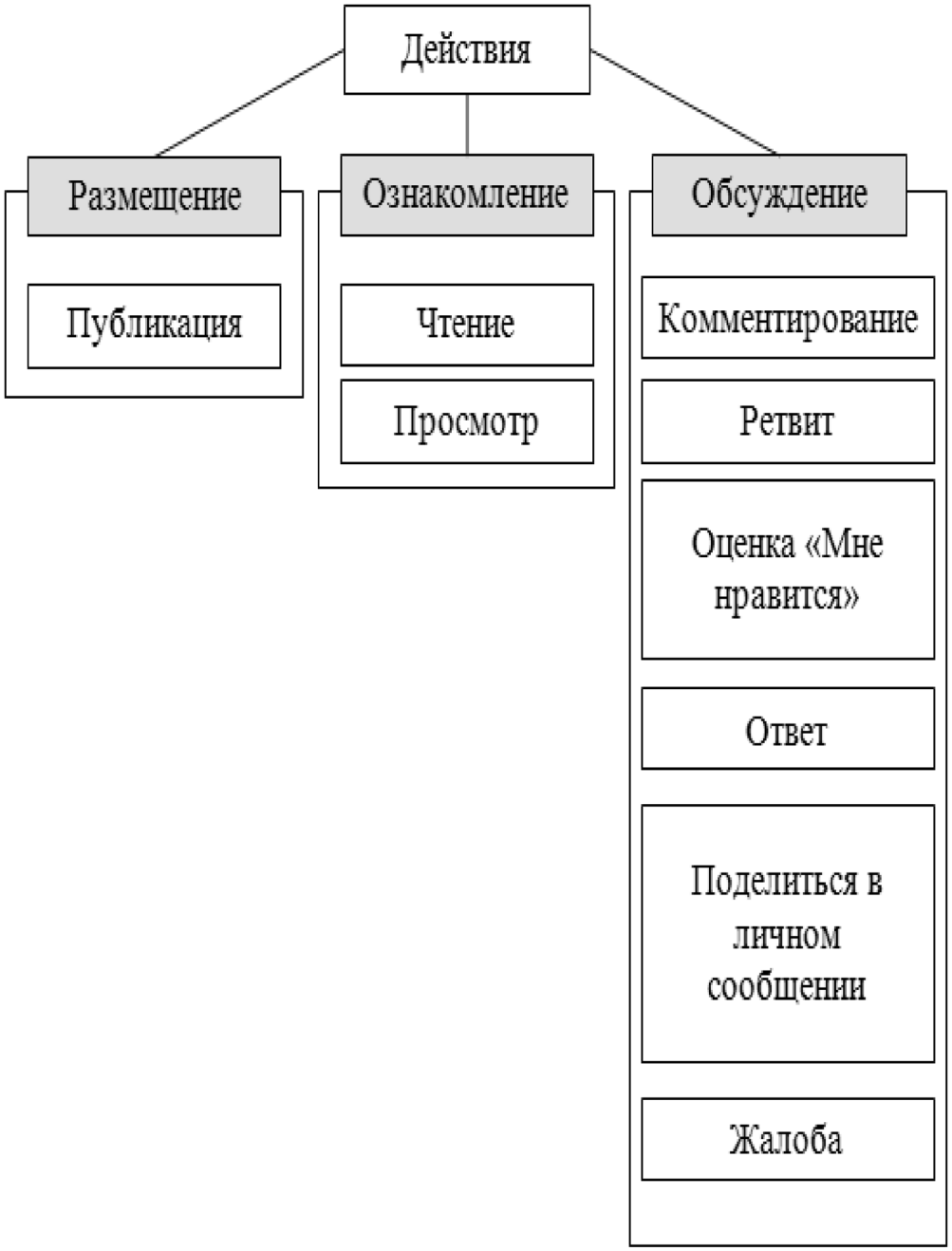


Рис. 3. Классификация действий субъектов в социальной сети Twitter

В социальной сети Twitter пользователи имеют определенный набор действий (функций), позволяющий им взаимодействовать с контентом.

Все возможности, доступные субъектам данной социальной сети целесообразно классифицировать по функциональным возможностям (рис. 4).

В данной классификации действия доступные пользователям были разделены на три основные категории: размещение, ознакомление и обсуждение[9].

В категории «размещение», пользователям доступно только одно действие – публикация.

В категории «ознакомление», пользователи имеют возможность прочитать

твит, перейти по ссылке в твите или посмотреть видео или изображение.

Весь функционал пользователя социальной сети Twitter сосредоточен в категории «обсуждение». В данную категорию попали действия, которые следуют после ознакомления пользователей с предложенным контентом [8].

Проведенная классификация способна дать представление о полном списке возможностей всех субъектов социальной сети Twitter, что позволит проанализировать их действие и поведение в различных ситуациях, а также построить структурно-функциональную модель социальной сети Twitter.

С учетом полученных классификаций контента, субъектов и их действий, а также сетевых ресурсов данной социальной сети предоставляется возможным построить структурно-функциональную модель социальной сети Twitter с учетом всех ее особенностей (рис. 4).

В данной модели функциональные связи представляют собой сложную структуру взаимодействия контента, сетевых ресурсов и субъектов, функционирующих в заданном сетевом пространстве.

Перейдем к статистическим данным, характеризующим количество узлов (пользователей) и ребер (дружеские связи, основанные на обмене закладками, комментировании статей и их публикации) в социальной сети.

Табл. 1

Статистические данные социальной сети Twitter

Характеристические данные	Показатели
Количество узлов	465017
Количество ребер	834797
Диаметр	8
Средняя длина пути	3,59
Коэффициент кластеризации (%)	0, 0613
Распределение плотности сетей	$3{,}8605\times10^{1}$
Средний диаметр сети	4,96

В отличие от аналогов, при моделировании диффузии в сети, будет осуществляться учет динамики развития сети.

Для этого проведено исследование, в процессе которого определено, как изменялось количество пользователей с течением времени.

По состоянию на 4 квартал 2016, сеть имеет порядка 319 млн. активных пользователей в месяц, и более 100 млн. в день.

Счетчик уникальных посещений насчитывает более 1 млрд [10].

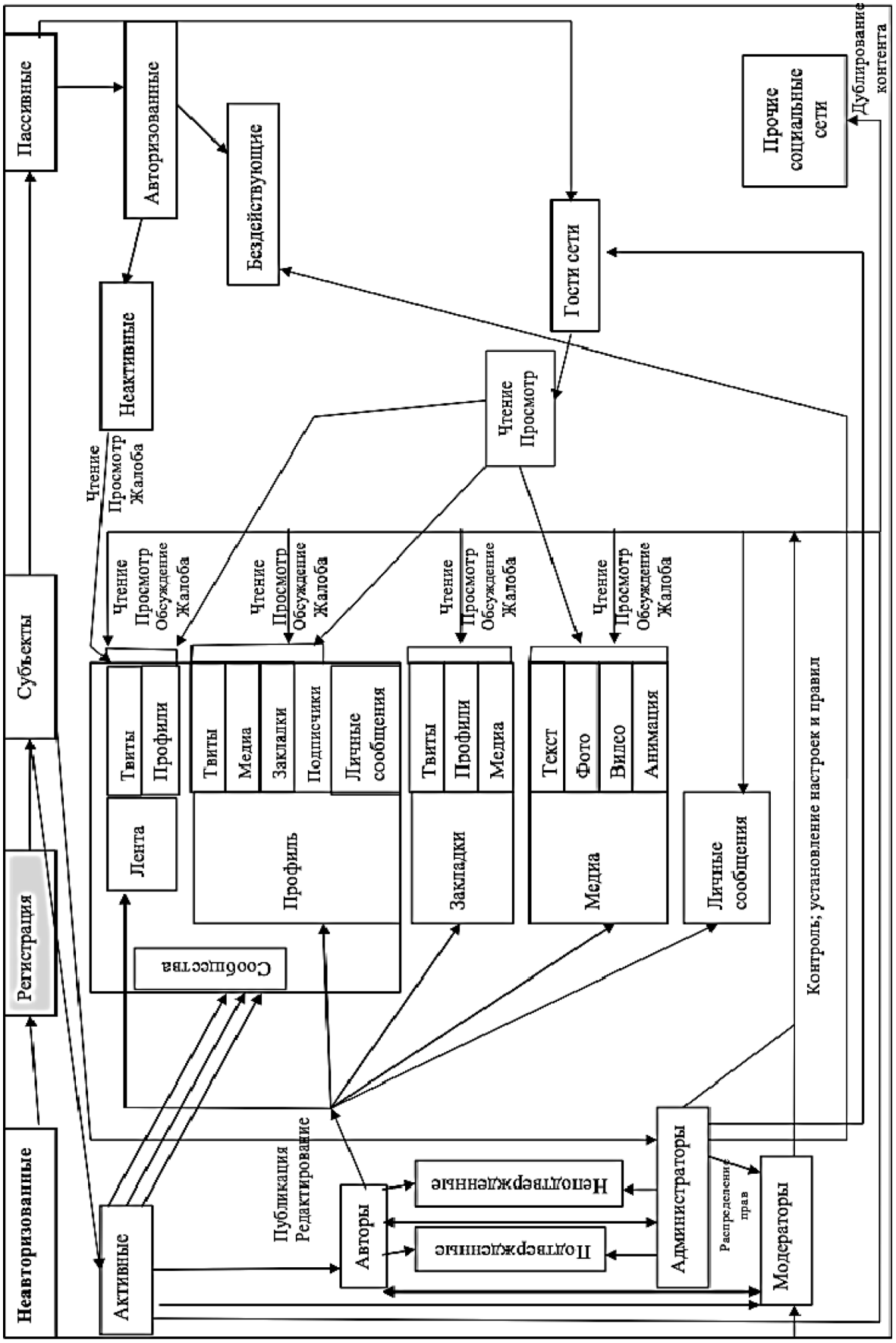


Рис. 4. Структурно-функциональная схема сети Twitter

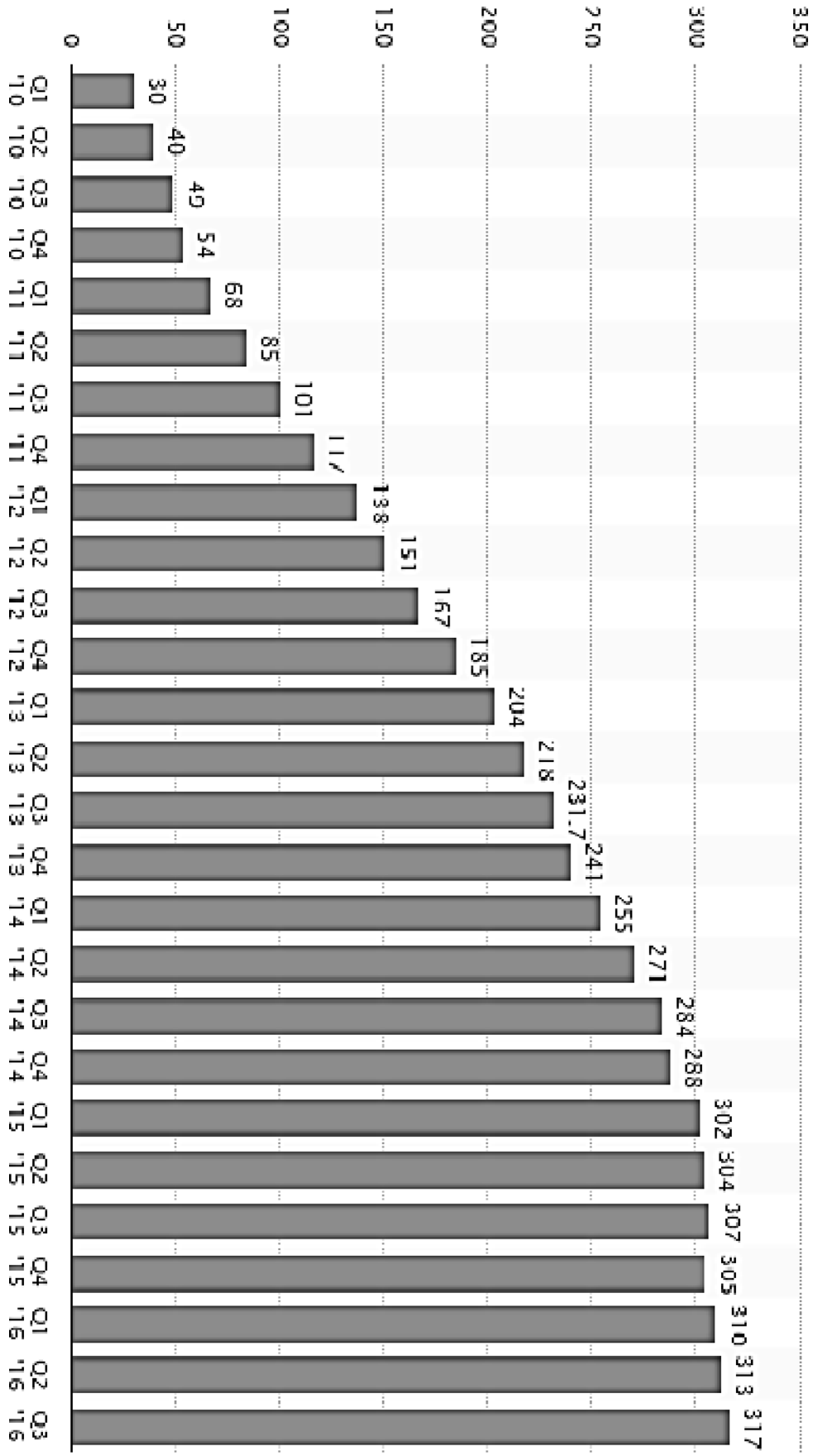


Рис. 5. Количество пользователей с 1 квартала 2010 по 3-й 2016

По приведенному выше рисунку нетрудно заметить, что количество пользователей из года в год активно растет, что позволяет сделать вывод о высокой востребованности сети Twitter. Однако в последнее время заметно, что рост замедлился, по сравнению с предыдущими годами. Динамический рост учитывался не по отдельным странам, по всему миру. На основе представленных статистических данных был получен коэффициент прироста социальной сети Twitter за последние 2 года, значение которого равно 0,012%. Численное значение коэффициента прироста усреднено, так как рост сети не является постоянным процессом. Наряду с появлением новых пользователей, происходит удаление (блокировка) старых аккаунтов. В связи с этим, значение динамического роста колеблется, но не уходит в отрицательные значения.

Литература

1. Социальная сеть Twitter. – Электрон. Дан. - Режим доступа: http://twitter.com.

2. Словарь терминов социальной сети Twitter. – Электрон. Дан. – Режим доступа: http://otwi.ru/dictionary/

3. Аналитически данные сети Twitter. – Электрон. Дан. – Режим доступа: https://analytics.twitter.com/

4. Alan E. Mislove. Online Social Networks: Measurement, Analysis, and Applications to Distributed Information Systems. Houston, Texas: RICE University, 2009.

5. Panagiotis Karampelas. Techniques and Tools for Designing an Online Social Network Platform. New Hampshire: Hellenic American University, 2013.

6. Jennifer Golbeck. Introduction to Social Media Investigation: A Hands-on Approach. Waltham: Elsevier Inc., 2015.

7. Valerio Arnaboldi, Andrea Passarella, Marco Conti, Robin I.M. Dunbar. Online Social Networks: Human Cognitive Constraints in book and Twitter Personal Graphs. Waltham: Elsevier Inc., 2015.

8. Barbara Carminati, Elena Ferrari, Marco Viviani. Security and Trust in Online Social Networks. Morgan&Claypool, 2014.

9. Черняк Л.М. Сервисы и теории социальных сетей / Л. М. Черняк // Открытые системы. СУБД, 2008. – № 8. – 78

10. Optimization of expert methods used to analyze information security risk in modern wireless networks / S.A. Ermakov, A.S. Zavorykin, N.S. Kolenbet, A.G. Ostapenko, A.O Kalashnikov // Life Science Journal. – 2014. – № 11(10s). – P. 511-514.

11. Peak risk assessing the process of information epidemics expansion / N.M. Radko, A.G.Ostapenko, S.V.Mashin, O.A. Ostapenko, A.S. Avdeev // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (Spl.End). – P. 251-255.

12. Гмурман В.Е., Теория вероятностей и математическая статистика. Учебное пособие. Высшее образование. – Москва, 2006 – С. 243.

13. Аналитически данные сети Twitter. – Электрон. Дан. – Режим доступа: https://analytics.twitter.com/.

14. Яндекс метрика. – Электрон. Дан. – Режим доступа: https://metrika.yandex.ru/.

Воронежский научно-образовательный центр управления информационными рисками
Voronezh Research and Education Center for Information Risk Management
Пан-Европейский Университет
Pan-European University

SOCIAL NETWORK TWITTER: STRUCTURAL - FUNCTIONAL ANALYSIS OF PROCESSES DISTRIBUTION OF THE CONTENT

A.N. Razgonyaev, E.S. Sokolova, S.S. Kulikov, D.N. Rakhmanin, Yu. Stefanovic

Discusses the social network analysis bookmarks in the context of the spread of destructive content, including the implementation of specific procedures required for managing the risk

Key words: risk, chance, social network, bookmarks, content, resources, objects, subjects

УДК 004.056.57

СОЦИАЛЬНАЯ СЕТЬ ДЛЯ КОЛЛЕКТИВНЫХ ОБСУЖДЕНИЙ REDDIT. МЕЖСЕТЕВОЕ ВЗАИМОДЕЙСТВИЕ МЕЖДУ REDDIT И YOUTUBE В РАМКАХ РАСПРОСТРАНЕНИЯ ЭПИДЕМИЧЕСКОГО ПРОЦЕССА С УЧЕТОМ ДИНАМИЧЕСКОГО РОСТА СЕТИ

А.В. Алтухов, И.В. Шевченко, А.Г. Остапенко, В.М. Питолин, Й. Воришек

В данной статье производится анализ крупнейшей сети для коллективных обсуждений REDDIT, симулируется эпидемический процесс распространения деструктивного контента между сетями и делаются соответствующие выводы

Ключевые слова: социальные сети, эпидемии

Социальная сеть для коллективных обсуждений Reddit является огромным форумом, где пользователи делятся друг с другом публикациями, представляющими собой ссылки на внешние источники. Предоставляемый контент является предметом обсуждения среди заинтересовавшихся, а также становится инструментом зарабатывания рейтинга среди пользователей. Так, за публикуемый пост автору путем голосования отходит определенное количество карма-очков (рейтинг в сети Reddit измеряется в таких единицах). Кроме этого, участник сети способен зарабатывать эти очки благодаря своим «качественным» комментариям [1]. Отобразим это на схеме (рис. 2). Здесь отображено то, каким образом контент в сети способен распространяться среди её участников. Как можно заметить, благодаря системе накопления очков, рейтинг пользователей меняется и при получении большего их количества, открывается возможность стать модератором Reddit [2]. Такого рода уникальные особенности сети стимулируют её участников размещать качественный контент, который заслужил бы уважение и вызвал бы интерес у как можно большего количества пользователей. Следовательно, они стараются размещать ссылки из надежных, достоверных источников.

Ссылки из данных источников могут носить за собой разный характер.

Это могут быть политические новости, событие дня, развлекательные материалы. Всё из этого, и не только, может предоставлять социальная сеть для обмена медиа-контентом YouTube в формате видеофайлов. По данным [3] значительную часть контента в сети занимает YouTube (рис. 1):

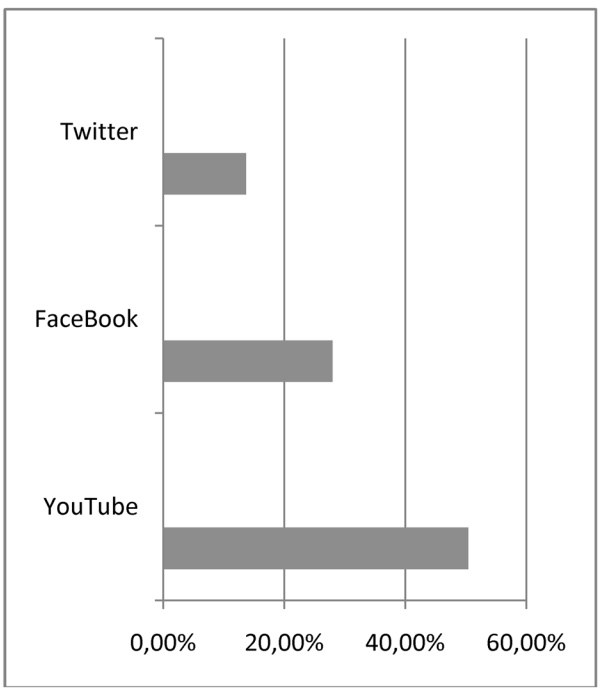


Рис. 1. Фрагмент статистики контента Reddit

Алтухов Александр Владимирович–ВГТУ, д-р техн.наук, профессор, e-mail: altuchov@yandex.ru
Шевченко Игорь Викторович– ВГТУ, аспирант, e-mail: mnac@comch.ru
Остапенко Александр Григорьевич – ВГТУ, докт. техн. наук, профессор, зав.каф., e-mail: mnac@comch.ru
Питолин Владимир Михайлович –ВГТУ д.т.н., профессор, e-mail: mnac@comch.ru
Воришек Йири - Пан-Европейский Университет (Словакия), к.т.н., профессор, научный сотрудник, e-mail: jiri.vorisek@paneurouni.com

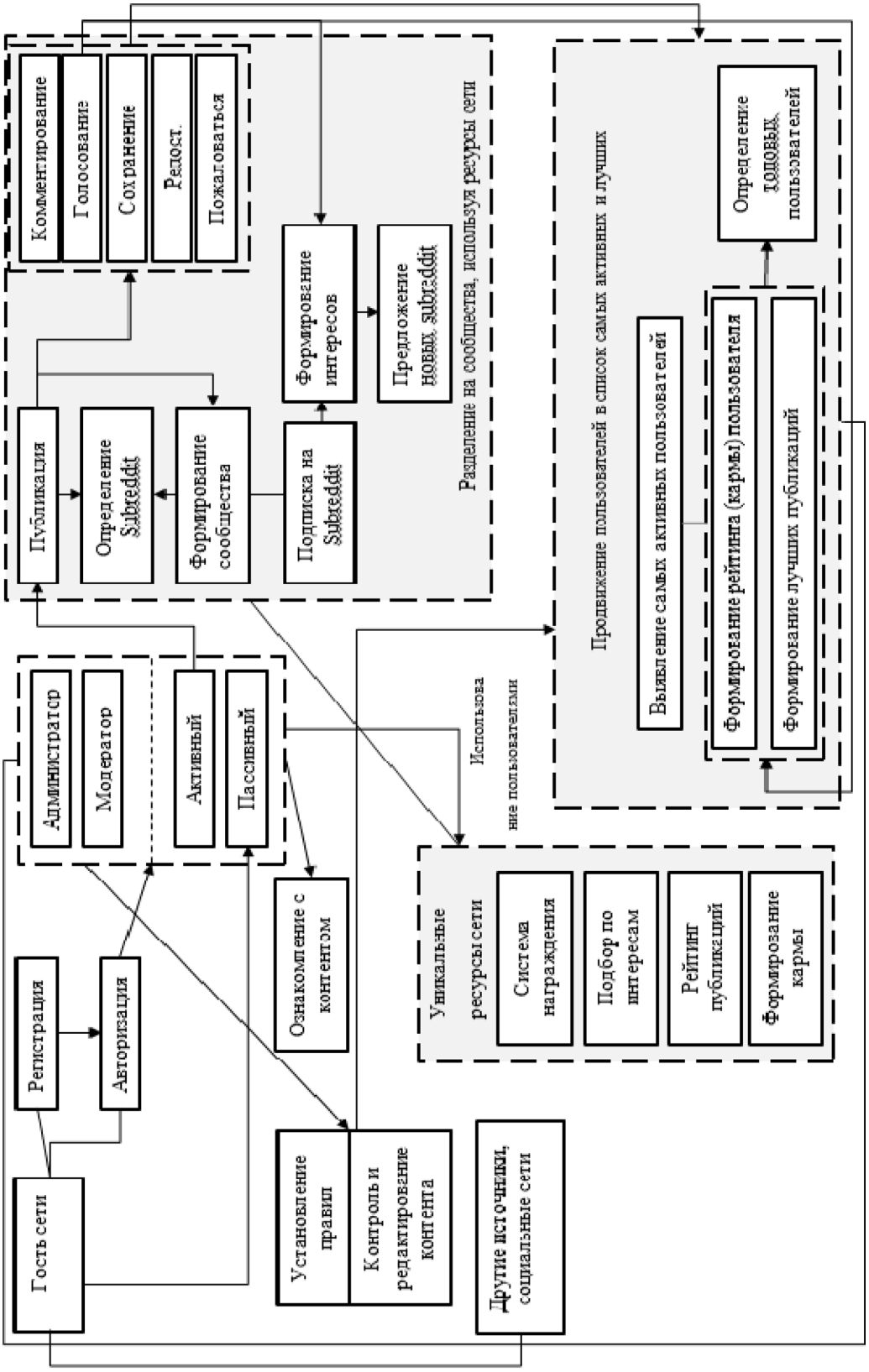


Рис. 2. Структурно-функциональная схема социальной сети Reddit

Посмотрим рейтинг разделов сети Reddit (рис. 3):

Rank	Reddit	Subscribers
1	/r/AskReddit	17,210,054
2	/r/funny	17,070,798
3	/r/todayilearned	17,037,412
4	/r/science	16,895,121
5	/r/worldnews	16,852,062
6	/r/pics	16,800,663
7	/r/IAmA	16,661,290
8	/r/announcements	16,481,937
9	/r/gaming	16,153,667
10	/r/videos	16,107,226
11	/r/movies	15,827,295
12	/r/blog	15,592,663
13	/r/Music	15,381,171
14	/r/aww	15,223,132

Рис. 3. Самые популярные разделы

Из рис. 3 видно, что самые популярные сообщества не так сильно и принципиально отличаются числом подписчиков.

Также отсюда известно, что существует отдельный раздел \videos, где пользователи сети делятся друг с другом ссылками исключительно на видеоматериалы, где 90% всех публикаций составляют ссылки на YouTube[1, 5].

Логично предположить, что взаимодействие между этими социальными сетями (Reddit и YouTube) реально, учитывая также то, что сеть для обмена медиаконтентом имеет возможность распространять информацию через другие социальные сети, включая Reddit (рис. 4).

Таким образом, их связь является двунаправленной [5, 7], где пользователи Reddit представляют на всеобщее обозрение ссылки на видео, а участники YouTube могут поделиться ссылкой в другой сети, имея в ней аккаунт.

Здесь важно отметить то, что в случае с Reddit необязательно иметь профили в других сетях для осуществления публикации ссылки.

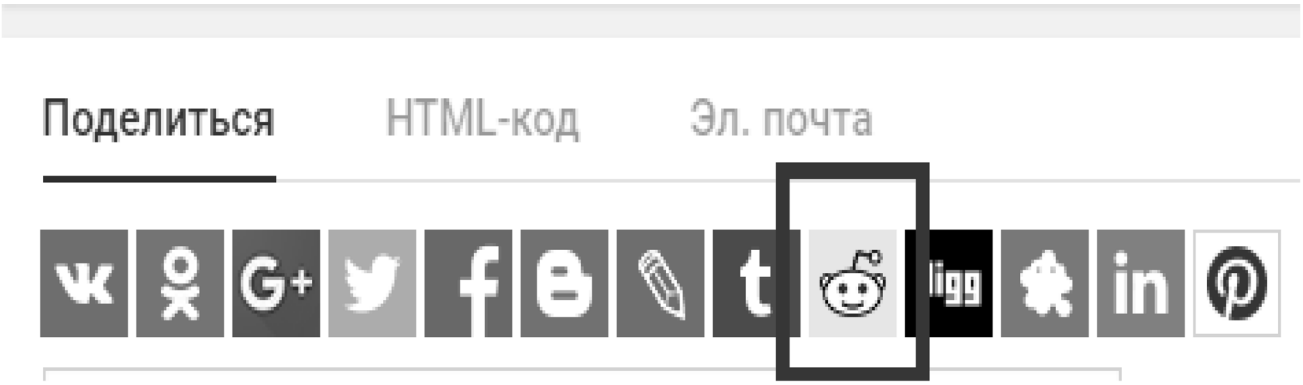


Рис. 4. Возможность YouTube делиться контентом

Со стороны сети YouTube все наоборот, необходимо иметь аккаунты в других социальных сетях.

Далее рассмотрим взаимодействие между двумя социальными сетями Reddit и YouTube, представив выборку каждой из них в виде графов, где вершинами будут являться пользователи, а связи будут характеризовать общее количество переданного трафика.

Т.к. строится общая модель, а не рассматривается какой-либо конкретный случай, то будет логичным, чтобы ребро образовывалось между централизованными узлами разных сетей, имеющих наибольшее количество связей.

Сначала, перед межсетевым взаимодействием рассчитаем коэффициенты прироста сетей, т.к. все социальные сети не являются статичными, а имеют особенность расти и развиваться.

Учитывая процент роста пользователей сети Reddit в период 01.2015 – 04.2017, определим среднее значение [4]:

Табл. 1

Фрагмент статистики роста сети

Месяц	Год	Процент прироста пользователей
Январь	2015	1,9
Февраль	2015	2,23
Март	2015	2,1
…	…	…
Февраль	2017	0,9
Март	2017	0,87
Апрель	2017	1,62

Зная среднее значения среди 28 перечисленных месяцев и коэффициент роста сети, получим 0,13.

Имея в наличии сети выборки Reddit и YouTube, а также их коэффициенты роста (для второй сети данные были взяты с [6]),

соединим между собой вершины с наибольшей связностью.

Коэффициент роста между известным 0,13 у Reddit и вычисленным 0,07 у YouTube равняется 0,1.

Симулируя эпидемический процесс распространения деструктивного контента между двумя сетями, заразим эпидемией хабы обеих и зададим длительность процесса заражения в 30 дней.

Получим (рис. 5):

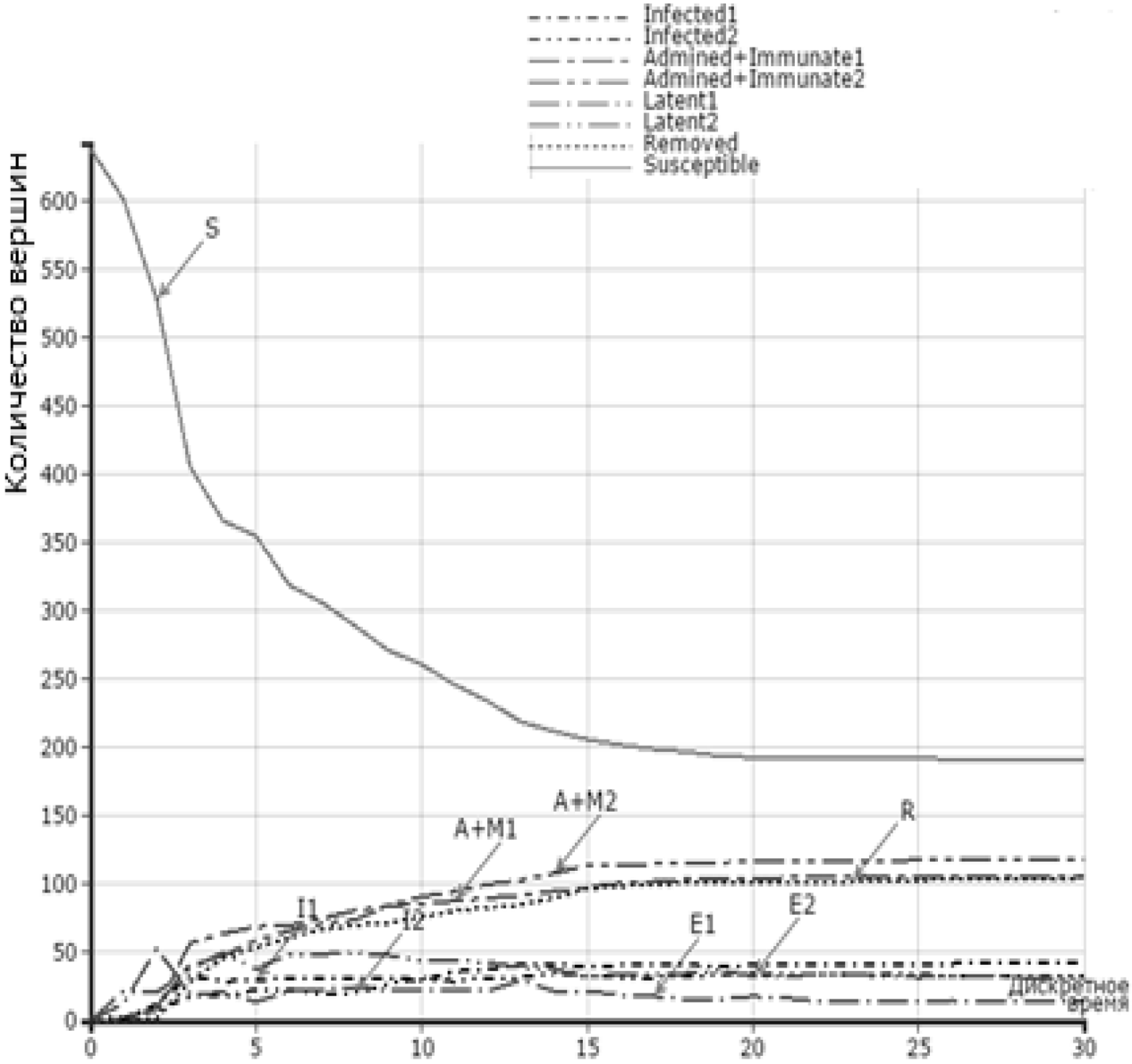


Рис. 5. График отображения изменения состояния вершин.

Из рис. 5 видно, что число действующих, неудаленных вершин значительно сократилось (график S).

В то же время узлы, не подверженные атаке, находятся в значительном меньшинстве (графики E1 и E2) так же, как и число не удаленных из сети, включая инфицированные и получившие иммунитет (A+M1 и A+M2, I1 и I2).

Исходя из проделанного опыта, можно сделать вывод, что конфликт, возникающий между двумя сетями, позволяет рассматривать противостояние как взаимодействие между двумя кластерами одной сети.

Однако, при рассмотрении других случаев всегда следует понимать и учитывать возможность передачи пользователями общего трафика, иначе

межсетевое взаимодействие невозможно реализовать.

Например, рассмотреть взаимодействие между Reddit и Twitter будет проблематично (рис. 2), т.к. наполняемый контент второй сети не совпадает с первым.

Это происходит за счет уникальных ресурсов Twitter, таких как:

—твиты;

—ретвиты;

—посты;

—обращения к пользователям;

—внутреннее размещение медиа-материалов.

Они регулируют размещение и циркуляцию контента внутри сети. Процесс распространения информации в Reddit значительно отличается от этого, как и сама структура сети.

Таким образом, выходит, что не всегда возможно представить две принципиально разные социальные сети как ограниченно взаимодействующие.

Литература

1. Социальная сеть. [Электронный ресурс] – Режим доступа: http://reddit.com.

2. Sundaresan V. Subreddit Recommendations within Reddit Communities / V. Sundaresan, I. Hsu, D. Chang // Stanford University, Department of Computer Science. – 2015. – P. 6-10.

3. Статистические данные [Электронный ресурс]– Режим доступа: https://www.similarweb.com/website/reddit.com#social

4. Исследование и анализ, статистика интернет ресурсов[Электронный ресурс]-.. Режим доступа: http://statista.com.

5. Ресурс, содержащий свежую статистику сети Reddit.com. [Электронный ресурс]- Режим доступа: http://redditlist.com.

6. Статистические данные социальных сетей. [Электронный ресурс]- Режим доступа: http://konect.uni-koblenz.de/networks.

7. Социальная сеть обмена медиа-контентом YouTube.[Электронный ресурс]- Режим доступа: https://www.youtube.com.

8. Paolo, Massa, Martino Salvetti, and DaniloTomasoni. Bowling alone and trust decline in social network sites. In Proc. Int. Conf. Dependable, Autonomic and Secure Computing, pages 658-663, 2016.

9. Tsvetovat, M. Social Network Analysis for network interests: Finding Connections on the Social Web / M. Tsvetovat, A. Kouznetsov // O'Reilly.-2011. — P. 45. — 192 c.

10. Maxwell, A. Pretzlav Last.fm Explorer: An Interactive Visualization of Hierarchical Time-Series Data / A. Maxwell // – 2008. – 11p.

11. Konstas, I. On Social Networks and Collaborative Recommendation / I. Konstas // – 2012. – 5p.

Воронежский научно-образовательный центр управления информационными рисками
Voronezh Science and Education Management Center Information risks
Пан-Европейский Университет
Pan-European University

SOCIAL NETWORK FOR COLLECTIVE DISCUSSIONS REDDIT. INTERNET-NET INTERACTION BETWEEN REDDIT AND YOUTUBE IN THE FRAMEWORK OF THE DISTRIBUTION OF THE EPIDEMIC PROCESS DISSEMINATION WITH THE DYNAMIC NETWORK GROWTH

A.V. Altukhov, I.V. Shevchenko, A.G. Ostapenko, V.M. Pitolin, J. Vorisek

This article analyzes the main problems for collective discussions. REDDIT, simulates the epidemic process of distributing destructive content between networks and results

Key words: social networks, epidemics

УДК 004.021/003.63/004.5

СЕТИ, БОЛЬШИЕ ДАННЫЕ (BIGDATA), ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ (DATAMINING) И ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ

П.Ю. Филяк

Рассматриваются подходы к решению проблем обеспечения информационной безопасности в информационного общества, когда на переднем плане в качестве ключевых факторов начинают выступать такие явления как сетевая глобализация, работа с большими объемами (массивами) структурированных и неструктурированных данных «большими данными» (BigData) и становится очевидной необходимость интеллектуального анализа данных (DataMining)

Ключевые слова: безопасность, информационная безопасность, данные, структурированные данные, неструктурированные данные, информация, анализ, информационно - аналитические системы (ИАС), большие данные (BigData), интеллектуальный анализ данных (DataMining)

В Указе Президента Российской Федерации от 31.12.2015. N 683 «О Стратегии национальной безопасности Российской Федерации» дается определение национальной безопасности, которая включает в себя оборону страны и все виды безопасности, предусмотренные Конституцией Российской Федерации и законодательством Российской Федерации, прежде всего государственную, общественную, информационную, экологическую, экономическую, транспортную, энергетическую безопасность, безопасность личности.

В развитие указанной стратегии в новой редакции Доктрины информационной безопасности Российской Федерации [1] дается четкое определение обеспечения информационной безопасности. Это осуществление взаимоувязанных правовых, организационных, оперативно-розыскных, разведывательных, контрразведывательных, научно-технических, информационно-аналитических, кадровых, экономических и иных мер по прогнозированию, обнаружению, сдерживанию, предотвращению, отражению информационных угроз и ликвидации последствий их проявления. Вопросам информационной сферы посвящены также и другие законодательные и нормативные правовые акты, в частности, - Указы Президента России «О Стратегии развития информационного общества в Российской Федерации на 2017 - 2030 годы» от 09.05.2017. N 203 и «О Стратегии экономической безопасности Российской Федерации на период до 2030 года» от 13.05.2017.N 208.

Совершенно очевидно, что современное развитие предполагает создание информационного общества, что с одной стороны предусматривает всестороннее и комплексное развитие информационных и коммуникационных технологий, обеспечивающих прогресс цивилизации, а с другой стороны означает столкновение с новыми угрозами и вызовами, связанными с применением таких технологий. XXI век можно с уверенностью назвать веком сетей – начиная от реальных физических сетей рек, автомобильных и железнодорожных магистралей, сетевого распространения эпидемий, сетевого маркетинга, сетевых финансовых и платежных систем, сетевого распространения информации, что стало обыденным в условиях всемирной паутины Internet, и заканчивая сетевой виртуализацией, связанной с появлением таких понятий как виртуальная личность, виртуальные сообщества, виртуальное массовое и общественное сознание и другие.

Причем, если с процессами, происходящими на физическом уровне, например распространением компьютерных вирусов, на настоящий момент уже как-то научились разбираться, - чему способствует

Филяк Петр Юрьевич – ФГБОУ ВПО СГУ, канд. техн. наук, доцент, e-mail: parallax-1@yandex.ru

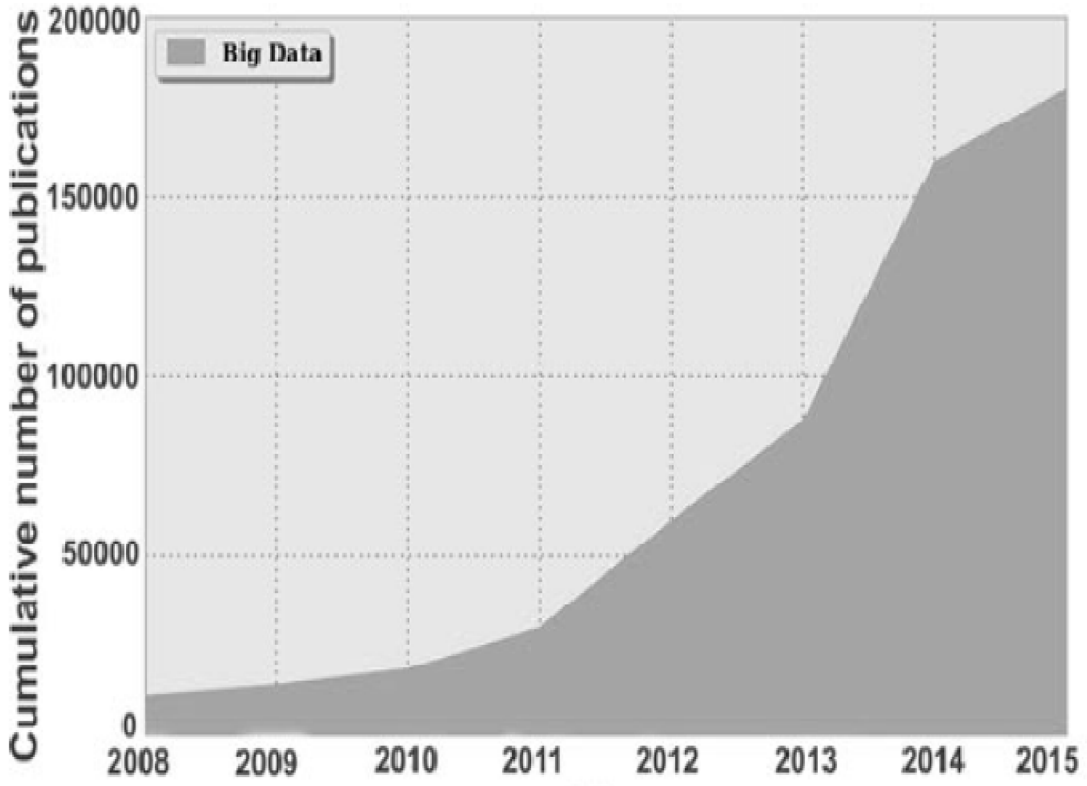


Рис. 1. Возрастание количества публикаций

накопленные статистика и опыт, то с виртуальными и виртуализированными явлениями и процессами ситуацию можно охарактеризовать только как начало пути.

В работах коллектива ученых - Остапенко А.Г., Калашникова А.О., Остапенко Г.А.,Остапенко О.А., Остапенко А.А., Плотникова Д.Г., Паринова А.В., Бабаджанова Р.К., Париновой Л.В., Радько Н.М., Пономаренко Е.Н., Гузева Ю.Н., Соколовой Е.С., Шварцкопфа Е.А., Шевченко И.В. и других представлен широчайший спектр авторских исследований в этой области [2-8], в которых предлагаются конкретные решения проблем и поставленных задач.

Какие дополнительные проблемы всплывают в связи с наступлением эры сетей, помимо обозначенных выше?

Прежде всего это проблема колоссального возрастания объемов данных и информации, с которой приходится сталкиваться и оперировать ею.

По словам Эрика Шмидта - председателя совета директоров компании Google: «Пять экзабайт информации создано человечеством с момента зарождения цивилизации до 2003 года, но столько же сейчас создаётся каждые два дня, и скорость увеличивается».

В настоящее время количество накопленной информации уже измеряется зеттабайтами, приближаясь к йоттабайтам.

Человеческий мозг может вместить примерно 35 экзабайт информации.

То есть, превзойден барьер, когда человеческий мозг уже в принципе не может впитать всю имеющуюся в мире информацию и объять необъятное [9,10].

Нарастание объемов информации можно сравнить с пандемией, что носит чаще всего экспоненциальный характер [9] (рис. 1).

Чем обусловлен такой рост объемов? Прежде всего структурой любой сети, которая, как известно, предполагает наличие узлов, так называемых Hub-ов, и топологических связей между ними (в расчетных схематизациях программного обеспечения, моделирующего сети, они чаще всего называются ребрами).

Конечно же каждый узел может генерировать и репродуцировать очень большое количество информации/данных, а сетевой характер организации

взаимодействия может увеличивать сгенерированную информацию в разы, на порядки.

Но, - возникает закономерный вопрос, - а возрастает ли пропорционально количество знаний в условиях явной информационной перегрузки и количество правильных управленческих решений?

Не является ли производство информации самодостаточным искусственным процессом, замкнутым на себя, живущим по законам и на основе коммерческих интересов производителя информации?

Обоснованы ли объективной потребностью общества такие объемы информации и как они коррелируют с объемами знаний общества?

Как оперировать такими объемами информации/данных и принимать в условиях информационной перегрузки управленческие решения? Давайте вспомним пирамиду семантической иерархии DIKW (**d**ata, **i**nformation, **k**nowledge, **w**isdom– **д**анные, **и**нформация, **з**нания, **м**удрость) Рассела Акоффа, предложенную им в 1989 году, которая предполагает следующую градацию:

- в основании находится уровень данных;
- информация добавляет контекст;
- знание добавляет «как» (механизм использования);
- мудрость добавляет «когда» (условия использования).

Рис. 2. Соотношение понятий «данные», «информация», «знания», «мудрость»

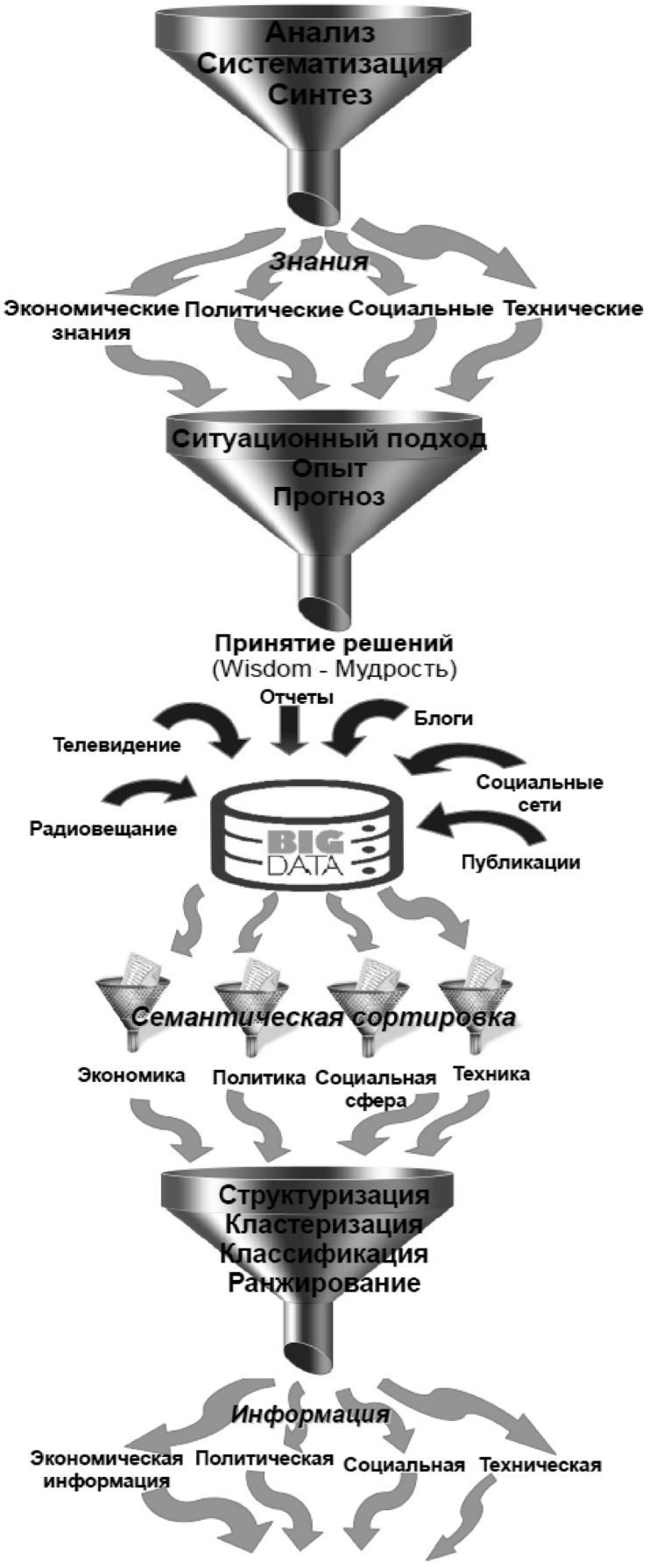


Рис. 3. Схема DataMining
(Интеллектуальный анализ данных)

Таким образом, по Акоффу, знание выступает в качестве ценности информации. Это то, что превращает информацию в инструкции (рецепты), далее эти «рецепты» уже могут быть использованы по своему собственному разумению.

