На правах рукописи

АБРАМОВ Константин Германович

# Модели угрозы распространения запрещенной информации в информационно-телекоммуникационных сетях

Специальность 05.12.13 – Системы, сети и устройства телекоммуникаций

**АВТОРЕФЕРАТ**

**диссертации на соискание ученой степени**

**кандидата технических наук**

Владимир 2014

Работа выполнена на кафедре «Информатика и защита и информации» в Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (ВлГУ).

**Научный руководитель:** доктор технических наук, профессор, МОНАХОВ Михаил Юрьевич

**Официальные оппоненты:** доктор технических наук, профессор, ведущий научный сотрудник отдела общесистемных исследований НИЦ МОУ «Институт инженерной физики» Данилюк Сергей Григорьевич,

кандидат технических наук, доцент, начальник факультета подготовки научно-педагогических кадров ФКОУ ВПО «Владимирский юридический институт федеральной службы исполнения наказаний» Курысев Константин Николаевич

**Ведущая организация:** ОАО «Владимирское конструкторское бюро Радиосвязи»

Защита диссертации состоится «23» сентября 2014 г. в «14» часов в ауд. 301-3 на заседании диссертационного совета Д212.025.04 при Владимирском государственном университете имени Александра Григорьевича и Николай Григорьевича Столетовых по адресу: 600000, г. Владимир, ул. Горького, 87, ВлГУ, ФРЭМТ.

С диссертацией можно ознакомиться в научной библиотеке университета.

Автореферат разослан «2» июля 2014г.

Отзывы на автореферат, заверенные печатью, просим направлять по адресу: 600000, г. Владимир, ул. Горького, д. 87, ВлГУ, ФРЭМТ.

Ученый секретарь диссертационного совета, доктор технических наук, профессор  Самойлов Александр Георгиевич

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность работы.** Информационно-телекоммуникационные сети (ИТКС) обеспечивают практически полный спектр возможностей для обмена информацией между пользователями - сетевыми абонентами. Современной проблемой таких систем является их низкий уровень информационной безопасности. Для обеспечения защиты информации в телекоммуникационных сетях, включая Интернет, разработано множество методов и средств, предложенных в трудах В.А. Герасименко, С.П. Расторгуева, П.Д. Зегжды, В.И. Завгороднего, А.А. Малюка, А.А. Грушо, В.В. Домарева, Р. Брэтта, К. Касперски, С. Норкатта, В. Столингса. Тем не менее, эффективной защиты абонентов от угроз распространения запрещенной информации, в частности в условиях широкого использования индивидуально-ориентированных сервисов и связанных с ними протоколов и технологий (SOAP, CORBA, REST и др.), не существует. Среди множества функций защиты принципиальной в отношении данных систем является функция предупреждения проявления запрещенной информации. Она реализуется за счет механизмов прогнозирования угрозы распространения и рассылки сообщений с предупреждениями о последствиях действий с запрещенным контентом. Использование других функций (предупреждения, обнаружения, локализации и ликвидации угрозы) предполагает наличие полного контроля над системой, что в настоящих условиях невозможно.

Одним из подходов к прогнозированию угрозы распространения запрещенной информации (УгЗИ) является моделирование, например, с использованием моделей влияния, моделей просачивания и заражения (Д.А. Губанов, Д.А. Новиков и А.Г. Чхартишвили, J. Leveille, D. Watts и S. Strogatz, R. Albert и A. Barabasi, J. Leskovec, M. Gjoka, S.N. Dorogovtsev, M.E.J. Newman и R. M. Ziff, J.O. Kephart и S.R. White и др.). Данные модели, как правило, не учитывают топологические особенности сети (распределение степеней связности, кластерный коэффициент, средняя длина пути). Взаимодействие между абонентами в рамках этих математических моделей описывается преимущественно гомогенным графом, что при моделировании крупномасштабных сетей (более 10 млн. узлов) может дать погрешность прогнозирования УгЗИ более 30%. Кроме того, данные подходы носят в основном теоретический характер, практика их использования не выходит за рамки экспериментов. Таким образом, исследования, направленные на создание моделей и алгоритмов УгЗИ, актуальны и имеют теоретическое и практическое значение в решении проблемы обеспечения информационной безопасности в системах и сетях телекоммуникаций.

**Объектом исследования** являются информационно-телекоммуникационные сети, находящиеся под воздействием угрозы распространения запрещенной информации.

**Предметом исследования** являются модели угрозы распространения запрещенной информации в информационно-телекоммуникационных сетях.

**Цель работы** заключается в повышении точности прогнозирования угрозы распространения запрещенной информации в информационно-телекоммуникационных сетях.

Для достижения цели работы необходимо решить следующие **задачи:**

1. Провести информационный обзор и эксперименты для выявления существенных характеристик объекта и внешних факторов, влияющие на процесс реализации УгЗИ. Выполнить анализ основных подходов к моделированию УгЗИ.
2. Разработать имитационную модель УгЗИ в ИТКС.
3. Синтезировать и показать адекватность аналитической модели УгЗИ в ИТКС.
4. Разработать методику формирования топологии ИТКС.
5. Смоделировать процесс реализации УгЗИ на топологии реальной крупномасштабной ИТКС с использованием разработанного программного обеспечения для супер-ЭВМ «Скиф-Мономах». Провести экспериментальное исследование по полученным результатам.