Следовательно, необходима та тонкая грань, которая превращает данные и информацию в знания, для чего необходимы смысловые зависимости (семантические метрики) и формализация, реализуемая в каких-либо из принятых моделях, - фреймовые модели, нейронные сети и т.д.

Это можно наглядно проиллюстрировать повернув «пирамиду Акоффа» на 180^{o} – пирамида превратится в воронку.

А что делать с большими объемами данных «большими данными» (BigData)? [11,12] - Применить «воронку» и «профильтровать» их с помощью «специальных фильтров».

Подобно тому, как в сепараторе из молока получают масло и другие молочные продукты, отбрасывая в итоге пахту (сыворотку) или же осуществляют крекинг нефти. (рис.3)

Что и кто могут выступать в качестве фильтров?

Традиционно исторически роль таких фильтров выполняли мудрецы, знатоки, жрецы, пропуская через себя все возможные объемы информации, предоставляя «на выхлопе» пользователю только ту усеченную рафинированную информацию, которая по их мнению именно и нужна была потребителю, зачастую очень высокопоставленному.

В настоящее время роль «фильтров» выполняют преподаватели, родители, «прокачивая» через себя Гига-, а то Терра-байты «первички», чтобы в ограниченное учебным или воспитательным процессом время сформировать у студентов/детей правильную картину событий или «картину мира».

В условиях информационного общества роль фильтров выполняют информационно-аналитические системы (ИАС), позволяющие осуществлять интеллектуальный анлиз данных (DataMining – получение знаний из данных) и когнитивные технологии, реализованные в тех или иных алгоритмах и когнитивных схемах, - фреймовые модели, нейронные сети, системы искусственного интеллекта [11,12].

Примером одной из эффективных реализаций подобного подхода можно назвать разработку корпорации IBM - компьютер Watson, некоторое время назад ставший уже серийным продуктом данной компании, предполагающим настройку этого компьютера, машинное его обучение и тренировки с накоплением статистики.

Литература

1. Указ Президента РФ от 05.12.2016 №646 «Об утверждении Доктрины информационной безопасности Российской Федерации»

2. Остапенко, А.Г. Теория сетевых войн. Социальные сети, как инструмент «цветных революций»/ А.Г. Отапенко, А.О. Калашников, О.А. Остапенко, Е.А. Шварцкопф, Е.С. Соколова, А.А. Резов, А.А. Ломов. - В.: ФГБОУ ВПО «ВГТУ», 2015. – 106 с.

3. Остапенко, А.Г. Сетевое противоборство социотехнических систем/ А.Г. Остапенко, А.О.Калашников, О.А. Остапенко, П.В. Бровченко, М.Г. Борисова, И.В. Шевченко, И.Г. Морозов. –В.: ФГБОУ ВПО «ВГТУ», 2015. – 110 с.

4. Остапенко, А.Г. Теория сетевых воин: глобальное информационное управление/А.Г. Остапенко, А.О. Калашников, О.А. Остапенко, Д.М. Баранов, М.А Тарелкин. -.В.: ФГБОУ ВПО «ВГТУ», 2016. – 140 с.

5. Остапенко, А.Г. Теория сетевых войн. Живучесть атакуемыхсетей/ А.Г. Остапенко, А.О. Калашников, Г.А. Остапенко, Д.Г. Плотников, О.В. Доросевич,Ю.Г. Стародубцева, С.В.Чернышова.. – В.: ФГБОУ ВПО «ВГТУ», 2016. – 157 с

6. Остапенко, А.Г. Статические и динамические параметры взвешенных сетей [Текст]/ А.Г. Остапенко, Д.Г. Плотников, Ю.Н.Гузев// Информация и безопасность: Регион.науч-техн. журнал.-В. 2016. -Т. 19. № 1. -С. 100-105.

7. Остапенко, Г.А.Стратегии сетевого противоборства[Текст]/ Г.А. Остапенко, Д.Г. Плотников, Ю.Н.Гузев//Информация и безопасность: Регион.науч-техн. журнал.-. В.2016. -Т. 19. № 2.- С. 250-253.

8. Остапенко, А.Г. Программный комплекс моделирования эпидемических процессов в социальных сетях[Текст]./ А.Г. Остапенко, Е.А. Шварцкопф, Д.А. Савинов, Е.В. Гусев, Д.В.Гусев// Информация и безопасность: Регион.науч-техн. журнал.-. В. 2017. -Т. 20. № 1-1 (4).- С.39-48.

9. Безопасность и конфиденциальность централизованной системы профилактики заболеваний. [Электронный ресурс]. URL: https://yadi.sk/d/F-CaJt3J3DgeS4

10. Филяк, П.Ю. Информационная безопасность и комплексная система безопасности: анализ, подходы[Текст]/П.Ю. Филяк//Информация и безопасность: Регион.науч-техн. журнал.-. В. 2016. -Т. 19. №. 1- С.72-79.

11. Майер-Шенбергер, В.Большие данныеРеволюция, котораяизменит то, как мы живем, работаем и мыслим. / В. Майер-Шенбергер, К.Кукьер. – М. : Манн, Иванов и Фербер, 2014. – 240 с.

12. Большие данные (BigData): изменение будущего человечества [Электронный ресурс] – Режим доступа: http://www.kitaichina.com/se/txt/2013-03/28/content_530849.htm

13. Филяк, П.Ю. Актуальность обеспечения информационной и экономической безопасности в условиях информационного общества[Текст]./ П.Ю. Филяк//Известия Тульского государственного университета. Технические науки:В. 2013.-Т.9. № 3. -С. 262-267.

14. Филяк, П.Ю. Проектирование с учетом обеспечения безопасности[Текст]/П.Ю. Филяк//Информация и безопасность.-В. 2015. -Т. 18. № 1.- С. 101-106.

15. Филяк, П.Ю. Информационная безопасность и комплексная система безопасности: анализ, подходы. [Текст]/ П.Ю. Филяк// Информация и безопасность.-В. 2016. -Т. 19. № 1. -С. 72-79

16. Agarwal A. Sentiment analysis of Twitter data / A.Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau// LSM ’11 Proceedings of the Workshop on Languages in Social Media, Association for Computational Linguistics.–2011.– P. 623.

17. Amna D. Social Relevance for review probabilities / D. Amna,H. Hatem // International Conference on Control, Engineering & Information Technology (CEIT'13) Proceedings Engineering & Technology. –2013. – Vol.1.

18. Bild D. Aggregate Characterization of User Behavior in Twitter and Analysis of the Retweet Graph / D.Bild, R. Dick, Y.Liu, Z.Mao //Texas CSSI.–2014.– Vol. 2.

19. Cha M. Flash Floods and Ripples: The Spread of Media Content through the Blogosphere / M.Cha, J. A. N.Perez, H.Haddadi // 3rd Int'l AAAI Conference on Weblogs and Social Media (ICWSM) Data Challenge Workshop. – 2009.– Vol. 12.

20. Fragkiskos D. Mining Social and Information Networks / D.Fragkiskos//UCSD Artificial Intelligence Seminar. – 2016.– P. 123-145.

21. Gao B. Topic-Level Social NetworkSearch / B.Gao, J. Tang , Y.Wan, S.Wu // 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining in New York– 2014 – P. 769–772.

ФГБОУ ВО «Сыктывкарский государственный университет имени Питирима Сорокина»
Syktyvkar State University them Pitirim Sorokin

NETWORKS, BIG DATA, DATA MINING, DATA MINING AND SECURITY

P.Yu. Filyak

Examines the approaches to resolving the problems of information security in knowledgeable society, when in the foreground as key factors are beginning to speak of phenomena such as network globalization, working with Big Data (big arrays of structured and unstructured data)

Key words: security, information security, data, structured data, unstructured data information, analysis, information - analytical system (IAS), Big Data, Data Mining

УДК 004.056.57

СОЦИАЛЬНЫЕ СЕТИ И НАУЧНО-ТЕХНИЧЕСКИЕ ПРЕДПОСЫЛКИ ПРОГРАММЫ «БЕЗОПАСНЫЙ ИНТЕРНЕТ»

А.Г. Остапенко, А.А. Акинина, Г.А. Остапенко, Е.Ю. Чапурин, Н.Ю. Щербакова

Рассматривается содержание и предпосылки развертывания региональной программы «Безопасный Интернет»

Ключевые слова: соцсети, программа, риски

Социальные сети – сегодня самый мощный вещатель (миллиарды активных и пассивных пользователей). Пользователь, как субъект сети, объективно заинтересован в достижении максимальной популярности в социуме и в получении самой полезной информации через сетевые сервисы. Этим объясняется феномен лавинообразного роста количества как самих социальных сетей различного назначения, так и их пользователей, а также – ожесточенной конкурентной борьбы в этом сегменте информационного пространства. Фактически речь идет о битве контентов за информационно-психологические предпочтения массы социально и экономически активных людей, охватывающей треть Человечества. Ставки весьма высоки и поэтому технологическая изощрённость конструирования и распространения контента в социальных сетях превышает все ожидания. И здесь краеугольным выступает вопрос об инструментарии – комплексе средств пошагового моделирования процесса диффузии контента в социальных сетях, исходя из вероятностных параметров его восприятия. То есть остро необходим инструмент прогнозирования сетевого успеха контента, который желают использовать как всякий активный пользователь, так и многочисленные исследователи социальных сетей. Ожидаемую востребованность такого комплекса (в том числе при исполнении его в виде мобильного приложения) трудно переоценить.

В плане импортозамещения здесь весьма полезными оказались отечественные наработки [1-23] в области дискретного риск-моделирования информационных процессов. Дело в том, что традиционный инструментарий моделирования диффузии наполнителя в сетевых структурах, широко используемый для описания медико-биологических эпидемий, в силу своего аналогового характера практически исчерпал свои возможности, оставив множество вопросов.

В этой связи возникает необходимость создания инструментария (математического и программного обеспечения) дискретного моделирования процессов распространения контента в гетерогенных информационных сетях, разрешающего вышеперечисленные противоречия в интересах более адекватного прогнозирования развития данных процессов в условиях нарастающего сетевого противоборства [14, 23].

Условно данный проект можно назвать «калькулятор сетевого успеха». Он нацелен на прогнозное моделирование процессов распространения контента в социальных сетях. Оценивая шансы восприятия сгенерированного контента сетевой структурой, пользователь соцсети измеряет возможности своего интеллектуального продукта. Амбиции автора контента могут

Остапенко Александр Григорьевич – ВГТУ, заведующий кафедрой СИБ, д-р техн. наук, профессор, e-mail: mnac@comch.ru
Акинина Анна Александровна – ВГТУ, студент, mnac@comch.ru
Остапенко Григорий Александрович –ВГТУ, профессор, д-р техн. наук, профессор, e-mail: mnac@comch.ru
Чапурин Евгений Юрьевич – ВГТУ, аспирант, e-mail: mnac@comch.ru
Щербакова Наталья Юрьевна – ВГПУ, ст. преподаватель, e-mail: mnac@comch.ru

успешно реализоваться в данном случае через сервисы социальных сетей. При этом затратная часть в основном сконцентрирована в области генерации качественного контента, а главной целью современного пользователя (в широком смысле) является самореализация через завоевание популярности в сети. Он готов многократно совершать попытки по достижению данной цели и «калькулятор сетевого успеха» может оказаться удобным подспорьем в сокращении количества подобных попыток.

Соответствующий программный комплекс «Netepidemic» создавался и развивается для моделирования процессов сетевого распространения деструктивного контента, т.е. оценки эпистойкости сети в контексте обеспечения её безопасности. Однако, он применим как к контенту со знаком «минус» (деструктивному), так и контенту со знаком «плюс» (позитивному). Здесь всё зависит от намерений пользователя, генерирующего контент. И чем искуснее создан этот продукт, тем больше пользы или вреда (ущерба) он принесет сети и её социальным элементам.

Вместе с тем, Интернет-пространство все еще остается одним из наименее зарегламентированных сегментов сетевого мироустройства, дающему шансы распространения самого разнообразного контента, продвижения широкого спектра мнений и отстаивания различных позиций. Здесь отказ от толерантности и монополизация пока затруднены принципами создания и функционирования сети, ибо подрыв сетевой демократии повлечет резкое падение количества пользователей, а, следовательно, и доходов владельцев ресурсов. Этим, к сожалению, пользуются и злоумышленники (от фейковых ньюсмейкеров до террористов), но цензура постепенно входит в это пространство, и регламентация его функционирования, очевидно, будет нарастать. Этого требует, прежде всего, необходимость обеспечения кибербезопасности и снижения прочих Интернет-рисков. На государственном и региональном уровнях будут появляться проекты, которые условно можно обозначить термином «Безопасный Интернет». В перспективе эта тенденция найдет свое отражение и в международном праве.

Несмотря на приведенные выше прогнозы, пока Интернет по-прежнему во многом остается территорией личной свободы. Этот личностный акцент глобальной сети обуславливает как ее популярность, так и многочисленные угрозы, проистекающие из нескончаемых попыток управления личностью, обществом, а иногда и государством. В этой дуальности будут происходить ее дальнейшее развитие и сетевое информационное противоборство.

Увы, мультисетевое пространство социальных сетей стало ареной ожесточенной политической борьбы за право управлять массовым сознанием. Это еще раз подчеркивает значимость сетевых исследований в плане оценки возникающих в данном случае рисков и шансов информационного влияния, управления и противоборства в интересах обеспечения национальной безопасности и устойчивого развития каждой страны.

Сетевая экспансия неотвратима. Сети создаются и развиваются в рамках предприятий, регионов, государств и континентов. Они проявляют свойство интеграции и взаимопроникновения.

Таким образом можно говорить об устойчивой мультисетевой организации современного общества.

При этом сети несут как благо, так и вред, и борьба за сетевую безопасность становится одной из самых острых проблем.

Всякий регион, являющийся научным и вузовским концентратом, может и должен активно включиться в эту борьбу, скажем, в рамках пилотного проекта, который условно можно назвать «Безопасный Интернет».

Привлекая все компетентные и заинтересованные структуры на территории региона, представляется возможным организовать систематическое и широкомасштабное сканирование и анализ Интернет-пространства, имея в виду концентрацию (по вузовским

специальностям) студентов, преподавателей и научных работников на правовых, психологических и пр. аспектах информационной защиты индивидуума в глобальных сетях, а также – на технической, социальной и др. специфике обеспечения сетевой безопасности населения региона, представленного в соответствующих Интернет-ресурсах.

Ориентировочное содержание работ и ожидаемые результаты по этапам проекта представлены в Табл. 1.

Табл. 1

План проекта «Безопасный Интернет»

№ этапа	Содержание работ по этапу	Ожидаемый результат работ по этапу
1	На основе психоанализа и социологических исследований: - выявление Интернет-предпочтений, устремлений и ожиданий различных категорий и групп населения; - установление популярности и режима пользования для Интернет-ресурсов, востребованных различными категориями и группами населения региона; - вероятностная оценка восприятия и распространения разнообразного по содержанию и форме контента, циркулирующего в интернет-пространстве, для соответствующих категорий и групп населения региона.	Модели регионального Интернет-пространства и его пользователя по соответствующим группам и категориям населения региона
2	На основе разработанной модели регионального Интернет-пользователя: - исследование Интернет-ресурсов, используемых жителями области, в контексте обеспечения информационной безопасности личности и общества; - выявление угроз нарушения региональной Интернет-безопасности для различных категорий и групп населения региона; - оценка риска реализации выявленных информационных угроз в отношении населения региона.	Модели Интернет-угроз для населения региона
3	На основе разработанных моделей: - формирование информационного обеспечения, необходимого для их программной реализации; - разработка математического обеспечения для моделирования Интернет-отношений и информационных рисков в структуре населения региона; - создание алгоритмического обеспечения для реализации, программно-технического комплекса Интернет-моделирования.	Методики и алгоритмы моделирования Интернет-процессов в структуре населения региона

№ этапа	Содержание работ по этапу	Ожидаемый результат работ по этапу
4	В соответствии с разработанными алгоритмами: - выбор необходимого лингвистического обеспечения, необходимого для их программной реализации; - создание программно-технического комплекса моделирования Интернет-процессов и информационных рисков в структуре населения региона; - тестирование и коррекция стартовой версии программно-технического комплекса Интернет-моделирования.	Стартовая версия программно-технического комплекса моделирования Интернет-процессов и информационных рисков в структуре населения региона
5	С учетом рекомендаций региональных органов власти и территориальных органов заинтересованных федеральных ведомств: - модернизация стартовой версии программно-технического комплекса моделирования Интернет-процессов и информационных рисков в структуре населения региона; - проведение межрегиональной научно-практической конференции по тематике программы и результатам проведенных в ней исследований и реализованных разработок; - публикация результатов работы и открытие соответствующих образовательных программ.	Рабочая версия программно-технического комплекса моделирования Интернет-процессов
6	Создание инструментария управления информационными рисками (с использованием разработанного программно-технического комплекса), включая: - прогнозирование деструктивных Интернет-процессов в структуре населения региона; - рекомендации и методики противодействия Интернет-угрозам, реализуемым в отношении жителей региона; - их практическое внедрение в реальных информационных контрмерах, реализуемых совместно с компетентными органами власти на территории региона.	Стартовая версия инструментария управления Интернет-рисками региона

Венцом проекта (на основе результатов вышеуказанного исследования) должны стать методики и алгоритмы управления Интернет-рисками, реализованные в рамках единого программно-технического комплекса и внедренные в практику деятельности компетентных органов власти.

Литература

1. Denial of service in components of information telecommunication systems through the example of “network storm” attacks / A.G. Ostapenko, S.S. Kulikov, N.N. Tolstykh, Y.G. Pasternak, L.G. Popova // World Applied Sciences Journal. – 2013. – 25 (3). – P. 404-409.

2. The usefulness and viability of systems: Assessment methodology taking into account possible damages / A.G. Ostapenko, E.F. Ivankin, V.S. Zarubin, A.V. Zaryaev // World Applied Sciences Journal. – 2013. – 25 (4). – P. 675-679.

3. Analytical estimation of the component viability of distribution automated information data system / G.A. Ostapenko, D.G. Plotnicov, O.Y Makarov, N.M. Tikhomirov, V.G. Yurasov // World Applied Sciences Journal. – 2013. – 25 (3). – P. 416-420.

4. 4. Analytical models of information-psychological impact of social information networks on users / G.A. Ostapenko, L.V. Parinova, V.I. Belonozhkin, I.L. Bataronov, K.V. Simonov // World Applied Sciences Journal. – 2013. – 25 (3). – P. 410-415.

5. 5. Ensuring the security of critically important objects and trends in the development of information technology / A.O. Kalashnikov, Y.V. Yermilov, O.N. Choporov, K.A. Razinkin, N.I. Barannikov // World Applied Sciences Journal. – 2013. - № 25 (3). – P. 399-403.

6. Optimization of expert methods used to analyze information security risk in modern wireless networks / S.A. Ermakov, A.S. Zavorykin, N.S. Kolenbet, A.G. Ostapenko, A.O Kalashnikov // Life Science Journal. – 2014. – № 11(10s). – P. 511-514.

7. Email-flooder attacks: The estimation and regulation of damage / V.V. Butuzov, A.G. Ostapenko, P.A. Parinov, G.A. Ostapenko // Life Science Journal. – 2014. – 11 (7s). – P. 213-218.

8. Assessment of the system’s EPI-resistance under conditions of information epidemic expansion / N.M. Radko, A.G. Ostapenko, S.V. Mashin, O.A. Ostapenko, D.V. Gusev // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (3). – P. 1781-1784.

9. Peak risk assessing the process of information epidemics expansion / N.M. Radko, A.G. Ostapenko, S.V. Mashin, O.A. Ostapenko, A.S. Avdeev // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (Spl.End). – P. 251-255.

10. Flood-attacks within the hypertext information transfer protocol: damage assessment and management / A.G. Ostapenko, M.V. Bursa, G.A. Ostapenko, D.O. Butrik // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (Spl.End). – P. 173-176.

11.Discreet risk-models of the process of the development of virus epidemics in non-uniform networks / V.V. Islamgulova, A.G. Ostapenko,,N.M. Radko, R.K. Babadzhanov, O.A. Ostapenko // Journal of Theoretical and Applied Information Technology. – 2016. – Vol. 86. – No.2. – P. 306-315.

12. Algorithm of Generation of Scale-Free Network at Realization Virus Attacks on Model Chiang Lu. / E. S. Sokolova, N. I. Barannikov, I. L. Bataronov, V.I.Belonozhkin, Research Journal of Pharmaceutical, Biological and Chemical Sciences. – 2016. – Vol. 7. – No.4. – P. 2438-2447.

13. Modeling of layering growth virus epidemic and spread of harmful content on Poisson networks / E.A. Shvartskopf, A.V. Zaryaev, L.V. Parinova, L.G. Popova. / Research Journal of Pharmaceutical, Biological and Chemical Sciences. – 2016. – Vol. 7. – No.4. – P. 2321-2331.

14. Discrete risk models of the process of viral epidemics development in homogenous information and telecommunication networks / E.N. Ponomarenko, V.N. Kostrova, R.K. Babadzhanov, Y.N. Guzev, V.S. Zarubin // Journal of Theoretical and Applied Information Technology. – 2016. – Vol. 92. – No.2. – P. 235-252.

15. Denial of service in components of information telecommunication systems through the example of “network storm” attacks / A.G. Ostapenko, S.S. Kulikov, N.N. Tolstykh, Y.G. Pasternak, L.G. Popova // World Applied Sciences Journal. – 2013. – 25 (3). – P. 404-409.

16. The usefulness and viability of systems: Assessment methodology taking into account possible damages / A.G. Ostapenko,

E.F. Ivankin, V.S. Zarubin, A.V. Zaryaev // World Applied Sciences Journal. – 2013. – 25 (4). – P. 675-679.

17. Analytical estimation of the component viability of distribution automated information data system / G.A. Ostapenko, D.G. Plotnicov, O.Y Makarov, N.M. Tikhomirov, V.G. Yurasov // World Applied Sciences Journal. – 2013. – 25 (3). – P. 416-420.

18. Analytical models of information-psychological impact of social information networks on users / G.A. Ostapenko, L.V. Parinova, V.I. Belonozhkin, I.L. Bataronov, K.V. Simonov // World Applied Sciences Journal. – 2013. – 25 (3). – P. 410-415.

19. Email-flooder attacks: The estimation and regulation of damage / V.V. Butuzov, A.G. Ostapenko, P.A. Parinov, G.A. Ostapenko // Life Science Journal. – 2014. – 11 (7s). – P. 213-218.

20. Assessment of the system's EPI-resistance under conditions of infor-mation epidemic expansion / N.M. Radko, A.G. Ostapenko, S.V. Mashin, O.A. Ostapenko, D.V. Gusev // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (3). – P. 1781-1784.

21. Peak risk assessing the process of information epidemics expansion / N.M. Radko, A.G. Ostapenko, S.V. Mashin, O.A. Ostapenko, A.S. Avdeev // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (Spl.End). – P. 251-255.

22. Flood-attacks within the hypertext information transfer protocol: damage assessment and management / A.G. Ostapenko, M.V. Bursa, G.A. Ostapenko, D.O. Butrik // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (Spl.End). – P. 173-176.

23. Discreet risk-models of the process of the development of virus epidemics in non-uniform networks / V.V. Islamgulova, A.G. Ostapenko,,N.M. Radko, R.K. Babadzhanov, O.A. Ostapenko // Journal of Theoretical and Applied Information Technology. – 2016. – Vol. 86. – No.2. – P. 306-315.

ФГБОУ ВО «Воронежский государственный технический университет»
Voronezh State Technical University
ФГБОУ ВО «Воронежский государственный педагогический университет»
Voronezh State Pedagogical University

GLOBAL WEB SITE: RISKS AND CHANCES OF THE MULTI-NETWORK PEACEKEEPING WORLDWIDE

A.G. Ostapenko, A.A. Akinina, G.A. Ostapenko, E.Yu. Chapurin, N. Yu. Shcherbakova

The content and prerequisites of the deployment of the regional program "Secure Internet"
Key words: social network, program, risks

УДК 004.056

ОЦЕНКА УРОВНЯ ЗАЩИЩЁННОСТИ ИНФОРМАЦИОННЫХ РЕСУРСОВ НА ОСНОВЕ ТЕСТОВ НА ПРОНИКНОВЕНИЕ

Р.Р. Галимов, В.П. Членов

В данной статье предлагается методика оценки защищенности информационных ресурсов на основе проведения тестовых атак с использованием ПО Kali Linux

Ключевые слова: оценка уровня защищенности, тесты на проникновение

Эффективность работы современного предприятия во многом зависит от возможности обеспечить требуемый уровень информационной безопасности. Сложность современных информационных систем организаций обусловливает актуальность задачи оценки текущего уровня защищённости информационных ресурсов. В частности, по результатам исследования, из 18-ти информационных систем, принадлежащих крупным промышленным банковским и IT-организациям, только одна оказалась действительно хорошо защищённой [1]. Причём в 67% случаев получение несанкционированного доступа оказалось возможным от лица любого внешнего злоумышленника, а в 27% достаточно было иметь пользовательский доступ к внутренней сети.

Данный результат во многом определяется тем, что вне зависимости от качества спроектированной системы защиты информационной системы, нельзя исключать такие факторы, как наличие уязвимостей или отсутствие актуальных обновлений безопасности в программном обеспечении, ошибочные действия пользователей, неправильная конфигурация средств защиты. Обнаружить такие уязвимости можно только посредством проведения тщательной проверки информационной системы квалифицированным специалистом при помощи специализированных аппаратно-программных средств. В связи с этим существует необходимость в средствах автоматизации, позволяющих повысить достоверность оценки уровня защищённости информационных ресурсов (ИР), характеризующейся небольшими стоимостными затратами.

Вопросам оценки уровня защищённости посвящено множество работ. В частности, в работе Осовецкого Л.Г. предлагается оценивать уровень защищённости автоматизированной системы исходя из рейтинга стойкости отдельных механизмов защиты [2]. В свою очередь, в статье Мукминова В.А, Лобузько А.В. и Хуцишвили В.М. защищённость компьютерной сети оценивается по результатам тестов, проведённых при помощи программ, используемых нарушителями [3]. Несмотря на достоинства данных работ, необходимо отметить, что в них в недостаточной степени рассмотрены такие вопросы, как конвертация результатов тестов в количественную оценку состояния защищённости информационного ресурса, не учитывается неравномерность распределения ИР по вычислительным средствам ИС. В связи с этим существует необходимость в разработке методики оценки защищённости, позволяющей получить количественные показатели риска информационной безопасности, учитывающей неравнозначность узлов ИС, на которых хранятся и обрабатываются информационные ресурсы.

Целью работы является повышение достоверности оценки защищённости информационных ресурсов.

Предлагаемая в данной работе методика основывается на подходе оценки уровня защищённости ИС в результате проведения тестов на проникновение. Данный подход гарантирует наибольшую

Галимов Ринат Равилевич-ОГУ, канд. техн. наук, доцент,e-mail:rin-galimov@yandex.ru
Членов Владимир Петрович- ОГУ, магистрант,e-mail: bob0451@rambler.ru

достоверность по сравнению с другими методами оценки защищённости, так как позволяет обнаружить бреши в защите автоматизированной системы, некорректные настройки, неисправности в аппаратных средствах защиты, использование устаревших либо ненадёжных версий программ[6]. В данной статье под информационным ресурсом понимается документы, представляющие собой персональные данные, конфиденциальную информацию, служебную и коммерческую тайну, которые распределены по узлам информационной системы. Объём информации, хранимой на компьютере, и её стоимость могут изменяться в зависимости от функций вычислительных средств. Например, серверный ПК обычно содержит больший объём конфиденциальной информации, чем рабочая станция сотрудника. Данное обстоятельство определяет необходимость учитывать неравнозначность компьютеров информационной системы для оценки её уровня защищённости.

Входными данными модели оценки уровня защищённости ИР являются: перечень информационных ресурсов в системе (***R***), перечень потенциальных угроз безопасности (***A***), полученные экспертным путём. На выходе формируются оценки защищённости информационных ресурсов: показатель уязвимости для каждого ПК в ИС организации (***Zk***) и риск информационной безопасности для каждого ресурса в системе (***ZR***). Каждый ***i***-ый информационный ресурс характеризуется следующими параметрами:

$$r_i : < C, L> \quad (1)$$

$$L=\{<l_1, s_1>, <l_1, s_1>, \ldots < l_n, s_n>\} \quad (2)$$

где ***C*** – стоимостная оценка ресурса; ***L*** – множество средств ВТ, по которым распределен информационный ресурс в системе; l_i – идентификатор ***i***-го компьютера, на котором хранится часть данного ресурса; s_i –доля ***i***-ого ресурса, размещенного ***i***-ом компьютере.

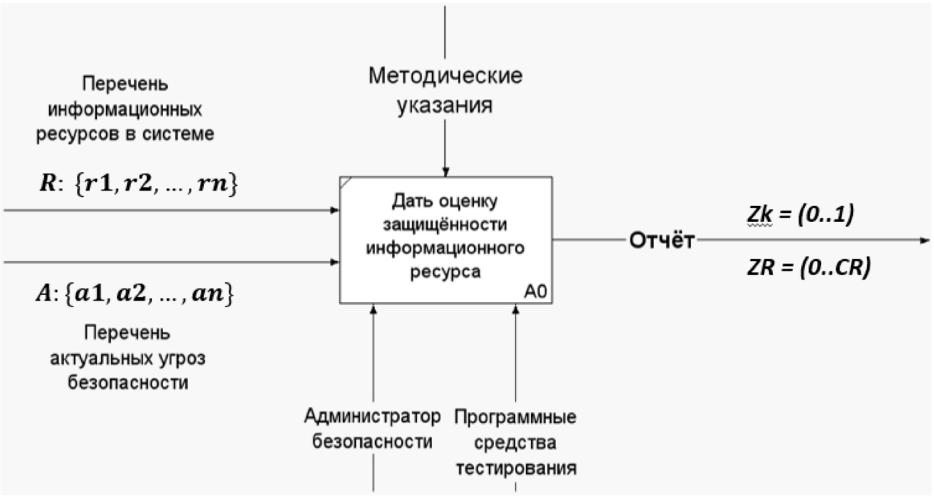


Рис. 1. Функциональная модель оценки уровня защищённости ИР

Коэффициент s_i может принимать следующие значения:

- 1 – в случае, когда на рабочей станции расположен весь ресурс целиком, либо его полная копия;
- 0,5 – рабочая станция хранит значительную часть ресурса, свыше 25% от полного объёма;
- 0,25 – рабочая станция хранит незначительную долю ресурса, ниже 25% от полного объёма.

Каждой ***i***-ой актуальной угрозе соответствует тестовая атака, определяющаяся следующими параметрами:

$$a_i : <n_i, v_i, u_i> \quad (3)$$

где n_i – идентификатор тестовой атаки; v_i – весовой коэффициент, характеризующий величину ущерба в случае реализации атаки; u_i – уязвимость ПК, наличие которой позволяет реализовать данную атаку.

Весовой коэффициент ***v*** может принимать следующие значения:

- *0,25* – в случае, когда злоумышленником получена информация о компьютере (установленных программах, хранящихся на ПК информационных ресурсах), однако доступ к нему не был получен;
- *0,5* – злоумышленнику удалось вмешаться в работу машины, вызвать сбой программы или отключение одного из компонентов системы защиты, но не удалось получить доступ к защищаемой информации;
- *1* – был получен полный доступ к защищаемой информации.

По результатам тестов определяется показатель уязвимости отдельного протестированного ПК (диапазон значений для данной величины от 0 до 1):

$$Z_{k_n} = \sum_{i=1}^{M} \frac{p_i v_i}{v_i} \quad (4)$$

где ***pi*** – результат прохождения ***i***-го теста из множества определённых атак ***A***, который может принимать значение 1 при успешной реализации, иначе 0; ***vi*** – весовой коэффициент i-ого теста; ***M*** – количество проведённых тестов.

Оценка уровня защищенности для конкретного информационного ресурса определяется по формуле:

$$Z_{r_i} = \sum_{n=1}^{N} \frac{Z_{k_n} * C_{r_i} * S_{nr_i}}{N} \quad (5)$$

где ***Zkn*** – показатель уязвимости ***n***-го ПК, на котором расположен данный ресурс; ***Cri*** – стоимость конкретного информационного ресурса; ***Snri*** – долевой коэффициент ***i***-ого ресурса для ***n***-ого компьютера; ***N*** – количество компьютеров, на которых полностью либо частично хранится ресурс r_i.

Таким образом, оценка защищённости информационного ресурса определяется как риски информационной безопасности, которые несёт владелец информации. С учётом разработанной модели был предложен алгоритм оценки уровня защищённости, представленный на рис. 2.

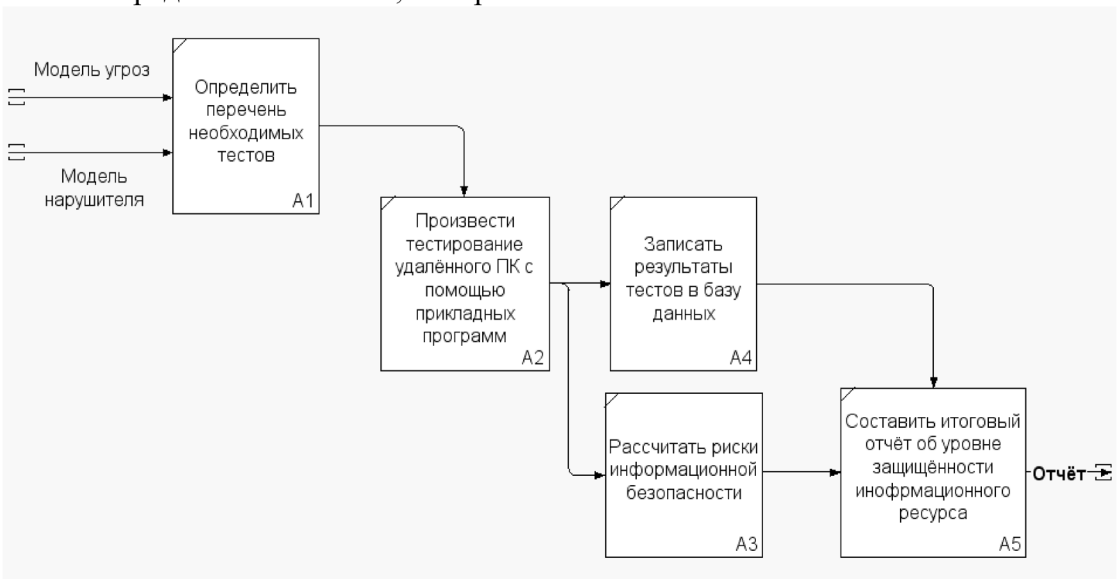


Рис. 2. Модель IDEF1 определения уровня защищённости ресурса

Основными шагами данного алгоритма являются:

1) Разрабатывается модель угроз информационной системы предприятия в соответствии с руководящими документами и методическими указаниями. На основе полученной модели угроз составляется перечень тестов, проверяющих компьютеры информационной системы на уязвимость к наиболее актуальным угрозам ИБ;

2) Производится тестирование рабочих станций с помощью специальных программных средств;

3) По формулам (1)-(5) проводятся вычисления количественной оценки состояния защищённости информационных

ресурсов;

4) Составляется итоговый отчёт, содержащий качественную оценку защищённости информационных ресурсов.

Структурная схема системы оценки уровня защищённости представлена на рис. 3. Администратор выступает в роли инициатора процесса тестирования, направленного на обнаружение уязвимостей. В своей работе он использует прикладные программы и утилиты, предлагаемые Kali Linux, такие как Nmap, Nessus, Metasploit Framework и другие. Объектом тестирования является элемент автоматизированной системы, выполняющий функции по обработке, хранению и передаче информации. Наиболее важными объектами тестирования являются операционная система ПК, запущенные службы, прикладные программы и непосредственно файлы, представляющие собой информационный ресурс. После проведения необходимых тестов программой оценки уровня защищённости выполняется процесс обработки результатов тестирования и занесения их в базу данных. В завершение своей работы программа формирует итоговый отчёт, который позже изучается администратором безопасности.

Рис. 3. Структурная схема системы оценки уровня защищённости информационных ресурсов

В качестве средства проведения тестирования был выбран Kali Linux. Достоинствами данного дистрибутива являются его свободное распространение, большой (около 600+) набор утилит, отличная документированность. С целью автоматизации процесса было разработано программное средство «Оценка защищённости информационного ресурса» [5], реализующее функции обработки результатов тестирования, расчёта показателей уязвимости персональных компьютеров и рисков безопасности ресурсов, занесение полученных значений в базу данных. Программа предназначена для работы в ОС Kali Linux версии 1.0.9 с графической оболочкой KDE версии 3.0 и выше. На заключительном этапе программа формирует отчёт, который содержит: список протестированных компьютеров с указанием имён и IP-адресов, результаты тестов для каждого проверенного ПК, оценку защищённости для каждого информационного ресурса, указанного пользователем в качестве входных данных программы, список обнаруженных уязвимостей и рекомендации по их устранению.

На рис. 4 представлен результат проведения оценки уровня защищённости 3 информационных ресурсов, расположенных на двух разных компьютерах. Каждый ресурс расположен на компьютере в полном объёме и является единственным

экземпляром ресурса в информационной системе. Исходные данные для расчёта уровня защищённости ИР представлены в таблице.

Табл.1

Тесты на атаки, наиболее актуальные для данной ИС

Условное обозначение	Весовой коэффициент теста	Проверяемая уязвимость
A1	0.25	Сетевая разведка
A2	1	Перехват сетевого трафика
A3	0.5	Уязвимость в протоколе RDP Windows
A4	0.25	Слабые пароли
A5	1	Получение удалённого доступа

В результате тестирования было успешно реализованы тесты A1, A3, A5 для первого ПК, и только тест A1 для второго. На рис. 6 представлены сформированные отчеты результатов оценки защищенности ИР. Выявлен высокий показатель уязвимости первого компьютера вследствие успешного теста A5 на получение удаленного доступа. С учетом значительной стоимости ресурса $\boldsymbol{R}1$ определен её низкий уровень защищенности. Качественная оценка уровня защищенности определяется на основе сравнения введенных пользователем пороговых значений информационного риска для организации с рассчитанными значениями программой.

Оценка защищённости получена для следующих ресурсов:

Наименование ресурса	Дата последнего тестирования	Установленная ценность	Задействованные ПК	Риск	Состояние защищённости
Лицензионные ключи Microsoft Office	2016-08-14	15000	<10.0.0.11> ;	8700	Неудовлетворительное
Цифровые сертификаты	2016-08-14	20000	<10.0.0.12> ;	1600	Удовлетворительное
Отчёты о работе подразделения	2016-08-14	2500	<10.0.0.11> ;	1450	Удовлетворительное

В процессе оценки были протестированы следующие компьютеры:

IP-адрес компьтера: 10.0.0.11

Показатель уязвимости для данного ПК: 0.58

При количестве выполненных тестов равном 5

В процессе тестирования на данном ПК были выявлены следующие уязвимости:

Наименование обнаруженной уязвимости	Описание уязвимости	Рекомендации по устранению
Отсутствует защита от сетевой разведки	Наличие подобного рода уязвимости не наносит системе ущерба как такового, но в то же время является отправной точкой для нарушителя, предоставляя ему всю необходимую информацию о вашей сети, такую как IP-адреса и доменные имена компьютеров, версии используемых Операционных систем, используемое на компьютерах программное обеспечение и открытые порты. Лишив злоумышленника информации можно предотвратить множество потенциальных атак.	Возможные меры: 1) Установка межсетевого экрана; 2) Настройка системы обнаружения вторжений с целью выявлять наличие множественных ping-запросов и ICMP-пакетов внутри сети; 3) Запрет ICMP-сообщений
Получение злоумышленником удалённого доступа через вредоносное ПО	Пользователи часто запускают на своём ПК посторонние файлы, не связанные с их рабочей деятельностью. Однажды такой файл может оказаться вирусом, предоставляющим нарушителю доступ к системе с правами пользователя, запустившего файл.	Ограничить привилегии пользователя, оставив ему возможность запускать только те исполняемые файлы, которые необходимы ему для осуществления должностных обязанностей. Также возможно стоит удостовериться, что антивирусное ПО активировано и его вирусные базы обновлены до актуальной версии.
Уязвимость в протоколе RDP Windows	Использование службы Windows для организации удалённого доступа к рабочему столу не является безопасным. Если данная служба запущена, злоумышленник имеет возможность вывести компьютер из строя и вызвать принудительную перезагрузку системы	Отключите службу Windows, предоставляющую доступ к удалённому рабочему столу. Панель управления -> Система -> Настройка удалённого доступа -> Запретить. Используйте для удалённого доступа более надёжные средства (например, AmmyAdmin или teamViewer)

Рис. 4. Отчёт оценки уровня защищенности ИР

Таким образом, разработанная методика позволяет определить уровень защищённости каждого конкретного информационного ресурса, зарегистрированного в автоматизированной системе, исходя из значения рисков информационной безопасности, полученных в процессе обработки результатов тестов, выполненных в рамках проведения тестирования на проникновение. Применение разработанной методики должно существенно повысить достоверность оценки уровня защищённости информационных ресурсов предприятия за счет выявления уязвимостей, которые не были определены в результате традиционного аудита безопасности ИС.

Литература

1. «Главные уязвимости корпоративных информационных систем в 2014 году: веб-приложения, пароли и сотрудники». Статья. // Блог команды Positive Technologies – Режим доступа: http://habrahabr.ru/company/pt/blog/255681/, яз. русский;

2 Осовецкий Л. Оценка защищённости сетей и систем/ Л. Осовецкий, В. Шевченко// Экспресс электроника. 2002. № 2-3. С. 20-24;

3 Мукминов В.А.Методика оценки реального уровня защищённости автоматизированных систем/ В.А. Мукминов, А.В. Лобузько, В.М. Хуцишвили // Международный журнал «Программные продукты и системы», Выпуск №1 – 2012 г. – с. 39-42;

4 ГОСТ Р 51275-99 «Защита информации. Объект информатизации. Факторы, воздействующие на информацию. Общие положения»;

5 Свидетельство о регистрации электронного ресурса «Оценка защищённости информационного ресурса» - Оренбургский государственный университет – 23.06.2016 – рег. номер 1299.

6. Penetration Testing: Assessing Your Overall Security Before Attackers Do. SANS Institute InfoSec Reading Room. June 2006. https://www.sans.org/reading-room/whitepapers/testing/. Accessed on July 3, 2015.

Оренбургский государственный университет
Orenburg State University

ASSESSMENT LEVEL SECURITY OF INFORMATION RESOURCES BASED ON THE PENETRATION TEST

R.R. Galimov, V.P. Chlenov

The technique of information resources security assessment on the basis of the test attacks using software Kali Linux

Key words: assessment of the level of security, penetration tests

УДК 004.056.57

ИССЛЕДОВАНИЕ СТРУКТУРНО-ФУНКЦИОНАЛЬНОЙ СХЕМЫ СОЦИАЛЬНОЙ СЕТИ ДЛЯ ОБЩЕНИЯ GOOGLE PLUS

В.А. Колесников, А.Е. Дешина, Е. Ружицкий

В данной работе производится анализ статистических данных сети для общения Google Plus. Были выведены и представлены практически полезные результаты

Ключевые слова: социальная сеть, граф

Google Plus — социальная сеть для общения в Интернет от компании Google. Основополагающими принципами действия сервиса являются: пользователи, приватность и живое общение внутри сообществ [5]. Пользователи могут проводить ознакомление с контентом, анонимно и открыто.

Аудитория данной сети представима в виде трёх групп пользователей: модераторы, авторизованные и неавторизованные пользователи. Весь контент, распространяющийся в данной социальной сети можно классифицировать на положительный и негативный[4-5]. Также его можно классифицировать по содержанию, по представлению и по критерию безопасности просмотра и использования. Действия, доступные пользователям, были разделены на три основные категории: размещение, ознакомление и реагирование [2-3].

С учетом данных классификаций предоставляется возможным построить структурно-функциональная модель социальной сети Google Plus с учетом всех ее особенностей.

Для дальнейшего анализа сети Google Plus необходимо воспользоваться открытой базой данных инциденции вершин и дуг данной сети.

$$\Gamma(x_i, a_{ij}, x_j) \Leftrightarrow \Gamma(i, \delta(a_{ij}), j), \quad (1)$$

где i и j – номера вершин x_i и x_j в сети, $\delta(a_{ij})$ – вес дуги a_{ij}, связывающей x_i и x_j, и направленной от i к j. Под весом дуги подразумевается динамический ресурс (2):

$$\delta(a_{ij}) = \frac{d[CV]}{dt} = \langle C \rangle V', \quad (2)$$

т.е. передачу определенного объема V и ценности $\langle C \rangle$ наполнителя сети в единицу времени. Усредненная ценность $\langle C \rangle$ связывается с популярностью и другими параметрами пользователей i и j, V' относится к интенсивности обмена контентом.

Такой формат позволяет построить звездную матрицу, элементы строки которой соответствуют дугам, входящим в данную вершину, а элементы столбца – дугам, исходящим из вершины.

Далее определяется степень исхода для каждого узла социальной сети.

Для последующей репрезентативной выборки из мультиразмерной сети необходимо знать удельный вес ее вершин и дуг.

Для этого предлагается нормировка их весов по сумме весов всех дуг сети (3):

$$\sum_{\substack{i,j \\ i \neq j}} \delta(a_{ij}), \quad (3)$$

т.е. – по суммарному трафику сети. Тогда нормированная величина, в данном случае (4):

$$\delta(\bar{a}_{ij}) = \delta(a_{ij}) / \sum_{\substack{i,j \\ i \neq j}} \delta(a_{ij}) \quad (4)$$

будет показывать удельный вес трафика в дуге a_{ij} к суммарному трафику сети.

Она и будет характеризовать степень ее взвешенной (по трафику) центральности.

Следует заметить, что суммарный трафик сети не разделяет входящие и исходящие дуги. Поэтому для определения взвешенной центральности вершины x_s можно использовать сумму (5):

Колесников Владислав Александрович– ВГТУ, студент, e-mail: mnac@comch.ru
Дешина Анна Евгеньевна – ВГТУ, ст. преподаватель, e-mail: mnac@comch.ru
Ружицкий Евгений – Пан-Европейский Университет (Словакия), к.т.н., декан, доцент,
e-mail: eugen.ruzicky@paneurouni.com

$$\sum_i \delta(a_{si}) + \sum_j \delta(a_{js}), \quad (5)$$

которую далее следует пронормировать по суммарному трафику сети.

В результате получим нормированную величину (6):

$$\delta(\bar{x}_s) = \left[\sum_i \delta(a_{si}) + \sum_j \delta(a_{js})\right] / \sum_{\substack{i,j \\ i \neq j}} \delta(a_{ij}), \quad (6)$$

которая будет характеризовать удельный вес трафика, проходящего через вершину x_s, по отношению ко всему трафику сети. В итоге, следуя формуле (6) получим матрицу взвешенной центральности элементов исходной сети (табл. 1).

Табл. 1

Матрица взвешенной центральности

Номер вершины	2	3	…	2192	2193	N
2	0	0,023	…	0	0,04	…
3	0,088	0,032	…	0	0	…
4	0	0,036	…	0	0,056	…
…	…	…	…	…	…	…
2190	0	0	…	0	0	…
2191	0,045	0	…	0	0	…
N	…	…	…	…	…	…

Последний этап алгоритма преобразования исходных данных сети заключается в нахождении и построении диагональной матрицы удельного баланса трафика в вершинах социальной сети Google Plus.

На следующем этапе просуммируем починные значения для получения необходимого результата (точности модели). В нашем случае до значения 0.95, так как для нас допустима 5% потеря трафика, полученная сумма должна быть не меньше 0,95. Этим критерием и ограничится репрезентативная (с точки зрения трафика) выборка:

$$\delta(\bar{x}_s) = 0{,}088 + 0{,}086 + \cdots + 0{,}0042 + \cdots + 0{,}063457 = 0{,}9519$$

На основе новых данных, с помощью программного обеспечения Gephi [6] построим усеченный граф социальной сети (рис. 1).

Стоит отметить, что у данной социальной сети явная кластерность не прослеживается. Из рис. 2 видно, что распределения генеральной и выборочной совокупности имеют схожую структуру. Основанием проверки будет принадлежность генеральной совокупности и полученной выборки к одному степенному закону с функцией плотности распределения

$$\varphi(x) = \alpha x^{-\beta}, \quad x \geq 1, \ \alpha > 0, \beta > 0.$$

На основе данных выборки методом максимального правдоподобия получены оценки параметров степенного распределения: $\alpha = 0.4975$, $\beta = 1.385$. В

результате функция плотности распределения примет вид:

$$\varphi(x) = 0.4975x^{-1.385}. \quad (7)$$

Для доказательства подобия воспользуемся критерием Пирсона.

Выдвинем нулевую гипотезу H_0- генеральная совокупность распределена по степенному закону с функцией плотности распределения $\varphi(x) = 0.4975x^{-1.385}$. Тогда конкурирующая гипотеза H_1 – генеральная совокупность не распределена по данному закону.

Значение критерия будем вычислять согласно формуле (8):

$$\chi^2_{\text{набл}} = \sum_{i=1}^{k} \frac{(n_i - n_i{}')^2}{n_i{}'}, \quad (8)$$

где n – объем выборки, n_i - вероятность попадания случайной величины в $n_i{}'$ - теоретическая вероятность попадания в рассматриваемый интервал.

Рис. 1. Выборка социальной сети Google Plus

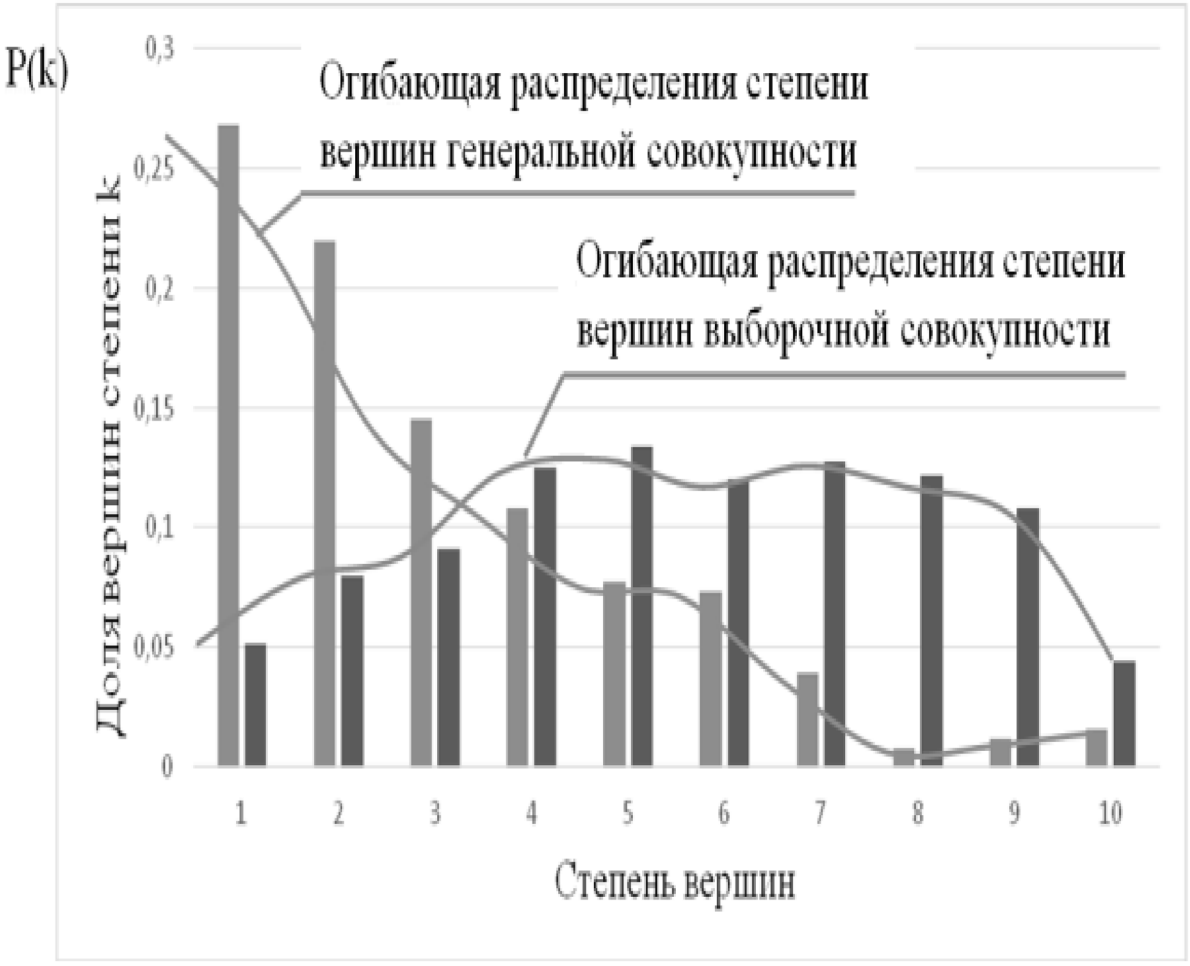


Рис. 2. Распределение количества вершин по степеням

Тогда следуя формуле (8), рассчитаем значение $\chi^2_{\text{набл}}$.

Для рассматриваемого случая, значение χ^2 для гипотезы H_0: $\chi^2_{\text{набл}} = 14.12$. Критическое значение коэффициента Пирсона для уровня значимости $\alpha = 0{,}05$ и $\nu = 10$ – степеней свободы равно $\chi^2{}_{\text{кр}} = 18.1324$.Так как $\chi^2_{\text{набл}} < \chi^2{}_{\text{кр}}$,то нет оснований отвергать нулевую гипотезу. Наблюдаемое значение χ^2 для гипотез не попадает в критическую область. Таким образом, генеральная и выборочная совокупность распределена согласно степенному закону (7).

Найдем значения среднеквадратического отклонения σ_{XY} двух совокупностей X и Y, $cov(X,Y)$ – коэффициент ковариации совокупностей, r_{XY} – коэффициент корреляции[7].

Рассчитанные данные представим в табл. 2.

Табл. 2

Сравнение параметров распределения степеней исходного графа с выборочными данными

Коэффициент	Значение
СКО	0,06882
Коэффициент корреляции	0,84778

При полученном значении коэффициента корреляции значение СКО является допустимым. Следовательно, выборка репрезентативная [1-3].

Таким образом, после анализа и систематизации исходных статистических данных социальной сети для общения Google Plus, удалось построить структурно-функциональную модель, учитывающую особенности рассматриваемой социальной сети, визуализировать модель усеченной сети, получить структурированные данные, описывающие распространение трафика внутри сети.

Литература

1. Flood-attacks within the hypertext information transfer protocol: damage assessment and management / A.G. Ostapenko, M.V. Bursa, G.A. Ostapenko, D.O. Butrik // Biosciences Biotechnology Research Asia. – 2014. – Vol. 11 (Spl.End). – P. 173-176.

2. Freeman L. C. The Development of Social Network Analysis / L.C. Freeman//Empirical Press. – 2004. – 30 p.

3. Optimization of expert methods used to analyze information security risk in modern wireless networks information analyze / S.A. Ermakov, A.S. Zavorykin, N.S. Kolenbet, A.G. Ostapenko, A.O Kalashnikov // Life Science Journal. – 2014. – № 11(10s). – P. 511-514.

4. Распространение нежелательной информации в социальных сетях Интернета /Абрамов К. Г., К.Г. Абрамов, Ю.М. // Диссертация – С.45-48.

5. Социальная сеть Google Plus. - Электрон. Дан. – Режим доступа: http://google.com

6. Средство визуализации данных. - Электрон. Дан. – Режим доступа: https://gephi.org

7. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб. пособие. – М.: Издательство Юрайт, 2010. – 479 с.

8. Maxwell, A. Pretzlav Last.fm Explorer: An Interactive Visualization of Hierarchical Time-Series Data / A. Maxwell // – 2008. – 11p.

9. Konstas, I. On Social Networks and Collaborative Recommendation / I. Konstas // – 2012. – 5p.

10. Byrd K. War with many unknowns / K. Byrd//Computerra. – M.: 2009. – No.

11. Grinyaev, S. Russia in global information society: threats, risks and possible ways of their neutralization / S. Grinyaev, – Electron. it is given. – Access mode:http://www.noravank.am/upload/pdf/419_ru.pdf.

12. Liu X., Tse C. K., Small M. Complex network structure of musical compositions: Algorithmic generation of appealing music, (2010). – P. 126–132.

13. Johnson, S. Entropic origin of disassortativity in complex networks / S. Johnson, J.J. Torres, M.A. Muñoz / Physical Review Letters. – 2010. – 4 p.

14. Филяк, П.Ю. Информационная безопасность и комплексная система безопасности: анализ, подходы//Информация и безопасность. 2016. Т. 19. №. 1 С. 72-79.

15. Гришина, Н.В. Организация комплексной системы защиты информации. – М.: Гелиос АРВ, 2007. – 256 с.,ил.

16. Остапенко, А.Г., Ермилов Е.В., Калашников А.О. Построение функций ущерба и риска для компьютерных атак, приводящих к нарушению доступности к информации//Информация и безопасность. 2013. Т. 16, № 2. – с. 207 – 210.

17. Остапенко А.Г., Ермилов Е.В., Калашников А.О. Риски ущербности, шансы полезности и жизнестойкость компонент автоматизированных систем в условиях воздействия на них информационных угроз//Информация и безопасность. 2013. Т. 16, № 2. – с. 215 – 218.

18. Остапенко, А.Г., Шершень А.Н., Калашников А.О. Концептуальный подход к расчету и регулированию рисков нарушения актуальности информации в элементах критической информационной инфраструктуры //Информация и безопасность. 2013. Т. 16, № 2. – с. 239 – 242.

19. Вишняков, Я.Д. Общая теория рисков / Я.Д. Вишняков, Н.Н. Радаев – 2-е изд., испр. – М.: Издательский центр «Академия», 2008. – 368 с.

20. Чернов, Г.В., Кудрявцев А.А. Управление рисками. - М.: ТК Велби, Изд-во Проспект, 2007. – 160 с.

21. Астахов, А.М. Искусство управления информационными рисками. – М.:ДМК Пресс, 2010 – 312 с., ил.

22. Филяк, П.Ю. Теория аналитики//Материалы первой всероссийской конференции «Аналитика

развития и безопасности страны: реалии и перспективы» - М.: ООО «Агентство печати «Столица», 2014. – С. 213-225.

23. Филяк, П.Ю., Федирко С.Н. Обеспечение информационной безопасности с помощью технологии управления знаниями «Brain»//Информация и безопасность. 2016. Т. 19. № 2. С. 238-243.

24. Филяк, П.Ю., Мишарин Г.Д., Уразов О.М., Золотарев В.В. Обеспечение информационной безопасности организации методами моделирования//Информация и безопасность. 2015. Т. 18. № 4. С. 560-563.

25. Михеев, В.А., Шевырев А.В., Шаламова Н.Г., Федотова М.А. Визуальное мышление в аналитике: Проблемы, возможные подходы и способы овладения//Материалы первой всероссийской конференции «Аналитика развития и безопасности страны: реалии и перспективы» - М.: ООО «Агентство печати «Столица», 2014. – С. 260-269.

26. Тузовский, А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии) / Под общ.ред. В.З. Ямпольского. – Томск: Изд-во НТЛ, 2005. – 260 с.

27. Гаврилова, Т.А., Хорошевский Ф.В. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. – 384с.

28. NEO4J. Официальный сайт Neo4j. [Электронный ресурс] URL: https://neo4j.com (Дата обращения: 06.01.2017).

29. THEBRAIN. Официальный сайт BRAIN. [Электронный ресурс] URL: http://www.thebrain.com (Дата обращения 07.01.2017).

30. Joeran Beel and Bela Gipp, Academic search engine spam and google scholar’s resilience against it. Journal of Electronic Publishing, 13(3), 2010. – P. 91–112.

31. Fabrício Benevenuto, Tiago Rodrigues, Virgílio Almeida, Jussara Almeida and Marcos Gonçalves. Detecting Spammers and Content Promoters in Online Video Social Networks. In ACM SIGIR Conference, Boston, MA, USA, 2009. – P. 61–71.

32. Konstas, I. On Social Networks and Collaborative Recommendation / I. Konstas // – 2012. – 5p.

33. Byrd K. War with many unknowns / K. Byrd//Computerra. – M.: 2009. – No.

34. Grinyaev, S. Russia in global information society: threats, risks and possible ways of their neutralization / S. Grinyaev, – Electron. it is given. – Access mode:http://www.noravank.am/upload/pdf/419_ru.pdf.

ФГБОУ ВО «Воронежский государственный технический университет»
Voronezh State Technical University
Пан-Европейский Университет
Pan-European University

STUDY OF THE STRUCTURAL-FUNCTIONAL SCHEME OF SOCIAL NETWORK FOR COMMUNICATION GOOGLEPLUS

V.A. Kolesnikov, A.E. Deshina, E. Ruzicky

In this paper, an analysis of the statistical data of the network for communicating GooglePlus is made. Practically useful results were withdrawn and presented

Key words: social network, graph

УДК 004.056.74

АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ВЗАИМОДЕЙСТВИЯ РАЗЛИЧНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ В ОБЩЕМ ИНФОРМАЦИОННОМ ПРОСТРАНСТВЕ

В. А. Кургузкин, А. В. Паринов, Д.Г. Плотников, Ю. Штефанович

В статье описаны общие способы обеспечения взаимодействия различных информационных систем. Кроме того, в работе представлены основные характеристики подобных алгоритмов и даны примеры

Ключевые слова: информационные системы, взаимодействия, диффузия

Известны исследования [1, 2] информационной диффузии в социальных сетях, в которых было установлено, как она происходит в различного рода информационных сетях и при каких условиях начинается информационная эпидемия. При этом даже внутренние кластеры, которые появлялись зачастую в неоднородных сетях, не рассматривались и не исследовались отдельно несмотря на то, что влияние узлов, которые входят в данные кластеры, и тех узлов, которые находятся на границах кластеров и входят одновременно в несколько кластеров, различное.

Исследование межсетевого взаимодействия с помощью моделирования в программном комплексе является очень актуальным направлением в связи с тем, что любая сеть сегодня (очень часто даже неинформационная) не может существовать по отдельности, в некотором вакууме. Любое масштабное предприятие сегодня представляет собой систему, которая бывает иногда столь сложной, что рациональнее рассматривать каждую из ее подсистем по отдельности. К примеру, добычу сырьевых ресурсов можно представить как совокупность, состоящую в первую очередь из сырьевой сети (СС), которую представляет непосредственно инфраструктура добычи сырьевого ресурса,

Кургузкин Владимир Александрович – ВГТУ, студент, e-mail: mnac@comch.ru
Паринов Александр Владимирович – ВГТУ, соискатель, e-mail: mnac@comch.ru
Плотников Денис Геннадьевич– ВГТУ, доцент, e-mail: mnac@comch.ru
Штефанович Юрий - Пан-Европейский Университет (Словакия), к.т.н., зам. декана, e-mail: juraj.stefanovic@paneurouni.com

а также ресурсы по добыче (в том числе человеческий). Кроме того, данная система тесно связана с транспортной сетью (ТС), которая также входит в данную совокупность. В нее входит транспортная инфраструктура и другие элементы, связанные с транспортом, а ее основной целью в контексте данной совокупности является сообщение сырьевого ресурса между остальными участниками предприятия по добыче сырья.

Кроме того, имеет место некоторый промышленный сегмент, занимающийся обработкой сырьевого продукта, перегонкой его в продукт потребления, который можно уже непосредственно реализовать. Назовем этот компонент промышленной сетью (ПС). Предварительно данные подсистемы схематично показаны на рис. 1. Конечно, это лишь первое приближение. На рис. 1 упущена крупная подсистема, которая пронизывает все остальные и постоянно с ними сообщается.

Это финансовая сеть, ведь изначально без нее невозможно начать добычу, обработку и транспортировку сырья, все финансовые процессы связаны с той или иной подсистемами и обеспечивают их стабильную работу и связность.

Можно модифицировать предложенную схему, добавив в нее финансовую подсистему (рис. 2).

На представленной схеме (рис. 3) стрелками показаны сообщения между сетями, которые означают наличие некоторых узлов, одновременно принадлежащих и той, и другой системе.

При наличии таких узлов становится понятно, что их бездействие может оказать

в некоторой степени иное влияние, чем бездействие узла из одной только подсети. Важным моментом в определении поведения системы, состоящей из смежных систем, является обнаружение пограничных узлов и определение степени их влияние на оба смежных кластера.

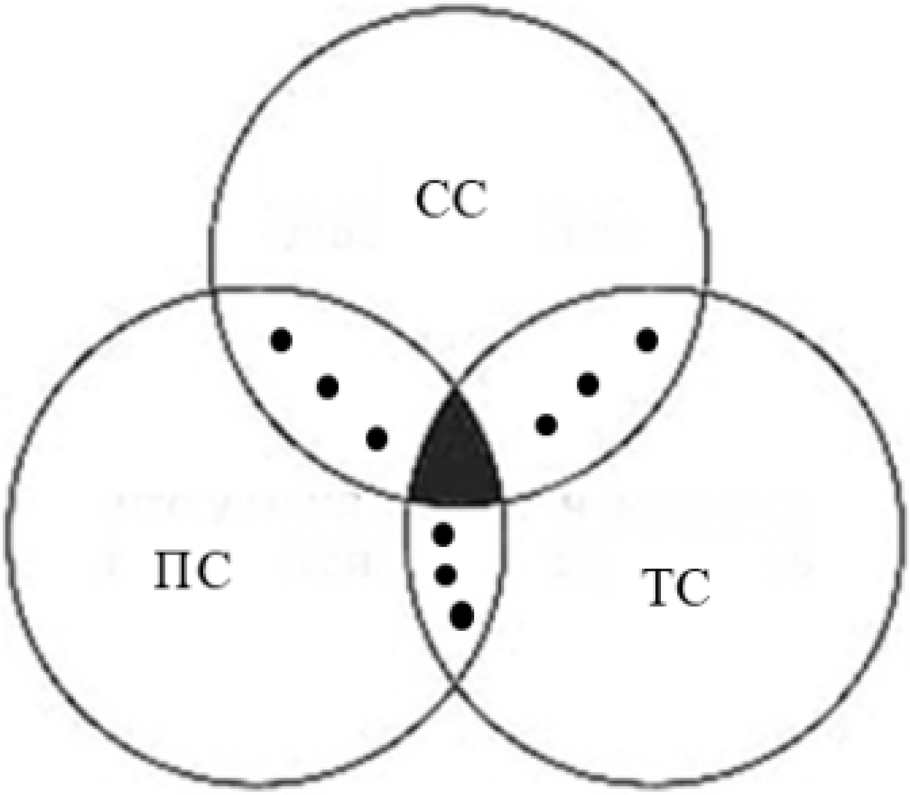


Рис. 1. Простейшая схема сырьевого комплекса

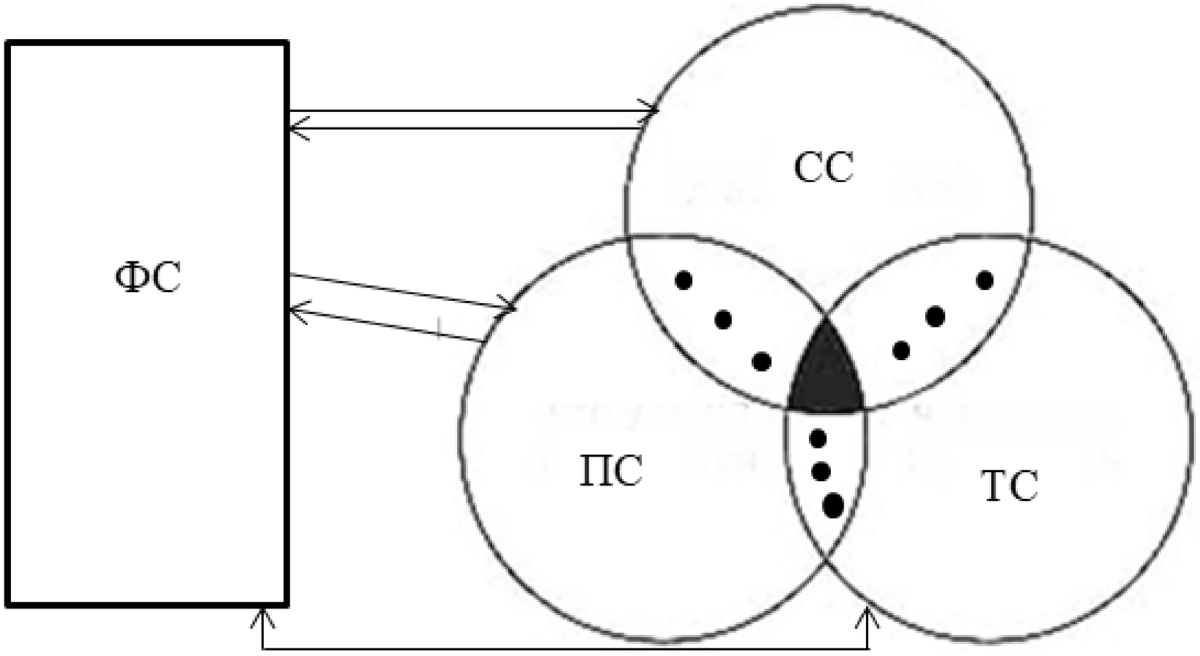


Рис. 2. Модифицированная схема комплекса

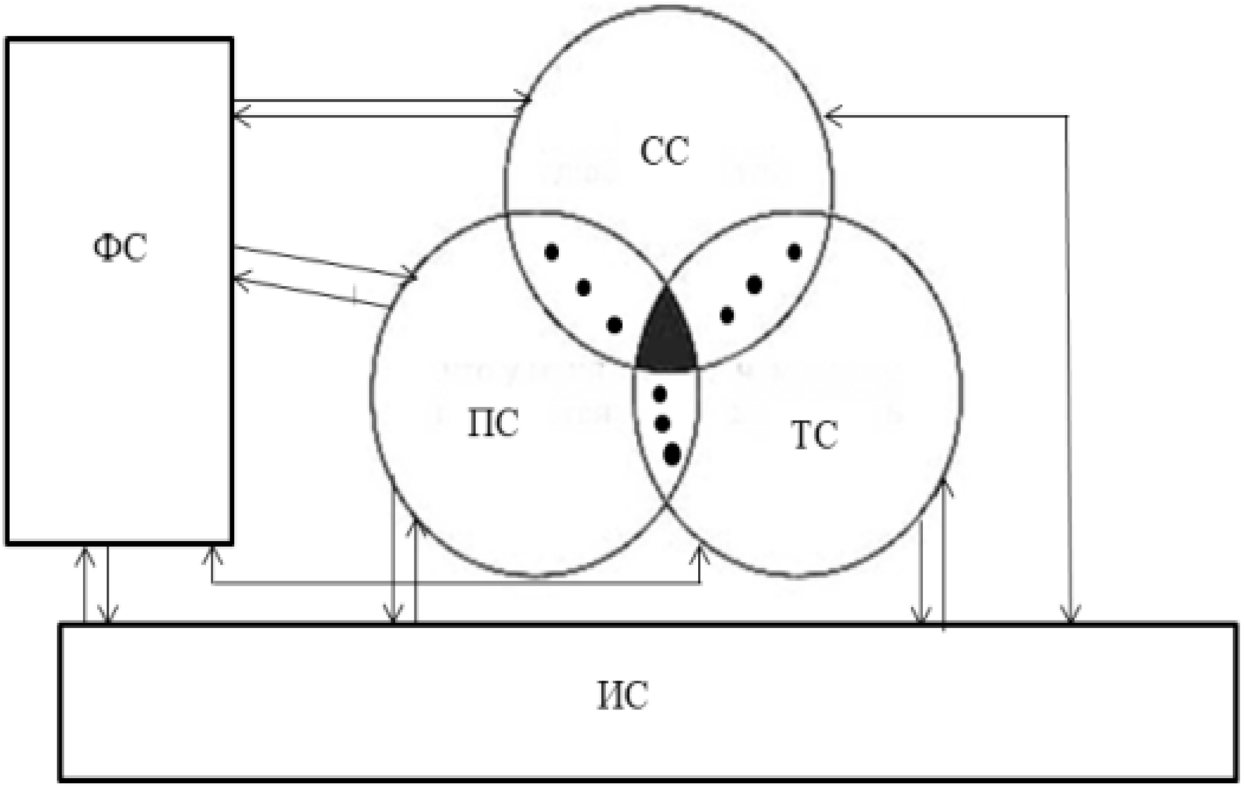


Рис. 3. Полная схема структуры комплекса

Далее рассматриваются только информационные сети в силу их присутствия практически во всех сферах жизни человека. Но по сути своей сети любого рода представляют одно и то же, так как в них есть узлы, генерирующие, получающие и хранящие наполнитель, а также ребра, соединяющие эти узлы, по которым наполнитель передается.

При рассмотрении различных видов систем меняется только вид наполнителя. Для того чтобы привести к некоторому обобщению разнородные системы, возможно объявить некоторый эквивалент ценности наполнителя сети и его объема, а также - присвоить соответствующие индексы узлам и ребрам.

После этого можно ориентироваться в данных сетях только в общих условных единицах.

Проблемной частью (при наличии модели для исследования распространения вредоносного контента) может оказаться поиск самих сообществ в общем информационном пространстве.

Для решения подобной проблемы можно воспользоваться алгоритмом, приведенным в работе [3]. В частности в алгоритме [3] учитываются сообщества, возникающие из социальных сетей в Интернете, они наблюдают множество случаев, когда появляется крупная и разнообразная структура сообщества.

Ребра, соединяющие эти отдельно кластеризованные группы узлов, которые называются граничными узлами, называются также граничными.

Далее реализуется алгоритм измерения близости граничного узла.

Этот алгоритм можно модифицировать для случая анализа взаимодействующих сетей.

Цель алгоритма состоит в том, чтобы иметь возможность ранжировать узлы в сети в соответствии с их способностью влиять на другие узлы в разных сообществах с помощью информации (контента), которой они обмениваются.

Это покажет пограничные узлы, которые играют ключевую роль в обмене информацией между различными

сообществами, и окружающие их узлы (узлы в их окрестностях).

Алгоритм основан на предпосылке, что информация перемещается, случайно блуждая[3], а не по кратчайшему пути, что вполне естественно.

Предложенный алгоритм[3] выполняет итерации через множество граничных узлов и выполняет набор независимых усеченных случайных ходов, начинающихся в каждом пограничном узле, до тех пор, пока не будет достигнута конвергенция в распределении посещений узлов в окрестности каждого граничного узла.

Это подсчет случайных переходов на пограничных узлах и узлах того же сообщества в окрестностях данных узлов, который позволяет ранжировать с точки зрения возможности влиять на другое сообщество посредством распространения контента. Описание алгоритма в шагах с 1 по 4 приведены в табл. 1. На первом шаге алгоритма набор связанных графов извлекается из исходного сетевого графа, G, используя первый поиск по ширине (ППШ). ППШ может выполняться с линейным временем по числу ребер и вершин, O (N + M), и сложности по используемой памяти, линейной по отношению к N. Учитывая, что большинство социальных сетей будут сетями «малого мира» и безмасштабными, количество ребер не будет расти слишком быстро при росте числа вершин. Относительно небольшое подмножество высокостепенных узлов отвечает за соединения в сети. Распределение степеней вершин по степенному закону [4] означает, что большинство узлов будет иметь небольшое количество ребер.

Второй шаг – маркировать каждый узел в соответствии с сообществом, к которому он принадлежит, и может быть выполнен в пределах O (N).

В работе [3] исследователи используют метод Лувена [5]. Время выполнения такого алгоритма равно Ω (NlogN), и доступна эффективная его реализация.

Тестирование выполняется с использованием этого метода, работающего с миллионами узлов в течение 2 минут на стандартном ПК.

Другие алгоритмы для обнаружения сообществ были опробованы и дали аналогичные результаты в наборах данных, где разделение сообществ было четким. В ситуациях, когда различные подходы производят различные маркировки для сообществ, алгоритм продолжает фокусироваться на другом наборе узлов, который, вероятно, будет иметь большое количество наложений результатов.

Третий шаг выделяет набор граничных узлов, которые управляют обменом информацией между разными сообществами, поскольку они являются единственными узлами, которые непосредственно соединяются с узлами в другом сообществе[3].

Табл. 1

Алгоритм определения удаленности граничного узла

1	Извлечь множество связных графов из исходного графа.
2	Для каждого связанного компонента получить маркировку сообщества графа.
3	Получить множество граничных узлов.
4	Измерить локальную окрестность каждого из граничных узлов при помощи метода случайного блуждания с заданной длиной и собрать все значения для графа в некоторое нормированное значение.

Заключительный шаг по табл. 1 запускает ряд так называемых случайных проходов из каждого пограничного узла до тех пор, пока критерий пересечений не будет удовлетворен на основании количества посещений узлов в сети. Число посещений каждого узла подсчитывается и является мерой способности распространять информацию через граничные границы и влиять на различные сообщества. Шаги 3 и 4 более подробно описаны в следующем подразделе. Этот алгоритм можно назвать алгоритмом граничной близости [3].

Для определения граничных узлов / ребер связного графа уместно использовать вектор маркировок сообществ на множестве

узлов в сети C и поиск соседних узлов с отличными ярлыками сообщества.

В списке ребер L для матрицы смежности данного графа каждое ребро представляется как номер строки и номер столбца в матрице.

Если две метки различаются по строкам относительно этого списка ребер, значит, обнаружено граничное ребро [3].

Таким образом, алгоритм представляется интересным с точки зрения применения его для создания программного обеспечения, для поиска смежных подсетей в крупной сети или исследования взаимодействия сетей в глобальном информационном пространстве. Определенные шаги, которые показывают значимость пограничных узлов могут быть использованы при оценке трафика, протекающего через данные узлы. Эта характеристика сможет показать популярность данного пограничного узла и отразит возможность протекания вредоносной информации через данный узел в смежные сети.

Вполне возможно реализовать (подобно вышеописанным процедурам) некоторые структуры данных, в которых будут храниться пограничные узлы.

При попадании в них информации, составляющей вредоносный контент, будут запускаться механизмы расчета вероятности заражения, а после при распространении информации вероятности передачи информации в одну либо в другую сеть.

При моделировании могут использоваться 3 структуры данных для двух сетей. Две из них - для вершин каждой из исследуемых сетей, а еще одна – для вершин пограничной области. Алгоритм исследования одиночной сети требует адаптации и переработки для двух сетей или сообществ в общем информационном пространстве. Пересмотра также требует и форма входных данных алгоритма, в которой представляются рассматриваемые информационные сети.

Имеющаяся модель не отвечает необходимым требованиям для исследования взаимодействия вершин сети, находящихся одновременно в нескольких информационных сетях.

Трехместный предикат необходимо развить до пятиместного, состоящий из двух вершин, соединенных ребром, его весовой характеристики и двух элементов, описывающих принадлежность каждой вершины к конкретной сети: $\Gamma(x_i, x_j, \delta(a_{ij}), \{n_i\}, \{n_j\})$, (1) где i и j – номера вершин x_i и x_j в сети; $\delta(a_{ij})$ - вес дуги a_{ij}, связывающей x_i и x_j, и направленной от i к j; n_i – сети, к которым принадлежит вершина x_i, n_j –сеты, к которым принадлежит вершинаx_j. Трехместный предикат (1) можно представить в виде, показанном на рис. 4.

x_i	x_j	$\delta(a_{ij})$	$\{n_i\}$	$\{n_j\}$
x_k	x_l	$\delta(a_{kl})$	n_t, n_y	n_y
⋮	⋮	⋮	⋮	⋮
x_{k+s}	x_{l+r}	$\delta(a_{k+s,l+r})$	n_t, n_y, n_u	n_o, n_u
⋮	⋮	⋮	⋮	⋮
x_n	x_m	$\delta(a_{nm})$	n_p	n_t

Рис. 4. Модифицированный предикат сети

Здесь n_t, n_y, n_u, n_p, n_t, n_o – примеры различных информационных сетей, участниками которых являются вершины x_i и x_j, индексы обозначают условные обозначения соответствующих сетей.

Вышеприведённая модификация алгоритма для одиночной модели, а также трехместного предиката, используемого в исследованиях информационных сетей, могут быть эффективно использованы при последующих исследованиях процессов распространения вредоносной информации в социальных и других сетях, а также - при модификации программного обеспечения «NetEpidemic» для моделирования информационной диффузии в смежных сетях.

Литература

1. Woo J. Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog / J. Woo, H. Chen // Graduate School of Information Security, Korea University, Anamro, Seoul, Korea. – 2016. – P. 19.

2. Cannarella J. Epidemical modeling of online social network dynamics / J. Cannarella, J.A. Spechler // Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA. – 2014. – P. 66.

3. Mantzaris A. V. Uncovering nodes that spread information between communities in social networks / A. V. Mantzaris // EPJ Data Science. – 2014.. – Vol.286. – P. 509–512.

4. Blondel V. D. Fast unfolding of communities in large networks / Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre // Journal of statistical mechanics, Theory and experiment. – 2008. – Vol. 2008

5. Kauai, HI: IEEE. Freeman, L. C. (1979). Centrality in social networks conceptual clarification. Social Network. 1(3). P. 215–239.

6. Antsupov, A. Ya. Conflictology: the textbook for higher education institutions / A.Ya. Antsupov, A. I. Shilov. – 3rd prod., reslave. and additional – SPb.: St. Petersburg, 2007. – 591 p.

7. Voldstad, R. Community Detection on Last.fm Artist Data / R. Voldstad // – 2014.

8. Panagiotis, Karampelas. Techniques and Tools for Designing an Online Social Network Platform. New Hampshire: Hellenic American University / – M.: World, 1992. – 400 p.

9. Paolo, Massa, Martino Salvetti, and DaniloTomasoni. Bowling alone and trust decline in social network sites. In Proc. Int. Conf. Dependable, Autonomic and Secure Computing, pages 658-663, 2016.

10. Tsvetovat, M. Social Network Analysis for network interests: Finding Connections on the Social Web / M. Tsvetovat, A. Kouznetsov // O'Reilly.-2011. — P. 45. — 192 c.

11. Newman, M. E. (2004). Coauthorship networks and patterns of Scientific Collaboration. Proceedings of the National Academy of Sciences, 101(suppl 1): 5200-5205.

12. Ahn, Y. Analysis of topological characteristics of huge onlane social networking services Advogato/ Y. Ahn, S. Han, H. Knak, S. Moon, H. Jeong // 16th International Conference on the World Wide Web. – 2007. – P. 835-844.

Воронежский научно-образовательный центр управления информационными рисками
Voronezh Science and Education Management Center Information risks
Пан-Европейский Университет
Pan-European University

ALGORITHMIC SUPPORT OF INTERACTION OF VARIOUS INFORMATIONAL SYSTEMS IN THE GENERAL INFORMATION SPACE

V.A. Kurguzkin, A.V. Parinov, D.G. Plotnikov, Yu. Stefanovic

In the article, general ways of ensuring the interaction of various information systems. In addition, the paper presents the main characteristics of such algorithms and gives examples

Key words: information systems, interactions, diffusion

УДК 004.5/004.021/003.63

ПРИМЕНЕНИЕ RAPID MINER И ОТКРЫТЫХ СРЕД КАК ИНСТРУМЕНТОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ ДЛЯ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ

П.Ю. Филяк, М.А. Виноградов

Рассматривается интеллектуальный анализ данных (DataMinig) с помощью информационно-аналитической системы RapidMiner, которая по сути является «опенсорсной» средой (OpenSource) – с открытым исходным кодом и, несмотря на то, что не является проприетарным программным обеспечением, предоставляет широкие возможности даже не искушенному пользователю и позволяет решать очень серьезные задачи при работе с «большими данными» (BigData)

Ключевые слова: анализ, аналитика, «большие данные» (BigData), интеллектуальный анализ данных (DataMinig), среда Orange, RapidMiner, универсальный язык программирования R, машинное обучение, GUI

В настоящее время решение проблем выработки и принятия правильных и эффективных управленческих решений равно, как и решение проблем обеспечения безопасности организации, не нуждаются в доказательстве его актуальности [1,2]. Совершенно очевидно, что принимать такие решения приходится в условиях информационной перегрузки и непрерывного оперирования большими объемами данных (*BigData)* [3]. Для этих целей разработан целый ряд программных продуктов, позволяющих эффективно осуществлять интеллектуальный анализ данных (*DataMinig)* [4], в частности, - «Семантический архив», «Галактика Zoom», «Web-Observer», «Intellectum.BIS», RCO «FactExtractor», «Астарта», «Тренд», «GetNews», «Невод», «ИСКРА», «Statistica», ИАС языке программировании Python, у него есть «Контур Стандарт», ИАС «PolyAnalyst» и другие. Большинство из этих продуктов являются проприетарными и предполагают распространение исключительно на коммерческих условиях и требуют пользователей соответствующей квалификации. В итоге, на данный момент существует множество организаций, у которых высокая стоимость продукта и проблема подготовки кадров в данной области, вынуждает руководство отказаться от идеи создания аналитической системы в организации. Но сегодня есть различные проекты в области аналитики: продукты для интеллектуального анализа данных и машинного обучения.Таких систем довольно много, но особое внимание стоит уделить основополагающим и условно-бесплатным программным продуктам, среди них стоит выделить:

- **Язык R** — универсальный язык программирования, ориентированный на статистическую обработку данных и высокоуровневую графику, он может быть расширен различными пакетами. При этом расширениями могут быть и пакеты, созданные пользователем.
- **Orange** — программный продукт с открытым исходным кодом, написанный на функции визуализации и анализа данных. Интеллектуальный анализ данных производится путем визуального программирования и с помощью Python - сценариев.
- **RapidMiner** — программный продукт, написанный на языке программировании Java, ранее он был известен как YALE, может использоваться для решения широкого спектра задач, таких как:
  - визуализация данных;
  - машинное обучение;
  - проведение экспериментов;
  - интеллектуальный анализ данных;

Филяк Петр Юрьевич – ФГБОУ ВПО СГУ, канд. техн. наук, доцент, e-mail: parallax-1@yandex.ru
Виноградов Михаил Александрович – ФГБОУ ВПО СГУ, студент, e-mail: arkaij@ro.ru

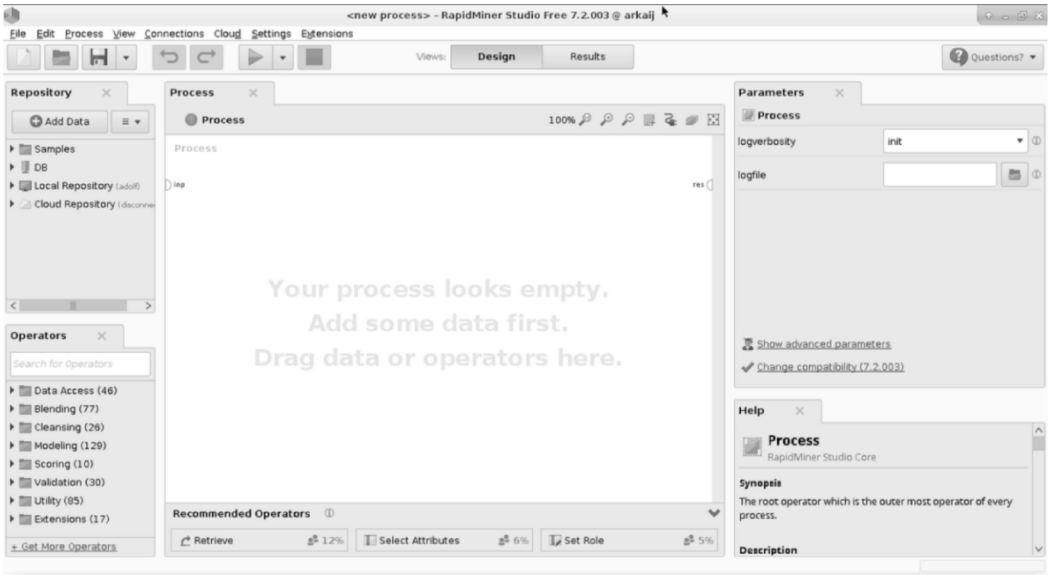


Рис. 1. ИнтерфейсRapidMinerв OS Linux

Остановимся и рассмотрим RapidMiner (RM) более подробно [5].

У любого инструмента имеются достоинства и недостатки.

Достоинства:

- Понятный **GUI** – интерфейс (Graphical user interface - графический интерфейс пользователя (ГИП)) – рис. 1.
- Расширяемость: можно задействовать язык R, т.е можно создавать свои собственные функции и операторы, на более низком уровне.
- Кроме **IDE** (Integrated Development Environment - интегрированная среда разработки), представленной выше, ещё есть сервер.
- RapidMinerStudio [6] создаёт процессы, а сервер в свою очередь управляет характеристиками процессов:
- частотой запуска определенного процесса, временем его работы и выделением ресурсов для него.
- Удобная система подсказок и предложений, т.е. RapidMiner строит статистику - какие операции, атрибуты или роли применяют в определенной ситуации (рис. 2).
- Качественная документация на английском языке.
- Многофункциональный.
- Для работы с ПО не требуется знание низкоуровневых языков, но желательно.

Рис. 2. Рекомендованные операции, атрибуты роли

Можно сказать, - это идеальный инструмент для работы с данными, но есть недостатки, из-за которых не каждый возьмется работать с ним:

- Большая нагрузка на оперативную память: поскольку RapidMiner написан на Java, исходя их этого возникает проблема с аппаратными мощностями для работы с ним.
- Бесплатная версия RapidMinerStudio последней версии сильно урезана по функционалу в отличие от предыдущего аналога. То есть, для наличия всех операций необходимо приобретать полноценный пакет или же использовать предыдущую версию с открытым исходным кодом.
- Высокая стоимость обучающих курсов, курсов по повышению квалификации и литературы по RapidMiner (RM).

Все эти недостатки были описаны с точки зрения руководителя. Работать с ним или нет, зависит от конкретно решаемой задачи. Но, основное достоинство RM — он

является готовым решением и на сегодня рядовой пользователь привык работать с GUI и визуальной средой.

Но RapidMiner не является единственным средством интеллектуального анализа, при его описании не раз упоминался язык R [7].

В отличие от RapidMiner, как было сказано выше, R—язык низкого уровня, т.е его можно применять в более обширных областях, в отличие от RM.

Среди его положительных и отрицательных сторон стоит выделить:

- Язык R как пластилин: из него можно получить любое средство обработки данных, но для этого необходимо потратить время для его изучения.
- Бесплатный.
- Нагрузка на оперативную память возможна при обработке колоссальных массивов неструктурированных данных.
- Он не является готовым решением, то есть многие функции необходимо создавать или искать среди дополнительных библиотек и пакетов.
- Нет графического интерфейса: все операции происходят в терминале или консоли, но при этом присутствуют библиотеки визуализации, которые позволяют создавать качественные графики и графы.
- Базовая сборка R занимает немного места на жестком диске, - около 22 МБ.

Ниже представлен пример для сравнения построение дерева принятия решений в RapidMiner (рис. 3) и в R (рис. 4), на основе данных, собранных о крушении круизного судна «Титаник».

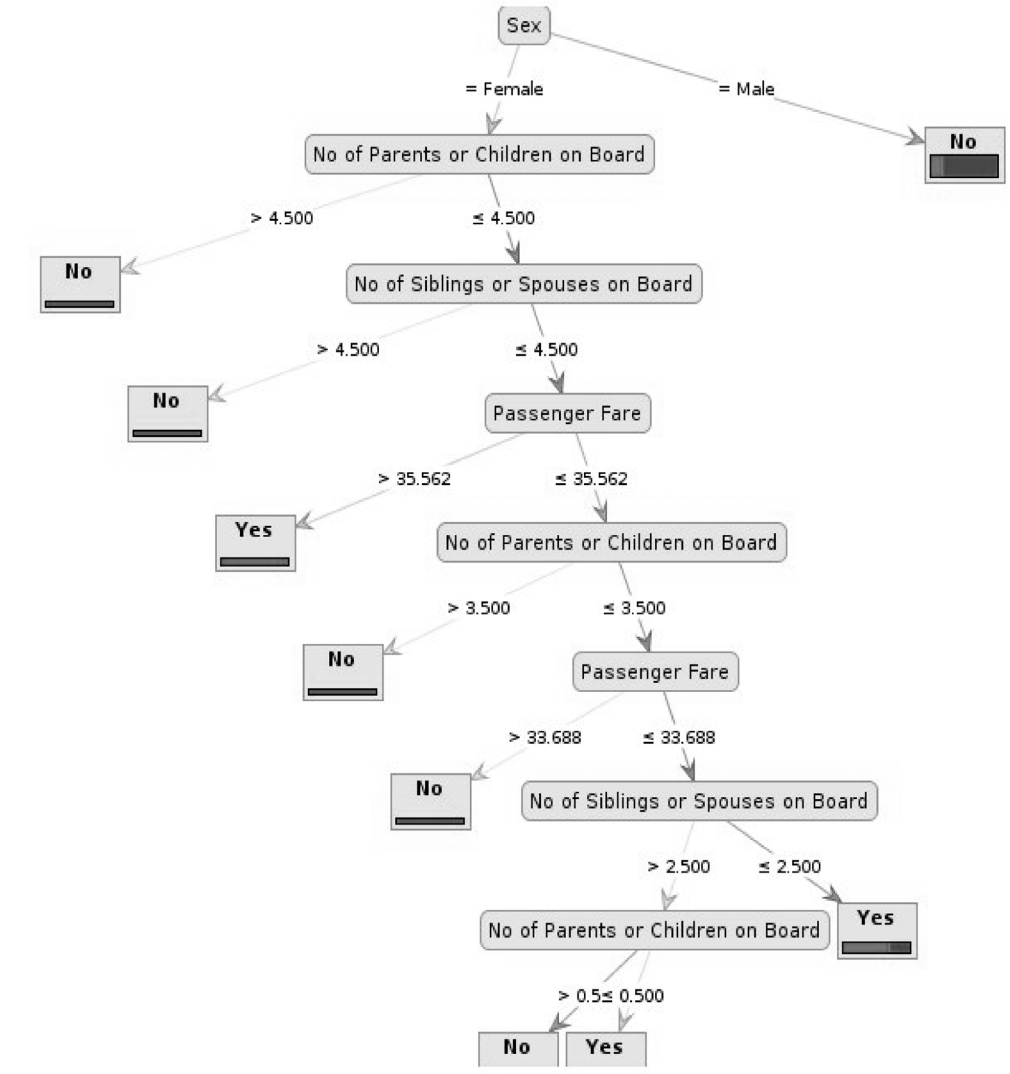


Рис. 3. Дерево принятия решений, построенное с помощью RapidMiner.

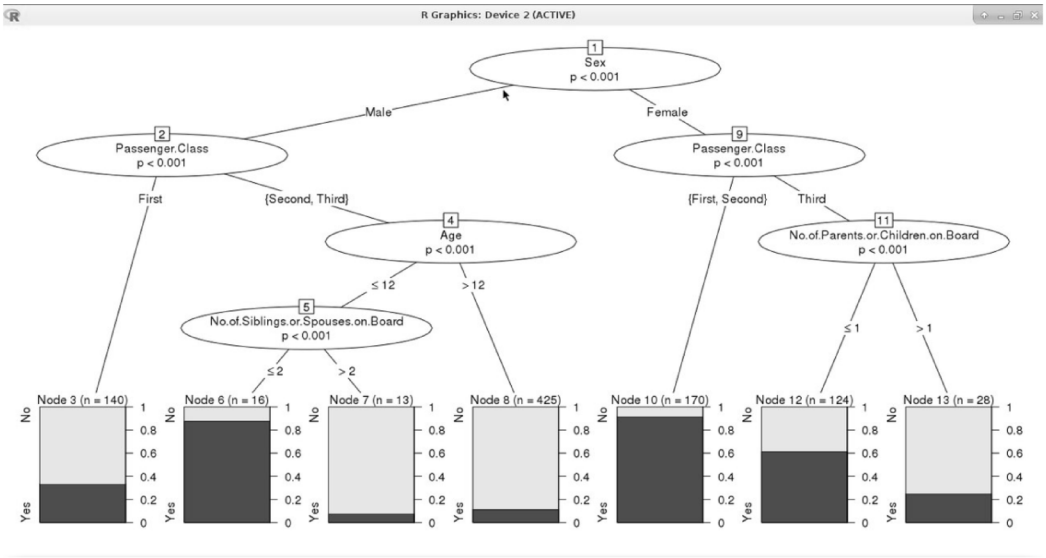


Рис. 4. Дерево принятия решений, построенное с помощью языка R

Литература

1. Филяк, П.Ю. Актуальность обеспечения информационной и экономической безопасности в условиях информационного общества[Текст]./ П.Ю. Филяк//Известия Тульского государственного университета. Технические науки:В. 2013.-Т.9. № 3. -С. 262-267.

2. Филяк, П.Ю. Проектирование с учетом обеспечения безопасности[Текст]/П.Ю. Филяк//Информация и безопасность.-В. 2015. -Т. 18. № 1.- С. 101-106.

3. Филяк, П.Ю. Информационная безопасность и комплексная система безопасности: анализ, подходы. [Текст]/ П.Ю. Филяк// Информация и безопасность.-В. 2016. -Т. 19. № 1. -С. 72-79.

4. Филяк, П.Ю. Применение инструментальных средств для работы с BigData и DataMining в решении проблем обеспечения безопасности организации.[Текст]/ П.Ю. Филяк, Э.Э.Байларли, В.И. Старченко//Информация и безопасность. – В.2017. - Т. 20. № 1-1 (4). - С. 133-136.