**Научная новизна работы**

1. Разработана имитационная модель реализации УгЗИ в ИТКС, учитывающая среднюю степень связности узлов, среднюю длину пути сети, коэффициент кластеризации сети, а также особенности информационного взаимодействия абонентов как человеко-машинных систем и позволяющая повысить точность представления процессов обеспечения информационной безопасности в крупномасштабных ИТКС.
2. Предложена аналитическая модель реализации УгЗИ, отличающаяся от классической эпидемиологической модели Кермака-Маккендрика учетом характеристик уязвимости ИТКС и позволяющая повысить точность оперативного прогноза, особенно в условиях неполноты исходных данных о топологии сети.
3. Разработана методика формирования топологии крупномасштабной ИТКС, включающая:

- алгоритм формирования графа доступной части сети, позволяющий произвести сбор данных о топологии с любого узла-абонента;
- алгоритм формирования полного графа сети, позволяющий в условиях неполноты исходных данных спрогнозировать топологию недостающей части сети.

Применение методики позволяет повысить точность представления модели топологии ИТКС.

**Практическая ценность работы**

1. Разработано программное обеспечение (свидетельство о государственной регистрации программы для ЭВМ №2013660757), автоматизирующее процесс поиска узлов – потенциальных

распространителей запрещенной информации в крупномасштабных информационно-телекоммуникационных сетях и позволяющее сократить время поиска таких узлов в 1,3 раза.

2. Разработана методика и программное обеспечение (свидетельство о государственной регистрации программы для ЭВМ № 2012610825) формирования топологии крупномасштабной информационно-телекоммуникационной сети, которые позволяют повысить защищенность организации за счет сокращения времени расследования инцидентов в рамках ликвидации последствий нарушения конфиденциальности.

**Достоверность и обоснованность результатов** подтверждается строгостью математических выкладок, статистическими и численными экспериментами, согласованностью результатов аналитического и имитационного моделирования.

**Реализация и внедрение результатов работы**

Результаты диссертационной работы внедрены и нашли практическое использование в организациях: ФГБОУ ВПО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (ВлГУ), федеральная служба по надзору в сфере связи, информационных технологий и массовых коммуникаций (РОСКОМНАДЗОР) по Владимирской области, ОАО «Владимирское производственное объединение «Точмаш». Внедрение результатов подтверждается соответствующими актами.

Исследования и практическая реализация результатов диссертационной работы проводилась в ВлГУ на кафедре «Информатика и защита информации» и использовались при выполнении х/д НИР №4013/10, г/б НИР №396/03, г/б НИР №848/13, г/б НИР №925/14.

**Апробация работы, публикации**

Результаты диссертационной работы апробированы на международной научно-технической конференции «Информационные системы и технологии ИСТ-2011» (г. Н.Новгород, 22 апреля 2011 года), всероссийской научно-технической конференции «Проблемы эффективности и безопасности функционирования сложных технических и информационных систем» (г. Серпухов, 2011 год), международной научно-технической конференции «Перспективные технологии в средствах передачи информации» (г. Владимир, 2011 год), международной научно-технической конференции «Проблемы информатики и моделирования» (Харьков-Ялта, 2011 год), российской научно-технической конференции «Новые информационные технологии в системах связи и управления» (Калуга, 1-2 июня 2011г.), всероссийской научно-практической конференции по имитационному моделированию и его применению в науке и промышленности «Имитационное моделирование. Теория и практика» ИММОД-2011 (г. Санкт-Петербург, 2011 год), научно-практической конференции «Математика и математическое моделирование» (г. Саранск, 13–14 октября 2011 года), международной

научно-практической конференции «Современные проблемы и пути их решения в науке, транспорте, производстве и образовании '2011» (Одесса: Черноморье, 2011)

По теме диссертации опубликовано более 15 статей, в том числе 3 статьи во включенных в перечень ВАК журналах.

**Структура и объем работы.** Основная часть диссертации объемом 117 страниц машинописного текста включает введение, четыре главы, заключение, список использованных источников из 139 наименований и содержит 58 рисунок и 8 таблиц. Объем приложений - 11 страниц.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обосновывается актуальность направления исследований, проводимых в данной работе, ставятся задачи исследования.