5. RapidMiner [Электронный ресурс]- Режим доступа: https://rapidminer.com.

6. RapidMinerStudioManual[Электронный ресурс]- Режим доступа:https://rapidminer.com.

7. R-bloggers [Электронный ресурс] – Режим доступа: https://www.r-bloggers.com

ФГБОУ ВО «Сыктывкарский государственный университет имени Питирима Сорокина»
Syktyvkar State University them.Pitirim Sorokin

APPLICATION OF RAPID MINER AND OPEN SOURCE ENVIRONMENTS AS DATA MINING TOOLS FOR SECURITY

P.Yu. Filyak, M.A. Vinogradov

Discusses Data Minig using RapidMiner - information-analytical system, which is essentially an Open Source Software with open source code and, in spite of the fact that is not proprietary software, gives ample opportunities user and allows to solve the very serious challenges when working with Big Data. Also is considered software Orange and universal programming language R with respect to tasks of analysis of Big Data and Data Minig

Key words: analysis,analytics, Big Data, Data Minig, software Orange, RapidMiner, universal programming language R, machine learning, GUI

УДК 004.056.55/003.26

КВАНТОВАЯ КРИПТОГРАФИЯ В ГРАФОВОЙ ИНТЕРПРЕТАЦИИ

П.Ю. Филяк, С.Н. Федирко, Ю.Н. Данилова

В статье рассматривается квантовая криптография, как один из новых методов криптографической защиты. Предложена наглядная интерпретация достаточно сложного метода инженерно-технической защиты информации, который в отличие от методов традиционной математической криптографии, требует оперирования определенными абстрактными категориями, которым очень важно дать некую их иллюстрацию, необходимую для понимания данной тематики

Ключевые слова: информационная безопасность, сертификат ключа проверки электронной подписи, криптография, квантовая криптография, квант, спин, суперпозиция, квантовое состояние, кубит, протокол шифрования, фотон, интерференция, квантовая запутанность

В Указе Президента России от 09.05.2017. N 203 «О Стратегии развития информационного общества в Российской Федерации на 2017 - 2030 годы» и Постановлении Правительства Российской Федерации от 15.04.2014**.** N 313 «Об утверждении государственной программы Российской Федерации «Информационное общество (2011 - 2020 годы)» представлена развернутая картина построения и развития информационного общества в нашей стране, что в полной мере отражает и даже опережает мировые тенденции в данной сфере. Одним из важнейших этапов этого длительного пути является перевод документооборота и предоставление большинства услуг как физическим, так и юридическим лицам, в электронную форму, что с учетом масштабов подразумевает высокую степень защиты информационных ресурсов и технологий, а это в свою очередь невозможно без применения криптографических методов. К сожалению, широко используемые на сегодняшний день традиционные методы криптозащиты, основанные на чисто математических методах, при опережающем развитии мощностей современной вычислительной техники не позволяют говорить о высокой стойкости даже самых совершенных алгоритмов шифрования.

Подтверждением тому служит, постоянное сокращение срока действия сертификат ключа проверки электронной подписи [1]. Эта проблема не уникальна и характерна для всех развитых стран, ставящих для себя целью построение информационного общества.

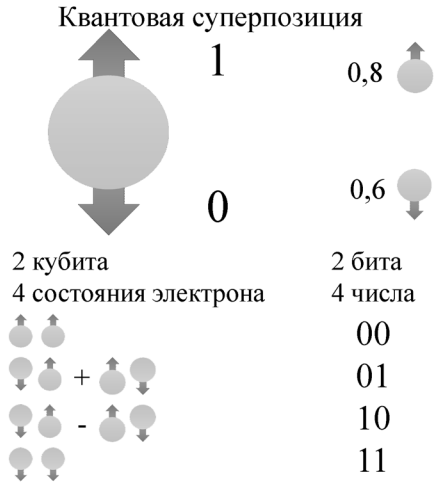


Рис. 1. Квантовая суперпозиция и кубит

Кардинально эту проблему можно решить путем применения принципиально новых методов криптографии, одним из которых является квантовая криптография [2-7], представляющая собой симбиоз квантовой физики (квантовой механики) и математических методов, где ключевыми являются такие понятия как фотон, квант, спин, суперпозиция, интерференция, квантовая запутанность, кубит, протокол

Филяк Петр Юрьевич – ФГБОУ ВПО СГУ, канд. техн. наук, доцент, e-mail: parallax-1@yandex.ru
Данилова Юлия Николаевна – ФГБОУ ВПО СГУ, студент, e-mail: bigdog375@yandex.ru
Федирко Станислав Николаевич– ФГБОУ ВПО СГУ, студент, e-mail: vladislav22009@yandex.ru

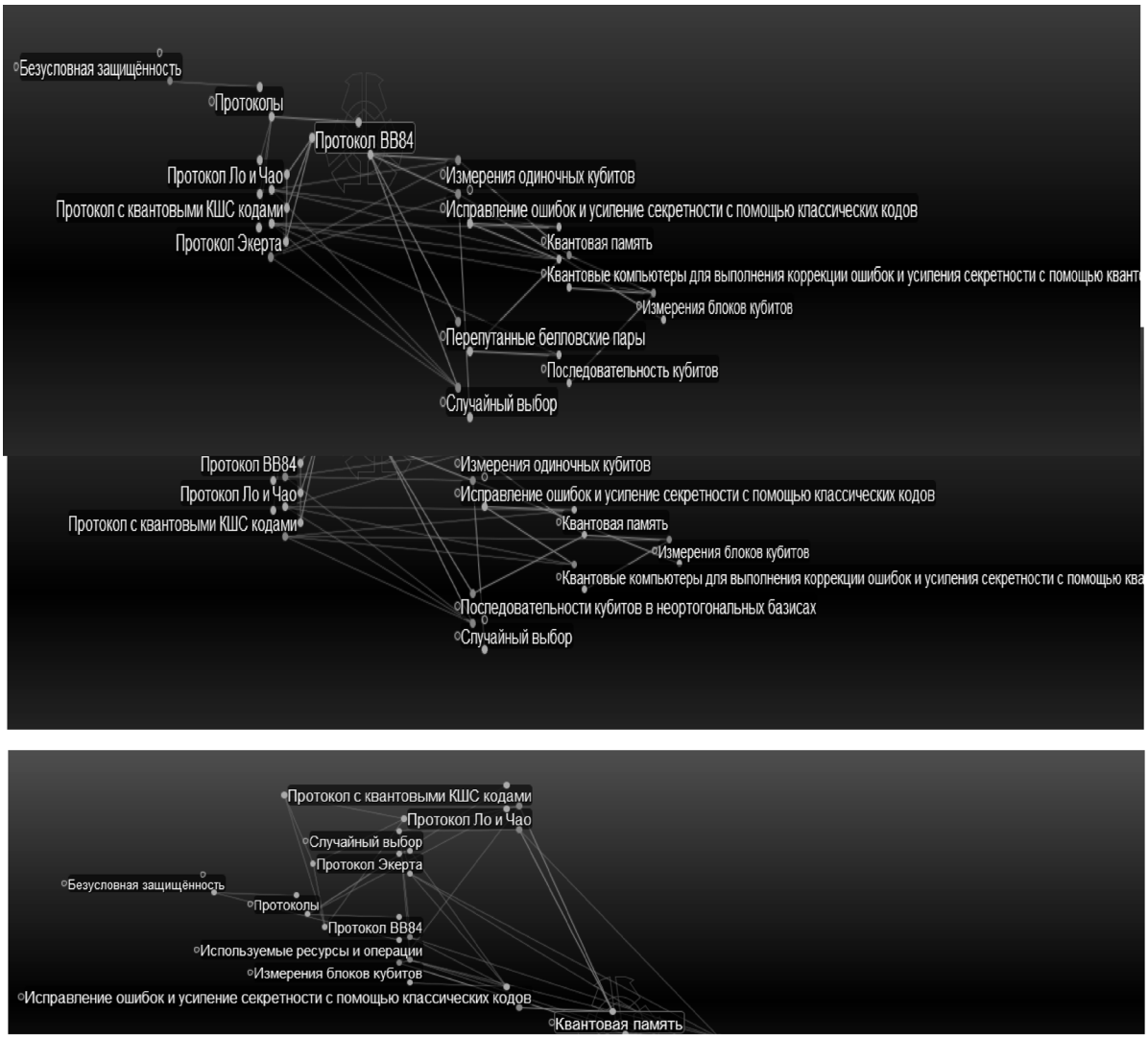


Рис. 2. Переход от BB84 к другим протоколам шифрования (Приложение 1)

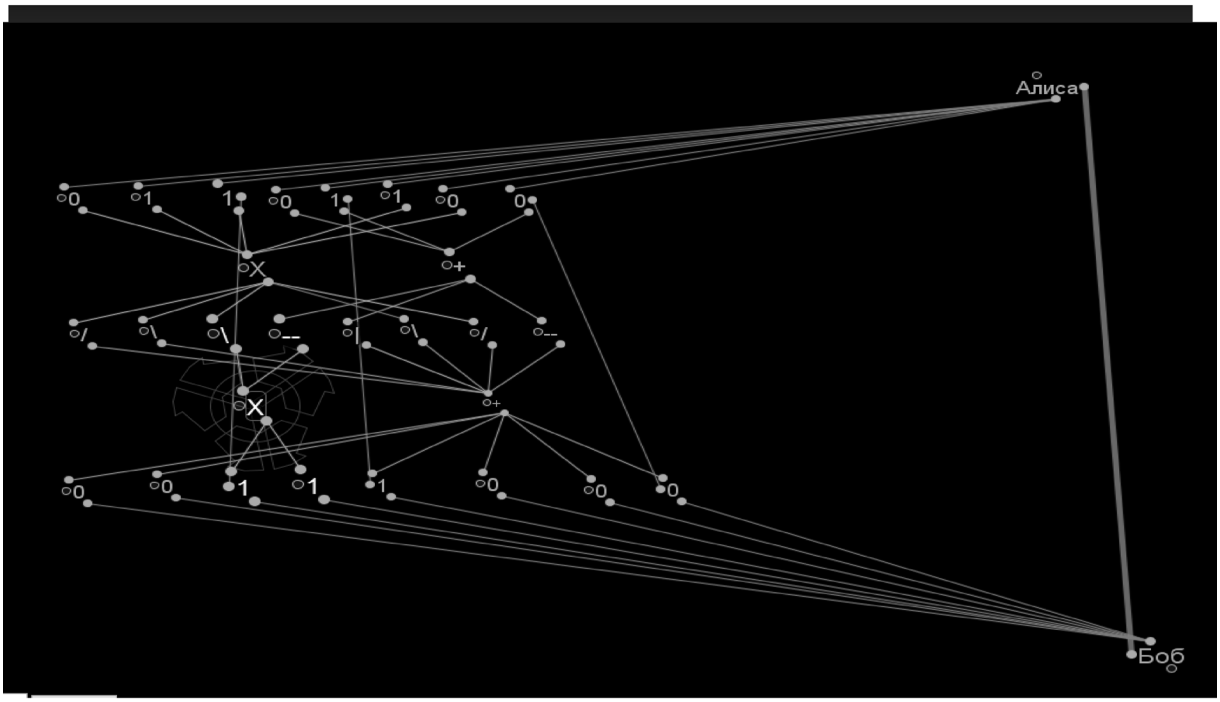


Рис. 3. Реализация протокола BB84 для идеальных условий

Как известно, N кубитов = 2^N бит. Вполне очевидно, что разобраться в квантовой криптографии задача нетривиальная – и дело не только в том, что приходится оперировать абстрактными категориями, но и в сущностной составляющей квантовой физики, которая предполагает и повсеместно предусматривает квантовую связанность и квантовую запутанность. шифрования (рис. 1). Математические методы, используемые в квантовой криптографии широко используют аппарат теории функций комплексной переменной (аналитических функций) и другие абстрактные категории для того, чтобы определить ключевые понятия, такие как квантовое состояние (состояние квантового объекта), квантовая связанность, квантовая запутанность (запутанность квантовых состояний), кубит (q-бит/кьюбит/*quantumbit*/квантовый бит), состояние кубитов.

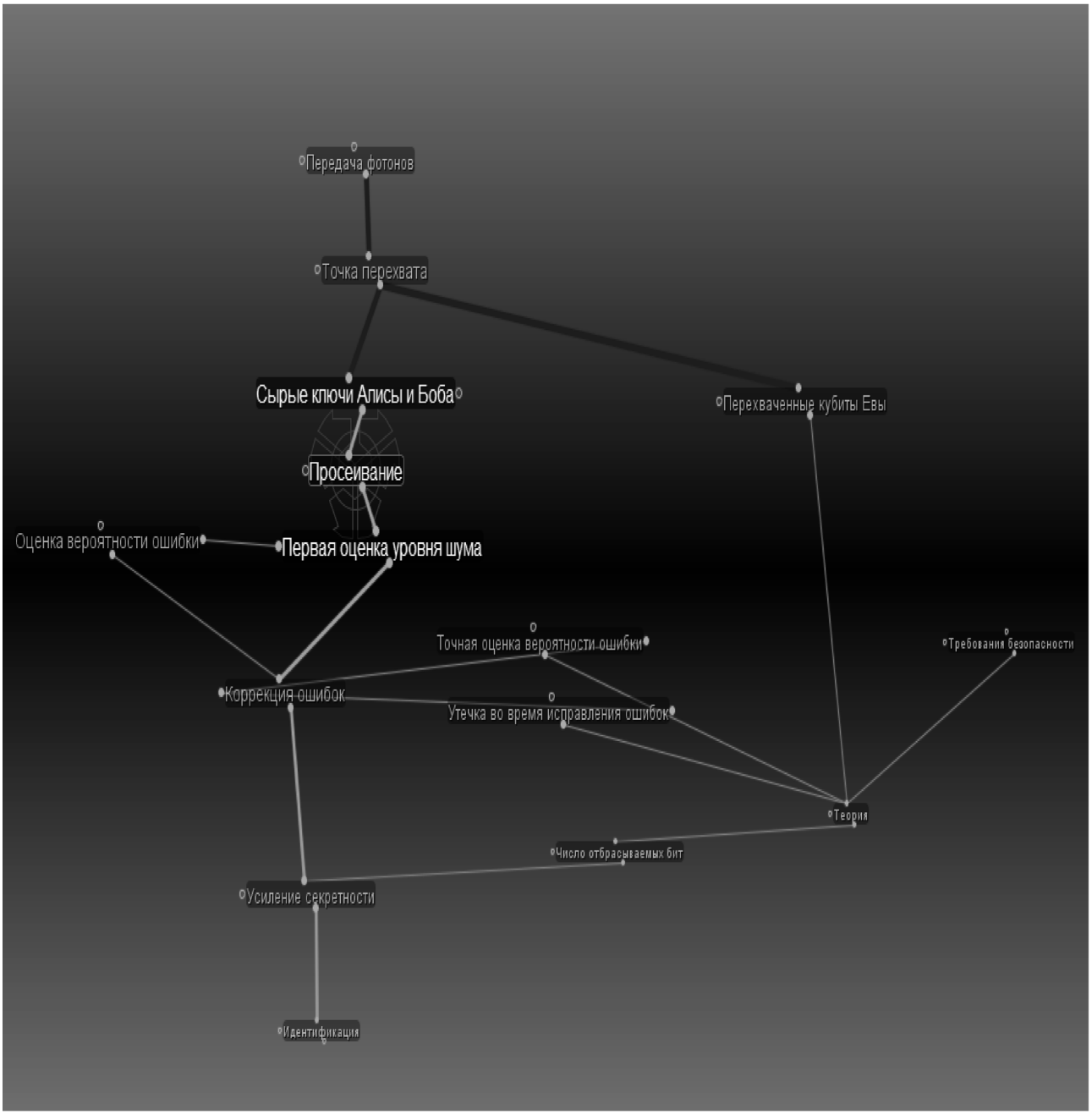


Рис. 4. Общая схема квантового распределения ключа (КРК)

Для воспроизведения этих состояний и иллюстрации их в трехмерном и в 2d – изображении необходим специальный инструмент, позволяющий совместить графику и интеллектуальную начинку, способную описать все многообразие состояний и связей.

В качестве такого эффективного инструмента очень хорошо подходит система управления знаниями Brain (PersonalBrain) [8] с мощным графическим интерфейсом. Применение TheBrain очень наглядно иллюстрирует протокол квантового распределения ключей BB84 и переход к другим протоколам (Рис. 2). На рис. 3 показана реализация протокола BB84 для идеальных условий, а на рис.4 представлена общая схема квантового распределения ключа.

Как видно из приведенных примеров, применение графовых методов позволяет наглядно иллюстрировать особенности квантовой криптографии и интерпретировать действия, связанные с квантовой криптозащитой, а также исследовать уязвимости квантовой криптографии и стойкость различных квантовых протоколов.

Литература

1. Федеральный закон «Об электронной подписи» от 06.04.2011. N 63-ФЗ

2. Килин, С. Я. Квантовая криптография: идеи и практика / С. Я. Килин, Д.Б. Хорошко, А. П.Низовцев – В.:2007. –159, 192, 230с.

3. Румянцев, К.Е. Квантовая криптография: принципы, протоколы, системы/ К.Е. Румянцев, Д.М. Голубчиков— В.:2008. — Р. 10—17.

4. Хорошко, Д.Б.Квантовая криптография: квантовое распределение ключа посредством кодирования через временные сдвиги/Д.Б. Хорошко, Д.И. Пустоход, В.Н. Чижевский, С.Я. Килин. // Материалы XIV Международной конференции "Комплексная защита информации" (Могилев, Беларусь, 19-22 мая 2009 г.):2009. – С. 188-189.

5. Слабое место в квантовой криптографии [Электронный ресурс]//Портал Хабрахабр: URL: https://habrahabr.ru/post/122282/

6. Красавин В. Квантовая криптография / В. Красавин // Подводная лодка — М.:2000 — №8.

7. Замена двоичной логики — увеличит ли это производительность? [Электронный ресурс]//Портал Хабрахабр: https://habrahabr.ru/post/166679/

8. Филяк, П.Ю. Обеспечение информационной безопасности с помощью технологии управления знаниями «BRAIN»/ П.Ю. Филяк, С.Н. Федирко//Информация и безопасность-В.: 2016. -Т. 19. № 2.- С. 238-243

9. Das, S. Anonymizing Edge-Weighted Social Network Graphs / S. Das, O. Egecioglu, A. El Abbadi // Computer Science, UC Santa Barbara, Tech. Rep. – 2009.

ФГБОУ ВО «Сыктывкарский государственный университет имени Питирима Сорокина»
Syktyvkar State University them. Pitirim Sorokin

QUANTUM CRYPTOGRAPHY IN A GRAPH INTERPRETATION

P.Yu. Filyak, St.N. Fedirko, Yu.N. Danilova

The article discusses quantum cryptography, as one of the new methods of cryptographic protection. Offered a visual interpretation of the rather complex engineering method of information protection, which, unlike traditional methods, mathematical cryptography, requires manipulating certain abstract categories, which is very important to give some of their illustration, needed to understand the subject

Keywords: information security, electronic signature verification key certificate, cryptography, quantum cryptography, quant,spin, superposition, quantum state,qubit, protocol encryption, photon, interference, quantum entanglement

УДК 004.5/004.021/003.63

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА POLYANALYST В ОБЕСПЕЧЕНИИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

П.Ю. Филяк, Ю.Н. Данилова, В.В. Растворов

Предлагается подход, позволяющей эффективно осуществлять комплексный интеллектуальный анализ (DataMining) при работе с «большими данными» (BigData) в целях обеспечения информационной безопасности, - на базе использования отечественной информационно - аналитической системы (ИАС) PolyAnalyst, что предоставляет большие возможности прогнозирования и принятия управленческих решений для адекватного реагирования на угрозы и негативные тенденции

Ключевые слова: данные, информация, угрозы, безопасность, информационная безопасность, анализ, большие данные (BigData), интеллектуальный анализ данных (DataMining), информационно - аналитические системы (ИАС), ИАС PolyAnalyst

В настоящее время, когда объемы данных и информации (BigData), которыми приходится оперировать лицам, принимающим решения (ЛПР), стали значительно превышать величины, которые в состоянии переработать человеческий мозг [1], необходимо широко внедрять в практику подходы, позволяющие своевременно, и эффективно извлекать из огромных объемов структурированных и неструктурированных данных знания посредством интеллектуального анализа данных и информации - DataMining [2-3], что невозможно представить без широкого внедрения и использования современных информационно-аналитических систем, представляющих собой эффективный симбиоз математики, современных средств вычислительной техники, программного обеспечения и средствмультимедиа.

Задача заключается в выборе из широкого спектра информационно-аналитических систем (ИАС) такого инструмента, который бы позволял эффективно достигать поставленные цели и решать конкретные задачи при максимально оптимальной величине традиционного критерия – соотношение цена/качество, не имел завышенных требований к программно-аппаратной платформе, обладал соответствующим интерфейсом и эргономичностью, удобством работы с «входной» и «выходной» информацией (и ее представлением).

В этой связи, по мнению авторов статьи, особого внимания заслуживает ИАС PolyAnalyst [4], возможности которой были ими изучены при решении конкретных задач обеспечения безопасности.

PolyAnalyst – информационно-аналитическая система для получения знаний из большого количества данных в доступной пониманию форме и в виде оперативно применяемых моделей. Данная информационно-аналитическая система может применяться для решения проблем обеспечения информационной безопасности.

Источниками данных в PolyAnalyst могут служить: brandwatch, CSV файлы, e-mail, FTP сервер, facebook, JSON, Lotus, MicrosoftAccess, MicrosoftExcel, NumericalSequence, OpenDatabaseConnectivity, OLEDB, RSS, SDLSM2, SPSS, twitter, XML, интернет-ресурсы, копии таблиц, объединённый поиск, ссылки и файлы.

PolyAnalyst может производить операции с колонками, такие как: модификация и фильтрация колонок; замена, консолидация и нормализация категорий; замена и извлечение терминов; объединение и применение моделей; дискретизация, нормализация и разбиение. Операции со строками: сортировка и фильтрация строк;

Филяк Петр Юрьевич – ФГБОУ ВПО СГУ, канд. техн. наук, доцент, e-mail: parallax-1@yandex.ru
Данилова Юлия Николаевна – ФГБОУ ВПО СГУ, студент, e-mail: danilova.yulia2@tandex.ru
Растворов Владислав Владимирович – ФГБОУ ВПО СГУ, студент, e-mail: vladislav22009@yandex.ru

суммы и ранги; выборка, обогащение данных, уникальные записи. Операции с таблицами: развёртка и свёртка транзакций; экспорт в ODBC и в файл; ABCXYZ анализ, агрегирование, конкатенация, объединение, отправка писем, производная таблица; структурная организация.

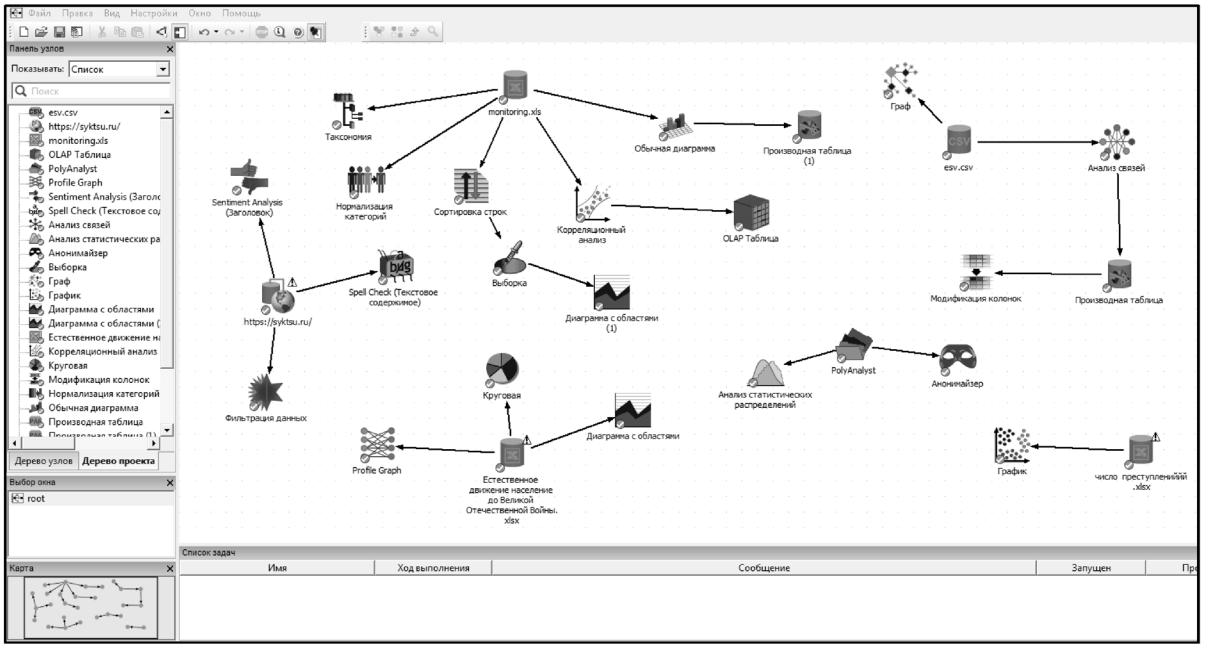


Рис. 1. Интерфейс PolyAnalyst

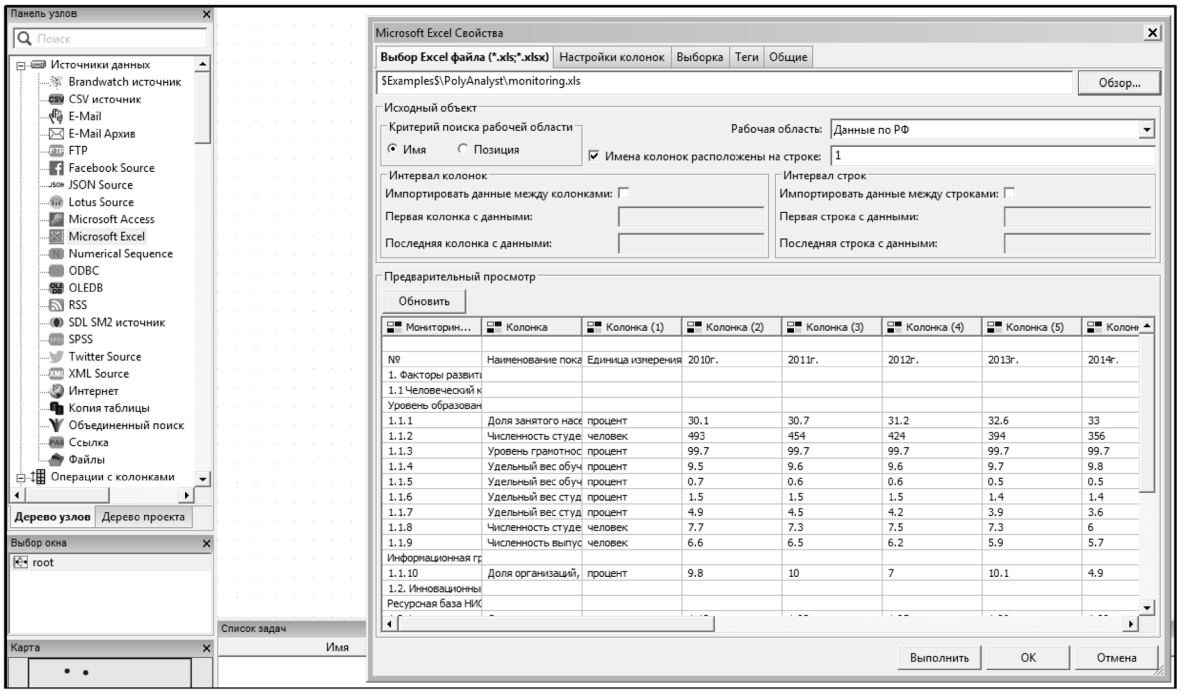


Рис. 2. Загрузка источника данных

Информационно-аналитическая система PolyAnalyst имеет возможности анализа данных, анализа текста и многомерного анализа. К анализу данных можно отнести: классификация Байеса, сеть Байеса, нейронная сеть ARIMA, CHAID-анализ, таблица решений, дерево решений, MediCop, SVM, аддитивная модель, анализ временных рядов, анализ покупательских корзин, транзакционный анализ покупательских корзин, дискриминантный анализ корзин, анализ связей, анализ социальной сети,

анализ статистических распределений, корреляционный анализ, аудит данных, ближайшие соседи, временные связи, кластеризация, линейная регрессия, логистическая регрессия, случайный лес решений, тестирование модели, факторный анализ и фильтрация данных. К текстовому анализу относится перевод текста, анализ тональности текста, извлечение ключевых слов, определение языка, поисковый запрос, проверка орфографии, проверка грамматики, связь терминов, текстовая кластеризация и удаление фрагментов. Многомерный анализ подразделяется на таксономию, OLAP - таблицу, многомерную матрицу и фильтрацию таблицы.

Также PolyAnalyst может представлять данные в виде карт, графов, графиков и диаграмм. На рис. 1 представлен интерфейс системы PolyAnalyst. Рабочая область представляется в виде графов, где каждый узел имеет свою смысловую нагрузку. Корневой узел – это, как правило, источник данных, а остальные узлы – это анализ, операция, или графическое представление данных.

Для начала работы выбираем источник данных, к примеруMicrosoftExcel. Выбираем в левом меню необходимый объект и начинаем его настраивать. Привязываем к объекту конкретный документ, как это представлено на рис. 2. В нашем случае статистика, - «Мониторинг развития информационного общества в Российской Федерации» - взятая с официального сайта Федеральной службы государственной статистики. Далее от источника данных (выбранной статистики) производим сортировку по нужному параметру.

В системе PolyAnalyst возможно построение дерева решений. На рис. 3 приведена матрица ошибок дерева решений. При выборе тестового источника данных, можно произвести тестовый анализ, который представлен на рис. 4.

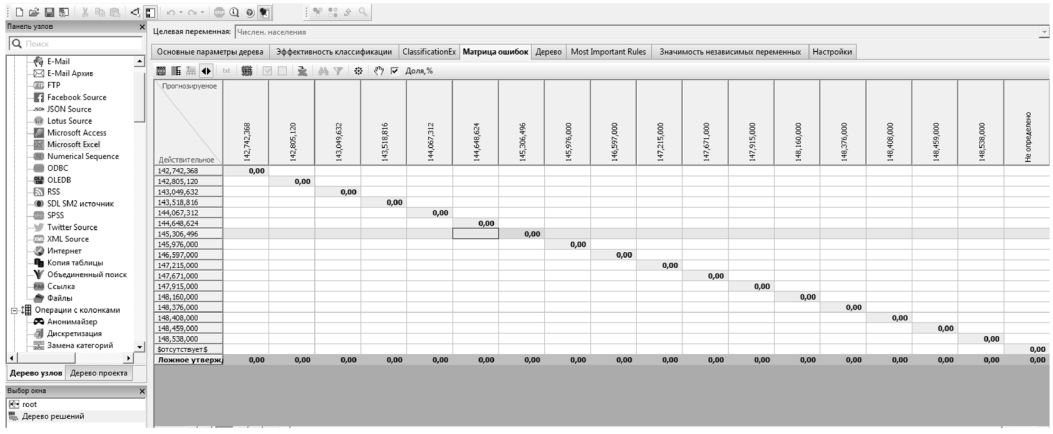


Рис. 3. Матрица ошибок в дереве решений

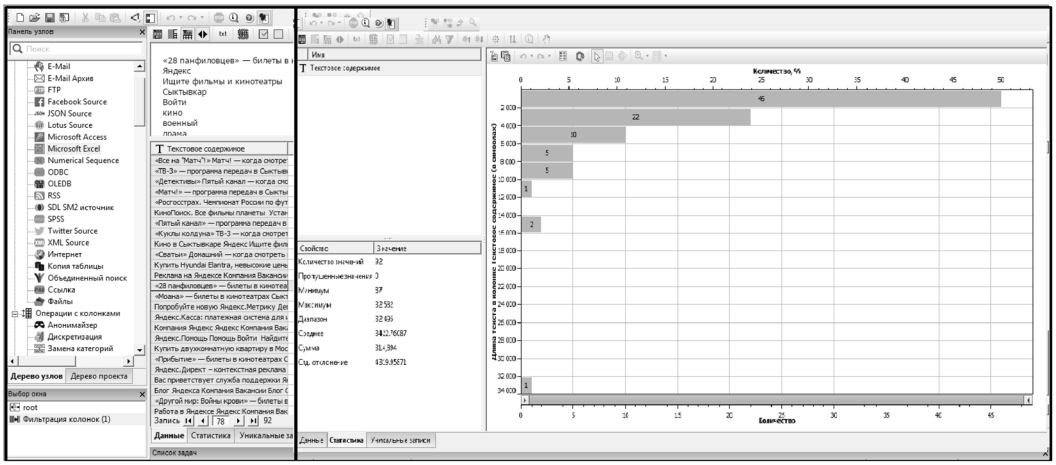


Рис. 4. Анализ текста

Ещё один вид графического представления информации – диаграмма. На рис. 5 представлен график по числу зарегистрированных преступлений на 100тыс. человек по субъектам Российской Федерации – так называемый коэффициент преступности.

Рис. 5. Диаграмма статистики

Таким образом, даже малая часть функционала информационно-аналитической системы PolyAnalyst, раскрытая в статье, говорит о колоссальных возможностях и потенциале указанного программного продукта. PolyAnalyst хорошо подходит для мониторинга, анализа, прогнозирования и принятия управленческих решений при решении проблем обеспечения безопасности на основе работы с большими массивами данных и информации.

Литература

1. Филяк, П.Ю. Информационная безопасность и комплексная система безопасности: анализ, подходы[Текст]/П.Ю. Филяк// Информация и безопасность.-В. 2016. -Т. 19. №. 1 -С.72-79.

2. Майер-Шенбергер, В. Большие данные. Революция, котораяизменит то, как мы живем, работаем и мыслим/В. Майер-Шенбергер, К. Кукьер– М. : Манн, Иванов и Фербер, 2014. – 240 с.

3. Большие данные (BigData): изменение будущего человечества [Электронный ресурс] – Режим доступа :http://www.kitaichina.com/se/txt/2013-03/28/content_530849.htm(Дата обращения: 01.03.2017).

4. Портал компании Megaputer – разработчика ИАС PolyAnalyst [Электронный ресурс] –Режим доступа: http://www.megaputer.ru/polyanalyst.php(Дата обращения:10.04.2017).

ФГБОУ ВО «Сыктывкарский государственный университет имени Питирима Сорокина»
Syktyvkar State University them.Pitirim Sorokin

INFORMATION-ANALYTICAL SYSTEM POLYANALYST IN ENSURING INFORMATION SECURITY

P.Yu. Filyak, Yu.N. Danilova, V.V. Rastvorov

An approach that allows to effectively implement a comprehensive data mining (Data Mining) when working with «Big Data» in order to ensure information security, based on the use of domestic information - analytical systems (IAS) PolyAnalyst, that provides great opportunities of forecasting and decision-making to respond adequately to the threat and negative trends

Key words: data, information, threats, security, information security, analysis, Big Data, data mining (Data Mining), information-analytical systems (IAS), the IAS PolyAnalys

УДК004.056

ПРЕДВАРИТЕЛЬНОЕ ФОРМИРОВАНИЕ РУКОПИСНЫХ ТЕКСТОВ В ПРОЦЕДУРЕ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ

В.П. Лось, Е.Д. Тышук

В статье дается краткий обзор современных методов биометрической аутентификации, основанных на использовании рукописной подписи. Эти методы предполагают использование изображений подписей каждого пользователя для решения задач создания эталонов и верификации, под которой понимается проверка подлинности подписи, что означает соответствие подписи конкретному пользователю, предъявившему эту подпись. Делается предположение о том, что подбор рукописных текстов может привести к более качественной аутентификации пользователей

Ключевые слова: аутентификация, рукописная подпись, верификация

В литературе приводятся два основных способа использования рукописной подписи для решения задач аутентификации: статический и динамический.

Статический или офлайновый (offline signature) способ предполагает использование статических изображений подписей наряду с дополнительными характеристиками, которые потенциально могут быть получены из таких изображений: направление движения пера, толщина линии и другие.

Динамический или онлайновый (online signature) способ является более эффективным по сравнению с предыдущим, поскольку позволяет наряду со статическими характеристиками использовать динамические: изменение величины давления пера, скорость перемещения пера, угол наклона ручки и другие.

В обоих случаях процедура верификации традиционно характеризуется ошибкой первого рода FAR и ошибкой второго рода FRR:

- FAR (False Acceptance Rate) - частота ложной аутентификации;

- FRR (False Rejection Rate) - частота ложных отказов.

Кроме этих ошибок рассматривается понятие EER, или ERR (Equal Error Rate, или ERror Rate) - частота ошибок. Если при каких-то настройках системы FAR=FRR, то это и является значением ERR.

В известных работах Шуба Д.А /1-3/, Сорокина И.А./4,5/ и Ложникова П.С./6,7/, посвященных проблеме аутентификации при использовании рукописной подписи, недостаточное внимание уделяется формированию исходных текстов, на основе которых строится образ подписей. Априорно считается целесообразность использования в этих целях подписи, которая используется в повседневной жизни и которой удостоверяется на бумажном носителе факты передачи денег, получения документов, отсутствия претензий и иные обстоятельства. Вместе с тем, в существующих работах по рассматриваемой тематике не доказывается, что именно рукописная подпись в ее классическом понимании наиболее целесообразна в системах аутентификации. Действительно, рассмотрим процедуру верификации подписи.

Верификацию подписи можно представить как последовательное выполнение следующих операций:

1. Определение сегментов подписи и их параметров.

2. Определение оценки схожести, путем использования оператора преобразования пары введенной подписи и образа подписей в оценку схожести.

3. Решение о признании предъявленной подписи подлинной или ложной (подделкой).

Лось Владимир Павлович – МИРЭА, д.в.н., профессор, e-mail: los-vladimir@mirea.ru
Тышук Екатерина Дмитриевна – МИРЭА, e-mail: dolina@mirea.ru

Содержание перечисленных операций никак не ориентировано на обязательное использование подписи в ее классическом понимании. Отсюда можно сделать предположение (гипотезу) о том, что для конкретного пользователя могут существовать слова, рукописное воспроизведение которых и использование в процедуре аутентификации, может характеризоваться меньшими ошибками FAR и FRR по сравнению со случаем классической рукописной подписи. Данная идея была описана авторами в работе [8].

Теоретическое подтверждение данной гипотезы в силу понятных причин затруднительно. Можно привести алгоритм машинного моделирования для оценки данной гипотезы. Этот алгоритм для каждого пользователя включает следующие этапы:

1. Сбор исходной информации для формирования множества эталонных подмножеств рукописных слов, включая подмножество подлинных подписей.

2. Моделирование одним из известных методов процесса аутентификации для каждого из эталонных подмножеств рукописных слов, включая подмножество подлинных подписей.

3. Оценка ошибок FAR и FRR.

4. Выявление подмножеств рукописных слов, использование которых приводит к меньшим ошибкам по сравнению с использованием подмножества подлинных подписей.

Матрица эталонов изображений рукописных подписей в классической процедуре аутентификации имеет вид

$$\mathrm{M}=\begin{bmatrix} z_{11} & \cdots & z_{1J} \\ \vdots & \ddots & \vdots \\ z_{I1} & \cdots & z_{IJ} \end{bmatrix},$$

где в обозначении z_{ij} i= $\overline{1,I}$; j= $\overline{1,J}$; I - количество пользователей системы, J – число реализаций рукописных подписей для каждого пользователя.

При предъявлении i-м пользователем изображения подписи Z_i^r оцениваются степени близости Z_i^r и эталонов i-ой строки:

$\alpha_j^i = \xi\,(Z_i^r, z_{ij})$, $j= \overline{1,J}$.(1),

где ξ - оператор преобразования пары предъявленного изображения и эталона (Z_i^r, z_{ij})в оценку степени близости α_j^i.

Если среди оценок α_j^i находится такая, что

$$\max_j\left(\alpha_j^i\right) \geq \alpha_{\text{доп}}, \qquad (2)$$

где $\alpha_{\text{доп}}$ – установленное значение степени близости, то принимается положительное решение о прохождении пользователем процедуры аутентификации. В противоположном случае принимается отрицательное решение и имеет место ошибка первого рода.

Если изображение подписи Z_l^r предъявляется l-м пользователем, пытающимся себя выдать за i-го пользователя, причем $l \neq i$, то в случае нахождения среди оценок близости

$\alpha_j^l = \xi\,(Z_l^r, z_{ij})$, j= $\overline{1,J}$

такой, что

$\max_j(\alpha_j^l) \geq \alpha_{\text{доп}}$,

имеет место ошибка второго рода.

Увеличение исходного базиса изображений, в качестве которых используются не только изображения подписи, но и другие рукописные тексты, имеет целью снижение ошибок первого и второго рода.

В рассматриваемом случае каждому пользователю сопоставляется матрица изображений

$$M_i=\begin{bmatrix} z_{11}^i & \cdots & z_{1J}^i \\ \vdots & \ddots & \vdots \\ z_{K1}^i & \cdots & z_{KJ}^i \end{bmatrix},$$

где в обозначении z_{kj}^i: $i= \overline{1,I}$; j= $\overline{1,J}$; $k= \overline{1,K}$,I- количество пользователей системы, J – число реализаций рукописного текста одного и того же слова; K- количество используемых слов.

Например, при наличии в системе пяти пользователей (I = 5), шести реализаций рукописного текста одного и того же слова (J = 6) и семи используемых слов (K= 7) получим следующие пять (по числу пользователей) матриц изображений:

$$M_1=\begin{bmatrix} z_{11}^1 & \cdots & z_{16}^1 \\ \vdots & \ddots & \vdots \\ z_{71}^1 & \cdots & z_{76}^1 \end{bmatrix},$$

$$M_2=\begin{bmatrix} z_{11}^2 & \cdots & z_{16}^2 \\ \vdots & \ddots & \vdots \\ z_{71}^2 & \cdots & z_{76}^2 \end{bmatrix},$$

$$M_3=\begin{bmatrix} z_{11}^3 & \cdots & z_{16}^3 \\ \vdots & \ddots & \vdots \\ z_{71}^3 & \cdots & z_{76}^3 \end{bmatrix},$$

$$M_4=\begin{bmatrix} z_{11}^4 & \cdots & z_{16}^4 \\ \vdots & \ddots & \vdots \\ z_{71}^4 & \cdots & z_{76}^4 \end{bmatrix},$$

$$M_5=\begin{bmatrix} z_{11}^5 & \cdots & z_{16}^5 \\ \vdots & \ddots & \vdots \\ z_{71}^5 & \cdots & z_{76}^5 \end{bmatrix}.$$

Заметим, что в общем случае

$z_{kj}^i \neq z_{kj}^l$ при $i \neq l$, $i, l = \overline{1,I}$.

При моделировании на ЭВМ из матрицы M_i, $i = \overline{1,I}$, используется только одна строка. Таким образом, на *s*- ом шаге моделирования в качестве матрицы эталонов изображений $M_{эs}$ выступает матрица, первая строка которой является одной из строк матрицы M_1, вторая строка - одной из строк матрицыM_2и так далее. Последняя строка является одной из строк матрицы M_I. Для приведенного выше примера матрица эталонов изображений $M_{эs}$ на *s*- ом шаге моделирования может иметь следующие строки:

$$M_{эs}^1=[z_{11}^1 \ldots z_{16}^1];$$
$$M_{эs}^2=[z_{11}^2 \ldots z_{16}^2];$$
$$M_{эs}^3=[z_{11}^3 \ldots z_{16}^3];$$
$$M_{эs}^4=[z_{11}^4 \ldots z_{16}^4];$$
$$M_{эs}^5=[z_{11}^5 \ldots z_{16}^5].$$

При оценке ошибки первого рода используется i – ая строка матрицы $M_{эs}$и в качестве предъявляемого системе изображения используется один из элементов i – ой строки матрицы $M_{эs}$, который исключается из процедуры подсчета оценки близости.

Для приведенного примера при выборе в качестве предъявляемого системе изображения z_{11}^i,получим в соответствии с (1):

$\alpha_1^i = \xi\,(z_{11}^i, z_{11}^i)$,
$\alpha_2^i = \xi\,(z_{11}^i, z_{12}^i)$,
$\alpha_3^i = \xi\,(z_{11}^i, z_{13}^i)$,
$\alpha_4^i = \xi\,(z_{11}^i, z_{14}^i)$,
$\alpha_5^i = \xi\,(z_{11}^i, z_{15}^i)$,
$\alpha_6^i = \xi\,(z_{11}^i, z_{16}^i)$.

Если все полученные оценки α_j^i, $j = 1,\ldots,6$ удовлетворяют требованию (2)

$$\alpha_j^i \geq \alpha_{\text{доп}},$$

строка $M_{эs}^i$ считается приемлемой для аутентификации с точки зрения ошибки первого рода.

Если среди полученных оценок α_j^i, $j = 1,\ldots,6$ находятся такие, что

$$\alpha_j^i < \alpha_{\text{доп}},$$

соответствующие строки матрицы M_i или элемент этой строки, приводящий к ошибке 1 – го рода исключаются из рассмотрения и на следующем шаге заменяются другими.

При оценке ошибки второго рода используется i – ая строка матрицы $M_{эs}$ и поочередно какой- либо элемент матрицы $M_{эs}$, не принадлежащий i – ой строке.

Для примера рассмотрим случай, когда рассматривается строка $M_{эs}^1=[z_{11}^1 \ldots z_{16}^1]$ и элементы этой строки сравниваются поочередно с элементами других строк. Если в результате такого сравнения находятся оценки близости такие, что

$\xi\,(z_{1j}^i, z_{1k}^l) \geq \alpha_{\text{доп}}$, $j = \overline{1,6}$, $l, i = \overline{1,5}$; $j = \overline{1,6}$, $k = \overline{1,7}, l \neq i$,

то l-ая строка или элемент z_{1k}^l исключаются из матрицы $M_{эs}$ на шаге *s*.

Количество шагов при моделировании на ЭВМ зависит от числа используемых слов K и количества различающихся между собой матриц $M_{эs}$.

По результатам моделирования определяется матрица $M_{эs}$, которой соответствуют меньшие ошибки первого и второго рода.

На практике обычно используется одно из следующих правил:

- при фиксированном значении ошибки первого рода минимизировать значение ошибки второго рода;

- либо при фиксированном значении ошибки второго рода минимизировать значение ошибки первого рода.

Для получения оптимальных соотношений между ошибками первого и второго рода используют различные классические критерии, в том числе:

- критерий минимального среднего риска (критерий Байеса);
- критерий максимума апостериорной вероятности (максимального правдоподобия);
- двухпороговый критерий Вальда;
- критерий Неймана-Пирсона.

Литература

1. Шуб Д.А. Модель процесса создания рукописного текста //Труды Института систжемного анализа Российской академии наук (ИСА РАН). – 2009. – т. 42(1) – С.264-270.

2. Шуб Д.А. Технология получения последовательности равностоящих отсчетов из последовательности неравностоящих отсчетов при анализе рукописного текста // Наукоемкие технологии. –2009. – т. 10, №8. – С. 37-42.

3. Андрианова Е.Г., Шуб Д.А. Повышение эффективности и защищенности дистанционного обучения путем применения современных средств анализа почерка. // Труды XVI -го Международного Симпозиума «Новые технологии в образовании, науке и экономике» / Под ред. Г. К. Сафаралиева, А. Н. Андреева – М.: Информационно-издательский центр Фонда поддержки вузов, 2007. С. 128 – 131.

4. Сорокин И.А. Компенсация искажений, вызванных нестабильностью угла наклона подписи в биометрических системах аутентификации. // Безопасность информационных технологий. Труды научно-технической конференции. — 2002. - Пенза, Изд-во ПНИЭИ. Том 3. с.87-90.

5. Сорокин И.А. Анализ траектории начертания почерка при проведении биометрической идентификации личности. / Оленин Ю.А., Сорокин И.А.// Современные технологии безопасности. - 2003 г. - №3 (6) - с. 12-17.

6. Ложников П.С. Использование сети многомерных функционалов Байеса для нейросетевого преобразования рукописной подписи человека в секретный ключ его электронной подписи. //Труды Межвузовской научно-практической конференции «Актуальные проблемы обеспечения информационной безопасности». – 2017.- Самара, Изд-во Инсома-пресс, с.124-128.