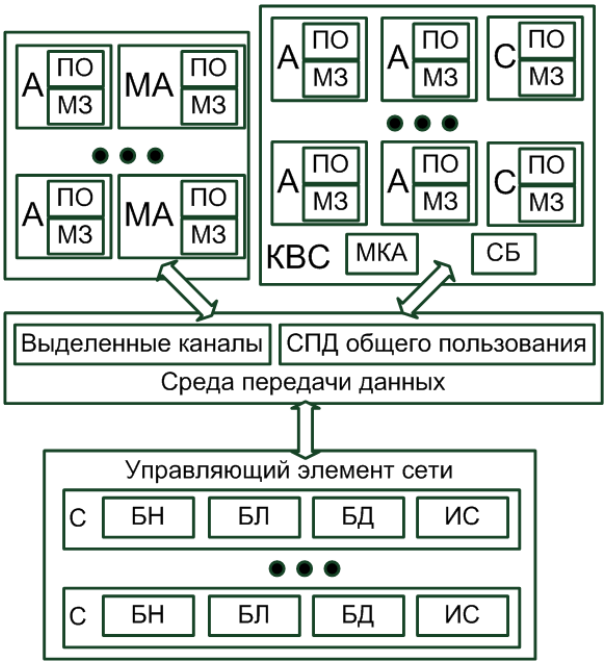


Рисунок 1 – Структурная схема ИТКС

**В первой** главе диссертации дается обзор современного состояния социальных сетей Интернета, как распределенных информационно-телекоммуникационных систем.

Обобщенная структурная схема ИТКС приведена на рисунке 1. Ее состав в общем случае образуют следующие элементы:

- абоненты (А) - человеко-машинные системы, состоящие из устройства, через которое осуществляется доступ к сети, и непосредственно пользователя. Абоненты могут быть объединены в корпоративную вычислительную сеть (КВС), включают в себя модули (информационной) защиты (МЗ) и программное обеспечение (браузер) для взаимодействия с управляющим элементом;
- мобильные абоненты (МА). Пользователи, использующие мобильные устройства (смартфоны, планшеты и тд.), для доступа к сети. Также используют программное обеспечение (специальное приложение) и МЗ;
- серверы (С). В КВС находятся информационные серверы различного функционального назначения, которые участвуют в информационном взаимодействии (например, прокси-сервера).
- КВС включает в себя кроме абонентов и серверов, также средства маршрутизации, коммутации и администрирования (МКА), систему безопасности (СБ), включающую механизмы защиты КВС;
- средства телекоммуникации, обеспечивающие взаимодействие между абонентами;

- управляющий элемент технически представляет собой совокупность коммутирующего и серверного оборудования, реализующего основные функции системы. Включает в себя серверы, содержащие в общем случае: балансировщики нагрузки (БН), элемент бизнес-логики (БЛ), базы данных (БД), инфраструктурные системы (ИС) (системы статистики, конфигурации, мониторинга и тд.).

Особенности ИТКС:

1) Абонентами являются человеко-машинные системы. Способ взаимодействия абонентов – смысловые сообщения. Решение о взаимодействии принимается пользователями.
2) Постоянно изменяющееся число абонентов и связей между ними.
3) Сложность процедуры идентификации абонентом и отнесения сообщения к запрещенной информации.
4) Сложность реализации защиты. Основной способ ее реализации – предупреждение абонента об ответственности за распространение запрещенной информации.
5) Особенности программно-логической организации ИТКС (например, социальные сети «ВКонтакте» и «Facebook»), которые приводят к неотвратимости получения сообщения абонентом при наличии связи между ним и абонентом-злоумышленником.
6) Крупномасштабность – ИТКС, как правило, содержат миллионы абонентов.
7) Основной проблемой ИТКС, кроме проблем, связанных с использованием глобальной сети Интернет, является проблема запрещенной информации.

Функционирование ИТКС, находящейся под воздействием УгЗИ, осуществляется по следующему алгоритму.

Шаг 1. Распространение запрещенной информации (ЗИ) (далее процесс «атаки») инициирует какой-либо абонент-злоумышленник, распространяя сообщения с ЗИ (реализует угрозу) по его списку контактов. Атаку может начинать один злоумышленник или группа.

Шаг 2. Абоненты-получатели, приняв сообщение с ЗИ, читают его и включаются в процесс атаки, распространяя ее дальше по своему списку контактов, либо игнорируют или вообще удаляют сообщение, т.е. в атаке не участвуют. Процесс атаки обычно идет лавинообразно. Атакующие абоненты не заканчивают атаку, единожды передав сообщение с запрещенной информацией. Окно атаки, как правило, продолжается в течение довольно значительного промежутка времени и зависит от типа подачи ЗИ в сообщении, заинтересованности абонента и тд.

Шаг 3. Абоненты-злоумышленники могут перестать распространять и, соответственно, воспринимать ЗИ (далее процесс «защиты»), вследствие воздействия механизмов защиты (например, предупреждение о ней), поэтому сообщения с ЗИ от атакующих абонентов будут постоянно отвергаться.

Шаг 4. Процесс продолжается пока в сети есть абоненты-злоумышленники, либо есть потенциально уязвимые узлы, если отсутствует процесс защиты.

Один из ключевых подходов при решении проблемы ЗИ - создание моделей и алгоритмов УгЗИ. Проведенный анализ показывает, что существующие решения малоэффективны. При моделировании УгЗИ не учитывается топология ИТКС (модель сети – полносвязный граф). А, если топология учитывается, то, как правило, используется простейшая SIS модель, а структура сети отражается Scale-Free или Small world сетью. При моделировании УгЗИ важно иметь топологию, отражающую структуру связей реальной сети, а также использовать корректную модель информационного взаимодействия узлов.