7. Ложников П.С., Сулавко А.Е., Еременко А.В., Волков Д.А. Экспериментальная оценка надежности верификации подписи сетями квадратичных форм, нечеткими экстракторами и персептронами // Информационно-управляющие системы.2016, №5 (84). – С. 73-85.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский технологический университет»
Moscow Technological University

PRELIMINARY FORMATION OF HANDWRITTEN TEXTS IN THE PROCEDURE OF USER AUTHENTICATION

V.P. Los, E.D. Tyshuk

The article gives a brief overview of modern methods of biometric authentication, based on the use of handwritten signatures. These methods assume the use of signature images of each user to solve the tasks of creating standards and verification, which is understood as the verification of the authenticity of the signature, which means that the signature corresponds to the specific user who presented this signature. It is suggested that the selection of handwritten texts can lead to better authentication of users

Key words: authentication, handwritten signature, verification

УДК 004.622

МОДЕЛЬ ПРОЦЕССА МЕЖВЕДОМСТВЕННОГО ОБМЕНА ИНФОРМАЦИЕЙ С ОГРАНИЧЕННЫМ ДОСТУПОМ

А.С. Пахомова, А.П. Пахомов, Й. Воришек

Представлена структурно-параметрическая модель процесса межведомственного информационного обмена, реализующая алгоритм вероятностного подтверждения данных как предварительного этапа официального запроса информации с ограниченным доступом

Ключевые слова: информация с ограниченным доступом, межведомственный информационный обмен, алгоритм вероятностного подтверждения данных

В современных условиях практически во всех процессах межведомственного обмена информацией используются современные информационные технологии (автоматизированные базы данных, автоматизированные средства обработки запросов и подготовки ответов). Использование таких информационных технологий позволяет существенно сократить время на решение многих задач за счет получения недостающей информации из сторонних организаций, в частности, из других ведомств. Однако существенным препятствием для полного использования возможностей информационных технологий служат правовые ограничения различных уровней (от федерального до организационного), направленные на выполнение требований по ограничению доступа к персональным данным, другим видам тайн, а также на выполнение локальных организационных правил предоставления данных сторонним организациям. При этом успех решения ряда задач связан не только с получением официальных данных, но и с подтверждением их наличия в той или иной сторонней организации. Решение задачи, требующей такого подтверждения, может быть существенно ускорено, если запрашивающая сторона как можно быстрее получит информацию о правильности своей гипотезы о наличии определенных данных в некоторой организации (ее базе данных). К таким задачам относятся задачи проведения финансовых расследований, связанные с необходимостью проверки множества проведенных подозреваемым финансовых операций, в том числе и через иностранные финансовые организации. Сохранение тайны следствия вкупе со строгими ограничениями на передачу персональных данных существенно тормозят процессы финансовых расследований, иногда приводя их к неактуальности за счет невозможности предотвратить следующие неправомерные действия [1]. Теоретические методы, позволяющие решить задачу ускорения информационного обмена при наличии организационных ограничений на официальное получение данных с ограниченным доступом, в настоящее время в научной литературе не представлены. Для разрешения противоречия между необходимостью повышения эффективности межведомственного информационного обмена и необходимостью обеспечить выполнение требований по правилам получения доступа к данным определенного содержания авторами предлагается использовать современные метод, состоящий в том, что предварительно на основе обмена данными с нулевым знанием может быть организован межведомственный обмен без предоставления данных ограниченного доступа, но с предоставлением данных, с некоторой вероятностью подтверждающих наличие либо отсутствие этих данных.

Для внедрения предложенной технологии требуются определенные изменения в организации информационных

Пахомова Анна Степановна — ВГТУ, канд. техн.наук, доцент, e-mail: mnac@comch.ru
Пахомов Алексей Павлович – РГУ нефти и газа им. Губкина, аспирант, e-mail: apal12@mail.ru
Воришек Йири - Пан-Европейский Университет (Словакия), к.т.н., профессор, научный сотрудник, e-mail: jiri.vorisek@paneurouni.com

процессов, поэтому для их обоснования руководству организаций должны быть представлены доказательства эффективности.

В настоящей статье предлагается математическая модель межведомственного информационного обмена, позволяющая оценить эффективность информационных процессов, основанных на существующих алгоритмах запросов информации, и предложенного алгоритма.

Для достижения цели настоящей статьи рассматривается ограниченный во времени фрагмент информационного процесса, соответствующий периоду проведения расследования, то есть периоду от выдвижения первоначальной гипотезы (на базе первичной исходной информации) до получения подтверждения правильности последней гипотезы на основе всей накопленной информации (в том числе не подтвердившей промежуточные гипотезы). Главным элементом модели является модель процесса запроса в стороннюю организацию. Структурная схема этого процесса приведена на рис. 1.

Рис.1. Структурная модель процесса запроса в стороннюю организацию

Такой процесс может быть представлен сетью – совокупностью последовательно наступающих событий, соответствующих подтверждению или отклонению данных (рис. 2). Переход между состояниями происходит с некоторыми вероятностями, величины которых зависят от правильности выдвинутых гипотез.

Сценарий реализации процесса состоит в следующем. Модель информационного процесса запроса-ответа включает сторону, обеспечивающую процесс подтверждения гипотезы и сбора данных(A). Объект интереса (v), в отношении свойств которого выдвигается гипотеза H(v), релевантная в течение времени tmav. Сторона (B), у которой может иметься искомая информация R(i(x)), необходимая для подтверждения гипотезы H(v) . Также у стороны B имеются наборы объектов со связанной с ними информацией v_b. Для идентификации объекта используется однозначно идентифицирующая объект v информация i(v). Информационный запрос-ответ Request(v) должен включать информацию (x), необходимую для идентификации объекта x, информации (H(x)) и необходимыми для запроса дополнительные данные, в частности R(i(x)). При условии выполнения предиката оценки достаточности информации в запросе Enough(H(x))=1 и для данных Legal(R(x))=1 будет выполняться запрос на стороне B.

Основными параметрами процесса, помимо отмеченных на рис. 2, являются следующие:

- B-множество информации, доступное B с мощностью |B|;

- ReqestHyposesis(R(i(x)))- информация, необходимая для уточнения вероятности(достоверности) гипотезы H(x), то есть это информация, необходимая для увеличения вероятности p(e), увеличения вероятности p(l), открытия стороне A новых

данных (l_new), для которых существует путь до v, открытия стороне A новых данных (e_new), для которых существует путь до v, обновления времени жизни Tmav(e,t),Tmav(l,t);

- (x) - информация, уникально идентифицирующая x для всего множества X (не

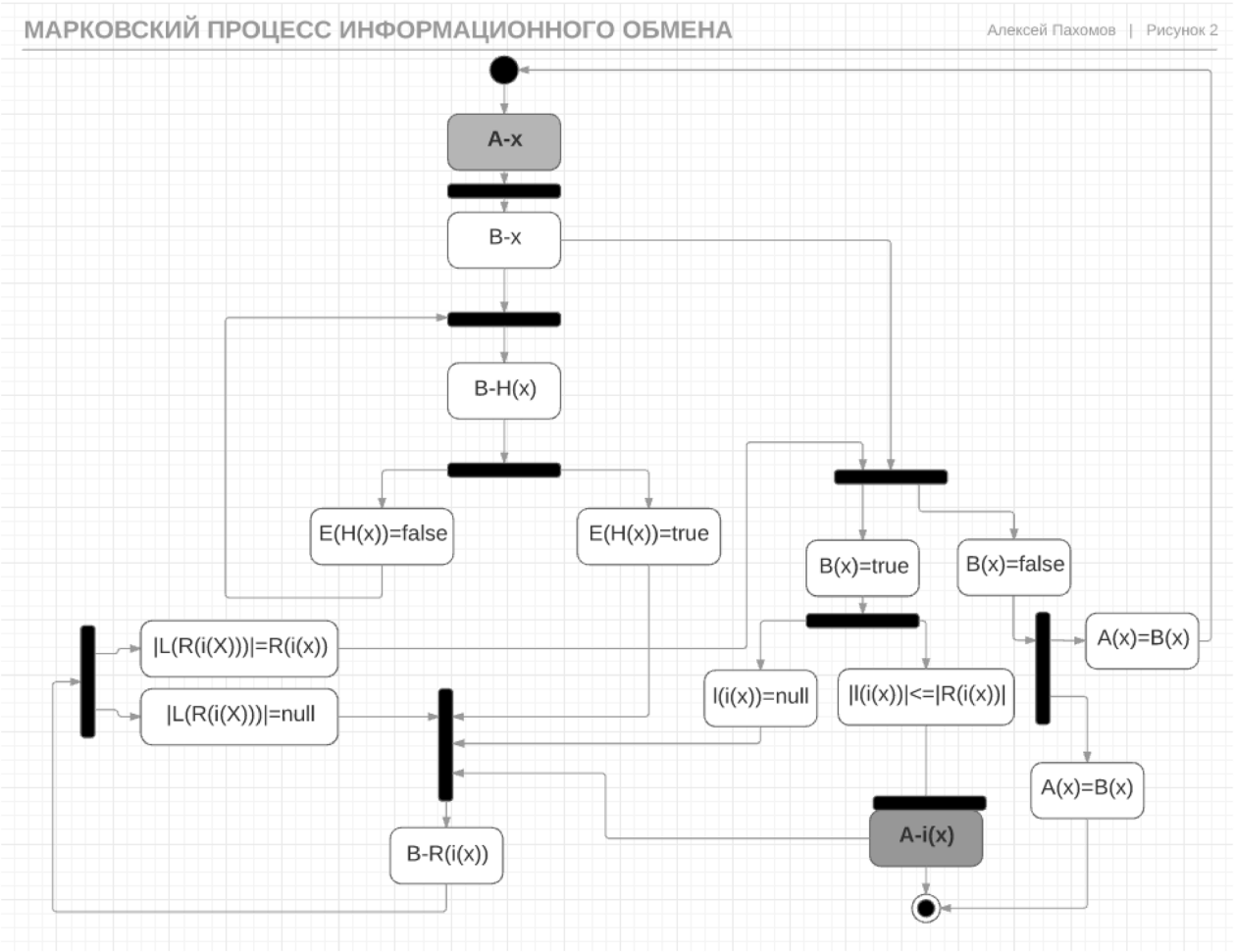


Рис.2. Формализованная модель запроса в стороннюю организацию

существует такой x’ из x для которого Identety(x)=Identity(x’));

- RequestInfo(i(x))|R(i(x)) - информация, необходимая для идентификации информации i(x);

- ReverseRequest(Request) - информация необходимая для выполнения предиката Enough();

- RequestReject(Request) – информацияобоснования Legal(i(x))=null;

- i(x) -информация об элементах гипотезы и дополнительных данных связанных с x обьектами;

- RequestNull(v) - информация о Search(v)=null;

- Legal|L(A,R(i(x))) - предикат, определяющий возможность легальной(соответствующей законодательству передаче ответа на запрос R(i(x));

- LegalInfo|I(i(v))-предикат, определяющий возможность легальной(соответствующей законодательству передаче данных;

- Enough(RequestHyposesis (v))|E(H(x)) - предикат, определяющий достаточность предоставленных данных для подготовки ответа;

- Search(Identity(v),B)|S(x) - функция поиска данных об объекте x в базе B, возвращающая информацию об объекте i(x).

Процесс запроса-ответа направлен на достижение следующего результата.

Если B обладает необходимой информацией(H(x)) и при этом для B предоставлена информация, необходимая для запроса, и сторона B имеет легальные основания передать данные A , то i(x) будут предоставлены A за время T. На основе приведенного параметрического описания разработана сеть Петри, использование которой позволяет рассчитать показатель эффективности применения предложенного алгоритма межведомственного обмена информацией – вероятность подтверждения правильной гипотезы за время t [2].

Разработанная в работе модель позволила провести оптимизацию процесса обмена информации и рассчитать прирост выигрыша во времени при использовании процедур предварительного запроса. Оценка прироста может быть осуществлена с помощью следующей формулы:

$$\Delta t = P_{AvailableInfo} * T_{exchange} * (1 + P_A/(1 - P_A)) + P_{NotAvailableInfo} * T_{exchange} - P_{AvailableInfo} * (t_{AnonymousExchange} * ((1+P_{error}/(1-P_{error})) + T_{exchange} * (P_{shell}/(1 - P_{shell}))) + P_{NotAvailableInfo} * t_{AnonymousExchange}$$

$$\Delta t = P_{AvailableInfo} * T_{exchange} + P_{error}/(1-P_{error}) * T_{exchange} + P_{NotAvailableInfo} * T_{exchange}$$

Из формулы видно, что эффективность обмена зависит от оценки вероятности отсутствия информации об объекте. Таким образом, исходя из реальных значений обмена данными, можно ценить необходимость внедрения запроса данных с нулевым знанием.

Литература

1. Кравчук, А.В. Модель процесса удаленного анализа защищенности информационных систем и методы повышения его результативности [Текст]/ А.В. Кравчук // Труды СПИИРАН. – М.:2015. -№1 (38). – с. 73-95.

2. Пахомова, А.С. Граф-модели процессов скрытия компьютерного шпионажа [Текст]/ А.С. Пахомова //Информационная безопасность –В.:2015.- Т.18. № 2. – С. 248-253.

3. Antsupov, A. Ya. Conflictology: the textbook for higher education institutions / A.Ya. Antsupov, A. I. Shilov. – 3rd prod., reslave. and additional – SPb.: St. Petersburg, 2007. – 591 p.

4.Voldstad, R. Community Detection on Last.fm Artist Data / R. Voldstad // – 2014.

5. Panagiotis, Karampelas. Techniques and Tools for Designing an Online Social Network Platform. New Hampshire: Hellenic American University / – M.: World, 1992. – 400 p.

ФГБОУ ВО «Воронежский государственный технический университет»
Voronezh State Technical University
Пан-Европейский Университет
Pan-European University

THE MODEL OF INTERAGENCY SHARING OF INFORMATION WITH RESTRICTED ACCESS

A.S. Pakhomova, A.P. Pakhomov, J. Vorisek

This paper presents structural parameterized model of intergovernmental information exchange. It was created for evaluation of effectiveness boost in informational exchange after implementation of anonymous exchange using zero-knowledge protocol. It prevents risks of information leak and reduces organizational and human costs of the process while conforming legal laws protecting personal information

Key words: information with restricted access, interagency information sharing, zero-knowledge protocol

УДК 004.622

МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ ОЦЕНКИ ЭФФЕКТИВНОСТИ ПРОЦЕССОВ РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ С ОГРАНИЧЕННЫМ ДОСТУПОМ

А.С. Пахомова, А.П. Пахомов, Е. Ружицкий

Статья посвящена вопросам обеспечения эффективности информационных процессов с ограничениями на распространение информации в области деятельности по противодействию отмывания доходов и финансированию терроризма. Предлагается решение задачи улучшения взаимодействия специалистов по финансовым расследованиям на основе применения формальных моделей take-grant, Штейнера, моделей теории игр, а также протоколов передачи информации с нулевым знанием

Ключевые слова: модель информационного обмена, условия доступа, модель take-grant, модель Штейнера, передача информации с нулевым знанием

В настоящее время информационные процессы, реализуемые с использованием информационных технологий автоматизированной обработки данных, становятся основой функционирования все новых объектов и целых сфер деятельности, в том числе связанных с обеспечением безопасности государств. При этом достаточно большая группа информационных процессов предназначена для обеспечения обмена информацией органами государственного управления и организациями с соблюдением ограничений, накладываемых законодательством о порядке распространения информации, а также ограничений, связанных с интересами отдельных организаций. Эффективность таких информационных процессов оказывается меньше требуемой вследствие того, что при проектировании информационных систем не учитывались упомянутые выше ограничения, в полной мере проявляющиеся в процессе эксплуатации систем. В связи с этим одной из важнейших проблем, решаемых в настоящее время, является модернизация информационных процессов с учетом правовых ограничений на предоставление информации. Эта проблема решается разными путями на государственном и ведомственном уровнях в ведущих зарубежных странах [1 - 3]. Имеется также ряд научных публикаций, посвященных внедрению новых технологий информационного взаимодействия [4, 5]. В то же время, в области финансовых расследований в сфере ПОД/ФТ эта проблема стоит достаточно остро. Поэтому тема настоящей статьи, посвященной рассмотрению математических методов и реализующего их программного обеспечения, предназначенного для обоснования рациональных информационных процессов финансовых расследований (ФР), представляется вполне актуальной. Задачей настоящей статьи является описание совокупности математических методов и программного обеспечения, которые позволяют осуществить модернизацию информационных процессов взаимодействия органов и организаций системы противодействия отмывания доходов и финансирования терроризма (ПОД/ФТ) в процессе проведения ФР. Эти процессы имеют как общие характеристики с другими информационными процессами, которые и вызвали существование проблемы модернизации, так и специфические характеристики, которые позволяют предложить решение этой проблемы для рассматриваемого объекта исследования.

В основу решения данной задачи положен

Пахомова Анна Степановна — ВГТУ, канд. техн.наук, доцент, e-mail: mnac@comch.ru
Пахомов Алексей Павлович – РГУ нефти и газа им. Губкина, аспирант, e-mail: apal12@mail.ru
Ружицкий Евгений – Пан-Европейский Университет (Словакия), к.т.н., декан, доцент,
e-mail: eugen.ruzicky@paneurouni.com

системный подход к анализу системы ПОД/ФТ как организации, целевой функцией которой является снижение риска осуществления деятельности по ОД и ФТ. Основной процесс ПОД/ФТ, в результате которого снижается риск осуществления деятельности по ОД/ФТ, представлен на рис.1.

Рис. 1. Технологическая схема процесса снижения риска ПОД/ФТ

Необходимо отметить, что научное обоснование эффективности ПОД/ФТ до настоящего времени в основном проводилось методами исследования экономических и юридических характеристик системы ПОД/ФТ и весьма ограниченно – методами исследования технических характеристик этой системы, обеспечивающих возможность реализации экономических и юридических мер.

Основным направлением повышения эффективности и результативности ФР за счет совершенствования характеристик технических систем, применяемых для проведения таких расследований, в настоящее время остается разработка математических методик первичной обработки данных финансового мониторинга, реализованных в виде программных продуктов информационных систем, используемых для оценки риска в отношении конкретных объектов финансового мониторинга.

Такие методики позволяют повысить эффективность решения одной частной задачи (функция 2 на рис. 1) по отношению к ограниченным группам объектов.

При этом весь процесс расследования (рис. 1), составляющий основу деятельности по предотвращению риска, не рассматривается.

Математически снижение риск в результате действий по ПОД/ФТ можно представить формулой

$$\Delta R=\Delta R_к+\Delta R_р,$$

где $\Delta R_к$ – снижение риска за счет выявления организаций и начала проведения против них расследований;

$\Delta R_р$ - снижение риска за счет создания оснований для пресечения действий организаций.

Очевидно, что конечный результат ПОД/ФТ достигается только в результате завершения расследований (функция 3 и на рис. 1).

Причем снижение риска тем больше, чем быстрее проводятся расследования против организаций, действия которых наносят максимальный ущерб.

Реализуемый в настоящее время подход, направленный на увеличение первой составляющей снижения риска $\Delta R_к$, позволяет снизить риск в весьма ограниченных пределах, что приводят к необходимости дополнения ИТ ПОД/ФТ программного обеспечения, реализующего методический аппарат оптимизации процесса проведения расследований, обеспечивающий снижение риска на величину $\Delta R_р$.

В качестве основной характеристики эффективности процесса расследования примем время проведения расследования.

Время проведения расследования, в свою очередь, зависит от возможности получить необходимые данные.

Таким образом, основным процессом, подлежащим моделированию, является

процесс расследования как процесс получения данных.

В ФР этот процесс имеет следующие особенности:

данные, необходимые для проведения расследования, содержатся в различных хранилищах (в специализированных базах данных субъекта расследования сторонних организаций, в открытых информационных ресурсах); информационный процесс расследования является последовательностью событий, состоящих в получении очередного фрагмента данных, их переработке совместно с ранее полученными данными и формировании запроса на получение следующего фрагмента данных;

процесс является активным, т.к. каждый очередной запрос формируется в зависимости от результатов обработки полученных данных; данные могут быть получены несколькими путями: путем запросов, извлечения из баз данных, поиска в открытых источниках, каждому из которых соответствует определенный режим доступа;

данные, необходимые для решения отдельных задач расследования, могут быть представлены в форме подтверждения факта их наличия или в содержательной форме. Для того, чтобы оптимизировать процесс проведения расследования, необходимо учесть влияние на время расследования следующих факторов: режимов доступа субъекта проведения расследования к различным фрагментам данных об объекте расследования; возможности получить отказ в получении данных по запросу или существенного увеличения времени получения ответа; возможности не получить данные по причине их отсутствия.

Для решения поставленной задачи оптимизации процесса расследования объект исследования представляется формальной моделью графа состояний процесса расследования.

Переходы по графу состояний расследования происходят параллельно с переходами по узлам распределенной базы данных, которая описывается установленными правами доступа.

Решение задачи расследования состоит в переходе из начального состояния I0 в конечное состояние Ip при условии прохождения I узлов, таких, что сведения $\sum D_i$., накопленные при их прохождении с учетом обработки полученных данных Di, соответствуют одному из некоторых заранее заданных перечней (присущему определенному типу правонарушения).

Указанная модель реализуется следующей математической моделью $T=\sum t_i$.

Оптимизация процесса расследования состоит в нахождении минимума $T=\sum t_i$ по всем возможным путям перехода из состояния в состояние. Ограничением является достижение $\sum D_i$ =Ip. Эта задача относится к известному классу задач поиска кратчайшего пути на ориентированном графе. В отличие от классической задачи коммивояжера данная задача по своим параметрам соответствует задаче Штейнера. Ее особенностями являются: зависимость времени перехода от стоимости данных; вероятностный переход из состояния в состояние, т.к. данные могут быть не предоставлены; положительная и отрицательная длина ребра.

Решение данной задачи является алгоритмическим, а время вычисления может быть достаточно большим, т.к. оно экспоненциально зависит от размерности. Для того, чтобы уменьшить время вычислений и гарантировать практическую реализуемость оптимизации необходимо наложить дополнительные ограничения на возможности переходов, исключив переходы с малой вероятностью получения данных.

Для этого модель Штейнера дополняется моделями оценки вероятности переходов. В качестве таких моделей в данной работе предлагаются две: параметрическая и информационная. Параметрическая модель позволяет количественно оценить вероятность получения данных, в то время как информационная – заранее, до получения данных, установить их наличие.

Для всех выбранных моделей существует программное обеспечение, что обеспечивает их реализуемость средствами информационных технологий и быстрое внедрение в практику.

Рис. 2. Структура формализованной модели, реализующей алгоритм оптимизации времени расследования

Литература

1. National Information Exchange ModelSharing. [Электронный ресурс] – Режимдоступа: http://www.sparxsystems.com/ domains/niem/national-information-exchange-model-niem-solution-with-enterprise-architect.html1.

2. Tackling Money Laundering: Towards a New Model for Information Sharing. [Электронный ресурс] – Режим доступа: http://www.academia.edu/16089685/Tackling_Money_Laundering_Towards_a_New_Model_for_Infor- mation_ Sharing.

3. Multi-Agency Working and Information Sharing Project. Early Findings.[Электронный ресурс] – Режим доступа: https://www.gov.uk/ government/uploads/ system/uploads/ attachment_data/file/225012/ MASH_Product.pdf.

4. R. Sandhu, R. Krishnan, G. B. White. Towards Secure Information Sharing Models for Community Cyber Security.[Электронный ресурс] – Режим доступа: http://engineering.utsa.edu/~krishnan/ conferences/ collaboratecom10.pdf.

5. J. L. Hernandez-Ardieta, J.E. Tapiador. Information sharing models for cooperative cyber defence. [Электронный ресурс] – Режим доступа: https://www.researchgate. net/ publication /255741958_ Ifor.

6. .Johnson, S. Entropic origin of disassortativity in complex networks / S. Johnson, J.J. Torres, J. Marro, M.A. Muñoz / Physical Review Letters. – 2010. – 4 p.

7. Kauai, HI: IEEE. Freeman, L. C. (1979). Centrality in social networks conceptual clarification. Social Network. 1(3). P. 215–239

ФГБОУ ВО «Воронежский государственный технический университет»
Voronezh State Technical University
Пан-Европейский Университет
Pan-European University

MATEMATICAL METHODS AND MODELS FOR EFFICIENCY EVALUATION OF INFORMATION WITH RESTRICTED ACCESS SHARING

A.S. Pakhomova, A.P. Pakhomov, E. Ruzicky

The article is devoted to some questions of information processes efficiency under the constraints of information sharing in the field of countermeasures to money laundering and financial terrorism. The solving of the problem of robust decision making among professionals of financial investigation is presented on the base of formal models take-grant, Steiner problem, game theory models and zero- knowledge transfer protocols

Key words information sharing model restricted access take-grant model Steiner problem, game theory models, zero- knowledge transfer protocols

УДК 004.056.57

НАУЧНО-ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО СНИЖЕНИЮ ЭПИДЕМИЧЕСКИХ РИСКОВ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ С ОДНОРОДНЫМИ КЛАСТЕРАМИ

Е.Н. Пономаренко, Ю. Штефанович

В статье приводятся научно обоснованные практические рекомендации по снижению эпидемических рисков в корпоративной информационно-телекоммуникационной сети с гомогенными кластерами

Ключевые слова: информационно-телекоммуникационная сеть, гомогенные кластеры, эпидемический процесс, риск

Объектом исследования является информационно-телекоммуникационная сеть с ярко выраженной гомогенной кластеризацией, в которой имеет место распространение компьютерных вирусов [1-3].

Для разработки рекомендаций по снижения эпидемического риска в рассматриваемой сети введем следующие параметры: N_k – количество вершин k -го кластера; $P_{кз}$ - вероятность единичного (одного элемента) заражения; K_k - количество смежных с другими кластерами вершин -го кластера (нередко оно равно и количеству смежного с k-м кластеров). В первом приближении (в рамках биноминального распределения) ожидаемое количество вирусованных в k -м кластере вершин на первом этапе эпидемического процесса будет равно [3]:

$$I\,[1] = [P_{кз}\,N_k], \qquad (1)$$

где [.] - оператор определения целой части.

При этом количество зараженных смежных с другими кластерами вершин может быть определено по аналогии

$$I_k\,[1] = [P_{кз}\,K_k] = S_2, \quad (2)$$

Оно и определяет количество смежных с k -м кластеров, в которые поступила инфекция на втором этапе эпидемического процесса. Назовем их вторично инфицированными кластерами. Тогда на втором этапе эпидемического процесса суммарное ожидаемое количество инфицированных вершин будет

$$I\,[2] = \sum_{i=1}^{S}[P_{iз}\,N_i]. \qquad (3)$$

Подобную процедуру можно повторять и дальше для последующих этапов процесса [3]. В результате суммарное количество инфицированных элементов будет представлять собой ряд

$$I = [P_{кз}\,N_k] + \sum_{i=1}^{S2}[P_{iз}\,N_i] + \\ + \sum_{j=1}^{S3}\left[P_{jз}\,N_j\right] + \cdots. \qquad (4)$$

который на практике резко убывающий в своих членах (начиная с третьего), так как обычно уже $S_2 > S_3$, да и антивирусные средства делают свое дело.

Основываясь на выражениях (1) – (4), можно сделать вывод, что произведение $[P_{кз}\,N_k]$ может быть использовано как некий инженерный критерий оценки эпидемического риска для гомогенного кластера. К примеру, уравняв эти произведения для исследуемых кластеров (с помощью $P_{кз}$), выражение (4) можно резко упростить

$$I = r\,(1 + S_2 + S_3 + \cdots), \qquad (5)$$

где r будет меньше наименьшего количества вершин из рассматриваемых кластеров.

Отсюда следует, что

$$P_{кз} = \frac{r}{N_k}, \qquad (6)$$

Выражение (6) может быть положено в основу рекомендации по выбору антивирусных средств для гомогенных кластеров [3]. Иными словами, чем больше элементов в кластере, тем более эффективное (с меньшей вероятностью заражения) антивирусное средство следует применять. Тогда общий эпидемический риск сети примерно будет равен

Пономаренко Елена Николаевна – соискатель каф. СИБ, e-mail: mnac@comch.ru
Штефанович Юрий - Пан-Европейский Университет (Словакия), к.т.н., зам. декана,
e-mail: juraj.stefanovic@paneurouni.com

$$Risk = \frac{r}{N}\,(1 + S_2 + S_3 + \cdots), \quad (7)$$

где N - количество вершин всей сети; $r = [P_{кз}\, N_k]$ для всех участвующих в эпидемическом процессе гомогенных кластеров.

Выражение (7) может быть рекомендовано для практических действий администрации сети по снижению эпидемических рисков [2].

Можно предложить следующие общие рекомендации по организационно-технической минимизации рисков исследуемой сети при распространении в ней вируса:

1. Необходимо установить антивирусные средства (АВС) на всех компьютерах сети. Следует проводить обновление баз сигнатур вредоносного ПО каждые 4 часа рабочего времени.

2. Уместно обновлять системное ПО не реже одного раза в день до начала эксплуатации компьютеров работниками.

3. В сети необходимо установить систему обнаружения вторжений (СОВ) на центральных узлах сети или в ее кластерах. Соответственно калибровку СОВ следует проводить не реже одного раза в неделю в свободное от эксплуатации время.

4. Уместно установить межсетевые экраны (МЭ) с блокировкой внешних ресурсов, не относящихся к деятельности предприятия. Соответственно повторную настройку МЭ следует проводить через один рабочий день и при добавлении новых внешних ресурсов в список блокировок. Необходимо также внедрение системы динамической фильтрации и проведение ее повторной настройки аналогично настройке МЭ.

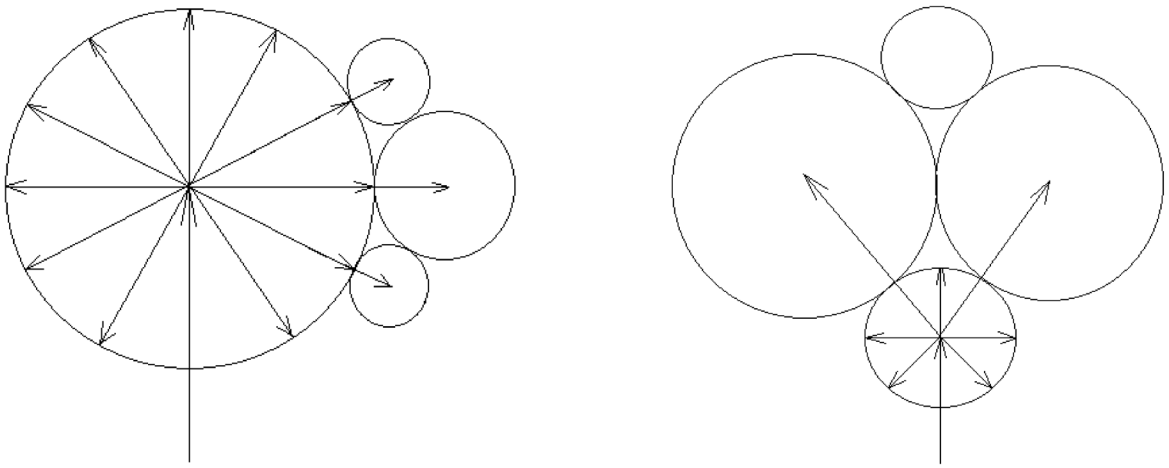
5. В начале каждой рабочей недели следует проводить инструктаж по вирусной безопасности среди сотрудников для повышения общего уровня «грамотности» пользователей в сфере информационной безопасности. Смена паролей пользователей должна производиться не реже одного раза в квартал.

6. Уместно проведение единовременной сегментации сети с учетом процессов ее кластеризации. Анализ и коррекцию структуры сети следует проводить каждый раз при добавлении в нее новых элементов и кластеров.

7. Необходимо единовременное установление минимальных необходимых привилегий пользователей в рамках существующих кластеров. Систематически необходимо проводить проверку и переопределение привилегий пользователей (не реже одного раза в рабочую неделю) и при изменении привилегий у существующих пользователей.

Предложенные рекомендации частично нашли свое применение в сетях с гомогенными кластерами.

Наряду с общими рекомендациями, администраторов значительно интересует параметрическая настройка ее элементов и кластеров. В этой связи рассмотрим обобщенно схему эпидемического процесса в сетях с гомогенными кластерами.



а) б)

Рис. 1. Иллюстрация интенсивности эпидемического процесса для вариантов заражения высокостепенного (а) и низкостепеннего (б) гомогенных кластеров сети

На рис.1 схематично проиллюстрирована география развития эпидемического процесса в сети с гомогенными кластерами для двух крайних вариантов заражения вирусом, когда:

а) заражение начинается с кластера, элементы которого гомогенны и имеют высокую степень (большое количество контактов внутри сети);

б) инфицирование сети реализуется через ее кластер с гомогенными элементами, имеющими малое количество внутрисетевых контактов.

Как видно из рисунка, последствия рассматриваемых вариантов различны.

Это связано с тем, что:

- эпидемический процесс в сетях с высокой степенью кластеризации во многом развивается внутри первично зараженного гомогенно кластера, и вне его инфекция диффундирует через немногочисленные (ввиду кластеризации) смежные элементы (вершины) сети;

- низкостепенные гомогенные кластеры обычно имеют незначительные размеры и минимальные контакты с другими кластерами, что также ограничивает развитие эпидемии в сети.

Такие выводы можно сделать при равновеликой вероятности единичного заражения элемента сети, из чего следует, что «уравнивание» вероятностей единичного заражения имеет место на практике обычно в рамках отдельно взятого гомогенного кластера, и установка различных антивирусных средств в разных кластерах сети открывает возможность снижения эпидемических рисков.

Представленная (1)-(7) иллюстрация базируется на приближении, когда количество успешных вирусных атак, имеющих место при внутрикластерных контактах однородных элементов, соответствует математическому ожиданию биномиального распределения.

Однако этот факт не следует рассматривать как ограничение предлагаемой методики настройки, т.к. представляется возможным сделать обобщение и на другие виды распределений через их ожидания:

$$I = Q_k + \sum_i Q_i , \qquad (8)$$

Q_k – ожидание заражений в кластере k;

Q_i – ожидания заражений в кластерах, контактирующих с первично зараженными ($k \neq i$).

К примеру, для пуассоновского приближения $P_{n,k}(k) = \frac{\lambda^k}{k!} e^{-\lambda}$ могут быть найдены ожидания λt определенного количества k успешных вирусных атак в кластере в интервале времени t с постоянной интенсивностью λ.

Подводя некоторые итоги в проведенных выше рассуждениях, уместно сделать следующие выводы по сформулированным рекомендациям:

1. Подобная стратегия снижения внутрикластерных рисков путем настройки АВС представляется вполне рациональной для регулирования эпидемических процессов в сетях с гомогенными кластерами. Позиции здесь могут быть дополнительно усилены, если между кластерами можно установить МЭ, где межкластерная диффузия вирусов, очевидно, будет ослаблена, и эпидемический процесс в основном замкнется внутри отдельно инфицированного кластера.

2. Внутрикластерный эпидемический процесс также может быть значительно заторможен за счет адаптивной подстройки АВС с учетом полученной от СОВ информации. В идеале это хотелось бы иметь автоматически и в реальном масштабе времени.

Однако в целях экономии такую процедуру возможно проводить в рамках периодической профилактики на основе обобщения администрацией сети зарегистрированных СВО обновлений сетевых вирусных атак на кластеры (за некоторый установленный интервал времени).

3. Практически наблюдаемая закономерность кластеризации корпоративных сетей в большинстве случаев свидетельствует о том, что с ростом коэффициента кластеризации снижается уровень стабилизации эпидемических рисков ввиду уменьшения количества N_k вершин кластера, их степени и количества K_k смежных с другими кластерами вершин (при фиксированном общем количестве пользователей N) сети.

Подобно переборкам между отсеками корабля, усиливающим его плавучесть при появлении пробоин в борту, МЭ между кластерами сети увеличивают её живучесть при вирусных атаках, поражающих эти кластеры. Следует заметить, что подобная кластеризация увеличивает также защиту конфиденциальности циркулирующей в кластерах информации, снижая её утечки и несанкционированный доступ.

Литература

1. Калашников, А.О. Атаки на информационно-технологическую инфраструктуру критически важных объектов: оценка и регулирование рисков [Текст]: монография / А.О. Калашников, Е.В. Ермилов, О.Н. Чопоров, К.А. Разинкин, Н.И. Баранников; под ред. чл.-корр. РАН Д.А. Новикова. – Воронеж: Издательство «Научная книга», 2013. – 160 с.

2. Остапенко, А.Г. Формализация процесса управления рисками в информационно-технологической инфраструктуре критически важного объекта [Текст] / А.Г. Остапенко, А.О. Калашников, Е.В. Ермилов, Н.Н. Корнеева // Информация и безопасность. – 2014. – Т. 17. – Вып. 2. – С. 164-179

3. Радько, Н.М. Некоторые оценки рисков, шансов и живучести сетей в условиях информационных атак вирусного характера [Текст] / Н.М. Радько, Л.В. Паринова, Ю.Г. Пастернак, К.А. Разинкин, Н.М. Тихомиров // Информация и безопасность. – 2013. – Т. 16. – Вып. 4. – С. 500-502.

ФГБОУ ВО «Воронежский государственный технический университет»
Voronezh State Technical University
Пан-Европейский Университет
Pan-European University

SCIENTIFIC AND PRACTICAL RECOMMENDATIONS ON REDUCING EPIDEMIC RISKS IN INFORMATION-TELECOMMUNICATION NETWORKS WITH HOMOGENEOUS CLUSTERS

E.N. Ponomarenko, Yu. Stefanovic

The article provides scientifically sound practical recommendations for reducing epidemic risks in a corporate information and telecommunications network with homogeneous clusters

Key words: information-telecommunication network, homogeneous clusters, epidemic process, risk

УДК 004.932

МОДЕЛЬ АНАЛИЗА ИЗОБРАЖЕНИЙ НА ОСНОВЕ МЕТОДА НЕЧЁТКОЙ КЛАСТЕРИЗАЦИИ

Д.В. Лакомов, В.В. Алексеев, Ю.В. Минин, Ю.В. Кулаков, Г.Н. Нурутдинов

В работе рассматриваются алгоритм нечётких С–средних, его параметры, а также его применение для анализа изображений

Ключевые слова: анализ, изображения, нечёткая логика, кластеризация, метод С-средних

В современных системах управления, принятия решений и обработки информации распознавание (идентификация) изображений затруднено тем, что воздействие негативных внешних и внутренних факторов вносит в этот процесс неопределённость, приводящую к размытости изображений. В связи с этим применяются алгоритмы и модели, позволяющие уменьшить влияние неопределённости при анализе изображений.

Анализ изображения – это процесс выделения нужной информации из изображения с помощью автоматических систем. В процессе анализа изображений обязательно возникает необходимость их сегментации, т. е. разделения пикселей изображения по группам в соответствии с определенными признаками. Методы сегментации можно представить как формализацию понятия извлечения объекта из фона или понятий, связанных с градиентом яркости [1].

Необходимо отметить, что при решении задач кластеризации наиболее популярны алгоритмы, которые основаны на оптимальном разбиении множества данных на кластеры. Подобные алгоритмы направлены на группировку данных в кластеры таким образом, чтобы целевая функция алгоритма разбиения достигала экстремума (минимума).

Лакомов Денис Вячеславович – ТГТУ, аспирант.
Алексеев Владимир Витальевич – ТГТУ, доктор техн. наук, профессор.
Минин Юрий Викторович - ТГТУ, канд. техн. наук, доцент.
Кулаков Юрий Владимирович - ТГТУ, канд. техн. наук, доцент.
Нурутдинов Геннадий Нурисламович - ТГТУ, канд. техн. наук.

Среди алгоритмов кластеризации стоит выделить алгоритм нечётких С – средних.

Исходной задачей для кластеризации является матрица наблюдений X, где l – число объектов, а n – число признаков (наблюдений) для каждого объекта.

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{l1} & \cdots & x_{ln} \end{pmatrix}$$

Задача кластеризации состоит в разбиении множества объектов на группы (кластеры) «похожих» между собой объектов. В n-мерном метрическом пространстве признаков мерой «сходства» двух объектов является расстояние между ними. При нечёткой кластеризации каждый объект принадлежит с различной степенью нескольким кластерам (иногда всем кластерам) [2].

Кластерная структура задаётся матрицей принадлежности M, где l – число объектов, c – число кластеров, а m_{ij} – степень принадлежности j – го элемента i – му кластеру.

$$M = \begin{pmatrix} m_{11} & \cdots & m_{1l} \\ \vdots & \ddots & \vdots \\ m_{c1} & \cdots & m_{cl} \end{pmatrix}$$

Матрица принадлежности должна удовлетворять условиям:

1) $m_{ij} \in [0,1] \quad i = \overline{1,c}, j = \overline{1,l}$,

2) $\sum_{i=1}^{c} m_{ij} = 1, \ j = \overline{1,l}$, т.е. каждый объект должен принадлежать всем кластерам,

3) $0 < \sum_{j=1}^{l} m_{ij} < l$, т.е. не должно быть пустых кластеров и кластеров, содержащих все элементы [3].

Для оценки качества разбиения используется критерий разброса, показывающий сумму расстояний от объектов до центров кластеров с

соответствующими степенями принадлежности J:

$$J = \sum_{i=1}^{c}\sum_{j=1}^{l}(m_{ij})^{\omega} d(v_i, x_j),$$

где $d(v_i,x_j)$ – расстояние в заданной метрике между j – м объектом $x_j = (x_{j1}, x_{j2}, \ldots, x_{jn})$ и i – м центром кластера $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})$, ω – экспоненциальный вес, определяющий размытость кластеров. Обычно применяется значение параметра $\omega = 2$ [4].

Для кластеризации изображения необходимо иметь критерий, чтобы сравнивать объекты для их разбиения на кластеры. Таким критерием является расстояние между объектами.

Рассмотрим меры расстояния, применимые в предлагаемой модели:

1) Евклидово расстояние – наиболее распространенное расстояние. Оно является геометрическим расстоянием в многомерном пространстве. Эта мера может применяться для вычисления расстояния между объектами, описанными количественными, качественными и дихотомическими признаками. Ее использование целесообразно, когда признаки однородны по смысловой нагрузке и одинаково важны для решаемой задачи.

$$D(v_i, x_j) = \sqrt{\sum_{k=1}^{n}(v_{ik} - x_{jk})^2} \quad i = \overline{1,c}, j = \overline{1,l}$$

В случаях, когда требуется придать большее значение более отдаленным друг от друга объектам используется квадрат евклидова расстояния.

$$D(v_i, x_j) = \sum_{k=1}^{n}(v_{ik} - x_{jk})^2 \quad i = \overline{1,c}, j = \overline{1,l}$$

2) Расстояние городских кварталов (манхэттенское расстояние). Это расстояние является средним разностей по координатам. В большинстве случаев эта мера расстояния приводит к таким же результатам, как и для обычного расстояния Евклида. Однако для этой меры влияние отдельных больших разностей (выбросов) уменьшается (так как они не возводятся в квадрат) [5].

$$D(v_i, x_j) = \sum_{k=1}^{n}(|v_{ik}| - |x_{jk}|) \quad i = \overline{1,c}, j = \overline{1,l}$$

3) Расстояние по Камберру. Предварительное задание весовых коэффициентов в формулах, определяющих расстояния, требует наличия определенной априорной информации и не всегда может быть сделано оптимальным образом. Поэтому особый интерес представляют расстояния, в которых заложена идея выравнивания весов слагаемых от различных компонент, если они существенно отличаются по своим абсолютным значениям. Примером такого расстояния является расстояние по Камберру.

$$D(v_i, x_j) = \sum_{k=1}^{n}\frac{|v_{ik} - x_{jk}|}{|v_{ik} + x_{jk}|} \quad i = \overline{1,c}, j = \overline{1,l}$$

4) Расстояние Хемминга**.** Данная мера наиболее часто используется для определения различий между объектами, задаваемыми дихотомическими признаками и интерпретируется как число несовпадений значений признаков у рассматриваемых объектов. Для дихотомических признаков она соответствует квадрату евклидова расстояния. Так же как и для евклидова расстояния, может применяться взвешенное расстояние Хэмминга [6].

$$D(v_i, x_j) = \frac{\sum_{k=1}^{n}(v_{ik} - x_{jk})^2}{n} \quad i = \overline{1,c}, j = \overline{1,l}$$

Алгоритм нечётких С – средних применяется для решения задачи нахождения матрицы *M*, минимизирующей критерий *J* [7].

Алгоритм С – средних проходит в несколько этапов:

1) Случайным образом генерируется матрица принадлежности *M.*

2) Построение матрицы координат центров кластеров – *V*, элементы которой определяются по формуле (1):

$$V = \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{c1} & \cdots & v_{cn} \end{pmatrix},$$

$$v_{ik} = \frac{\sum_{j=1}^{l}(m_{ij})^{\omega} x_{jk}}{\sum_{j=1}^{l}(m_{ij})^{\omega}} \quad k = \overline{1,n}, i = \overline{1,c} \qquad (1)$$

3) Вычисление расстояния от объектов множества *X* до центров представленных кластеров *V*. В данной статье рассмотрим Евклидову метрику

$$d_{ij} = \sqrt{\sum_{k=1}^{n}(v_{ik} - x_{jk})^2}, \; i = \overline{1,c}, j = \overline{1,l}$$

4) Перерасчёт элементов матрицы принадлежности M.

$$m_{ij} = \frac{1}{(d_{ij})^{\frac{2}{\omega-1}}}, \text{при } d_{ij} > 0$$

$$m_{ij} = \begin{cases} 1, \text{если } p = i \\ 0, \text{если } p \neq i \end{cases}, \text{при } d_{ij} = 0, \; p = \overline{1,c} \qquad (2)$$

Система (2) обозначает, что если расстояние $d_{ij} = 0$, то для j-го элемента $m_{ij} = 1$ для i – го кластера, а для всех остальных кластеров этой точки $m_{ij} = 0$.

5) Для каждого $j = \overline{1,l}$ суммируем элементы m_{ij}

6) Нормируем элементы m_{ij} , для этого разделим каждый m_{ij} из п.4 на соответствующую сумму из п.5 (в случае $m_{ij} = 1$, элементы нормированны по умолчанию)

7) Проверка условия $|J - J^*| < \varepsilon$, где J^* - критерий разброса предыдущей итерации алгоритма.

При обработке изображения с помощью алгоритма С-средних целесообразно использовать цветовую модель HSV. Модель HSV - цветовая модель, в которой координатами цвета являются: цветовой тон насыщенность и яркость. Модель HSV часто используется в программах компьютерной графики, так как она обеспечивает возможность явного задания требуемого оттенка цвета. Среди прочих используемых в настоящее время моделей, эта модель отражает физические свойства цвета и наиболее точно соответствует способу восприятия цвета человеческим глазом. Модель HSV также позволит уменьшить затраты ресурсов на обработку изображения [8].

Результат работы алгоритма нечёткой кластеризации С – средних с параметрами ε = 0,001, ω=2, c=7 представлен ниже:

Рис. 1. Исходное изображение

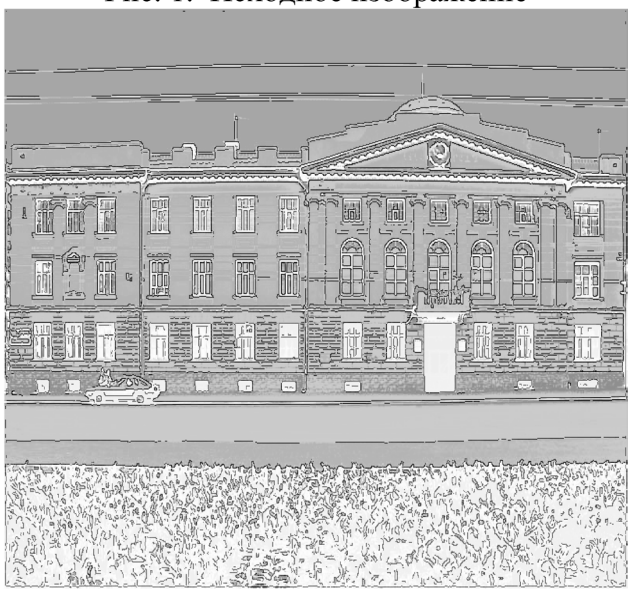


Рис. 2. Обработанное изображение

Процедура кластеризации предоставляет возможность детального анализа изображений, необходимого для определения их атрибутов. Комплексная процедура исследования (с применением более широкого аппарата) позволит фиксировать небольшие смещения или изменения размеров запечатленных на изображении объектов. Нечёткая кластеризация по методу С-средних – это удобный подход для выделения объектов на изображении, тесно связанных с заданными кластерами. Применяя его в комбинации с различными цветовыми моделями, метриками расстояний между объектами, можно найти близкое к оптимальному решение задачи кластеризации.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта No 17-48-680254.

Литература

1. Y. Yong, Z. Chongxun and L. Pan, “A Novel Fuzzy C-Means Clustering Algorithm for Image Thresholding” // Measurement Science Review, vol. 4, no.1, 2004.

2. R. Ravindraiah, K. Tejaswini, “A Survey of Image Segmentation Algorithms Based On Fuzzy Clustering” // IJCSMC, Vol. 2, Issue. 7, July 2013, pg.200 – 206.

3. Тараскина А.С., Черемушкин Е.С. Обработка микрочипированных данных с помощью алгоритма нечёткой кластеризации // Вычислительные методы и программирование: новые вычислительные технологии. 2006. №2 (7)

4. Мамедов А.С. Применение нечёткой кластеризации для детального анализа цветных изображений // Приволжский научный вестник. 2012,№ 1 (5)

5. Гороховатский В.А. Метрики на множествах ключевых точек изображений.// Бионика интеллекта. 2008. № 2 (69).

6. Виро О. Я., Иванов О. А., Нецветаев Н. Ю., Харламов В. М. Элементарная топология. М.: МЦНМО, 2010. 352 с.

7. Громов Ю.Ю. Методология дистанционной оценки пространственных распределений оптико-теплофизических параметров объектов, замаскированных под поверхностью грунта / В.В. Алексеев, Ю.Ю. Громов, Ю.А. Губсков, И.Н. Ищук. - М.: ООО «Научтехлитиздат», 2014. - 284 с.

8. Шикин Е. В., Боресков А. В. Компьютерная графика. Полигональные модели. – М.: ДИАЛОГ-МИФИ, 2000. – 464 с.