Концептуальная математическая модель информационного взаимодействия абонентов представляется графом, вершинами которого являются абоненты, а ребрами – связи между ними. Отметим свойства графа, принципиальные для настоящего исследования: большая размерность, гетерогенность, динамика связей и узлов, наличие групп узлов, имеющих большое количество связей внутри кластера и небольшое – между ними.

Из анализа предметной области можно сделать вывод о том, что на процесс реализации УгЗИ в ИТКС существенное влияние оказывают сетевые структурные характеристики (топология). Получение структуры социальной сети связано с выборкой узлов из нее, что само по себе уже является нетривиальной задачей, так как это выборка должна отражать свойства всей сети в целом, то есть быть репрезентативной.

Таким образом, для повышения точности прогнозирования УгЗИ в ИТКС нужно:

1. разработать имитационную модель УгЗИ, учитывающую топологические характеристики и особенности информационного взаимодействия абонентов как человеко-машинных систем;
2. разработать аналитическую модель реализации УгЗИ, учитывающую характеристики уязвимости ИТКС и позволяющую повысить точность оперативного прогноза в условиях неполноты исходных данных о топологии сети;
3. экспериментально подтвердить адекватность аналитической модели, смоделировав процесс реализации УгЗИ на топологии крупномасштабной сети (более 10 млн. узлов), для чего разработать специализированное программное обеспечение.

**Вторая глава** посвящена разработке и исследованию моделей УгЗИ в ИТКС.

*Имитационная модель УгЗИ*

Входные данные: $N$ - количество узлов, равное числу абонентов сети, $k$ - средняя степень связности узлов, $\alpha$ - параметр, отражающий среднюю длину пути и уровень сетевой кластеризации, $\beta$ – параметр, отражающий силу угрозы, вероятность осуществления атаки, $\gamma$ – параметр отражающий

степень противодействия угрозе, вероятность защиты абонента (в модели считается, что $\beta$ и $\gamma$ одинаковы для каждого абонента), $I_0$ (абоненты-злоумышленники - изначальные источники угрозы) $R_0$ (абоненты, изначально невосприимчивые к атакующим воздействиям). Выходные данные: $I(t)$, $R(t)$, $S(t)$ – численные массивы данных, описывающие динамический процесс реализации УгЗИ (количестве атакующих, защищенных и потенциально уязвимых узлов в каждую условную единицу времени соответственно).

Шаг 1. Создание топологии ИТКС – графа $G_{sw} = <V, E>$, где $G_{sw}$ – граф small-world сети (на основе модели Watts-Strogatz), $V = \{v_i\}$ – множество вершин, $E = \{e_{ij}\}$ – множество ребер, $i$=1..$N$, $j$=1..$N$. Данный шаг осуществляется с использованием свободно распространяемой программы Pajek, адаптированной под данную задачу, за счет задаваемых топологических параметров $N$, $k$, $\alpha$.

Шаг 2. Сформировать множество $V=\{V^I, V^S, V^R\}$, где $V^I = \{v_i^I\}$ – множество атакующих узлов ($|V^I|=I_0$), $V^R = \{v_i^R\}$ – множество защищенных узлов ($|V^R|=R_0$), $V^S = \{v_i^S\}$ – множество потенциально уязвимых узлов ($|V^S| = N - I_0 - R_0$).

Шаг 3. $\forall v_i^I$ если $\exists e_{ij}$ и $v_j \in V^S$ $j = 1..N$, то с вероятностью $\beta$ выполнить: $V^S \backslash v_j$ и $V^I \cup v_j$; с вероятностью $\gamma$ выполнить: $V^I \backslash v_i$, $V^R \cup v_i$.

Шаг 4. Если $V^I=\varnothing$ или $\gamma$=0 и $V^S=\varnothing$, то конец алгоритма, иначе перейти к шагу 3.

Данный алгоритм был реализован в разработанном ПО ModelGraph. На рисунках 2 и 3 приведены результаты имитационного моделирования УгЗИ на сети с параметрами ($N$=10000, $\beta$=0,2, $\gamma$=0,1, $I_0$=1, $R_0$=2000).

Анализируя процесс информационного взаимодействия абонентов при распространении запрещенной информации в ИТКС, можно сделать следующие выводы. Имеем дело с тремя типами абонентов: атакующие абоненты, которые распространяют запрещенную информацию, защищенные абоненты, характеризующиеся тем, что не принимают участие в распространении запрещенной информации и никогда не будут этим заниматься, и потенциально уязвимые абоненты, которые могут быть подвержены негативному влиянию со стороны атакующих узлов и могут начать распространять запрещенную информацию. При этом мы наблюдаем два противоборствующих подпроцесса атаки и защиты абонентов сети. Для моделирования таких явлений часто применяют эпидемиологические модели, в частности нашему описанию точно соответствует SIR-модель Кермака-Маккендрика. Характер графиков, полученных в результате имитационного моделирования (рисунки 2 и 3), схож с результатами, которая дает данная модель. По приведенным выше причинам данная модель была взята за основу в настоящем исследовании. При использовании системы дифференциальных уравнений SIR-модели для анализа УгЗИ в ИТКС получили результаты в виде графиков, которые

хотя и правильно описывают характер процесса, но не дают нужной точности прогноза.

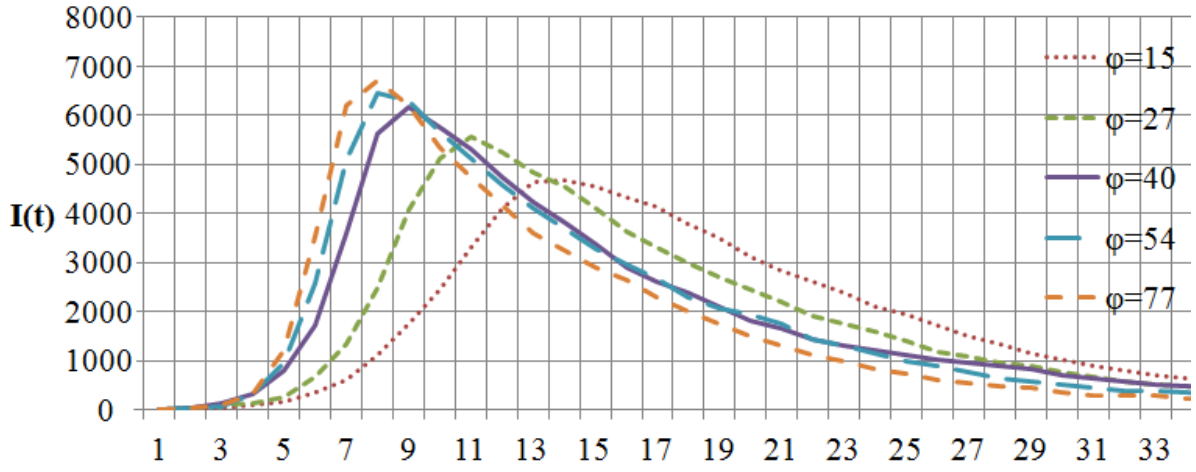


Рисунок 2 – Результаты имитационного моделирования УгЗИ

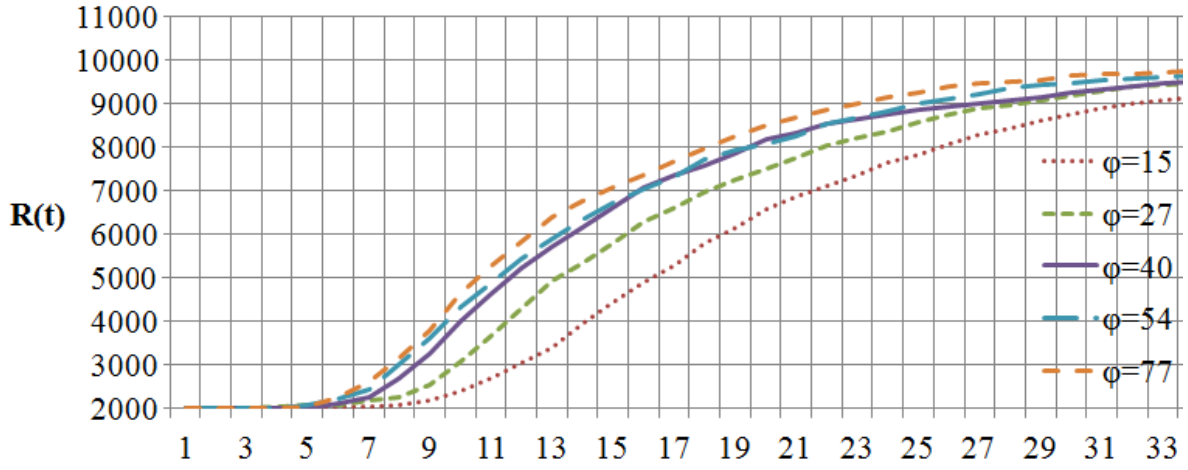


Рисунок 3 – Результаты имитационного моделирования УгЗИ

Была выдвинута гипотеза о том, что система не дает нужной точности в связи с тем, что в модели, которую она описывает, не учитываются топологические особенности сети. В связи с этой гипотезой была поставлена задача адаптирования системы путем интегрирования в нее параметра топологической уязвимости сети $\varphi$. В итоге получили аналитическую модель, описываемую системой дифференциальных уравнений:

$$\begin{cases} \dfrac{dI}{dt} = 2 \cdot \ln \varphi \cdot \beta \cdot \dfrac{S(t) \cdot I(t)}{N} - \gamma \cdot I(t) \\ \dfrac{dR}{dt} = \gamma \cdot I(t) \\ \dfrac{dS}{dt} = -2 \cdot \ln \varphi \cdot \beta \cdot \dfrac{S(t) \cdot I(t)}{N} \end{cases}, \qquad (1)$$

где $\varphi$ - коэффициент топологической уязвимости сети - показатель, который вычисляется по формуле:

$$\varphi = \frac{k \cdot (C+1)}{L},$$

$k$ – средняя степень связности узлов сети, $C$ – коэффициент кластеризации сети, $L$ – средняя длина пути сети.