Тамбовский государственный технический университет
Tambov state technical University

MODEL OF IMAGE ANALYSIS BASED ON FUZZY CLUSTERING

D.V. Lakomov, V.V. Alekseev, Yu.V. Minin, Yu.V. Kulakov, G.N. Nurutdinov

The paper considers the algorithm of fuzzy C–means, its parameters, and its application to image analysis

Keywords: analysis, images, fuzzy logic, clustering method C-means

УДК 004.056:061.68

ОДИН ИЗ ПОДХОДОВ К АВТОМАТИЗАЦИИ РАСПРЕДЕЛЕНИЯ РАБОЧИХ ПРОЦЕССОВ В ГИБРИДНОЙ СРЕДЕ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ

М.А. Попов, А.В. Царегородцев

В статье рассматривается метод, позволяющий на основании требований безопасности распределить критически важные активы организации в гибридной среде облачных вычислений, на основании карты размещений облачных сервисов между доступными компонентами облачной среды

Ключевые слова: информационная безопасность, облачные вычисления, гибридная облачная среда, рабочий процесс, транспортный сервис

Достижение целей информационной безопасности организации, является ключевым фактором для принятия решений об услугах аутсорсинга информационных технологий и, в частности, для принятия решения о миграции информационных активов организации на различные модели предоставления облачных сервисов [1].

Для построения формализованной модели безопасности рабочего процесса в среде облачных вычислений в качестве основы будем использовать модель децентрализованного управления метками безопасности с использованием ключевых элементов теории графов [2].

Представим процесс обработки данных в виде следующих элементов.

1. Задачи. Основной блок информационного потока.

2. Граф потока управления. Множество всех возможных путей исполнения процесса, представленное в виде графа.

3. Субъекты. Под субъектами будем понимать активные облачные сервисы, запущенные от имени пользователя.

4. Элементы данных. Элементы данных – записи в хранилище данных или файлы. Субъект выполняет задачу путём создания или использования уже существующих элементов данных.

Царегородцев Анатолий Валерьевич – Московский государственный лингвистический университет, д-р техн. наук, профессор, e-mail: academic_tsar@mail.ru
Попов Максим Анатольевич – Институт информационных наук и технологий безопасности РГГУ, аспирант, e-mail: maxmax@bk.ru

Процесс обработки данных представим в виде ориентированного графа, облачные сервисы и данные которого будут представлены в виде вершин. Примем за основу, что облачный сервис потребляет от нуля до нескольких элементов данных и создаёт один или более элементов новых данных. Рёбра графа представляют собой зависимость обрабатываемых данных.

1. Представим облачные сервисы, как T, а данные, как O. Определим владельца данных, как S, субъектов с правами на чтение данных, как U.

2. Определим набор действий A, который сервис T может выполнить от имени S и U с объектом O. Примем во внимание, что сервис рабочего информационного процесса оперирует с данными путём применения операций чтения и записи, где $A = \{r, \omega\}$.

3. Определим метки конфиденциальности данных, входящего и исходящего канала облачной среды, как $L: T \times O \rightarrow A$.

4. Определим карту текущих размещений блоков рабочего процесса, как $l: T + O \rightarrow L$.

Сформулируем ряд правил обработки данных для рабочего процесса, протекающего в среде облачных вычислений:

1) все действия, выполняемые сервисами, должны соответствовать политикам безопасности владельцев данных;

2) сервисы могут функционировать только на компоненте облачной среды с меткой конфиденциальности, имеющей, по меньшей мере, те же ограничения, что и сервис;

3) сервис не может прочитать данные, метки которых не содержат субъекта с

правами на чтение, от имени которого он запущен;

4) Сервис не может записать данные на узел с более низким уровнем ограничений, чем ограничения метки записываемых данных.

Для решения поставленной задачи рассмотрим, как модель децентрализованного управления метками [2], примененная к рабочим информационным процессам, может быть расширена для учёта требований безопасности при распределении процессов между компонентами среды облачных вычислений. Поскольку облачная архитектура позволяет выбрать для размещения более одного облака на выбор, необходимо принять решение относительно того, как следует распределить данные и облачные сервисы между компонентами облачной архитектуры с разными метками безопасности.

Рассмотрим гибридную архитектуру, в рамках которой развёрнуты два компонента облачной среды:

– частная облачная среда (ЧОС) с высоким уровнем доверия, расположенное в интрасети организации и удовлетворяющее повышенному уровню информационной безопасности по стандарту COBIT,

– общедоступная облачная среда (ООС) с меньшим уровнем доверия, удовлетворяющая базовым требованиям информационной безопасности. ООС представляет собой компонент, в рамках которого функционируют сервисы от имени не доверяющих друг другу субъектов.

Расширим модели управления доступом в целях обеспечения систематического принятия решения о том, где сервисы и данные рабочего информационного процесса могут быть развернуты в рамках гибридной защищенной облачной среды для обеспечения непрерывности бизнеса и соблюдения требований безопасности. Для решения этой задачи добавим в модель новые переменные.

1. Карта размещений рабочего процесса.

Карта размещения рабочего информационного процесса должна включать в себя доступные компоненты среды облачных вычислений, которые обозначим, как P:

$$l: T + O + P \rightarrow L.$$

2. Карта присвоений сервисов и данных к облаку.

Показатель H будет использоваться для описания присвоения каждого сервиса и данных в облаке:

$$H: T + O \rightarrow P.$$

Сформулируем новое правило: блок рабочего процесса (сервис или данные) может быть развернут на компоненте только в том случае, если метка безопасности компонента, по крайней мере, имеет те же ограничения, что и метка безопасности данных, обрабатываемые сервисом.

Если данные o_1, хранятся в компоненте p_a, сервис t_1 в компоненте p_b, данные o_2 в компоненте p_c , то должны выполняться следующие условия.

$$L(p_a) \sqsubseteq L(o_1), \quad (1)$$

$$L(p_b) \sqsubseteq L(t_1), \quad (2)$$

$$L(p_c) \sqsubseteq L(o_2), \quad (3)$$

Тогда, согласно [1]:

$$L(p_c) \sqsubseteq L(o_2) \sqsubseteq L(t_1). \quad (4)$$

Используя полученную модификацию модели доступа можно автоматически получить все допустимые варианты распределения рабочего процесса. Для этого определим новый вводный параметр для учёта разных компонентов облачной среды в виде (P) и примем во внимание набор сервисов (T), набор данных (O) и карту зон безопасности (l).

Определим варианты (V) перехода сервисов и данных между облаками.

$$V: T + O \rightarrow P, \quad (5)$$

$$V = \{b \rightarrow p \mid b \in T + O, p \in P, l(b) \leq l(p)\}. \quad (6)$$

После определения всех допустимых переходов сервисов и данных в облаках, определяется множество всех действительных развертываний рабочего процесса W. Алгоритмически W вычисляется путем перекрёстного произведения присвоения блоков на облака, содержащихся в V.

Однако, применение метода в текущей постановке практически реализовать в рамках распределенной системы затруднительно по следующим причинам:

1) сервис может создавать выходные данные непосредственно на другом облаке, без предварительного сохранения данных на компоненте, на котором он сам размещен;

2) сервис может использовать в качестве входной информации данные из другого облака без обязательного сохранения их на своём компоненте.

Для преодоления указанных ограничений, во-первых, введем понятие нового вида сервиса – транспортного, который будет передавать данные из одного облака в другое. Аналогом транспортного сервиса является оператор обмена в распределенной обработке запросов.

Переход будет осуществляться за счёт добавления в модель новых компонентов (входящих и исходящих каналов), функционирующих на облаке-источнике и облаке-получателе. Транспортный сервис принимает данные на одном облаке и создает её копию на другом. Все рабочие процессы из множества W трансформируются и включают в себя транспортные сервисы.

Введём четыре правила преобразования графа информационного потока:

$$o_j^a \to t_i^a \Rightarrow o_j^a \to t_i^a \quad (7)$$

$$o_j^a \to t_i^b \Rightarrow o_j^a \to \text{передатчик} \to o_j^b \to t_i^b \quad (8)$$

$$t_i^a \to o_j^a \Rightarrow t_i^a \to o_j^a \quad (9)$$

$$t_i^a \to o_j^b \Rightarrow t_i^a \to o_j^a \to \text{передатчик} \to o_j^b. \quad (10)$$

Преобразования (7) и (9) отражают тот факт, что если оба узла размещены на одном облаке, то включения дополнительных модификаций не требуется. В преобразованиях (8) и (10) вводится новый компонент (транспортный сервис) для передачи данных между облаками.

Создание новых копий данных с помощью правил (8) и (10) может привести к потенциальным проблемам раскрытия копируемых данных. При применении правила (8), необходимо удостовериться, что облако b имеет уровень конфиденциальности достаточный для хранения копии данных o_j, которая наследует уровень конфиденциальности оригинала. В силу этих причин должны быть соблюдены следующие правила:

$$L(p_b) \geq L(o_j). \quad (11)$$

Аналогично, для правила (10):

$$L(p_a) \geq L(o_j). \quad (12)$$

Если происходит нарушение любого приведенного условия, то, варианты, по которым происходит распределение рабочего процесса, перестают отвечать установленным требованиям безопасности, должны быть исключены из набора W надёжных переходов.

Рассмотрим данные, которые получаются в результате работы сервиса. Правило (12) может быть нарушено путем применения преобразований (10) в случае, когда сервис t_1 производит операцию записи данных, таким образом, выполняется условие $l(p_a) < l(o_j)$. Переход от $l(p_b)$ до $l(o_1)$ осуществляется по правилу (8), которое позволяет рабочему процессу создавать копию o_1. Переход от $l(p_b)$ до $l(o_2)$ совершается по правилу (10), которое делает возможным добавление копии o_2.

Один из вариантов отображения распределений рабочего процесса с использованием специальных компонентов передачи данных между облаками показан на рис. 1.

Для иллюстрации уровня безопасности 0 для общедоступной облачной среды (ООС), данных и сервисов использованы контуры, выделенные зеленым цветом, на уровне 1 (частной облачной среды, ЧОС) – красным цветом. Построенные диаграммы позволят специалисту по защите информации получить представление о возможных вариантах развёртывания рабочих процессов в рамках гибридной облачной архитектуры. Для примера описания предлагаемого подхода были использованы простые, линейные рабочие процессы, в то же время он может быть применен и для ориентированных графов, независимо от сложности их структуры.

Предлагаемый подход рассматривается в качестве основы для автоматизации процесса распределения рабочих процессов в рамках гибридной среды облачных вычислений.

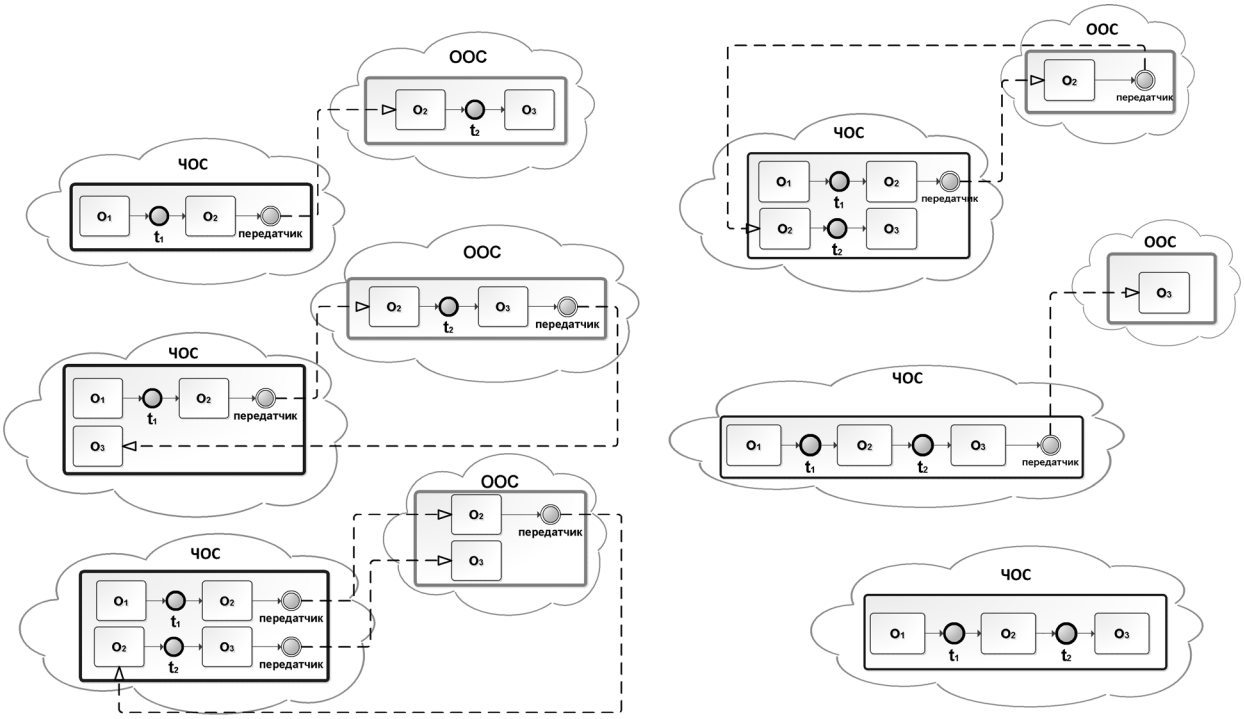


Рис. 1. Допустимые распределения рабочих процессов в гибридной среде облачных вычислений

Данный подход должен заменить процесс выбора администратором возможного варианта распределения процессов, который носит субъективный характер и может привести к ошибке. Рассмотренный подход, имеет преимущества, которые могут снизить, как потенциальные нарушения безопасности, так и расходы на информационную инфраструктуру организации.

Литература

1. Царегородцев, А.В. Методика построения защищенных информационно-телекоммуникационных систем на базе гибридной облачной среды [Текст] / Царегородцев, А.В., Мухин, И.Н., Белый, А.Ф. // Информация и безопасность. 2015. Т.18, №3. С.404-407.

2. Царегородцев, А.В. Формализованная модель безопасности рабочих процессов информационно-телекоммуникационных систем, функционирующих на основе технологии облачных вычислений [Текст] / Царегородцев, А.В. // Нелинейный мир. 2013. Т.11, №9. С.610-620.

ФГБОУ ВО «Московский государственный лингвистический университет»
Moscow State Linguistic University
ФГБОУ ВО «Российский государственный гуманитарный университет»
Russian State University for the Humanities

ONE OF APPROACHES TO THE AUTOMATION OF WORKFLOW ALLOCATION IN A HYBRID CLOUD COMPUTING ENVIRONMENT

M.A. Popov, A.V. Tsaregorodtsev

The article concerns a method on the basis of cloud deployment maps between available cloud components that allows distributing critical groups in a hybrid cloud computing environment based on security requirements

Keywords: information security, cloud computing, hybrid cloud environment, workflow, transport service

УДК 004.056.57

МИКРОМОДЕЛИРОВАНИЕ ВЗАИМНОГО ВЛИЯНИЯ ИНФОРМАЦИОННЫХ СЕТЕЙ, СУЩЕСТВУЮЩИХ В ОБЩЕМ ИНФОРМАЦИОННОМ ПРОСТРАНСТВЕ

В. А. Кургузкин, А. В. Паринов, А. Е. Дешина, Й. Воришек

Рассматриваются микромодели смежных сетевых структур, которые находятся в общем информационном пространстве с возможным взаимным распространением негативного контента через смежных агентов

Ключевые слова: смежные информационные сети, риск, шанс, контент, ресурсы, объекты, субъекты

На данный момент известны исследования [1, 2] диффузии информации в различного рода информационных сетях и возникающих в данном случае информационных эпидемий. При этом сети рассматривались как нечто целое в изоляции от какого-либо внешнего мира. Вместе с тем представляет реальный интерес рассмотрение межсетевого взаимодействия. При построении такой модели все также следует придерживаться определенных принципов, которые формулируются следующим образом:

1. Время дискретно. При этом одна итерация по продолжительности такова, что вершина меняет свое состояние на другое строго в течение ее длительности.

2. Сеть ограничена и связная. Таким образом, всегда есть вероятность взаимодействия любого пользователя с любым другим из данной сети. Обратный случай не представляет интереса для исследования, так как без возможности взаимного влияния между узлами не будет также возможности обмена информацией.

3. Моделирование начинается с построения модели сети в виде множества вершин, обозначающих пользователей данной сети. Построение модели осуществляется на основании некоторых исходных статистических данных. Каждой вершине присваивается некоторое обозначение, которое однозначно определяет текущее состояние данного узла в процессе информационной эпидемии. В сложившихся моделях вершина может быть зараженной, к примеру, латентной или иммунизированной. Кроме того, следует разграничивать вершины по их принадлежности к различным сетям. Точно такое же разграничение следует ввести для ребер, их соединяющих, иначе будет просто невозможно отличить принадлежность ребер между одними и теми же вершинами, а также принадлежность самих вершин к той или иной сети. При микромоделировании необходимо использовать фрактал[3], который отражает поведение отдельно взятой вершины и является вероятностной моделью процесса инфицирования. Ранее микро-фрактал строился одинаково для всех вершин и задавался по всей сети. Теперь необходимо изменить данный подход, дополнив его удвоенным микро-фракталом, каждый из которых будет отображать вероятности переходов в те или иные состояния для вершин, лежащих в одной из сетей. Кроме того, необходимо рассмотреть особый класс узлов сети, а именно те, которые лежат одновременно в двух сетях. Для них следует построить отдельный микро-фрактал, который учитывал бы специфику их расположения. Именно через данные вершины любая информация может просачиваться из одной сети в другую и наоборот. Обобщенно можно отобразить взаимодействие нескольких сетей схематично (рис. 1). На данном рисунке присутствуют три множества. Два из них

Кургузкин Владимир Александрович – ВГТУ, студент, e-mail: mnac@comch.ru
Паринов Александр Владимирович – ВГТУ, соискатель, e-mail: mnac@comch.ru
Дешина Анна Евгеньевна – ВГТУ, ст. преподаватель, e-mail: mnac@comch.ru
Воришек Йири - Пан-Европейский Университет (Словакия), к.т.н., профессор, научный сотрудник, e-mail: jiri.vorisek@paneurouni.com

отображают вершины, которые относятся каждая к своим сетям (С1, С2). Третье множество является множеством смежных вершин (СВ), которые принадлежат к обеим сетям сразу.

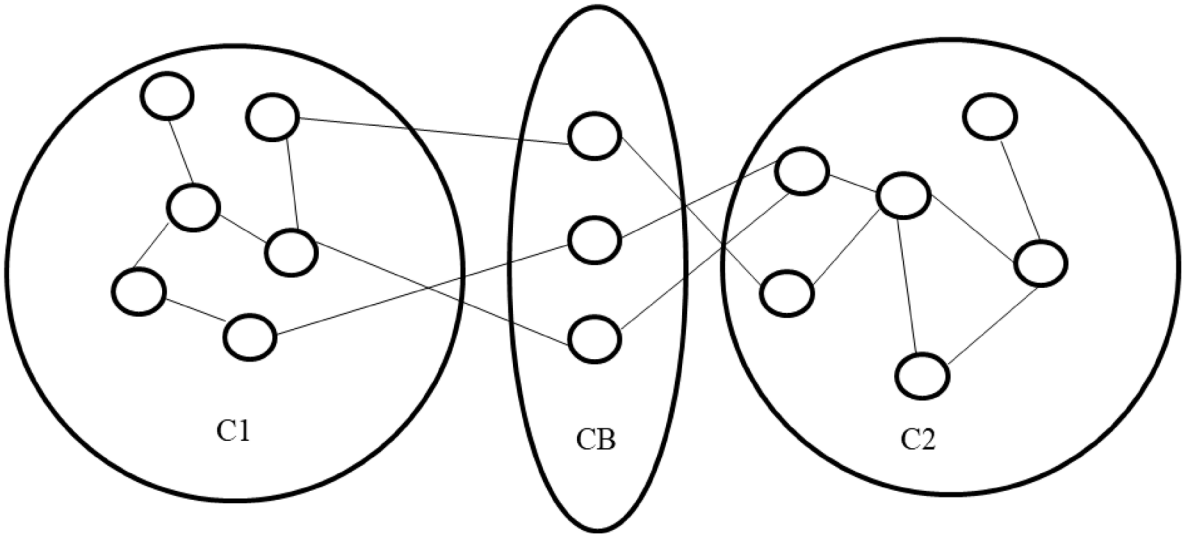


Рис. 1. Схематичное представление межсетевого взаимодействия

Так же, как и раньше, микро-фрактал представляется в виде графа, для которого вершины – это состояния эпидемического процесса, а ребра – вероятности перехода из одного состояния в другое. Состояния для различных сетей берутся одинаковые без учета специфики распространения контента в различных сетях. Формализация данной специфики может стать развитием данного направления исследования и привести к расширению существующих моделей и другим параметрам моделирования. Кроме того, мы будем опираться также на известные переходы [1] между состояниями, которые в формализованном виде выглядят следующим образом:

1. Знакомство с контентом с целью ее анализа и принятия для последующего использования. Здесь пользователь находится в нейтральном по отношению к контенту состоянии, но теоретически восприимчив к его содержанию (состояние S). Для двух сетей обозначения будут одинаковы, так как для упрощения рассматривается ситуация, когда контент не различается по форме и специфике для различных сетей.

2. Стадия, когда пользователь позитивно воспринимает контент (состояние латентного заражения E) или остается к нему равнодушным, т.е. пребывает в исходном состоянии S.

3. Далее из состояния E пользователь может перейти в другие состояния: иммунизированное, приобретенное самостоятельно (состояние M) или посредством модерации (состояние A); состояние, когда он становится распространителем инфекции (состояние I). При этом пользователь может остаться в состоянии E.

4. В отношении I-состояния следует заметить, что в нем остаются далеко не все инфицированные пользователи. Часть из них под действием внутренних (потеря работоспособности) и внешних (устранение модератором) факторов могут перейти в состояние R, которое характеризуется временной нежизнеспособностью индивидуального ресурса пользователя.

Для применения данной модели необходимо пояснить, что изначально микро-фракталы должны строиться для каждой сети в отдельности. Некоторые вершины одновременно с этим в разных сетях будут иметь различные состояния, но эти состояния должны иметь некоторую зависимость, так как такие вершины обычно представляют одного человека, который

может переносить информацию из одной сети в другую.

Вышеприведенные рассуждения позволяют построить вероятностную модель в виде графа (рис. 2), характеризующего первый шаг эпидемического процесса для нескольких взаимодействующих сетей, где, как и в случае описания микрофрактала для одной сети, P_M, P_I, P_R и P_E – вероятности возникновения переходов, а индексы 1..n определяют принадлежность вероятности к соответствующей сети. Достоинством данной модели является факт того, что в ней используются уже полученные вероятности, используемые в исследованиях эпидемических процессов для каждой из сетей по отдельности.

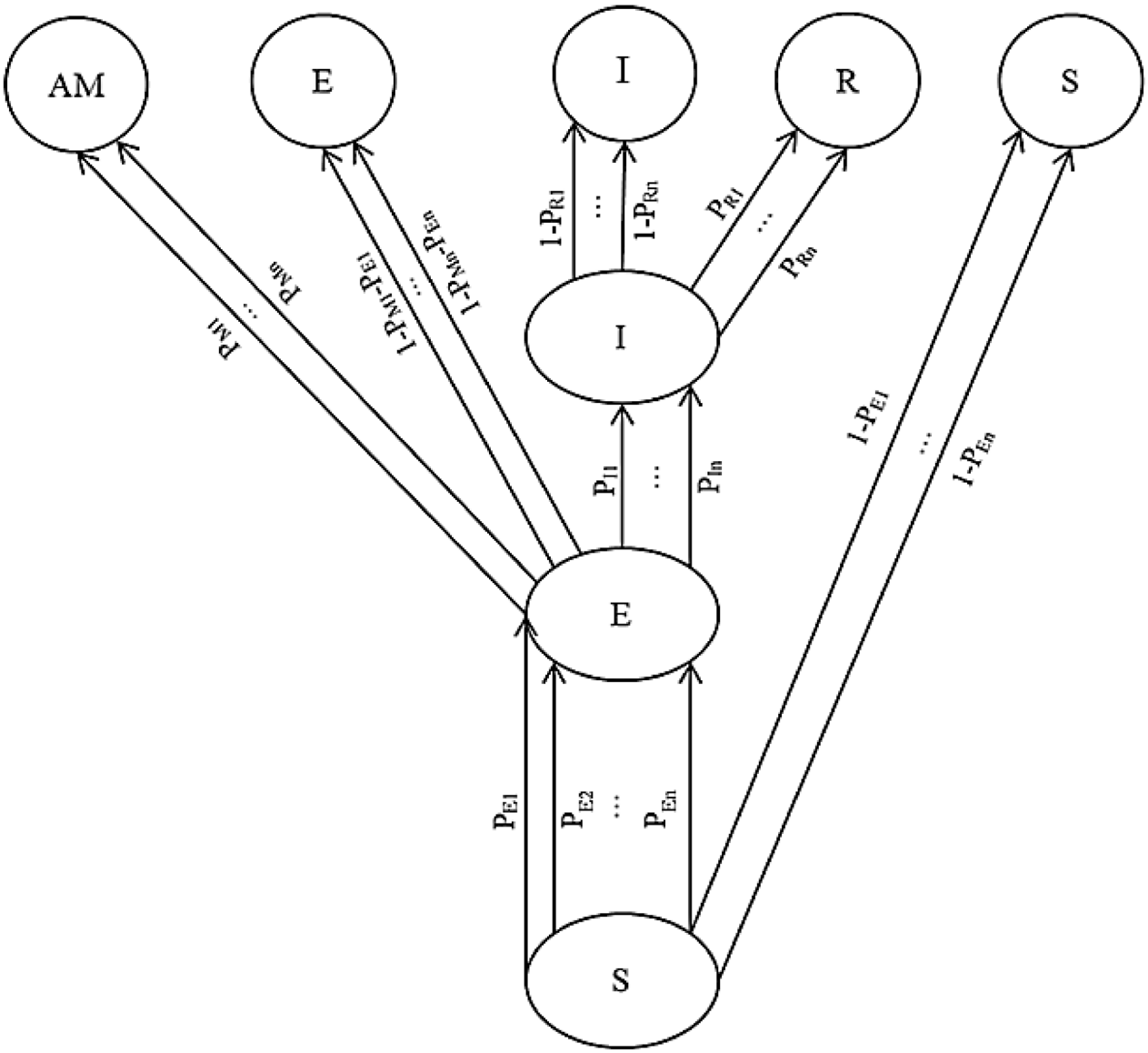


Рис. 2. Вероятностная модель распространения контента среди пользователей, находящихся в разных сетях

В данном случае рассматривается только первый шаг эпидемического процесса. Кроме того, необходимо учесть и те вершины, которые находятся на пересечении сетей, то есть они одновременно принадлежат им обеим.

Для них имеют место различные значения вероятностидля разных сетей. Имеет смысл найти некоторую результирующую вероятность при построении отдельного микро-фрактала для данных вершин.

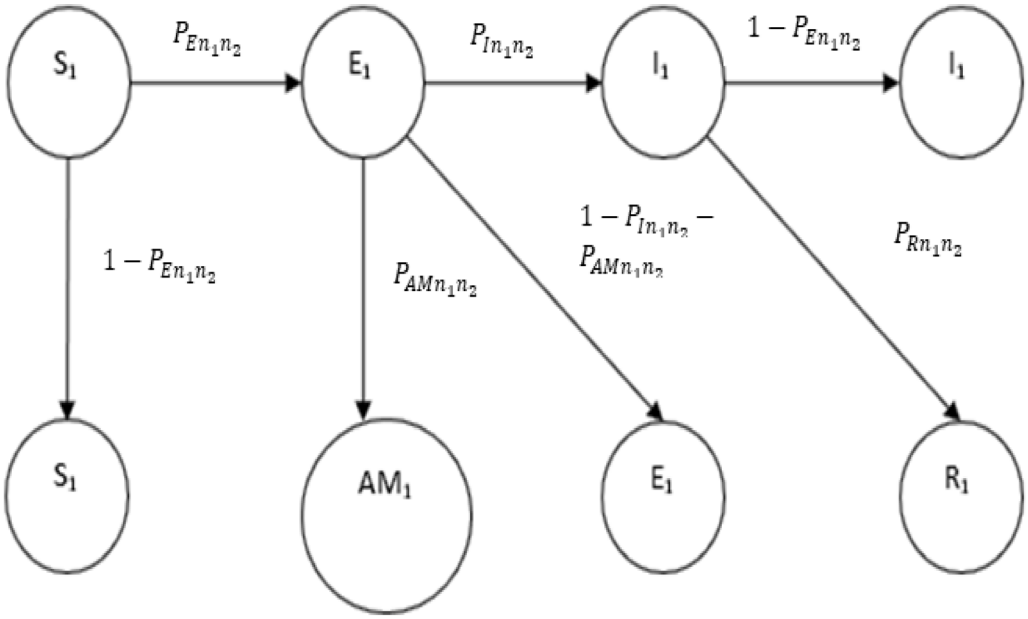


Рис. 3. Пример микро-фрактала эпидемического процесса для вершин, принадлежащих к нескольким сетям

Результирующая картина представлена на рис. 3 с соответствующими индексами сетей, к которым принадлежат вершины.

Необходимо также учитывать различия контента в рассматриваемых сетях, а также степень противоборства между сообществами.

Предложенные вероятности, возможно, следует определять экспертным методом либо на основе имеющейся уже статистики переходов и взаимодействий в сетях и между ними.

Вполне может пригодиться и опросный метод, в котором должно участвовать необходимо и достаточно великое число независимых опрашиваемых [4].

Литература

1. Woo J. Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog / J. Woo, H. Chen // Graduate School of Information Security, Korea University, Anamro, Seoul, Korea. – 2016. – P. 19.

2. Cannarella J. Epidemical modeling of online social network dynamics / J. Cannarella, J.A. Spechler // Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA. – 2014. – P. 66.

3. Федер Е.Фракталы: Пер. с англ. / Е.Федер // М.: «Мир», 1991. – 254 с.

4. Орлов А. И. Экспертные оценки. Учебное пособие / А. И. Орлов //М.: ИВСТЭ,2002. – 486 с.

ФГБОУ ВО «Воронежский государственный технический университет»
Voronezh State Technical University
Пан-Европейский Университет
Pan-European University

THEMICROMODELINGOF MUTUAL INFLUENCE OF INFORMATION NETWORKS EXISTING IN GENERAL INFORMATION SPACE

V. A. Kurguzkin, A. V. Parinov, A. E. Deshina, J. Vorisek

The micromodels of adjacent network structures that exist in the common information space with possible mutual distribution of negative content through related agents are considered

Key words: adjacent information networks, risk, chance, content, resources, objects, subjects

УДК 004

НАПИСАНИЕ КЛИЕНТ-СЕРВЕРНОГО ПРИЛОЖЕНИЯ ДЛЯ ОРГАНИЗАЦИИ МОНИТОРИНГА СЕТЕВЫХ УСТРОЙСТВ РАДИОЭЛЕКТРОННЫХ ОБЪЕКТОВ

Ю.Ю. Громов, В.Е. Дидрих, С.Н. Вихляев, К.Н. Банникова, П.А. Трефилов, А.Ю. Ковергина

Статья посвящена анализу процесса реализации клиентского приложения для осуществления клиент-серверного взаимодействия. В статье проведен анализ этапов проектирования приложения, а так же анализ существующих моделей взаимодействия по технологии клиент-сервер. В результате проведенного анализа в соответствии с поставленной задачей была спроектирована модель взаимодействия, которая удовлетворяет требованиям к написанию мобильного приложения для взаимодействия с системой сканирования удаленных портов, осуществляющей обработку и хранение результатов.

Ключевые слова: клиент-серверная архитектура, сервер БД, проектирование мобильных приложений

Современное общество, в прямом смысле слова, зависит от информации. Массивы данных даже в небольших организациях неуклонно растут. Фирмы-конкуренты или прочие злоумышленники стремятся завладеть информацией, следовательно, информация, информационные и вычислительные ресурсы должны быть защищены в обязательном порядке. Для анализа защищенности вычислительных систем в сети Интернет используют специальные утилиты и инструменты, позволяющие выявить и/или предупредить уязвимости. К таким программам относится «сканер портов», который используется для защиты сети, так как совершает анализ текущего состояния портов сервера.

Классический подход к сканированию портов заключается в разовом обращении к системам сканирования, однако более крупные организации, осуществляющие беспрерывное администрирование локальных вычислительных сетей (ЛВС), нуждаются в постоянном контроле определенного (или распределенного диапазона) *IP*-адресов. Следовательно, разумно осуществить разработку практичной и не менее функциональной системы сканирования и обработки результатов.

Как правило, администратор информационной безопасности устанавливает сканер портов на автоматизированное рабочее место (АРМ) и регулярно производит удаленный аудит защищаемой вычислительной сети, причем результаты сканирования не подвергаются сохранению. С точки зрения использования ресурсов это не практично, т.к. зачастую один и тот же хост подвергается множественному сканированию с получением статичных результатов. Студентами ФГБОУ ВО «Тамбовского государственного технического университета» в лабораторных условиях было реализовано взаимодействие размещенной базы данных (БД) об отсканированных хостах со сканирующим аппаратом, размещенным на АРМ. Теперь, когда сервер с БД позволяет совершать сканирование с сохранением результатов имеет смысл задуматься о написании клиентского приложения, позволяющего вести аудит защищаемой ВС удаленно.

Общеизвестными представителями сканеров являются *Masscan* и *Nmap*[1][2]. В рамках нашего исследования мы будем использовать уже имеющийся в разработке сканер, основой для проектирования которого явился Masscan. *Masscan* – самый быстрый из известных *SYN*-сканеров портов на сегодняшний день, скорость

Юрий Юрьевич Громов – ТГТУ, директор института ИАИТ
Валерий Евгеньевич Дидрих – ТГТУ, д.т.н., профессор
Сергей Николаевич Вихляев – ТГТУ, аспирант
Павел Александрович Трефилов – ТГТУ, аспирант
Кристина Николаевна Банникова – ТГТУ, студент
Ангелина Юрьевна Ковергина – ТГТУ, студент

сканирования так велика, что весь существующий диапазон Ipv4 сканируется за шесть минут. Такую скорость он достигает благодаря специальному драйверу *PF_RING*. Исходный код проекта *Masscan* выложен на *GitHub* в свободном доступе [3]. Также сканер не имеет функций UDP-сканирования и графического интерфейса, что сужает круг его использования, однако разработанный в лабораторных условиях ФГБОУ ВО Тамбовского Государственного Технического Университета «ТГТУ» сканер лишен этого недостатка, что делает его более универсальным [4-6].

В лабораторных условиях ФГБОУ ВО «ТГТУ» была спроектирована серверная часть, которая осуществляет сохранение и обработку результатов сканирования разработанного сканера, кроме того на языке программирования *php* была написана *HTML* страница, позволяющая пользователю обращаться к базе данных (БД) и, при необходимости, редактировать ее содержимое.

В ходе проектирования серверной части была поставлена и успешно решена задача создания интерфейса прикладного программирования (API), что уже само собой дает возможность проектирования и создания мобильного приложения, позволяющего удаленно осуществлять обработку результатов сканирования, хранящихся в БД, т.е. системный администратор может удаленно обратиться к данным об уязвимых портах и «на месте» локализировать уязвимость.

Целью работы является формирование набора требований к написанию мобильного приложения, которое будет взаимодействовать с системой сканирования удаленных портов, осуществляющей обработку и хранение результатов.

Для этого нам нужно решить следующие задачи:

1. Анализ технологии написания мобильных приложений.
2. Анализ моделей взаимодействия клиент-сервер.
3. Проектирование модели взаимодействия.

Первым этапом написания создания приложения является постановка технического задания. На данном этапе определяется, какие потребности пользователей и клиента будет разрешать данное приложение, а также формулируются его основные задачи. Эта часть этапа разработки очень важна, так как от нее зависит функциональность приложения.

Вторым шагом является проектирование UI/UX дизайна. На данном этапе нам нужно определить работу приложения, и произвести создание графической карты взаимодействия между экранами.

Следует продумать расположение кнопок на каждом экране, а так же связи между ними. Изначально прорабатывается дизайн 1-3 экранов, который закладывает всю основу приложения. Дизайн должен быть удобным для пользования, понятным и интерактивным.

UI дизайн (User Interface, или пользовательский интерфейс) является системой, обеспечивающей простой, приятный и не обременяющий способ взаимодействия пользователя с продуктом. Большую часть работы во время создания такого интерфейса составляет наблюдение за поведением пользователя, что позволяет принимать решения, основанные на собранных данных. Пользовательский интерфейс включает в себя два компонента: оборудование (физический компонент) и программное обеспечение (логический компонент).

UI являются средством для выполнения двух типов взаимодействия:

– ввод (Input), позволяющий пользователям управлять системой;

– вывод (Output), позволяющий системе демонстрировать эффект от произведенных пользователем манипуляций.

Существуют правила UI дизайна:

1. Организованность элементов интерфейса.
2. Все элементы должны быть логически структурированы и взаимосвязаны.
3. Группировка элементов интерфейса.

4. Выравнивание элементов интерфейса.

5. Наличие свободного пространства (это позволяет разграничивать информационные блоки).

6. Единый стиль элементов интерфейса.

UX дизайн (User Experience Design, что в переводе означает «опыт взаимодействия») включает в себя различные компоненты: информационную архитектуру, проектирование взаимодействия и графический дизайн. Он основывается на поведении, отношении и эмоциональном состоянии человека, вызванными и связанными с использованием продукта. Кроме того, UX дизайн включает в себя восприятие пользователем характеристик системы, таких как полезность, простота использования и эффективность. Опыт взаимодействия носит субъективный характер, потому что речь идет об индивидуальном восприятии и оценке системы [7].

Разработка UX начинается с определения базового слоя пользователей и его характеристик. Основываясь на этих знаниях, можно вывести специальные требования к разрабатываемому проекту. Разрабатывается информационная архитектура и проектируется иерархия содержимого. Далее выбирается наиболее оптимальный метод прототипирования, который должен быть достаточно экономичным, однако и эффективным настолько, чтобы была возможность сбора обратной связи в быстрой и легкой форме [8].

UI является частью UX, поэтому эти два понятия очень тесно связаны между собой. Тем не менее, UX может существовать и без пользовательского интерфейса.

На следующем этапе мы анализируем модели клиент-серверного взаимодействия.

Клиент-серверные технологии определяют собственные или используют имеющиеся правила взаимодействия между клиентом и сервером, которые называются протоколом обмена (протоколом взаимодействия).

Клиенты в системе клиент-сервер делятся на толстый и тонкий. Толстый клиент – это приложение, обеспечивающее полную функциональность и независимость от центрального сервера. Тонкий клиент – это компьютер с браузером, использующимся для работы с веб-приложениями[15][16]. Для написания мобильного приложения мы будем использовать толстый клиент.

В любой сети, построенной на современных сетевых технологиях, присутствуют элементы клиент-серверного взаимодействия, чаще всего на основе **двухзвенной архитектуры**, которая названа так из-за распределения трех базовых компонентов между двумя узлами (клиентской части и сервером) [7].

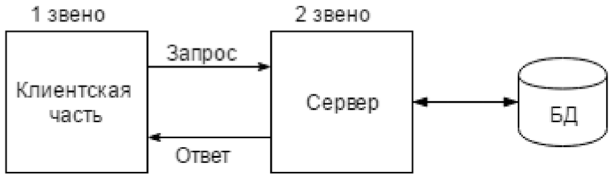


Рис. 1. Двухзвенная клиент-серверная архитектура

Двухзвенная архитектура используется в клиент-серверных системах, где сервер отвечает на клиентские запросы напрямую и в полном объеме. При этом он использует только собственные ресурсы, т.е. не вызывает сторонние сетевые приложения и не обращается к другим ресурсам для выполнения части запроса (рис. 1).

Расположение компонентов на стороне клиентской части или сервера определяет следующие основные модели их взаимодействия в рамках двухзвенной архитектуры:

- сервер терминалов – распределенное представление данных;

- файл-сервер – доступ к удаленной базе данных и файловым ресурсам;

- сервер БД – удаленное представление данных;

- сервер приложений – удаленное приложение.

Классификация двухзвенных моделей клиент-сервер представлена на рисунке 2. Выше пунктирной линии изображено приложение, снизу – сервер БД. Для нашего проекта нам потребуется модель

распределенного представления, где представление данных расположено как на сервере, так и на приложении, а прикладная логика и данные размещены только на сервере. Прикладная логика – это аспект, который рассматривает находящиеся взаимосвязи между методами разработки и логического взаимодействия.

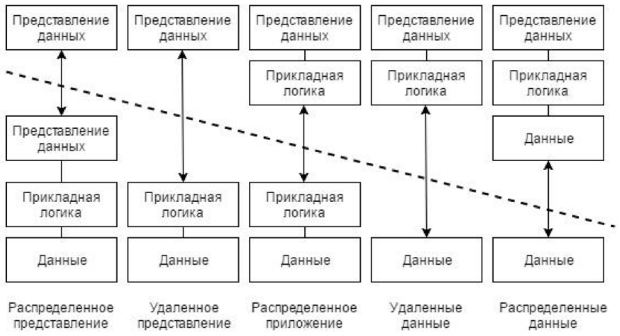


Рис. 2. Модели клиент-серверного взаимодействия

С появлением персональных компьютеров и локальных сетей, была реализована модель файлового сервера, которая предоставляет доступ к файловым ресурсам и к удаленной базе данных. В этом случае выделенный узел сети является файловым сервером, на котором размещены файлы базы данных. На клиентах выполняются приложения, в которых совмещены компонент представления и прикладной компонент (СУБД и прикладная программа), использующие подключенную удаленную базу как локальный файл. Протоколы обмена при этом представляют набор низкоуровневых вызовов операций файловой системы. Такая модель показала свою неэффективность ввиду того, что при активной работе с таблицами БД возникает большая нагрузка на сеть. Частичным решением является поддержка тиражирования (репликации) таблиц и запросов. В этом случае, например при изменении данных, обновляется не вся таблица, а только модифицированная ее часть.

С появлением специализированных СУБД появилась возможность реализации другой модели доступа к удаленной базе данных – модели сервера баз данных. В этом случае ядро СУБД функционирует на сервере, прикладная программа на клиентской части, а протокол обмена обеспечивается с помощью языка SQL. Такой подход по сравнению с файловым сервером ведет к уменьшению загрузки сети и унификации интерфейса «клиент-сервер». Однако сетевой трафик остается достаточно высоким, кроме того, по-прежнему невозможно удовлетворительное администрирование приложений, поскольку в одной программе совмещаются различные функции [9].

С разработкой и внедрением на уровне серверов баз данных механизма хранимых процедур появилась концепция активного сервера БД. В этом случае часть функции прикладного компонента реализованы в виде хранимых процедур, выполняемых на стороне сервера. Остальная прикладная логика выполняется на клиентской стороне. Протокол взаимодействия – соответствующий диалект языка SQL.

В рамках нашего проекта мы используем архитектуру «сервер БД».

Помимо двухзвенной архитектуры существует трехзвенная, третьим звеном в которой становится сервер приложений – это сервисная программа, которая обеспечивает доступ клиентской части к прикладным программам, выполняющимся на сервере. Перенос функций прикладного компонента на сервер снижает требования к конфигурации клиентской части и упрощает администрирование, но представляет повышенные требования к производительности, безопасности и надежности сервера [9][10].

Рис. 3. Трехзвенная клиент-серверная архитектура

Компоненты в трехзвенной архитектуре распределяются следующим образом(рис. 3):

- представление данных – на стороне клиентской части;

- прикладной компонент – на выделенном сервере приложений (как вариант, выполняющем функции промежуточного ПО);

- управление ресурсами - на сервере БД, который и представляет запрашиваемые данные.

Преимущества сервера приложения:

- целостность кода и данных;
- централизованное управление (изменения в конфигурации прикладных программ, такие как, например, смена сервера баз данных, выполняются централизованно);
- безопасность (централизованные средства, через которые поставщик услуг (сервис-провайдер) может управлять доступом к данным и компонентам приложения, позволяют выполнять проверку подлинности потенциально ненадежных клиентов в среднем слое и не затрагивать уровень базы данных. К таким средствам можно отнести сканер отпечатков пальцев);
- производительность (сервер приложений может решать задачи балансировки сетевого трафика и распределения нагрузки между другими физическими серверами системы).

Системы, построенные на основе сервера приложений, имеют один основной недостаток, присущий всем централизованным решениям – «падение» сервера приведет к недоступности программ для всех клиентов. К тому же эффекту приведут и неполадки в сетевом подключении[11][12].

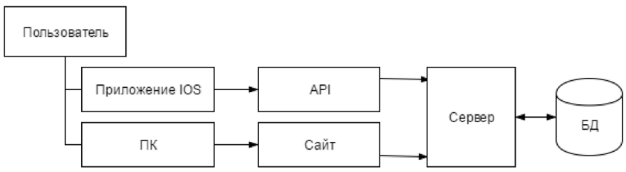


Рис. 4. Модель взаимодействия клиент-серверной системы с приложением

На рис. 4 представлена модель взаимодействия клиент-серверной системы с приложением, имеющим доступ к удаленной БД, которая может быть размещена на сервере приложения.

Сервер может располагаться как на хостинге, так и локально. В рамках нашего проекта для альфа и бета тестирования нам потребуется сайт, который мы будем размещать на локальном сервере.

Альфа тестирование - это первая стадия тестирования приложения, которое проходит внутри организации, где разрабатывается продукт. Его проводят специалисты. В течение нескольких дней приложение тестируется, после создается таблица с выявленными ошибками, формируется список всех недочетов и недоработок. Далее производится отладка, и приложение проходит второе тестирование. В рамках нашего института автоматики и информационных технологий мы будет тестировать наше приложение в узком кругу.

Бета тестирование – повторное тестирование приложения, в результате которого происходит проверка исправления ошибок и недочетов, проводится их устранение перед окончательным выходом на рынок, к массовому потребителю. Бета тестирование производится будущими пользователями данного продукта. Так же оно может быть открытым и закрытым, когда тестирование проводят сами разработчики или пользователи по приглашениям. Бета версия не является финальной стадией версии продукта, поэтому на данном этапе не исключено выявления новых ошибок, которые исправляются [13][14]. На этапе бета тестирования мы выложим код нашего сканера на *OpenSource*, а так же можем разместить на *GitHub*, где пользователи могут оценить работу нашего приложения.

Далее обычно производится создание иконки продукта. Сначала рисуется ее эскиз, затем он корректируется, прорисовывается и утверждается.

Для поддержки работы приложения создаются обновления, которые модернизируют продукт и улучшают его функциональные и потребительские свойства. Чаще всего происходит обновление базы данных, а дизайн приложения обновляется значительно реже.

В рамках поставленных задач была проанализирована и выбрана подходящая технология написания мобильных приложений. В процессе анализа мы спланировали предстоящие шаги проектирования, а так же провели анализ моделей взаимодействия и выбрали из них наиболее подходящую для нашего приложения. Помимо всего была

спроектирована система взаимодействия с учетом выбора модели размещения компонентов на стороне клиента и сервера.

Литература

1. MASSCAN: Mass IP port scanner. [Электронный ресурс] – Режим доступа: https://github.com/robertdavidgraham/masscan свободный.
2. Инструкция по использованию Masscan — лучшего массового сканера больших сетей. [Электронный ресурс] – Режим доступа: https://hackware.ru/?p=577, свободный.
3. Ализар А. Masscan: Сканирование диапазона IPV4 за шесть минут [Электронный ресурс]. – 2013. – Режим доступа: https://xakep.ru/2013/09/16/61262/, свободный.
4. Степанов Д. Для чего нужны сканеры портов [Электронный ресурс] // 10-Страйк. – 2013. – Режим доступа: http://infoarena.ru/articles/2121/0/48546, свободный.
5. Ермаков А.В. Использование сетевого сканера для повышения защищенности корпоративной информационно-вычислительной сети – Москва: ИПМ им. М.В.Келдыша РАН, 2001.
6. Различные приемы сканирования портов. [Электронный ресурс] – Режим доступа: https://nmap.org/man/ru/man-port-scanning-techniques.html
7. Что такое UX и UI дизайн – особенности и отличия [Электронный ресурс] – Режим доступа: http://www.kasper.by/help/chto-takoe-ux-i-ui-dizain/, свободный.
8. Что вы знаете о UI/UX, или зачем нужны дизайнеры интерфейсов? [Электронный ресурс] – Режим доступа: http://www.it-academy.by/news/articles/chto-vy-znaete-o-uiux-ili-zachem-nuzhny-dizaynery-interfeysov, свободный.
9. Введение в базы данных. [Электронный ресурс] – Режим доступа: http://www.mstu.edu.ru/study/materials/zelenkov/ch_7_1.html, свободный.
10. Компоненты сетевого приложения. Клиент-серверное взаимодействие и роли серверов. [Электронный ресурс] – Режим доступа: http://www.4stud.info/networking/lecture5.html, свободный.
11. Понятие клиент-серверных систем. [Электронный ресурс] – Режим доступа: http://bourabai.ru/dbt/client1.htm, свободный.
12. Технология "клиент – сервер". [Электронный ресурс] – Режим доступа: http://www.intuit.ru/studies/courses/508/364/lecture/8643?page=2, свободный.
13. Тестирование: Что такое альфа- и бета-тест? [Электронный ресурс] – Режим доступа: http://myblaze.ru/testirovanie-chto-takoe-alfa-i-beta-test/, свободный.
14. Сысоев Э.В., Бурцева Е.В.. Базы данных: лекции к курсу. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2007. – 48 с. – 70 экз.