Релевантность результатов аппроксимации подтверждена серией экспериментов на топологиях реальных сетей («ВКонтакте», «Facebook») с использованием имитационного моделирования. При этом погрешность для процесса защиты составила не более 10%, для процесса атаки - не более 15%.

**В третьей главе** разрабатывается методика формирования топологии крупномасштабной ИТКС.

Для моделирования УгЗИ необходимо иметь топологию реального объекта. Прямое получение этой информации затруднено в связи со следующим противоречием. Для повышения точности результатов моделирования необходимо иметь топологию всей сети. Получить такую информацию без прав администратора не представляется возможным. При сборе данных с правами абонента ИТКС имеем дело с двумя типами узлов: открытыми и закрытыми. Если в ходе сбора данных мы получаем идентификаторы (id) узла и смежных с ним узлов, то такой узел называем открытым. Если же получаем только id узла (абонент с помощью настроек скрыл информацию о своих контактах), то такой узел называем закрытым. Также в сети могут существовать узлы, которые соединены только с закрытыми узлами. В таком случае невозможно получить даже идентификатор узла. Таких узлов в сети незначительная часть. Эмпирически показано, что закрытых узлов на порядок больше, чем открытых, поэтому при сборе данных теряется значительная часть информации.

*Методика формирования топологии крупномасштабной ИТКС*

Методика состоит из двух последовательно применяемых алгоритмов: алгоритма формирования графа доступной части сети и алгоритма формирования полного графа сети с учетом добавления недоступной части на основе вычисленных прогнозируемых топологических характеристик (распределение степеней связности, средняя длина пути).

Граф доступной части сети – граф, содержащий открытые и закрытые узлы и связи между ними Полный граф сети – граф, содержащий открытые узлы и закрытые узлы, перешедшие в состояние открытых, и связи между ними. $V$ – множество вершин, включающее два подмножества ($W=\{w_i\}$ – подмножество открытых вершин, $U=\{u_i\}$ – подмножество закрытых вершин); $E$ – множество связей между узлами ($e_{ij} = e_{ji}$ – связь между $i$-м и $j$-м узлами); $A$ – массив, содержащий id пройденных узлов ($a_i$ – элементы массива); $k$ – счетчик узлов; $Z=\{z_i\}$ – множество соседних узлов $k$-го узла; $flag$ – флаг, определяющий статус узла ($flag$=1 – открытый, $flag$=0 – закрытый); $n$ – текущее значение длины массива $A$; $i$ – счетчик соседних узлов; $X$ – временное множество.

*Алгоритм формирования графа доступной части сети*

Шаг 1. Начальная установка. Обнулить множества вершин $V=\emptyset$ и связей $E=\emptyset$. Инициализировать счетчик узлов ($k$=1). Добавить вершину $v_1$ в множество $V$ ($V = V \cup v_1$), сделать ее текущей. Выполнить $a_k=id(v_k)$.

Шаг 2. Выполнить функцию *Get*($a_k$, *Z*, |*Z*|, *flag*) получения множества *Z* соседних узлов *k*-го узла, где $a_k$ – идентификатор *k*-го узла, *Z* - возвращаемое множество, |*Z*| - его мощность, *flag* – флаг, определяющий статус узла (открытый/закрытый). Если *flag*=1 (узел открытый), перейти к шагу 3, иначе (*flag*=0) – к шагу 5.

Шаг 3. Для $\forall z_i \in Z(i=1,..,|Z|)$ если $z_i = v_k$, то $Z = Z \setminus z_i$ и если $z_i \in U$, то $E = E \cup e_{k,z(i)}$.

Шаг 4. Определить длину массива *A* (*n* = *length*(*A*)). Для $\forall z_{n+i} \in Z(i=1,..,|Z|)$ добавить ребро с *k*-й вершиной $E = E \cup e_{k,n+i}$. Выполнить функцию *Get*($z_{n+i}$, *X*, |*X*|, *flag*). Если *flag*=1, то $V = V \cup w_{n+i}$, иначе (*flag*=0) $V = V \cup u_{n+i}$. Выполнить $a_{n+i}=id(z_i)$.

Шаг 5. Перейти к следующему узлу $k = k + 1$. Если $a_k$ = *NULL*, то конец алгоритма, иначе перейти к шагу 2.

*Алгоритм формирования полного графа сети*

Шаг 1. Вычислить среднюю длину пути *L* в графе *G*.

Шаг 2. Получить прогнозируемое распределение (гистограмму) степеней связности по закрытым узлам: массив $D = \|d[j]\|, d[j] = t \cdot |U| / |W|$, где *t* – число вершин со степенью связности *j* (*j*=1..$k_{max}$; $k_{max}$=max{$k_1$..$k_{|V|}$}; *k* – степень связности узла).

Шаг 3. Сформировать массив *N*=||*n*[*i*]||, *i*=1..|*U*| по правилу: в массив включаются значения *j* из массива *D* $d_j$ раз. Отсортировать *N* по убыванию.

Шаг 4. Сформировать двумерный массив *C*=||*c*[*i*]|| по правилу: $\forall i = 1..|U|$ $c[i,1] = u_i$, $c[i,2] = k(u_i)$. Отсортировать *C* по значениям *k* в порядке убывания.

Шаг 5. Сохранить исходную конфигурацию сети: *E**=*E*, *C**=*C*. Инициализировать переменную *L***=0 и множество *E***=Ø.

Шаг 6 . Получить новую конфигурацию сети:

Инициализировать счетчик узлов *i*=1. Для $\forall i = 1..|U|$ определить число добавляемых связей для *i*-го узла $r = n[i] - c[i,2]$, *d*=1. Пока *r*>0 и *i*+*d*≤|*U*|, найти узел для связи: если он существует *c*[*i*+*d*,2]<*n*[*i*+*d*], то добавить связь *c*[*i*,2]= *c*[*i*,2]+1, *c*[*i*+*d*,2]= *c*[*i*+*d*,2]+1, *E* = *E* U $e_{c[i,1],c[i+d,1]}$, *r*=*r*-1; *d*=*d*+1.

Шаг 7. Вычислить среднюю длину пути *L** для графа сети с новой конфигурацией.

Шаг 8. Если значение *L** удовлетворяет заданной точности *q* (|*L*-*L**|<*q*), то конец алгоритма.

Шаг 9. Если значение $L^*$ текущей конфигурации ближе к $L$, чем значение $L^{**}$ из предыдущих конфигураций ($|L-L^*| < |L-L^{**}|$), то сохранить лучшую конфигурацию ($L^{**}=L^*$, $E^{**}=E$). Восстановить исходную конфигурацию сети ($E=E^*$, $C=C^*$).

Шаг 10. Сгенерировать новый вариант расстановки узлов в массиве $C$. Если вариантов больше нет, то конец алгоритма, иначе перейти к шагу 6.

Представленная методика реализована в виде разработанного ПО и апробирована на двух фрагмента сетей («ВКонтакте», «Facebook» (алгоритм формирования полного графа)).

**Четвертая глава** посвящена экспериментальным исследованиям и особенностям внедрения.

Моделирование УгЗИ на крупномасштабной ИТКС является трудоемкой задачей. Ее решение в приемлемые сроки и получение актуальных результатов возможно только при использовании распределенных вычислительных ресурсов. При проведении экспериментальных исследований в данной работе была использована супер-ЭВМ «Скиф-Мономах». Экспериментальные исследования проводились на двух фрагментах ИТКС. Первый (фрагмент из 16270504 узлов социальной сети «ВКонтакте») получен в рамках данной научной работы, а второй (фрагмент из 16163521 узла социальной сети «Facebook») получен независимо американскими учеными Minas Gjoka, Maciej Kurant и др.

Эксперименты по моделированию УгЗИ проводились с разными начальными условиями. Примеры графиков результатов проведенного моделирования (на сети «ВКонтакте») приведены на рисунках 4 и 5.

В ходе экспериментальных исследований были получены результаты, касающиеся топологии рассматриваемых ИТКС. Значение средней длины пути для «ВКонтакте» получилось равным 3,32, а для «Facebook» - 4,48. Миланский университет и Facebook, проводя совместное исследование теории шести рукопожатий, получили значение 4,74. Расхождение в значениях объясняется количеством узлов в выборке. Для «ВКонтакте» также были проведены независимые исследования по подсчету средней длины пути. Цепочки оказываются короче (3-4 человека), что соответствует полученным данным в этой работе. Объясняется такое значение тем, что аудитория «ВКонтакте» ограничена (Россия и страны СНГ). Приведенные данные позволяют нам при исследовании крупномасштабных ИТКС использовать фиксированное значение средней длины пути.

Результаты экспериментальных исследований были использованы при проверке адекватности аналитической модели, которая была апробирована на данных, полученных компаниями «SMM3» и «YOUSCAN» в ходе экспресс-мониторинга негативных упоминаний бренда Nestle (дезинформация по поводу обнаружения стекла в детском питании Banana, 1–7.08.2011). Результаты распространения данной

дезинформации в сети «ВКонтакте» были аппроксимированы с помощью аналитической модели (система уравнений 1). Погрешность аппроксимации составила приблизительно 13%.