ФГБОУ ВО «Тамбовский Государственный Технический Университет»
Tambov State Technical University

CUSTOMER-CLIENT APPLICATION WRITTEN FOR MONITORING NETWORK DEVICES FOR RADIOELECTRONIC OBJECTS

Y.Y. Gromov, V.E. Didrih, S.N. Vikhlyaev, P.A. Trefilov, K.N. Bannikova, , A.Y. Kovergina

The article is devoted to the analysis of the implementation of the client application for the completion of client-server interaction. The article analyzes the projects related to client-server technology. As a result of the conducted analysis, in accordance with the task, an interaction model was designed that meets the requirements for the preservation of remote ports, processing and storing the results

Key words: client-server architecture, database server, design of mobile application

УДК 004.657

АНАЛИЗ СУЩЕСТВУЮЩИХ МОДЕЛЕЙ ПРЕДСТАВЛЕНИЯ ТЕМПОРАЛЬНЫХ ДАННЫХ

П.А. Трефилов, М.А. Ивановский, Н.Г. Шахов, А.И. Елисеев

Рассмотрены вопросы, касающиеся представления темпоральных сущностей реального мира с использованием механизмов реляционных баз данных (РБД). Проведен обзор и анализ существующих методов представления темпоральных данных, выявлены особенности, достоинства и недостатки

Ключевые слова: MySQL, реляционные базы данных, темпоральные данные, модель данных

Для долговременного хранения данных зачастую используют реляционные базы данных (РБД) под управлением систем управления базами данных (СУБД, СУРБД) [14, 18, 19]. Использование РБД позволяет минимизировать объем избыточной информации и предоставляет мощные возможности по организации процессов поиска и извлечения информации, соответствующей запросу пользователя.

Существенный недостаток традиционных РБД и реляционной модели в целом – невозможность оптимально организовать процессы хранения, поиска и извлечения темпоральных данных [17], т.е. данных, привязанных к определенному моменту времени. На сегодняшний день не существует полноценной темпоральной базы данных (ТБД) – существуют лишь различные модули и плагины к популярным коммерческим СУБД. Все эти надстройки обладают рядом недостатков, в частности – при проектировании темпоральной структуры не существует какого-либо универсального метода, позволяющего спроектировать гибкую логическую структуру, применимую сразу к множеству предметных областей.

На сегодняшний день существует достаточно много работ, посвященных темпоральным базам данных. Есть работы [1-8, 12], в которых была решена задача представления темпоральных данных с помощью традиционной реляционной модели.

Одним из ключевых периодов в области исследований темпоральных баз данных, временем ее «официального» представления считаются 1992–1993 гг [27]. В это время сначала Ричард Снодграсс (Richard Snodgrass) высказал идею о возможном темпоральном расширении стандарта языка запросов к реляционным базам данных SQL-92. Позже была образована комиссия по созданию спецификации подобного языка, в которую вошли Ричард Снодграсс, Илсу Ан (Ilsoo Ahn), Гэд Ариав (Gad Ariav), Дон Бэтори (Don Batory), Джеймс Клиффорд (James Clifford), Картис Дайрсон (Curtis E. Dyreson), Кристиан Йенсен (Christian S. Jensen), Рамес Элмасри (Ramez Elmasri), Фабио Гранди (Fabio Grandi), Вольфганг Кафер (Wolfgang Kaefer), Ник Клайн (Nick Kline), Кришна Кулкарни (Krishna Kulkarni), Тинг Клифф Леунг (Ting Y. Cliff Leung), Никос Лоренцос (Nikos Lorentzos), Джон Роддик (John F. Roddick), Эрай Серев (Arie Segev), Майкл Су (Michael D. Soo) и Суринарайяна Срипада (Surynarayana M. Sripada). После нескольких лет плодотворной работы в начале 1994 года появилась первая предварительная версия спецификации языка, а на основе полученных замечаний в сентябре того же года была выпущена окончательная спецификация языка запросов TSQL2.

Все рассмотренные ниже модели классифицируются по следующим признакам [12]:

1) По типу хранимого времени. Если отношение содержит только действительное время моделируемой реальности, то такое отношение называется историческим

Трефилов Павел Александрович – ТГТУ, аспирант.
Ивановский Михаил Андреевич - ТГТУ, канд. техн. наук, доцент.
Шахов Николай Гурьевич - ТГТУ, канд. техн. наук, доцент.
Елисеев Алексей Игоревич - ТГТУ, канд. техн. наук.

(historical). Если отношение содержит только транзакционное время, оно называется отношением возврата (rollback). Если отношение содержит обе метки времени, оно называется битемпоральным (bitemporal).

2) По типу нанесения временных меток (*timestamping*). Выделяют метки кортежей (*tuple timestamping*) и метки атрибутов (*attribute timestamping*).

3) По минимальной нормальной форме. Выделяют отношения на основе 1НФ и на основе Н1НФ. Как правило, Н1НФ соответствуют меткам атрибутов, а 1НФ – меткам кортежей.

4) По допустимому значению времени атрибутов. Различают однородные (homogeneous) и разнородные (heterogeneous). В однородных отношениях жизненный цикл атрибутов по времени должен быть одинаковым. К разнородным структурам таких требований не предъявляется.

5) По формату временной метки. Различают граничные точки (boundary points) и интервалы (interval). В случае с временными точками, метка времени содержит 2 значения – начала и конца интервала. В случае интервала метка времени представляется лишь одним значением.

Ричард Снодграсс внес огромный вклад в развитие темпоральных баз данных. Во-первых, он дал определения основным темпоральным терминам, используемым по сей день. Во-вторых, в его работах [5, 6, 13] приводится модель темпоральных данных, реализованная на основе традиционной реляционной модели и позволяющей оперировать данными посредством стандартного SQL. В этой модели для каждого темпорального события справедливо выражение: $T = (AS, AF)$, где

AS – момент актуализации значения,

AF – момент утраты актуальности.

Пример темпорального отношения приводится на рис. 1.

COMPANY	TRN	CN	VALID TIME	
			(FROM)	(TO)
Apple	Jack	5.2	2/11/1994	25/4/1995
Apple	Jack	5.2	7/8/1996	1/1/2010
Apple	Mark	3.3	2/1/1992	8/11/1996
Apple	Mark	3.5	30/4/1995	1/1/2010
IBM	Tim	5.2	19/3/1997	21/4/1997
IBM	Tim	5.0	17/12/1995	1/1/2010
Microsoft	Karen	3.3	25/6/1996	1/1/2010

Рис. 1. Отношение модели Снодграсса

Снодграсс и Йенсен сотрудничали в области развития темпоральных баз данных на протяжении нескольких лет. За это время ими было написано множество работ, посвященных семантике темпоральных данных [5, 6].

Главный вклад их совместной работы – новая модель представления темпоральных данных, которая получила название битемпоральной модели. Битемпоральная модель, помимо действительного времени, также поддерживает время транзакции, что дает более мощные механизмы управления часто изменяющимися данными и контроля целостности.

Модель представлена как: $T = (AS, TS, AF, TF)$, где

AS – момент актуализации значения,

TS – время записи факта в БД,

AF – момент утраты актуальности,

TF – время удаления факта из БД.

Рассмотренная модель, хоть и сохраняет простоту реляционной модели, уступает другим временным предложенным моделям с репрезентативной точки зрения из-за большого количества временных меток, которые содержит каждый кортеж, и представления временных меток как битемпоральных элементов. Однако данная модель на практике является самой часто используемой [22].

Пример отношения в модели Йенсена-Снодграсса выглядит следующим образом (рис. 2).

COMPANY	TRN	CN	T
Apple	Jack	5.2	{2/11/1994, 3/11/1994, ..., 24/4/1995, 7/8/1996, 8/8/1996, ..., 31/12/2009}
Apple	Mark	3.3	{2/1/1992, 3/1/1992, ..., 7/11/1996}
Apple	Mark	3.5	{30/4/1995, 1/5/1995, ..., 31/12/2009}
IBM	Tim	5.2	{19/3/1997, 20/3/1997, ..., 20/4/1997}
IBM	Tim	5.0	{17/12/1995, 18/12/1995, ..., 31/12/2009}
Microsoft	Karen	3.3	{25/6/1996, 26/6/1996, ..., 31/12/2009}

Рис. 2. Отношение в модели Йенсена-Снодграсса

Модель Гадия [3, 4] построена на основе H1НФ (Не Первой Нормальной Формы). H1НФ не предполагает атомарности атрибутов. Атрибуты в этой модели представляются как: $A = (AV,[T1,T2])$, где

AV – значение атрибута сущности;

T1 – начальное время;

T2 – конечное время.

Особенности H1НФ активно используются в документно-ориентированных СУБД [9]. С точки зрения проектирования таких отношений с помощью реляционных механизмов, данная модель является достаточно сложной.

COMPANY	TRN	CN
[2/1/1992, NOW] Apple	[2/11/1994, 24/4/1995] ∪ [7/8/1996, NOW] Jack [2/1/1992, NOW] Mark	[2/11/1994, 24/4/1995] ∪ [7/8/1996, NOW] 5.2 [2/1/1992, 7/11/1996] 3.3 [30/4/1995, NOW] 3.5
[17/12/1995, NOW] IBM	[17/12/1995, NOW] Tim	[19/3/1997, 20/4/1997] 5.2 [17/12/95, NOW] 5.0
[25/6/1996, NOW] Microsoft	[25/6/1996, NOW] Karen	[25/6/1996, NOW] 3.3

Рис. 3. Отношение в модели Гадия

В модели Бен-Зви [2] битемпоральное отношение *R* состоит из набора атрибутов {*A1*, …, *An*, *T*}, где *T* – темпоральный атрибут, определенный на множестве битемпоральных элементов. Тогда *R* будет представлено в модели Бен-Зви следующим образом:

R = (A1,...,An,Tes,Trs,Tee,Tre,Td).

В кортеже значение атрибута Tes (effectivestart) – это время, когда значение атрибута кортежа начинает быть актуальным [27]. Атрибут *Trs* (record start) хранит информацию о том, когда *Tes* было сохранено в БД. Аналогично, *Tre* хранит информацию о том, когда факт перестает быть актуальным в моделируемой реальности, а *Tee* – когда *Tre* было зафиксировано в БД. Последний атрибут *Td* указывает на время, когда запись была логически удалена из БД.

Недостатками этой модели является:

- поддержка только одного уровня вложенности времени;
- однородность;
- не поддерживаются NULL-значения;
- не формализованы реляционные операции;
- не поддерживаются JOINы.

Модель Тензеля представлена обладает следующими свойствами [1]:

- метки атрибутов;
- основана на H1НФ;
- поддерживается только действительное время;
- разнородная структура, возможность пометки только изменяющихся во времени атрибутов.

Математически, эта модель представлена как множество $T = [<t,v>]$, где

t – совокупность временных меток;

v – значение изменяющегося во времени атрибута.

COMPANY	TRAINER	
	TRN	CN-H
		CN
Apple	Jack	<{[2/11/1994, 25/4/1995) ∪ [7/8/1996, 1/1/2010)}, 5.2>
	Mark	<{[2/1/1992, 8/11/1996)}, 3.3> <{[30/4/1995, 1/1/2010)}, 3.5>
IBM	Tim	<{[19/3/1997, 21/4/1997)}, 5.2> <{[17/12/1995, 1/1/2010)}, 5.0>
Microsoft	Karen	<{[25/6/1996, 1/1/2010)}, 3.3>

Рис. 4. Отношение в модели Тензеля

Сравнительная таблица достоинств и недостатков известных моделей:

Табл. 1

Сравнение моделей представления темпоральных данных

	SQL	НФ	Битемпоральность	Метка времени
Снодграсс	ДА	1НФ	НЕТ	Кортеж
Йенсен-Снодграсс	ДА	1НФ	ДА	Кортеж
Гадия	НЕТ	Н1НФ	ДА	Атрибут
Бен-Зви	ДА	1НФ	ДА	Кортеж
Тензель	НЕТ	Н1НФ	НЕТ	Атрибут

Рассмотренные модели обладают одним общим недостатком – ни в одной модели не предполагается возможность обработки нечетких значений атрибутов кортежей. Существуют работы [15, 21, 23], посвященные организации хранения нечетких данных в реляционным базам данных, но приведенные в них методы нельзя применить к темпоральным моделям. Исключение может составить модель Снодграсса и Йенсена-Снодграсса, однако этот вопрос требует проработки.

Проанализируем модель Йенсена-Снодграсса, как наиболее применяемую в практических задачах, в работе.

Пусть $R = (a1,a2,a3,a4,t1,t2)$ – некое отношение, где $a1, a2, a3$ – статичные (не изменяющиеся во времени) атрибуты, а $t1$, $t2$ — темпоральные атрибуты. Пусть $t1$ — темпоральный момент времени, а $t2$ — темпоральный интервал. При использовании в качестве темпоральной метки $T = (AS,TS,AE,TE)$ структура будет одинакова как для момента времени, так и для временного интервала.

Поскольку речь идет о темпоральных атрибутах, любое добавление, изменение и удаление записи повлечет за собой добавление строки в исходную таблицу. При этом значения нетемпоральных атрибутов будут дублироваться, что повлечет за собой избыточность данных.

Для того, чтобы избавиться от избыточности, отношение R нормализовано и приведено к $R = (RS, Rt1, Rt2, ..., Rtn)$, где RS – отношение, состоящее из нетемпоральных атрибутов, $Rt1$, $Rt2$, …, Rtn – отношения, описывающие изменения темпоральных атрибутов.

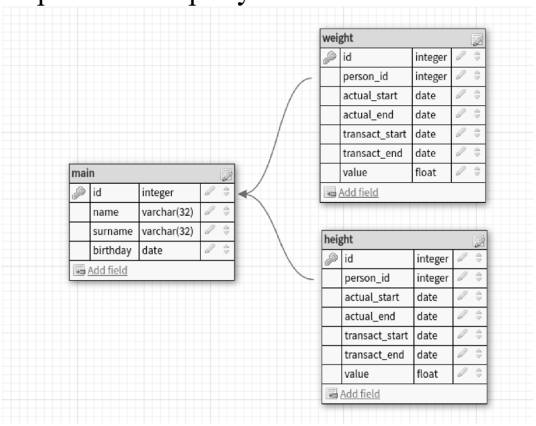


Рис. 5. Схема БД

В качестве демонстрационного примера возьмем базу данных, содержащую информацию о людях, их росте и весе в определенные интервалы времени. База данных представлена тремя отношениями: отношение «Main» (не темпоральное), отношение «Weight» (темпоральное), отношение «Height» (темпоральное). Схема базы данных будет выглядеть как показано на рис. 5.

Темпоральные SQL-запросы к разработанной будут построены с использованием традиционного синтаксиса SQL, без всяких модификаций языка. Например, запрос выборки роста и веса каждого человека за интервал «февраль-март» с учетом разработанной структуры БД выглядит следующим образом:

```
SELECT pm.name,pm.surname,ph.value AS
height,pw.value AS weight
FROM persons_main AS pm
JOIN person_height AS ph ON
(pm.id=ph.person_id)
JOIN person_weight AS pw ON
(pm.id=pw.person_id)
WHERE (ph.actual_start>='2017-02-01'
AND ph.actual_end<='2017-03-31') AND
(pw.actual_start>='2017-02-01' AND
pw.actual_end<='2017-03-31')
```

Особое внимание стоит обратить на запросы модификации данных – DELETE, INSERT, UPDATE. В случае удаления, согласно модели Снодграсса-Йенсена, запись может быть фактически удалена, но факт удаления должен быть отражен в темпоральном отношении. В нашем случае – это отношения «Weight» и «Height». Запрос добавления бывает двух видов: добавление нетемпорального факта (таблица «Main») и добавление темпорального факта. При этом, добавление строки в таблицу Main зачастую потребует добавления зависимых записей в темпоральные таблицы. Запрос обновления (UPDATE), согласно модели, возможен только для нетемпоральных отношений. Обновление зависимых от времени атрибутов – это, по сути, запрос добавление темпорального факта.

В качестве заключения подведем итог. Темпоральные данные – данные, изменяемые с течением времени. Темпоральные данные бывают двух видов: временной момент и временной интервал. Хранение темпоральных данных возможно реализовать с помощью традиционной реляционной модели и с использованием известных реляционных СУБД. При этом язык темпоральных запросов, как таковой, не требуется – возможно использование стандартного DML SQL.

В настоящее время идея представления темпоральных данных в традиционных реляционных моделях активно развивается. Среди последних работ имеются работы О.В. Ланкина и Петуховой Н.Ю. [20, 25, 26], где приводятся примеры практического применения ТБД в определенной сфере человеческой деятельности. Однако до сих пор не проработана идея введения в модели ТБД нечетких множеств, которые бы позволили оперировать лингвистическими переменными, более удобными для описания условий запросов.

Литература

1. Tansel A. Adding Time Dimension to Relational Model and Extending Relational Algebra. Information Systems, 11(4): p. 343-355, 1986.

2. Ben-Zvi, «The Time Relational Model», 1982, PhD thesis, Computer Science Department, UCLA.

3. Gadia S.K. and Nair S.S. Algebraic Identities and Query Optimisation in a Parametric Model for Relational Temporal Databases. IEEE Transactions on Knowledge and Data Engineering, 10(5), 793-807 (1998)

4. Gadia S.K., Nair S.S. and Poon Y.C. Incomplete Information in Relational Temporal Databases. Proceedings of the 18th International Conference on Very Large Data Bases, Vancouver, Canada, 395-406 (1992)

5. Jensen C. S., Snodgrass R., Soo M. D. Extending Normal Forms to Temporal Relations. Technical Report TR-92-17. Department of Computer Science, University of Arizona, Tucson, AZ, 1992.

6. Jensen C. S., Soo M. D., Snodgrass R. T. 1994: Unifying Temporal Data Models Via Conceptual Model // Information Systems V. 19, №. 7. P. 513-547.

7. M. D. Soo, R. T. Snodgrass, and C. S. Jensen. Efficient Evaluation of the Valid-Time Natural Join. In Proceedings of the IEEE International Conference on Data Engineering, pp. 282–292, Houston, Texas, February 1994.

8. McKenzie E., Snodgrass R. Supporting Valid Time in an Historical Relational Algebra: Proofs and Extensions. Technical Report TR-91-15. Department of Computer Science, University of Arizona, Tucson, AZ, 1991.

9. NoSQL – понимаем суть [Электронный ресурс]. URL: https://habrahabr.ru/post/152477 (дата обращения: 14.12.2016).

10. Optimizing database structure [Электронный ресурс]. URL: https://dev.mysql.com/doc/refman/5.7/en/optimizing-database-structure.html

11. Optimizing SQL statements [Электронный ресурс]. URL: https://dev.mysql.com/doc/refman/5.6/en/statement-optimization.html (Дата обращения: 15.06.2017)

12. Praveen Kumar Gupta, Rahul Rishi, Ranjit Biswas. A Comparative Analysis Of Temporal Data Models // International Journal of Advanced Computational Engineering and Networking, Vol-1, 2013, Oct

13. Snodgrass R. Developing Time-Oriented Database Applications in SQL. San Francisco: Morgan Kaufmann Publishers Inc., 1999. 541 p.

14. Д. Крёнке. Теория и практика построения баз данных. 8-е изд. СПб.: Питер, 2003. 800 с.: ил.

15. Естефеев В. И. Нечеткие запросы к реляционной базе данных в информационной системе подбора персонала [Текст] / В. И. Естефеев, И. Ю. Балашова // Образование и наука в современных условиях: материалы VII Междунар. науч.–практ. конф. (Чебоксары, 22 мая 2016 г.). В 2 т. Т. 2 / редкол.: О. Н. Широков [и др.]. — Чебоксары: ЦНС «Интерактив плюс», 2016. — № 2 (7). — С. 69–72. — ISSN 2412-0537

16. Интернет вещей – а что это? [Электронный ресурс]. URL: https://geektimes.ru/post/149593 (дата обращения: 12.12.2016).

17. История и актуальные проблемы темпоральных баз данных [Электронный ресурс]. URL: http://citforum.ru/database/articles/temporal/ (дата обращения: 10.04.2017).

18. История развития баз данных [Электронный ресурс]. URL: http://bourabai.ru/dbt/dbms/1.htm (дата обращения: 13.12.2016).

19. История развития СУБД [Электронный ресурс]. URL: http://generim.ru/istoriya-razvitiya-sistem-upravleniya-bazami-dannyih.html (дата обращения: 13.12.2016).

20. Ланкин О. В. Темпорально-реляционный подход к организации информационного обеспечения автоматизированных систем управления критического применения: Монография / О.В. Ланкин. - Воронеж: Воронежский ЦНТИ – филиал ФГБУ «РЭА Минэнерго России», 2013.- 246 с.

21. Математическое описание нечетких запросов к реляционным базам данных. [Электронный ресурс]. URL: http://rybanoff.narod.ru/bdat/bd_lection_12.pdf (Дата обращения: 16.12.2016)

22. Моделирование темпоральных (временных) данных в хранилищах данных [Электронный ресурс]. URL: http://www.intuit.ru/studies/courses/599/455/lecture/10165 (дата обращения: 12.12.2016).

23. Нечеткие запросы к реляционным базам данных [Электронный ресурс]. URL: http://www.compdoc.ru/bd/other/ill_defined_requests_to_bd/ (Дата обращения: 16.06.2017)

24. Оптимизация MySQL запросов [Электронный ресурс] – URL: https://habrahabr.ru/post/41968/

25. Петухова Н. Проблемы обеспечения информационной безопасности в темпоральных базах данных // Transport and Telecommunications. 2006. № 03. С. 30–32.

26. Петухова Н.Ю. Темпоральные модели данных в информационных системах на железнодорожном транспорте: автореферат дис. ... канд. техн. наук. Рига, 2010. 58 с.

27. Тоноян С.А., Сараев Д.В. Темпоральные модели базы данных и их свойства [Электронный ресурс]: статья, опубликованная в 12-м выпуске журнала «Инженерный журнал: наука и инновации» за 2014 г. URL: http://engjournal.ru/catalog/it/hidden/1333.html (дата обращения: 15.12.2016)

Тамбовский Государственный Технический Университет
Tambov State Technical University

ANALYSIS OF REPRESENTATION EXISTING MODELS OF TEMPORAL DATA

P.A. Trefilov, M.A. Ivanovskiy, N.G. Shakhov, A.I. Eliseev

The questions concerning the representation of temporal entities of the real world using the relational database mechanisms (RDB) are considered. The review and analysis of existing methods for presenting temporal data was carried out, features, advantages and disadvantages were revealed

Keywords: MySQL, relational database, temporal data, data model

УДК004.056.53

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ ПОДСИСТЕМЫ РАЗГРАНИЧЕНИЯ ДОСТУПА В ОПЕРАЦИОННЫХ СИСТЕМАХ LINUX

А.М. Каннер

Проводится исследование эффекта от применения разработанной автором подсистемы разграничения доступа к данным в Linux. Проведен ряд экспериментов для подтверждения следующих свойств данного средства защиты от несанкционированного доступа: гарантированное создание изолированной программной среды с невозможностью нарушения установленной политики управления доступом, исключение возможности «размыкания» данной среды, невозможность возникновения неразрешенных субъектов доступа, отсутствие влияния внедряемых функций защиты на производительность операционной системы.

Ключевые слова: Linux, подсистема разграничения доступа, экспериментальные исследования.

В опубликованных ранее работах автора [1-4] разработаны механизмы защиты информации, с использованием которых разработана подсистема разграничения доступа в операционной системе (ОС) Linux, не обладающая имеющими место недостатками существующих средств защиты информации от несанкционированного доступа (НСД). В [5] обосновано соответствие и корректность данного средства защиты информации от НСД относительно известных формальных моделей безопасности компьютерных систем (КС).

Задачей настоящей статьи является подтверждение ожидаемого (положительного) эффекта от использования предложенной подсистемы разграничения доступа в реальных КС в ходе экспериментальных исследований. В данных исследованиях необходимо экспериментально обосновать и подтвердить следующие свойства подсистемы разграничения доступа: гарантированное создание изолированной программной среды (ИПС) и невозможность ее «размыкания», невозможность нарушения установленной политики управления доступом, невозможность возникновения неразрешенных субъектов доступа, отсутствие влияния внедряемых функций защиты на производительность ОС.

В соответствии с [5] для разработанной подсистемы разграничения доступа существует возможность контролируемой загрузки ядра защиты и контроля целостности (КЦ) всех связанных с ним объектов до загрузки ОС (с использованием доверенной загрузки загрузчика и ОС [4], пошагового КЦ или обеспечения целостности компонент в ходе «ступенчатой» загрузки на технологическом уровне). В результате этого расширяется действие ИПС (см. для сравнения рис. 1 и 2), а процедуры контроля целостности, контроля порождения субъектов доступа и непосредственно разграничения доступа субъектов к объектам активируются с самого раннего этапа загрузки системы (в данный момент существует единственный реальный субъект – системный процесс с UID = 0 [5]) и работают непрерывно в течение любого последующего периода работы ОС. При этом достижимость абсолютной ИПС [5] с невозможностью нарушения установленной политики управления доступом, а также невозможность «размыкания» этой среды обеспечивается тем, что:

– ядро защиты гарантированно запускается (порождается), а целостность связанных с ним объектов либо контролируется (на аппаратном уровне), либо обеспечивается технологически до активации системы;

Каннер Андрей Михайлович – ЗАО «ОКБ САПР», разработчик
e-mail: kanner@okbsapr.ru

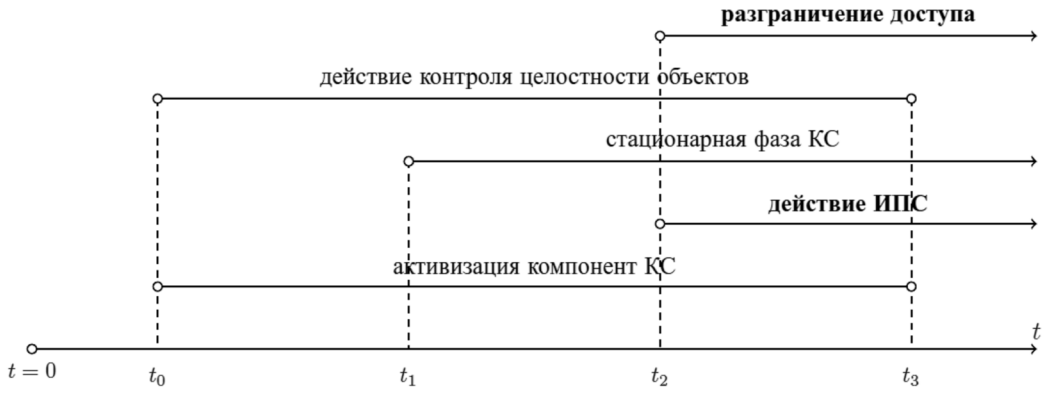


Рис. 1. Этапы функционирования КС формальной СО-модели ИПС

Рис. 2. Этапы функционирования КС с внедренной подсистемой разграничения доступа (с аппаратным КЦ компонент системы, контролем порождения субъектов и разграничением доступа)

– ядро защиты активируется с момента ранней загрузки ОС и сразу после своей инициализации реализует необходимые функции защиты (дальнейшая абсолютная корректность [5] новых субъектов относительно существующих обеспечивается уже за счет этих функций);

– после инициализации ядра защиты существует только два субъекта доступа – оно само и системные процессы с UID = 0 (представляющие один субъект доступа, который ответственен за загрузку ОС и порождение всех дальнейших разрешенных процессов и субъектов системы, см. рис. 3), которые абсолютно корректны относительно друг друга.

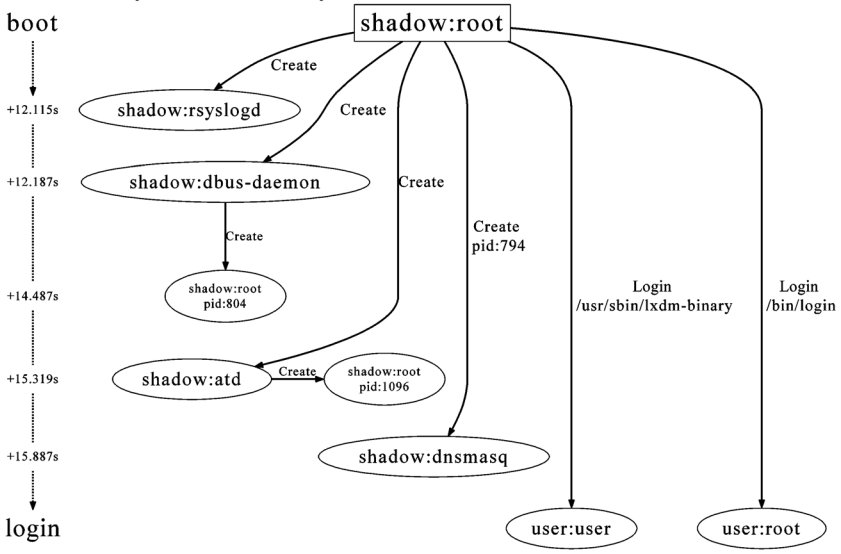


Рис. 3. Граф порождения (санкционированных) субъектов доступа в процессе загрузки одного из дистрибутивов Linux

За счет этого же обеспечивается невозможность возникновения неразрешенных субъектов доступа ОС и возможность порождения только санкционированных (созданных в «обучающем» режиме или вручную при настройке подсистемы разграничения доступа).

В ходе экспериментальных исследований показана изолированность пользовательских сессий (сред) различных субъектов доступа ОС. С использованием информации из журналов событий безопасности (которые при максимальном уровне детальности хранят данные о любых попытках доступа, а не только неразрешенных в соответствии с политикой управления доступом) построены графы санкционированных обращений субъектов к объектам (см. рис. 4). На основе этого выявлено, что в пересечении подграфов все вершины представляют собой общедоступные объекты, для которых кроме контроля доступа необходимо контролировать целостность (например, при любых попытках доступа к данным объектам в ходе динамического контроля целостности [3, 5]). С помощью дополнительного контроля целостности таких объектов становится возможным абсолютно изолировать сессии различных субъектов доступа через общедоступные объекты (в том числе, исключить взаимовлияние этих сессий даже при их каскадном создании – когда одна сессия активизируется из другой в ходе смены субъекта доступа). Однако, это не позволяет исключить взаимовлияние сессий одного и того же субъекта доступа, так как при создании множества таких сессий (удаленных или локальных, с разных устройств ввода/вывода и так далее) общие для них связанные объекты доступа в некоторых случаях доступны для изменения в одной или нескольких сессиях. В соответствии с этим, в подсистеме разграничения доступа предусмотрена возможность запрета создания множества одновременных (параллельных) сессий одного и того же субъекта доступа (что соответствует пункту УПД.9 мер ФСТЭК России по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных [6]).

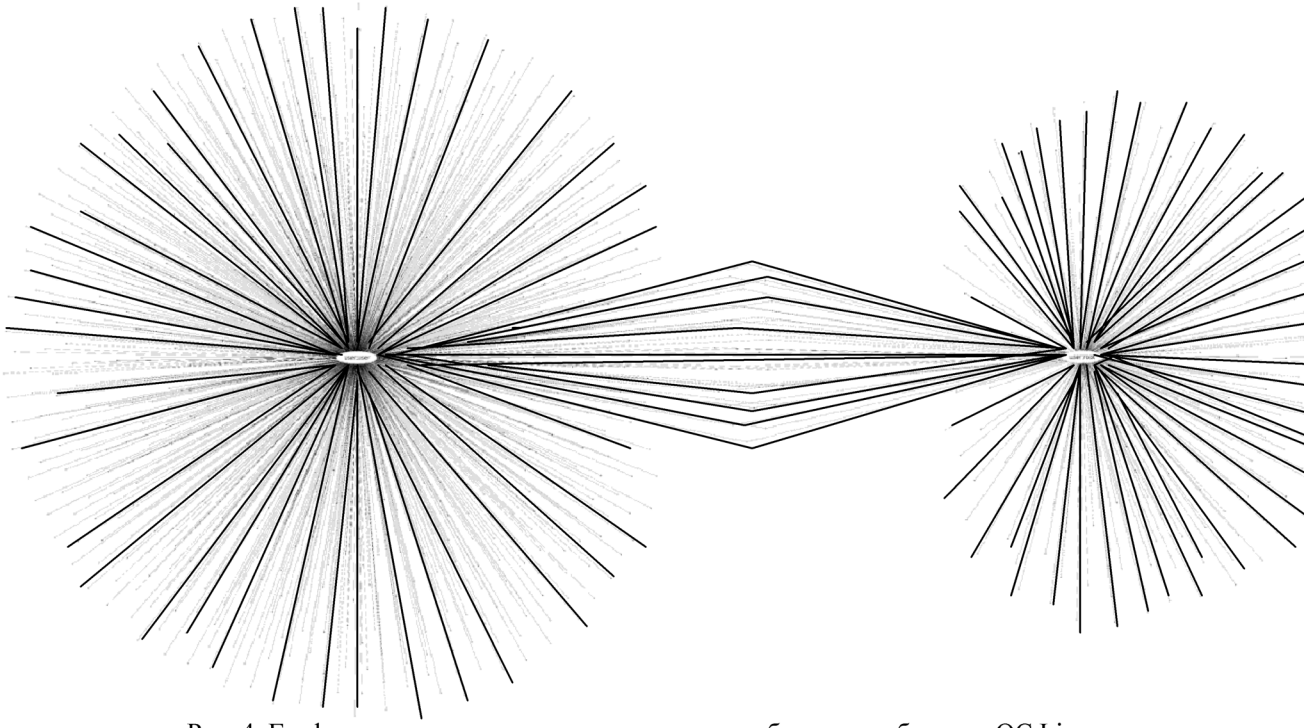


Рис. 4. Граф санкционированных доступов двух субъектов к объектам ОС Linux

Отметим, что ассоциированные [5] с ядром защиты объекты (списки санкционированных субъектов доступа, контроля целостности, контроля доступа и другие) играют ключевую роль в проектировании ИПС. При изменении данных объектов (в том числе санкционированном – то есть в процессе администрирования уполномоченным администратором) возможно, так называемое, «размыкание» ИПС. В разработанной на основе [1-4] подсистеме разграничения доступа ядро защиты функционирует на уровне ядра ОС, а доступ к ассоциированным с ядром защиты объектам имеет только выделенный субъект доступа (администратор безопасности). При этом само администрирование выполняется только в специальные промежутки времени, когда в системе не существует других субъектов доступа. Применение данных мер (то есть, так называемого, абсолютного разделения пользовательских и административных полномочий) позволяет исключить возможность «размыкания» ИПС.

Для исследования оказываемого влияния подсистемой разграничения доступа на производительность системы был проведен ряд benchmark-тестов. Тесты запускались в следующих условиях:

1. ОС со стандартными настройками сразу после «чистой» установки;
2. ОС с внедренной подсистемой разграничения доступа;
3. аналогично пункту выше, но с включенной функцией очистки оперативной памяти и остаточной информации.

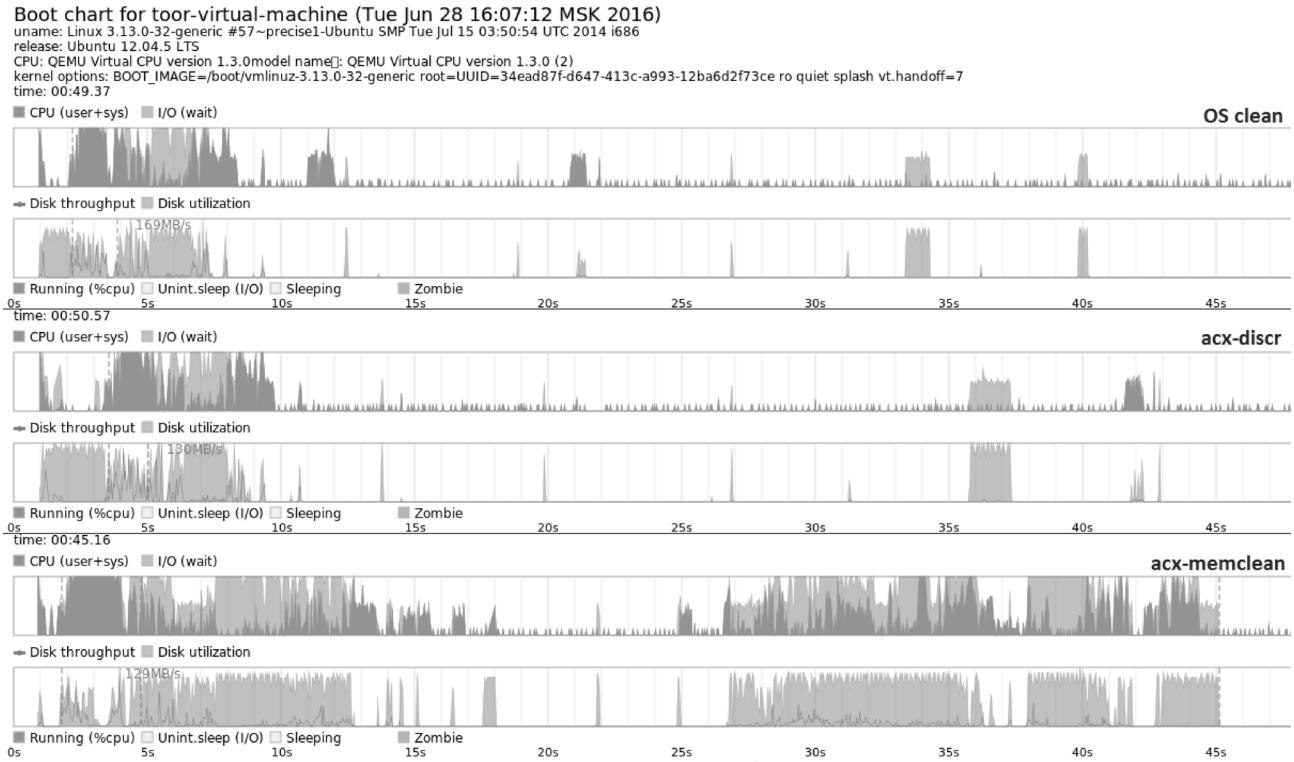


Рис. 5. Результаты тестов загрузки ОС и работы системных сервисов

Обозначения: *OS clean* – «чистая» установка ОС, *acx-discr* – ОС с внедренной подсистемой разграничения доступа, *acx-memclean* – дополнительное включение функций очистки памяти.

Сами тесты проводились в двух состояниях системы: процесс загрузки и стационарная фаза. В первом случае замерялось время работы всех системных сервисов и процессов в ходе загрузки ОС (краткие результаты см. на рис. 5). Нужно отметить, что загрузка ОС является наиболее критичным периодом работы за счет огромного количества различных выполняемых системных вызовов ядра ОС.

Так как выполнение каждого системного вызова влечет за собой выполнение определенной функции подсистемы разграничения доступа – это может негативно сказаться на производительности ОС.

Во втором случае запускались стресс-тесты из состава ПО *phoronix-test-suite*, непосредственно связанные с вводом-выводом и операциями с объектами доступа: *build-linux-kernel* и *unpack-linux* для сравнения времени компиляции и распаковки ядра, *compilebench* – компиляции произвольных проектов, *compress-gzip* – проведения операции сжатия, *iozone* в отношении эффективности различных операций ввода-вывода (краткие результаты см. на рис. 6).

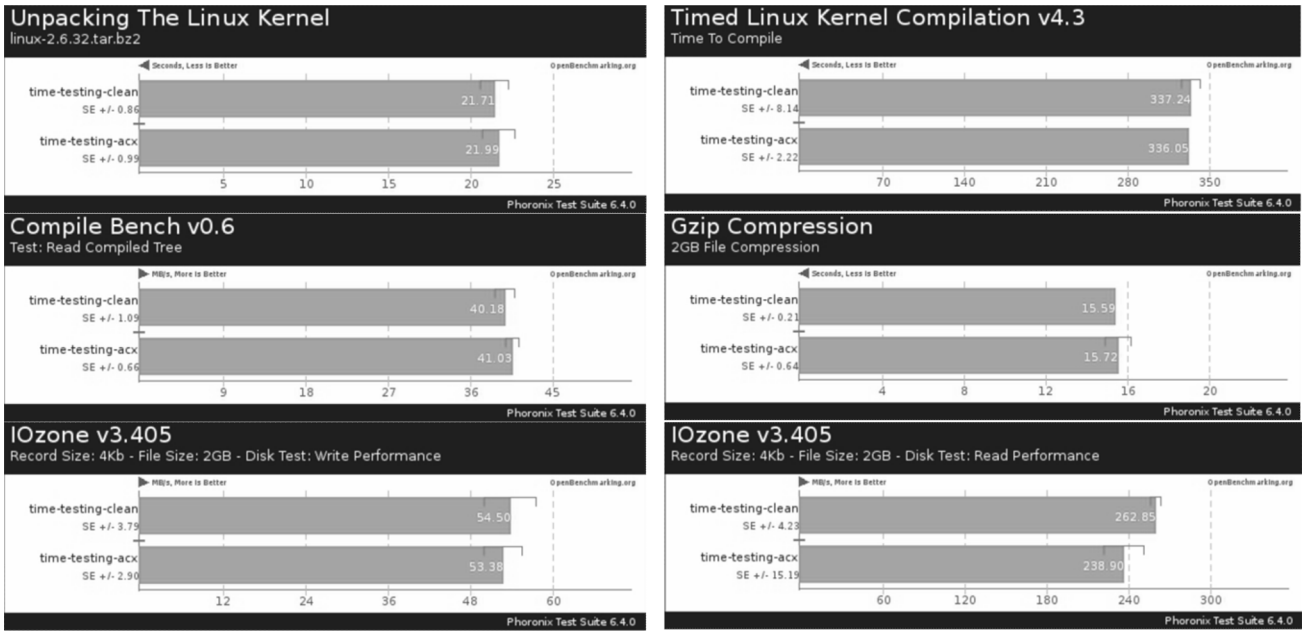


Рис. 6. Результаты тестов измерения производительности ОС без дополнительных средств защиты от НСД и с внедренной подсистемой разграничения доступа

В результате проведенных тестов можно сделать вывод о том, что разработанная на основе [1-4] подсистема разграничения доступа в целом не оказывает существенного влияния на производительность ОС Linux. Однако, при использовании механизма очистки оперативной памяти [7] была замечена потеря производительности приложений и ОС в среднем на 1-2%, что вполне удовлетворяет поставленной задаче по минимизации влияния на общую производительность системы. Механизм же очистки остаточной информации в уничтожаемых файлах понижает производительность операции удаления примерно в 2 раза. В соответствии с этим очистку остаточной информации целесообразно производить для удаления только защищаемых объектов доступа, с целью исключения потери производительности в масштабах всей системы.

Таким образом, проведенные исследования подтвердили как выполнение ожидаемых свойств ОС со встроенной подсистемой разграничения доступа (которые, несомненно, свидетельствуют о качественном увеличении защищенности данных), так и отсутствие существенного влияния внедренных функций защиты от НСД на производительность системы.

Литература

1. Каннер А.М., Ухлинов Л.М. Управление доступом в ОС GNU/Linux // Вопросы защиты информации. Научно-практический журнал. — М., 2012. — № 3 — С. 35–38

2. Каннер А.М. Linux: к вопросу о построении системы защиты на основе

абсолютных путей к объектам доступа // Комплексная защита информации. Электроника инфо. Материалы XVIII Международной конференции 21-24 мая 2013 года. — Брест (Республика Беларусь), 2013. — № 6. — С. 126–128

3. Каннер А.М. Linux: объекты контроля целостности // Комплексная защита информации. Электроника инфо. Материалы XVIII Международной конференции 21-24 мая 2013 года. — Брест (Республика Беларусь), 2013. — № 6. — С. 123–126

4. Каннер А.М. Linux: о доверенной загрузке загрузчика ОС // Безопасность нформационных технологий. — М., 2013. — № 2. — С. 41–46

5. Kanner A.M. Correctness of Data Security Tools for Protection against Unauthorized Access and their Interaction in GNU/Linux // Global Journal of Pure and Applied Mathematics. — 2016. — Vol. 12, no. 3. — Pp. 2479–2501

6. Приказ ФСТЭК России №21. Об утверждении состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных — 2013. — С. 20

7. Прокопов В.С., Каннер А.М. Особенности реализации механизма очистки освобождаемых областей оперативной памяти в GNU/Linux // Комплексная защита информации. Электроника инфо. Материалы XVIII Международной конференции 21-24 мая 2013 года, Брест (Республика Беларусь). — 2013. — № 6. — С. 120–123

Закрытое акционерное общество «ОКБ САПР»
Closed Joint Stock Company «OKB SAPR»

EXPERIMENTAL RESEARCH OF ACCESS CONTROL IN LINUX OPERATING SYSTEM

A.M. Kanner

Author considers the effect of using previously developed access control subsystem in Linux. A series of experiments is held to confirm some properties of this data security tool: guaranteed creation of isolated program environment with the impossibility of violation of the established security policy, the exclusion of the possibility of "opening" of such sandbox, inability of creation of unauthorized subjects and the lack of an impact of security functions on operating system performance.

Keywords: Linux, access control subsystem, experimental research.

УДК 004.891.3

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ДЕЯТЕЛЬНОСТИ ПРАВООХРАНИТЕЛЬНЫХ ОРГАНОВ ПО ОБЕСПЕЧЕНИЮ ОБЩЕСТВЕННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ

А.М. Ахметвалеев, А.С. Катасёв

В статье рассматривается проблема взаимодействия сотрудников полиции с гражданами при осуществлении ими функций по обеспечению общественной безопасности. Для ее решения предлагается и рассматриваются интеллектуальные технологии, основанные на интеллектуальном анализе видеоизображений

Ключевые слова: обеспечение общественной безопасности, интеллектуальные технологии, потенциально опасное лицо, видеоаналитика, целеуказание

В настоящее время всё большую актуальность перед правоохранительными органами приобретают задачи обеспечения общественной безопасности, в числе которых выделяют профилактику и предупреждение правонарушений, а также противодействие преступности. Как известно, наиболее эффективным инструментом в деятельности полиции является патрулирование местности и обеспечение правопорядка в общественных местах. Однако, несмотря на широкий спектр прав сотрудников полиции, их взаимодействие с гражданами, в том числе «подозрительными», жестко регламентировано Федеральным законом "О полиции" от 7 февраля 2011 г. N 3-ФЗ [1].

Так, правилами обращения к гражданам указанного закона установлено, что основаниями для проверки документов являются следующие обстоятельства:

- наличие данных, дающих основания подозревать граждан в совершении преступления;
- наличие данных считать, что гражданин находится в розыске;
- наличие повода к возбуждению в отношении гражданина дела об административном правонарушении;
- имеются основания для задержания.

Перечисленные обстоятельства, несмотря на их конституционность и гуманность, по факту являются ограничивающими факторами, существенно снижающими возможности полиции по работе с так называемым неблагополучным контингентом, среди которого вероятность совершения преступлений наиболее высока. Сотрудник полиции не имеет права превышать свои полномочия и не станет обращаться к гражданам для установления личности или досмотра, несмотря на очевидность такой необходимости, но отсутствии достаточных условий для реализации своих прав.

В то же время, согласно ст.11 того же Федерального закона, полиция в своей деятельности обязана использовать достижения науки и техники, информационные системы и т.д. На сегодняшний день активно внедряются и эксплуатируются системы городского видеомониторинга, ситуационные центры (см. рис. 1), элементы видеоаналитики, основанные на интеллектуальных технологиях (см. рис. 2) и другие технические средства. Любые внедряемые технологии в разной степени способствуют повышению эффективности деятельности полиции по обеспечению общественной безопасности [7], однако при их внедрении необходимо находить баланс между критериями стоимости и эффективности перспективных к внедрению систем.

На сегодняшний день получили распространение различные аналитические системы, основанные на видеонаблюдении. Они находят применение в сферах обеспечения общественной безопасности, безопасности дорожного движения и т.д.

Ахметвалеев Амир Муратович – КНИТУ-КАИ, аспирант, e-mail: amir1985@bk.ru
Катасёв Алексей Сергеевич – КНИТУ-КАИ, канд. технич. наук, доцент, e-mail: Kat_726@mail.ru

Рис. 1. Рабочие места операторов ситуационного центра

На рисунке представлен пример ситуационного центра городского видеомониторинга г. Сочи (Центр управления «Безопасный Сочи»). В данном Центре 40 операторами осуществляется круглосуточный мониторинг изображений с более 1200 городских камер наблюдения, установленных в наиболее оживленных местах города-курорта. При обнаружении признаков преступления, Центр в реальном времени обеспечивает координацию и взаимодействие со всеми силовыми структурами по дальнейшим действиям.

Очевидно, что при реализации данного подхода требуются существенные затраты, среди которых можно выделить первоначальные вложения в создание ситуационных центров и их техническое оснащение, эксплуатационные расходы на обслуживание инженерных и технических систем, заработную плату персоналу и т.д. Таким образом, следует вывод о невысокой экономической эффективности подобных решений, ввиду больших финансовых затрат из средств федерального, регионального и муниципального бюджетов.

Наиболее перспективными решениями для повышения эффективности деятельности правоохранительных органов по обеспечению общественной безопасности являются интеллектуальные технологии. Исследованиями различных авторов отмечаются высокие показатели эффективности внедряемых решений, возможности автоматизации процессов обнаружения правонарушений [9,12] и выявления потенциально опасных лиц [2,3,5,11]. Так, наибольший интерес вызывают интеллектуальные системы, использующие так называемое интеллектуальное целеуказание, при котором на кадре видеоизображения отмечаются предметы, действия или лица, способные вызывать интерес правоохранительных органов, не нарушая при этом положения регулирующего Федерального закона. На рисунке 2 представлены примеры подобного целеуказания.

а) б) в)

Рис. 2. Примеры целеуказания лиц в различных системах

На рисунке представлены примеры целеуказаний, полученных в результате обработки с применением технологий интеллектуального анализа видеоизображения. Так, в примере 2а интеллектуальная система выделила лежащего на полу человека, который в реальной ситуации оказался в данном положении после полученных во время драки травм. В данном случае интеллектуальный алгоритм выделяет значимые признаки на изображении и способен их отслеживать [8,10].

На рисунке 2б выделен человек, имеющий аномалии в двигательной (моторной) активности, что может говорить о его перевозбуждении, эмоциональном напряжении или психическом расстройстве. Здесь алгоритмом регистрируются вибрации тела человека, по частоте колебаний которого можно судить о физиологической норме или отклонении [11].

Пример 2в показывает возможность выявления потенциально опасных лиц, находящихся в состоянии алкогольного, наркотического или иного токсического опьянения. Интеллектуальный алгоритм выделяет в поле зрения видеокамер лица, имеющие функциональные отклонения в зрачково-двигательном рефлексе на изменение освещенности [3,4].

Последний подход основан на методе пупиллометрии – способе исследования зрачковой реакции, использующемся в медицинской практике для выявления патологий параметров цереброспинальных вегетативных центров, отражающих состояние организма, нервной системы, а также некоторых внутренних органов [14]. Методика выявления потенциально опасных лиц (ПОЛ), реализованная на пупиллометрическом подходе, в первую очередь направлена на содействие правоохранительным органам при решении задач обеспечения общественной безопасности. Выявление ПОЛ в потоке людей способствует более рациональному в условиях ограниченности ресурсов акцентированию внимания сотрудников полиции на отдельных субъектах, а не общей массе граждан. Также очевидна оперативность подхода и снижение временных затрат на выявление ПОЛ, поскольку анализ зрачковой реакции занимает не более 3 секунд [3, 6, 13], а дальнейшая непосредственная проверка лиц на состояние опьянения становится более избирательной и точечной.

Таким образом, при наличии доказанной эффективности интеллектуальных систем и комплексов, результат их работы можно использовать в качестве правового основания (к примеру, наличие повода к возбуждению административного дела) для обращения к гражданам с целью установления их личности или досмотра. Следовательно, можно утверждать о повышении эффективности проведения мероприятий, направленных на обеспечение общественной безопасности, за счет автоматизации выявления правонарушений, а также предварительного интеллектуального отбора и идентификации граждан, представляющих интерес для правоохранительных органов. В результате повысится не только уровень безопасности граждан, но и качество их жизни.

Литература

1. Федеральный закон от 07.02.2011 N 3-ФЗ (ред. от 18.06.2017) "О полиции" [Электронный ресурс] http://www.consultant.ru/document/cons_doc_LAW_110165/ (дата обращения 27.06.17).

2. Ахметвалеев А.М., Катасёв А.С. Выявление потенциально опасных лиц в системах обеспечения общественной безопасности // Информационная безопасность и защита персональных данных. Проблемы и пути их решения: материалы VII Межрегиональной научно-практической конференции. – Брянск: БГТУ, 2015. – С. 23-26

3. Ахметвалеев А.М., Катасёв А.С. Концепция бесконтактной идентификации лиц, представляющих угрозу общественной безопасности // Современные проблемы безопасности жизнедеятельности: материалы IV Международной научно-практической конференции / Под общей ред. д-ра техн. наук, проф. Р.Н. Минниханова. – Казань: ГБУ «Научный центр безопасности жизнедеятельности». – 2016. – С. 67-72.

4. Ахметвалеев А.М., Катасёв А.С., Кирпичников А.П. Редукция нейросетевых моделей на основе метода двухэтапной генетической оптимизации // Вестник Казанского технологического университета. – Т. 20. – № 9. – 2017. – С. 71-75

5. Ахметвалеев А.М., Катасёв А.С., Шлеймович М.П. Повышение эффективности обнаружения лица и глаз человека на видеоизображении в задачах бесконтактного выявления потенциально опасных лиц // Информация и безопасность. – Т. 19. – № 4 (4) – 2016. – С.519-522

6. Ахметвалеев А.М., Катасёв А.С., Шлеймович М.П. Проблема стимуляции направления взгляда человека в задачах бесконтактного выявления потенциально опасных лиц // Информационная безопасность и защита персональных данных. Проблемы и пути их решения: Материалы VIII Всероссийской научно-практической конференции / под ред. О.М. Голембиовской, М.Ю.Рытова. - Брянск: БГТУ, 2016. – С.5-8

7. Губанов Н.Н. Информатика и информационные технологии в профессиональной деятельности // Курс лекций. – Ставрополь: КУ МВД РФ (Ставропольский Филиал), 2014.

8. Дзенчарский Н.Л., Медведев М.В., Шлеймович М.П. Поиск изображений с выделением особых точек на основе вейвлет-преобразования // Вестник КГТУ им. А.Н. Туполева. Казань: Изд-во Казан. гос. техн. ун-та. – 2011. – № 1. – С. 131-135.

9. Катасёв А.С., Глова В.И., Шакиров Р.Х. Системы поддержки принятия решений при видеомониторинге подвижных объектов // Информационные технологии в науке, образовании и производстве: Материалы Всероссийской научной конференции. Казань: Изд-во Казан. гос. техн. ун-та, 2007. – С. 519-522.

10. Катасёв А.С., Макаров Д.А. Методика, алгоритмы и программный комплекс слежения за движущимся объектом в системах видеонаблюдения // Вестник КГТУ им. А.Н. Туполева. – №4. – 2010. – С. 145-150.

11. Конобеевский М.А. Способ бесконтактного оптического измерения параметров вибрации механизмов, конструкций и биологических объектов // Патент на изобретение № 2546714.

12. Синезис. Видеоанализ в системах защиты периметра. [Электронный ресурс] https://habrahabr.ru/company/synesis/blog/137006/ (дата обращения 27.06.17).

13. Фоменко В. Н., Куприянов А. С. Математические модели зрачковых реакций глаза человека (пупиллограмм) // Известия ПУТС. – СПб.: Петербургский гос.ун-т путей сообщения, 2010. – Вып. 4 (25). – С. 220-230.

14. Цимбал Ф.А., Цимбал М.В. Исследование порога чувствительности метода пупиллометрии при интоксикации фосфорорганическими соединениями // Токсикологический вестник. – 2007. – № 1. – С. 26-28.

ФГБОУ ВО «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ»
Kazan National Research Technical University

IMPROVING THE EFFICIENCY OF POLICE TO ENSURE PUBLIC SAFETY ON THE BASIS OF INTELLIGENT TECHNOLOGIES

A.M. Akhmetvaleev, A.S. Katasev

The article discusses the problem of police officers' interaction with citizens in the course of their functions to ensure public safety. To solve it, intelligent technologies based on intelligent analysis of video images are proposed and considered

Key words: public safety, intellectual technologies, potentially dangerous person, video analytics, target designation

УДК 004.056:061.68

МОДЕЛЬ ДЕЦЕНТРАЛИЗОВАННОГО УПРАВЛЕНИЯ МЕТКАМИ БЕЗОПАСНОСТИ ДАННЫХ В ОБЛАЧНОЙ СРЕДЕ

А.В. Царегородцев, М.А. Попов

В статье рассматривается модель управления доступом к данным информационного потока в сложной, децентрализованной, с взаимным недоверием среде, к которой относится среда облачных вычислений. Возможность управления доступом в недоверенной среде достигается за счет присвоения атрибутов владельца данных политике безопасности информационного потока

Ключевые слова: информационная безопасность, облачные вычисления, гибридная облачная среда, модель управления доступом, метки безопасности

Принимая во внимание парадигму облачных вычислений, организация отказывается от прямого контроля над многими аспектами безопасности и, тем самым, создаёт беспрецедентный уровень доверия облачному провайдеру [1]. В связи с этим возникает проблема обеспечения информационной безопасности при обработке данных в условиях среды облачных вычислений [2].

Известно, что управление информационными потоками обеспечивает прямую защиту конфиденциальных данных в отличие от дискреционного управления доступом, но существующие практические проблемы применения теории информационных потоков препятствуют их повсеместному распространению. Поддержка меток конфиденциальности обычно осуществляется динамически, что приводит к потере производительности и нарушению непрерывности бизнеса.

В связи с чем всё большую актуальность приобретает задача построения модели управления доступом в сложной, децентрализованной, с взаимным недоверием среде, к которой относятся облачные вычисления. И первое, что необходимо сделать – это разработать подход к маркированию информационного потока для среды облачных вычислений.

Царегородцев Анатолий Валерьевич – Московский государственный лингвистический университет, д-р техн. наук, профессор, e-mail: academic_tsar@mail.ru
Попов Максим Анатольевич – Институт информационных наук и технологий безопасности РГГУ, аспирант, e-mail: maxmax@bk.ru

Предлагается настройку меток секретности осуществлять децентрализовано для некоторого сообщества клиентов, использующих облачный сервис. Очевидно, что нельзя установить универсальный подход проставления значений меток, т.к. каждый субъект самостоятельно должен определить критичность обрабатываемых данных в рамках своего бизнес контекста. Традиционные модели управления потоками данных не поддерживают операцию уменьшения уровня секретности в недоверенной среде. Сформулируем ключевые требования к децентрализованной модели для обеспечения безопасной обработки данных в недоверенной среде.

1. Модель должна позволять субъектам присваивать собственные метки секретности к элементам данных. При этом политика безопасности информационного потока всех субъектов отражена в принципе маркирования данных и модель должна работать даже при условии, когда субъекты не доверяют друг другу.

2. Модель должна позволять субъектам понизить уровень секретности путем изменения политик в присваиваемой метке. При этом произвольное рассекречивание данных не возможно в силу учета политик безопасности других субъектов.

3. Модель должна предоставлять полный набор безопасных правил смены меток, которые будем называть перемаркировками.

4. Модель должна иметь формальное семантическое описание, которое позволит точно определить является ли

перемаркировка допустимой, с возможностью проверки на допустимость изменения политик безопасности других субъектов в метке данных.

5. Метки безопасности должны иметь решетчатую структуру, позволяющую проводить статическую проверку эффективности работы облачных сервисов.

6. Модель должна быть расширяема на соответствие требованиям целостности данных, что приведет к повышению уровня безопасности обрабатываемых данных.

Ключевыми элементами разработанной модели являются:

– владельцы данных, для которых следует обеспечивать требуемый уровень секретности;

– метки, с помощью которых субъекты-владельцы контролируют права доступа к своим данным;

– правила, которым необходимо следовать при обработке для предотвращения утечек данных, включая механизм безопасного понижения уровня секретности.

Под субъектами децентрализованной модели будем понимать активные сущности, которые совершают операции над данными в процессе обработки информационного потока. Под меткой безопасности будем понимать набор политик, описывающих требования безопасности. Таким образом, политика включает в себя две ключевые переменные: владельца (o) и множество субъектов (r), имеющих право на чтение данных, и может быть описана в виде {o; r}. Владелец данных – это субъект, который инициирует запрос на обработку данных с соответствующей меткой, включающей в себя политики безопасности.

Формально метку L можно представить в виде множества:

$$L = \{o_1: r_1 r_2; o_2: r_2 r_3\},$$

где $o_1, r_1 r_2, r_3$ – это субъекты, знак «;» разделяет две составные политики одной метки.

В модели каждая переменная имеет метку, которая применяется для всех значений этой переменной. Когда над значением переменной совершается операция чтения данных, переменная должна иметь ту же метку, что и значение. Когда новое значение сохраняется в переменную, метка значения стирается и перезаписывается меткой этой переменной. Таким образом, присвоение значения переменной вызывает операцию смены метки в копии значения. Чтобы избежать утечки информации, метка копируемого значения должна, по крайней мере, иметь то же ограничение, что и исходная метка значения. Такого рода операцию будем называть ограничением. Выражение $L_1 \sqsubseteq L_2$ означает, что метка L_1 имеет меньше ограничений или равна метке L_2 (альтернатива: метка L_2 по меньшей мере имеет то же ограничение, что и L_1). Используя это определение присвоение значения x в переменную v допустимо, если $L_x \sqsubseteq L_v$, где L_x и L_v соответствующие метки.

Смена метки ограничена условием, что все политики старой метки должны исполняться в новой метке. Политика J метки L_1 должна гарантировано учитываться в политике K, если обе политики имеют одного и того же владельца, а набор пользователей с доступом на чтение K является подмножеством набора J.

Поскольку метки в рассматриваемой модели содержат информацию о владельцах данных, эти владельцы могут сохранить контроль над распространением своих данных и в случае необходимости понизить уровень ограничения доступа. Этот способ понижения уровня секретности можно рассматривать, как второй вариант смены метки данных.

Таким образом, метка L_1 может быть изменена на L_2, если $L_1 \sqsubseteq L_2 \sqcup L_A$, где L_A – это метка, содержащая точный набор политик состояния {p:} для каждого субъекта (p) в рамках текущих полномочий. Данное предположение основывается на правиле смены метки вследствие ограничений. Подмножество правил для изменения метки L_1 на L_2 утверждает, что для всех политик J в L_1 должна быть политика K в метке L_2, которая, по меньшей мере, имеет те же ограничения.

При рассмотрении среды облачных вычислений все пользователи являются внешними субъектами по отношению к запускаемым облачным сервисам. Считаем,

что информация будет подвержена утечке только в случае попадания в исходящий канал компонента частной облачной среды. Поэтому входящий и исходящий каналы компонентов облачной среды необходимо также маркировать для предотвращения утечек данных.

При этом, в случае чтения данных из входящего канала значение принимает метку этого входящего канала. Аналогично значение может быть записано в исходящий канал только в том случае, если метка исходящего канала имеет, по меньшей мере, те же ограничения, что и метка передаваемого значения, в противном случае предполагается, что будет происходить утечка информации. Маркировку исходящих каналов можно представить в виде функции, которая принимает данные с меткой L и последовательно осуществляет следующие три операции.

1. Трансформация данных, например, шифрование с помощью открытого ключа.

2. Деклассификация (понижение уровня секретности) данных до метки { }.

3. Передача через исходящий канал компонента.

Расширим модель обеспечения требований конфиденциальности данных и на соблюдение требований целостности. В то время как обеспечение конфиденциальности защищает от неправомерного чтения при обработке данных со стороны недоверенного облачного сервиса, политика целостности защищает данные от неправомерного изменения.

Структура децентрализованной политики обеспечения целостности идентична структуре политики конфиденциальности данных. Аналогично вводятся понятия владельца, список субъектов с правами на запись данных, которым разрешено изменять данные владельца. Метка может содержать несколько политик целостности данных различных владельцев информации. Политика $\{o: w_1, w_2\}$ гарантирует владельцу o, что только субъекты w_1 и w_2 могут изменять значения его данных.

Допустимое изменение политики конфиденциальности может быть совершено с помощью 5 инкрементных шагов, аналогично введем правила и для изменения метки целостности – см. табл. 1.

Табл. 1

Зависимость правил политик целостности и конфиденциальности

Номер правила	**Описание правила политики целостности**	**Описание правила политики конфиденциальности**
1	Добавление пользователя с правами на запись данных	Разрешение удаления субъекта с правами на чтение
2	Удаление политики	Добавление произвольной политики
3	Замена субъекта w′ на w, где $w' \succcurlyeq w$.	Добавление субъекта r′ к списку разрешенных, если r также является разрешенным субъектом и $r' \succcurlyeq r$ и в этом случае r может быть удален
4	Добавление политики J идентичной существующей политике I с владельцем с более низким уровнем доступа ($oI \succcurlyeq oJ$).	Замена владельца J на oI, что делает две политики идентичными.
5	Удаление владельца политики из набора субъектов с правами на запись	Владелец политики может быть добавлен в набор субъектов с правами на чтение

Если L_1 и L_2 являются метками конфиденциальности и L_1 может быть изменена на L_2, то это изменение можно описать в виде последовательного применения правил обеспечения конфиденциальности. Предположим, что L_1^I и L_2^I являются метками целостности в той же модели безопасности. Тогда существует последовательность шагов для смены метки L_2^I на L_1^I; если $L_1 \sqsubseteq L_2$, то $L_1^I \sqsubseteq L_2^I$. Это свойство означает, что все правила обеспечения целостности могут быть

получены напрямую с помощью преобразований правил обеспечения конфиденциальности. Для меток конфиденциальности L_1 и L_2 и меток целостности L_1^I и L_2^I справедливо:

$$P \vdash L_1 \sqsubseteq L_2 \leftrightarrow P \vdash L_2^I \sqsubseteq L_1^I.$$

Также можно представить и правила для объединения и пересечения двух меток целостности в виде обратной функции правил конфиденциальности:

$$L_3 \approx L_1 \sqcup L_2 \leftrightarrow L_3^I \approx L_1^I \sqcap L_2^I,$$
$$L_3 \approx L_1 \sqcap L_2 \leftrightarrow L_3^I \approx L_1^I \sqcup L_2^I.$$

Таким образом, можно определить операцию для меток целостности, аналогичную операции понижения уровня секретности для меток конфиденциальности. Для меток конфиденциальности механизмы понижения уровня секретности позволяют осуществлять удаление политик. Обратное действие для политик целостности – это добавление новых политик в ситуации, когда данные имеют более высокие требования к целостности данных, чем предполагает анализ зависимостей.

Понижение уровня целостности для метки L_1 на метку L_2 разрешено, если $L_2 \sqcap L_A^I \sqsubseteq L_1$, где L_A^I метка целостности, в которой существует политика для каждого субъекта, участвующего в процессе.

Представим, как S_p набор меток конфиденциальности, S_I – набор меток целостности, каждая из которых может создавать соответствующие отношения $\sqsubseteq_p$ и $\sqsubseteq_I$. Эти два вида меток могут быть использованы для одновременно создания системы комбинированных меток, обеспечивающих ограничения на целостность и конфиденциальность.

Комбинированная метка может быть рассмотрена, как пара $\{L_P, L_I\}$ набора $S_P \times S_I$. Для комбинированных меток справедливы следующие виды отношений:

$$\{L_P, L_I\} \sqsubseteq \{L'_P, L'_I\} \equiv L_P \sqsubseteq_p L'_P \wedge L_I \sqsubseteq_I L'_I,$$
$$\{L_P, L_I\} \sqcup \{L'_P, L'_I\} \equiv L_P \sqcup_p L'_P \wedge L_I \sqcup_I L'_I,$$
$$\{L_P, L_I\} \sqcap \{L'_P, L'_I\} \equiv L_P \sqcap_p L'_P \wedge L_I \sqcap_I L'_I.$$

Аналогично для комбинированной метки $\{L_P, L_I\}$ может быть совершена операция понижения уровня.

Таким образом, в статье предложен подход к маркированию информационного потока для построения модели безопасной обработки данных в облачной среде, основанной на принципе децентрализации уровней секретности.

Литература

1. Царегородцев, А.В. Методика построения защищенных информационно-телекоммуникационных систем на базе гибридной облачной среды [Текст] / Царегородцев, А.В., Мухин, И.Н., Белый, А.Ф. // Информация и безопасность. 2015. Т.18, №3. С.404-407.

2. Царегородцев, А.В. Построение деревьев целей для идентификации требований безопасности среды облачных вычислений [Текст] / Царегородцев, А.В. // Национальная безопасность. 2013. №5(28). С.51-69.

ФГБОУ ВО «Московский государственный лингвистический университет»
Moscow State Linguistic University
ФГБОУ ВО «Российский государственный гуманитарный университет»
Russian State University for the Humanities

MODEL OF DECENTRALIZED MANAGEMENT OF DATA SECURITY MARKS IN THE CLOUD ENVIRONMENT

A.V. Tsaregorodtsev, M.A. Popov

The article deals with the model of access control to the information flow data in a complex, decentralized environment with mutual distrust, which includes the cloud computing environment. The ability to control access in the unverified environment is achieved by assigning data owner attributes to the information flow security policy

Keywords: information security, cloud computing, hybrid cloud environment, access control model, security marks

УДК 004.056.57

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО ОБЕСПЕЧЕНИЮ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ

В.И. Васильев, В.Е. Гвоздев, М.Б. Гузаиров, А.Д. Кириллова

В статье рассматривается проблема обеспечения информационной безопасности автоматизированной системы управления технологическими процессами (АСУ ТП). Приводится обзор основных нормативных документов по обеспечению безопасности АСУ ТП. Для оперативного решения задач обеспечения информационной безопасности АСУ ТП предлагается использовать интеллектуальную систему поддержки принятия решений (СППР). Рассмотрен пример построения и реализации решающих правил в составе СППР.

Ключевые слова: система поддержки принятия решений, информационная безопасность, автоматизированная система управления технологическими процессами, оценка риска, вероятность реализации угрозы, нечеткая логика, нейронная сеть

Большинство автоматизированных систем управления технологическими процессами (АСУ ТП) были разработаны в 80-х годах и основными объектами автоматизации являлись крупные предприятия. Вопросы обеспечения информационной безопасности (ИБ) не имели первостепенную важность, поскольку системы функционировали в изолированной среде, использовались индивидуальное программное обеспечение, специфичные сетевые протоколы. Это порождало ложное состояние защищенности.

К настоящему времени АСУ ТП вышли за рамки крупного производства и получили широкое применение в различных областях. Для любого предприятия или организации повышение эффективности производства в первую очередь определяется эффективностью существующей системы управления. Решение задач координации взаимодействия между всеми подразделениями, оперативной обработки и анализа получаемых данных сегодня невозможно без обеспечения информационной безопасности АСУ ТП [1].

Васильев Владимир Иванович – УГАТУ, д.т.н. профессор, e-mail: vasilyev@ugatu.ac.ru
Гвоздев Владимир Ефимович – УГАТУ, д.т.н., профессор, e-mail: wega55@mail.ru
Гузаиров Мурат Бакеевич – УГАТУ, д.т.н., профессор, e-mail: guzairov@ugatu.su
Кириллова Анастасия Дмитриевна – УГАТУ, магистр, e-mail: kirillova.andm@gmail.com

Актуальность проблемы ИБ АСУ ТП заключается в необходимости выполнения требований повышенной надежности систем автоматизации, поскольку современные автоматизированные системы управляют сложными и опасными технологическими процессами, сбой в которых может привести к авариям на производстве или техногенным катастрофам.

В целом, вопросам обеспечения ИБ АСУ ТП пока уделяется недостаточное внимание в законодательстве, регулирующем безопасность критически важных, потенциально опасных объектов на территории Российской Федерации, наравне с вопросами обеспечения промышленной безопасности АСУ ТП таких объектов.

Развитие нормативно-методической базы по обеспечению ИБ АСУ ТП началось с выпуска в 2007г. документов ФСТЭК России, регулирующих вопросы обеспечения ИБ в ключевых системах информационной инфраструктуры (КСИИ). Данные документы имеют гриф «Для служебного пользования», что подтверждает важность содержащейся в них информации и крайней необходимости защиты информационных инфраструктур. Ограничение доступа к документации такого рода затрудняет работу специалистов предприятий, принимающих решения о проведении мероприятий по обеспечению ИБ АСУ ТП.

В 2011г. был принят Федеральный закон № 256-ФЗ «О безопасности объектов топливно-энергетического комплекса» [2]. В

нем более детально проработаны вопросы обеспечения ИБ критически важных объектов и потенциально опасных объектов топливно-энергетического комплекса и, в частности, АСУ ТП, которые функционируют в составе данных объектов.

Развитие высокотехнологических систем на объектах повышенной опасности для жизни, здоровья человека и окружающей среды, а также информация о реализованных атаках на инфраструктуру критически важных объектов привели к принятию в 2014г. Приказа ФСТЭК России № 31 «Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей среды» [3].

Приказом определен гибкий подход по заданию требований к обеспечению защиты информации (ЗИ) в автоматизированных системах управления, который позволяет учитывать и нейтрализовать все актуальные угрозы на объекте, а также позволяет учитывать структурно-функциональные характеристики и особенности этого объекта [3].

В приказе рассмотрены следующие вопросы:

– требования к организации ЗИ в АСУ ТП;

– требования к мерам ЗИ в АСУ ТП;

– определение класса защищенности АСУ ТП;

– состав мер ЗИ и их базовые наборы для соответствующего класса защищенности;

– порядок формирования адаптированного перечня мер ЗИ, соответствующих определенному классу защищенности АСУ ТП.

В то же время, в связи с неоднородным составом нормативно-правовой базы по вопросам ИБ, возникают сложности и ошибки в определении актуальности угроз ИБ, формировании требований к обеспечению ИБ АСУ ТП, выборе и реализации мер защиты информации АСУ ТП.

Для решения задач обеспечения ИБ АСУ ТП в данной статье предлагается использование системы поддержки принятия решений (СППР).

Система поддержки принятия решений (Decision Support System, DSS) – это компьютерная автоматизированная система, целью которой является помощь лицам, принимающим решение (ЛПР), в сложных условиях для полного и объективного анализа предметной деятельности [4].

В настоящее время нет общепринятого определения СППР, так как архитектура СППР напрямую зависит от вида решаемых задач, от данных и знаний, используемых системой для принятия решений, а также от квалификации пользователя системы. Однако, в большинстве случаев СППР – это интерактивная автоматизированная информационно-аналитическая система, которая помогает лицу, принимающему решения, использовать данные и модели для решения его профессиональных задач [4].

Взаимодействие ЛПР и СППР изображено на рис. 1.

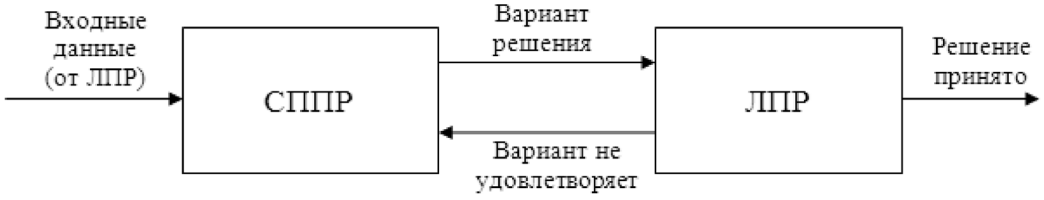


Рис. 1. Процесс принятия решения

ЛПР задает входные данные, СППР, в свою очередь, вырабатывает вариант действий, соответствующий этим данным, ЛПР оценивает вариант и принимает его или отвергает. Таким образом, рекомендации, сформированные СППР, являются всего лишь основой для принятия решения [5].

Современные СППР активно используются для своевременного и быстрого анализа больших объемов информации, на основе которого происходит принятие решений.

Решение задач обеспечения ИБ АСУ ТП имеет свои особенности. Это прежде всего:

– высокая неопределенность исходной информации и трудоемкость ее получения;

– необходимость учета многих требований к средствам ЗИ при оценке и выборе наилучших вариантов.

В результате анализа процессов обеспечения ИБ АСУ ТП можно выделить следующие основные задачи, решаемые с помощью СППР:

– накопление и систематизация информации об ИБ АСУ ТП;

– оценка риска при реализации угрозы ИБ АСУ ТП;

– помощь в выработке рекомендаций для минимизации возможных рисков ИБ АСУ ТП.

Показатели, используемые при разработке СППР по обеспечению ИБ АСУ ТП, могут быть как количественными, так и качественными, поэтому всегда имеет место появляется неопределенность в принятии решений по оценке рисков ИБ АСУ ТП. В данном случае для определения уровня риска ИБ АСУ ТП предлагается использовать технологии интеллектуального анализа данных с помощью модульной (ансамблевой) нейронной сети.

Общая архитектура СППР по обеспечению ИБ АСУ ТП представлена на рис. 2.

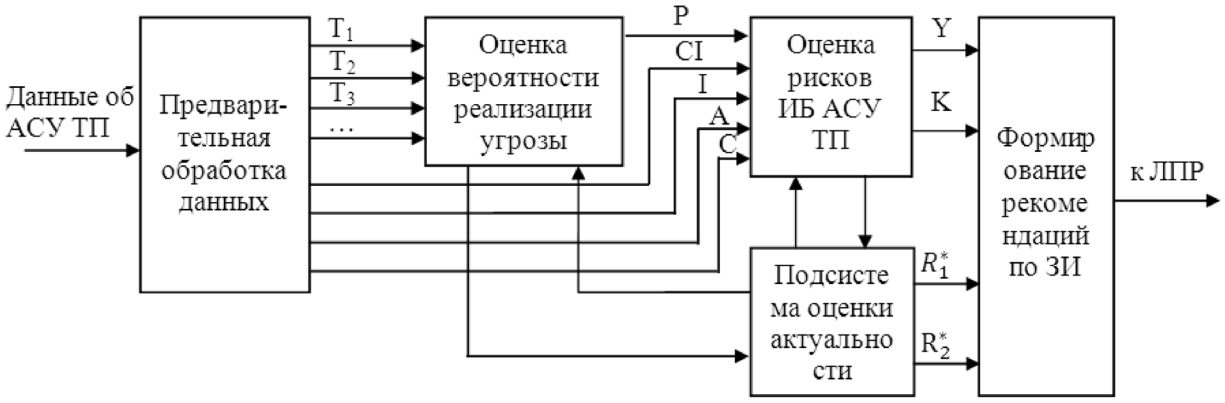


Рис. 2. Архитектура СППР

Модуль предварительной обработки данных об АСУ ТП приводит входные величины модульной нейронной сети к единому масштабу.

На входы нечеткой нейронной сети подаются показатели выявленных уязвимостей ИБ АСУ ТП ($T_1 \div T_4$), показатели ценности информации, содержащейся в системе (CI), а также степени возможного ущерба при нарушении конфиденциальности (C), целостности (I) или доступности (A). Нейронная сеть определяет вероятность реализации угрозы P, после чего на основе совокупности правил производит оценку риска Y и определяет класс защищенности АСУ ТП. Перечисленные показатели могут выражаться как в количественных, так и в качественных величинах.

Выходными данными подсистемы оценки актуальности правил являются: R_1^* - вектор оценок вклада правил в формирование оценки вероятности

реализации угроз, R_2^* - вектор оценок вклада правил в формирование оценки рисков ИБ.

Предполагается, что все правила срабатывают в той или иной мере, т.е. имеют разный уровень активности. Однако превышение некоторого порогового значения говорит о значительном вкладе определенных правил в результат. Выбранные правила могут показать, какие из посылок наиболее подходящие и, следовательно, приводят к полученному результату.

На основании полученной оценки риска Y с учетом вклада решающих правил в определение этой оценки формируются рекомендации по определению состава мер по ЗИ.

Принятие правильного и своевременного решения по обеспечению ИБ АСУ ТП напрямую зависит от полноты и корректности созданной базы правил, содержащей варианты решения той или иной проблемы по ИБ АСУ ТП, составленные на основании анализа предметной области и знаний экспертов. Поэтому создание базы правил при проектировании СППР является первостепенной задачей.

Правила принятия решений могут быть представлены в нечеткой базе правил в системе нечеткого вывода Мамдани и имеют следующий вид:

Π_j: Если X_1 есть A_1^j и X_2 есть A_2^j и … и X_n есть A_n^j, то Y_j есть B^j, где Π_j – j-е правило (j = 1,2,…,m); X_i – входные переменные, (i = 1,2,…,n); Y_j – результат применения j-го правила; A_i^j и B^j – термы (нечеткие подмножества).

Важной задачей исследования является отображение множества задач принятия решений по обеспечению ИБ АСУ ТП на множество правил принятия решений. Влияние уязвимости на реализацию конкретной угрозы находит свое отображение в правилах, имеющих следующую схему:

ЕСЛИ Уязвимость – ВЫСОКАЯ, ТО Вероятность реализации угрозы – ВЫСОКАЯ, и т.д.

Согласно этому принципу, количество правил будет зависеть от количества уязвимостей, дифференцированных по степени опасности и определяющих влияние данной угрозы.

В ходе работы составленная система правил была реализована в пакете математического моделирования Fuzzy Toolbox в среде Matlab.

Из рис. 3 видно, что для определения класса защищенности было введено три входных показателя конфиденциальности (C), целостности (I) и доступности (A), на выходе нейронной сети получаем класс защищенности АСУ ТП.

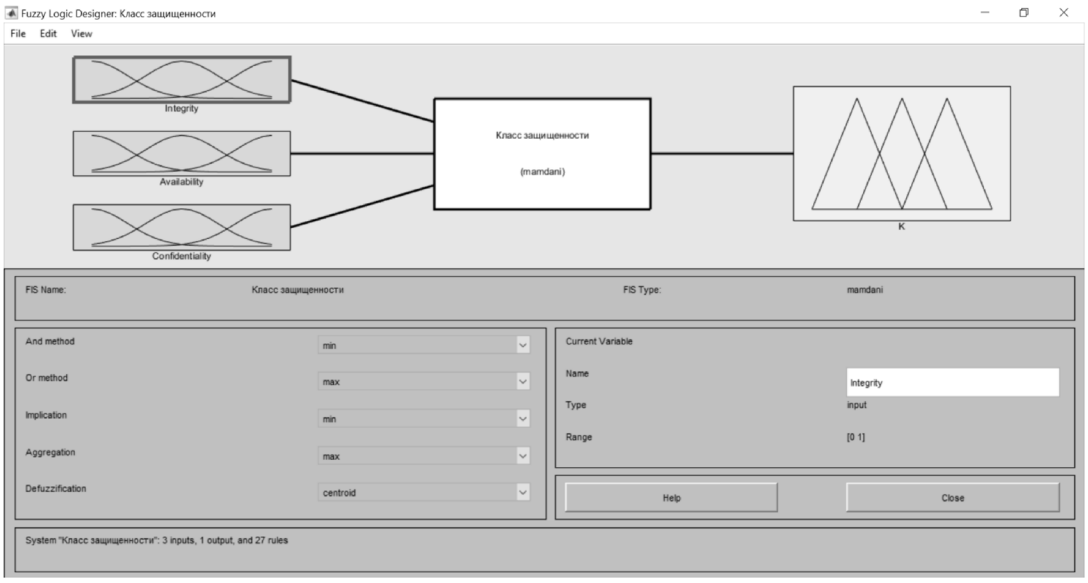


Рис. 3. Решающие правила для оценки класса защищенности АСУ ТП

Входные показатели конфиденциальности (C), целостности (I) и доступности (A) определены тремя лингвистическими термами, оценка которых производится экспертом по шкале от 0 до 1:

L – (0; 0,3) – «низкая степень ущерба»;

M – [0,3; 0,7] – «средняя степень ущерба»;

H – (0,7; 1] – «высокая степень ущерба».

Класс защищенности АСУ ТП (K), зависящий от показателей конфиденциальности, целостности и доступности, также определен тремя лингвистическими термами, значения которых определяются с помощью нейронной сети на основе правил, установленных экспертом:

L – [0; 0,3] – «Первый класс защищенности»;

M – (0,3; 0,7) – «Второй класс защищенности»;

H – [0,7; 1] – «Третий класс защищенности»

Правила, по которым определяется класс защищенности АСУ ТП, приведены в таблице 1. Так как на входе имеем три переменные I, A и C, определенные тремя лингвистическими термами L, M и H, то таблица правил содержит 3^3=27 правил.

Табл.1

Система правил определения класса защищенности

№	Входные показатели			Класс защищенности,
	I	A	C	K
1.	L	L	L	L
2.	L	L	M	M
3.	L	L	H	H
4.	L	M	L	M
5.	L	H	L	H
…	…	…	…	
27.	H	H	H	H

В результате исследований была построена СППР, позволяющая оценить уровень риска ИБ АСУ ТП и выдать рекомендации по его минимизации.

Рассмотрим особенности применения разработанной СППР на следующем примере. Предположим, что АСУ ТП имеет следующие уязвимости:

T_1 – Отсутствие идентификации и аутентификации субъектов и объектов доступа;

T_2 – Наличие незащищенных каналов удаленного доступа;

T_3 – Отсутствие защиты периметра АСУ ТП, сопряжение с корпоративными сетями и Интернет;

T_4 – Отсутствие защиты от атак типа отказ в обслуживании.

Входные показатели НС определим следующим образом: T_1=0,3 (M); T_2=0,1 (L); T_3=0,6 (M); T_4=0,95 (H).

При этом ценность информации, обрабатываемой и циркулирующей в сетях передачи данных АСУ ТП, определим, как CI=0,6.

На выходе нейронной сети получим значения вероятности реализации угрозы (P), оценку риска ИБ АСУ ТП (Y), а также вектор оценок вклада правил в формирование оценки вероятности реализации угрозы (R_1^*) и оценки риска(R_2^*).

Как показали расчеты, значение вероятности реализации угрозы равно 0,727. Это говорит о том, что вероятность реализации угрозы, действующей через данные уязвимости, выше среднего. В свою очередь, на входы нечеткой нейронной сети по определению уровня риска ИБ АСУ ТП подаются значения P=0,727 и CI=0,6. При таких входных показателях на выходе сети получаем значение уровня риска, равное 0,537, то есть уровень риска тоже выше среднего.

Формирование рекомендаций для снижения уровня риска ИБ АСУ ТП происходит следующим образом. Все правила в данном случае срабатывают в различной степени, но превышение порогового значения (в данном случае оно равно 0,95) позволяет отобрать правила с самым весомым вкладом в конечный результат. Эти выбранные правила позволяют показать, почему в результате получена такая оценка вероятности и риска, и выявить слабые места. На основе знаний о слабых местах рекомендуются соответствующие меры защиты информации АСУ ТП. Таким образом, СППР в конечном итоге выдает оценку рисков ИБ АСУ ТП, и рекомендации по обеспечению заданного уровня ИБ с указаниями, на что необходимо обратить первоочередное внимание. Применение предложенной СППР по обеспечению ИБ АСУ ТП позволит повысить уровень ИБ защищенности предприятий, поддерживая систему защиты информации АСУ ТП в актуальном состоянии, и оперативно предоставляя ЛПР рекомендации по повышению ИБ.

Литература

1. Лекция 9: Информационные технологии предприятий – [Электронный ресурс] – Режим доступа: http://www.intuit.ru/studies/courses/1055/271/lecture/6882

2. Федеральный закон от 21 июля 2011 г. № 256-ФЗ «О безопасности объектов топливно-энергетического комплекса»

3. Приказ ФСТЭК России от 14 марта 2014 г. № 31 Об утверждении Требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды.

4. Попов А.Л. Системы поддержки принятия решений: Учебно-метод. пособие / Попов А.Л. – Екатеринбург: Урал. гос. ун-т, 2008. – 80 с.

5. Логунова Е.А. Математические модели систем поддержки принятия решений // Физико-математические науки и информационные технологии: проблемы и тенденции развития: сб. ст. по матер. IV междунар. науч.-практ. конф. – Новосибирск: СибАК, 2012.

ФГБОУ ВО «Уфимский государственный авиационный технический университет»
Ufa State Aviation Technical University

SYSTEM OF DECISION MAKING SUPPORT ON INFORMATION SECURITY MAINTENANCE OF AUTOMATED TECHNOLIGAL PROCESSES SYSTEMS

Vasilyev V.I., Gvozdev V.E., Guzairov M.B., Kirillova A.D.

The article deals with the problem of ensuring information security of automated technological processes control system (ATPCS). The overview of basic regulations to ensure the security of ATPCS is presented. To facilitate the solution of the problems of ATPCS information security, it is offered to use the intelligent decision support system. The example of construction and implementation of decision rules is considered

Key words: decision support system, information security, automated technological processes control system, risk assessment, probability of threat realization, fuzzy logic, neural network

ПРАВИЛА

оформления и представления рукописей для публикации в журнале «Информация и безопасность»

В целях улучшения качества оформления настоящего издания редколлегия просит авторов направляемых материалов руководствоваться следующими правилами оформления:

1. Рукопись общим объемом 8 **полных** страниц для обзорной статьи (тезисов пленарного доклада), 4 **полных** страницы для статьи (тезисов доклада) и 2 **полных** страницы краткого сообщения (тезисов стендового доклада) представляют в отпечатанном виде на одной стороне листа формата А4 шрифтом Times New Roman Cyr через 1 интервал и на дискете 3,5 (в редакторе Word for Windows). Форматирование статьи для издания в Журнале производить по форме, аналогичной принятой Международной академической издательской компанией «Наука» (см., например, «Журнал неорганической химии» РАН и др.).
2. Страницы рукописи должны иметь следующие размеры полей: верхнее - 2,35 см, нижнее - 2,35, левое- 2,5 см, правое- 1,5 см. На первой странице текста располагают УДК (в левом углу листа от поля, размер шрифта 12), название статьи (заглавными буквами, размер шрифта 12), инициалы и фамилию автора (авторов) (размер шрифта 12), аннотацию и ключевые слова (не более 8 строк, размер шрифта 10, отступы слева и справа – 1,25 см, абзацный отступ – 0,8 см), сведения об авторах (сноской, размер шрифта 10). Далее следуют текст рукописи (размер шрифта 12) и цитируемая литература (размер шрифта 12). Текст рукописи и цитируемую литературу представляют на листе в две колонки шириной по 8,25 см каждая (межколоночное расстояние 0,5 см). **Обе колонки должны быть заполнены равномерно и полностью!**

Пример оформления текста:

Если приходит следующий (J_{lim}+1)-й пакет с запросом на соединение, то этот пакет отбрасывается.

Если атака обнаруживается до этого момента времени с некоторой вероятностью P_{det} и ее развитие блокируется, то вероятность реализации атаки может быть рассчитана по формуле:

$$P_u(t) = 1 - \frac{\lambda_{syn} \cdot \bar{\tau}_u \cdot e^{-\frac{(t-t_0)\cdot(1-P_{\det})}{\bar{\tau}_u}}}{\lambda_{syn} \cdot \bar{\tau}_u - (1-P_{\det})} + \frac{(1-P_{\det}) \cdot e^{-\lambda_{syn} \cdot (t-t_0) \cdot}}{1 - \bar{\tau}_u \cdot \lambda_{syn} \cdot (1-P_{\det})}, \quad (3)$$

где P_{det} – вероятность обнаружения атаки;

t_0 – время ожидания подтверждения сеанса связи.

Рассмотрим модель динамики реализации атаки – шторм ICMP – «эхо-ответов» (Smurf) [3]. Суть атаки заключается в посылке "эхо-запроса" по протоколу ICMP по широковещательному адресу с указанием в качестве адреса отправителя IP-адреса компьютера – цели атаки, ответить на которые может множество компьютеров.

После литературы указывается название организации, откуда поступил материал на русском и английском языках. Затем приводится резюме на английском языке (название статьи (заглавными буквами, размер шрифта 12), инициалы и фамилию автора (авторов) (размер шрифта 12), аннотацию (не более 8 строк, размер шрифта 10)). Название организации и резюме располагаются посередине страницы. **Номера страниц не проставляются!** На обороте последней страницы должны быть подписи всех авторов.

3. К рукописи **необходимо** приложить экспертное заключение о возможности ее публикации в открытой печати.
4. На отдельном листе следует приложить служебные и домашние адреса (с почтовым индексом), телефоны авторов статьи.
5. Таблицы располагают по тексту. Каждый элемент таблицы должен представлять собой отдельную ячейку. **Не допускается размещать колонку или строку с данными в одной ячейке!** Если в рукописи одна таблица, то слово «Таблица» в названии не пишут. Если в статье несколько таблиц, то перед названием таблицы справа пишут «Таблица 1 (2, 3 и т.д.)». Ссылку на таблицу оформляют следующим образом: «табл. 1 (2, 3 и т.д.)».
6. Оформление рисунков, не внедренных в документ Word, осуществляется в формате BMP. Подрисуночные подписи не входят в состав рисунков, а располагаются отдельным текстом с размером шрифта 10 под рисунками. Буквы и цифры на рисунке должны быть разборчивы. Тоновые фотографии представляют в двух экземплярах на белой матовой фотобумаге, пояснительные надписи на одной из этих фотографий должны отсутствовать. Если в рукописи несколько рисунков, то перед названием справа пишут «Рис. 1 (2, 3 и т.д.)». Ссылка на рисунок оформляется следующим образом «рис. 1 (2, 3 и т.д.)». Если в статье один рисунок, то слово «Рис.» в названии не пишут.
7. Абзацный отступ, равный 0,8 см, должен начинаться после ввода (автоматически**). Не допускается формирование абзацного отступа при помощи пробелов и табуляции!**
8. Используемые в работе термины, единицы измерения и условные обозначения должны быть общепринятыми. Все употребляемые авторами обозначения (за исключением общеизвестных констант) и аббревиатуры должны быть определены при их первом упоминании в тексте.
9. Формулы нумеруют в круглых скобках (2), литературные ссылки - в прямых [2], подстрочные примечания - арабскими цифрами.
10. Библиографические ссылки даются по следующим образцам:
    - Для книг – фамилия, инициалы автора; название книги; инициалы, фамилия автора; место издания; наименование издательства; год издания; номер тома; объем. Пример:
      Шульце Г. Металлофизика / Г. Шульце. – М.: Мир, 1971. – 503 с.
      Если авторов более одного - фамилия, инициалы первого автора; название книги; инициалы, фамилия всех авторов (включая первого); место издания; наименование издательства, год издания; номер тома; объем. Пример:
      Ландау Л.Д. Квантовая механика / Л.Д. Ландау, Е.М. Лифшиц. - М.: Физматгиз, 1963. - 25 с.
    - Для статей в сборнике (журнале) – фамилия, инициалы автора; название статьи [Текст]; инициалы, фамилия автора; название сборника, серии; год издания; том издания; номер издания; объем. Пример:
      Кузнецов, В.Ю. Немонотонный потенциал в обогащенных слоях [Текст] / В.Ю. Кузнецов // Изв. вузов. Сер. Химия (или Сер. физ.). - 1989. - Т. 43, № 5. - С. 106-111.
      Если авторов двое или трое – фамилия, инициалы первого автора; название статьи [Текст]; инициалы, фамилия каждого автора (включая первого); название сборника, серии; год издания; том издания; номер издания; объем. Пример:
      Моисеев, С.И. Динамическое торможение дислокаций в кристалле с межфазной границей [Текст] / С.И. Моисеев, В.Н. Нечаев // Вестник ВГТУ. Сер. Материаловедение. – 1997. – Вып.1.2. – С. 14-18.
      Если авторов более трех: название статьи ; инициалы, фамилия каждого автора; название сборника; место издания, название издательства; год издания; номер тома; объем. Пример:
      Системное проектирование образовательных программ на базе высоких технологий / В.К. Бойко, Ю.М. Белов, В.Н. Киселев, И.Д. Усов // Высокие технологии в технике, экономике и образовании: Сб.науч.тр. Воронеж: Изд-во ВГТУ, 2000. Ч. 3. С. 63-73.

- Для авторефератов и диссертаций – фамилия, инициалы автора; название работы; название вида работы; название ученой степени; место написания; год написания; объем. Пример:
  Недорезов, С.С. Особенности зарождения и структура пленок некоторых металлов при конденсации из ионного потока // Автореф. дис. … д-ра физ.-мат. наук/ ФТИНТ. - Харьков, 1985. - 16 с.
- Для авторских свидетельств и патентов – вид документа; его номер; название страны; индекс МКИ; название работы; инициалы и фамилия(и) автора(ов); регистрационный номер заявки; дата подачи заявки; дата публикации; издание, в котором опубликован документ; объем. Пример:
  А.с. 1381379 СССР, МКИ 0125 /18. Вычисление пьезомодулей квазиизотропных текстур/ В.Н. Чернышов и др. - Заявлено 25.06.86. Опубл. 30.03.89. Бюл. № 12. - 2 с.

Общие требования

Для публикации материалов в журнале авторам необходимо представить в редакцию:

- *два экземпляра рукописи, включающие в себя страницы с аннотацией, основной текст статьи, пронумерованные иллюстрации и таблицы, подписи к рисункам;*
- *электронную версию статьи;*
- *рецензию внешнего ведущего специалиста в области излагаемого материала;*
- *экспертное заключение, заверенное руководителем организации или его заместителем и печатью;*
- *ключевые слова статьи;*
- *сведения об авторах, включающие фамилию, имя, отчество, место работы и должность, ученую степень и звание, контактный телефон, почтовый (с индексом) и электронный адрес для переписки (если он есть).*

Редакционная коллегия оставляет за собой право осуществлять дополнительное рецензирование и техническое редактирование представленных работ.

Статья будет принята к рассмотрению только при условии выполнения всех требований!

Научное издание

ИНФОРМАЦИЯ И БЕЗОПАСНОСТЬ

Том 20. Выпуск 4. 2017.

Главный редактор А.Г. Остапенко

Журнал зарегистрирован в «Федеральной службе по надзору в сфере массовых коммуникаций, связи и охраны культурного наследия»
Рег. номер ПИ №ФС77-28802 от 29 июня 2007 г.

Компьютерная верстка А.А. Акинина

Подписано в печать 25.10.2017. Формат 60x84/8. Усл. печ. л. 37,5.
Тираж 500 экз. Заказ 954

Отпечатано с готового оригинала-макета
в типографии Издательско-полиграфического центра
Воронежского государственного университета
394000, г. Воронеж, ул. Пушкинская, 3.