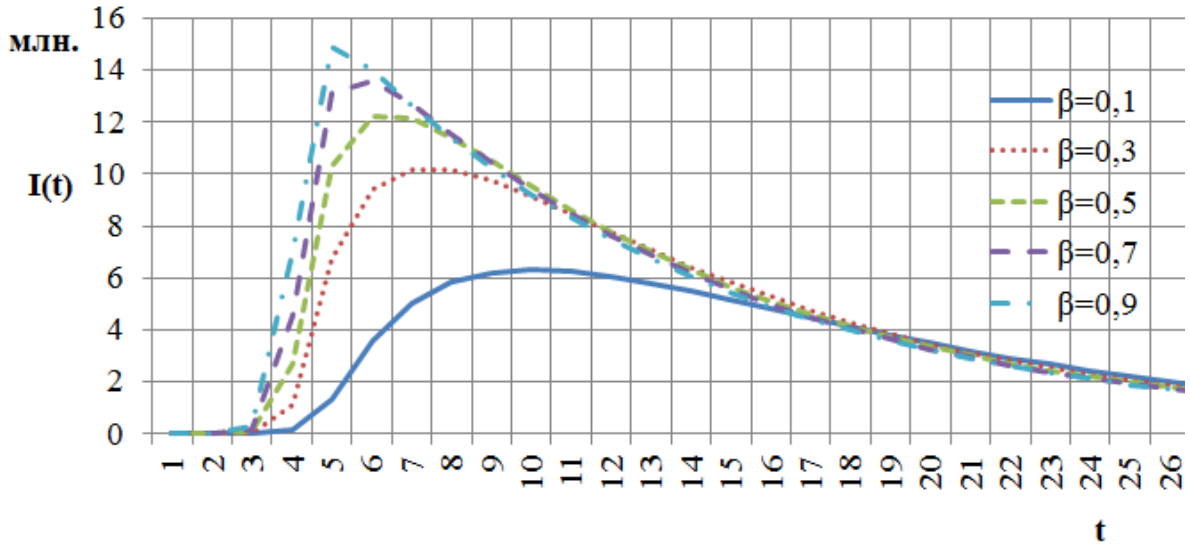


Рисунок 4 – Результаты моделирования с параметрами $\gamma = 0{,}1$, $I_0 = 1$, $R_0 = 0$

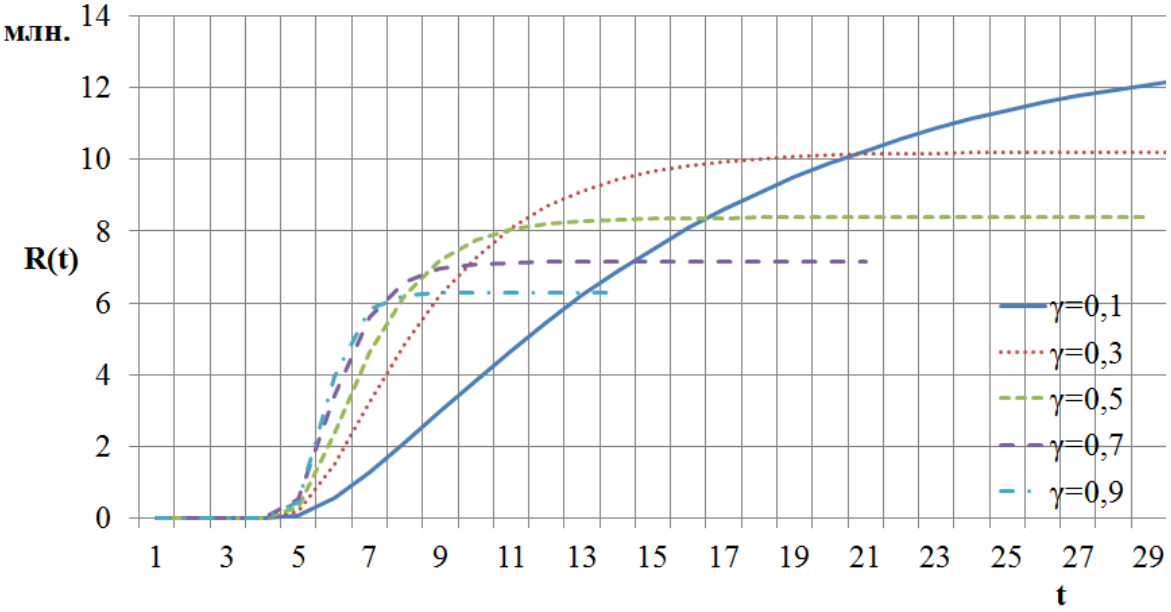


Рисунок 5 – Результаты моделирования с параметрами $\beta = 0{,}2$, $I_0 = 1$, $R_0 = 0$

При наличии административного ресурса можно реализовать автоматизированную систему противодействия угрозе распространения запрещенной информации. Обобщенный алгоритм работы такой системы представлен на рисунке 6. Рассмотренные функции реализуются с помощью типовых средств.

Шаг 1. Ввод данных - типовое сообщение, содержащее информацию, запрещенную к распространению. База данных таких сообщений формируется из федерального списка экстремистских материалов и единого реестра доменных имен, указателей страниц сайтов в сети «Интернет» и сетевых адресов,

позволяющих идентифицировать сайты в сети «Интернет», содержащие информацию, распространение которой в Российской Федерации запрещено.


Рисунок 6 – Обобщенный алгоритм противодействия распространению угрозы запрещенной информации

Шаг 2. Выявление «маркеров», то есть слов и словосочетаний, минимально изменяющихся в ходе переформулировки.

Шаг 3. Синтез формального описания «маркеров» с использованием регулярных выражений или контекстно-свободной грамматики.

Далее работа алгоритма разбивается на две, параллельно выполняющиеся процедуры предупреждения и устранения последствий угрозы.

Предупреждение

Шаг 4а. Составление правил фильтрации сообщений на основе формального описания. Осуществляется путем компиляции регулярных выражений при помощи средств, предназначенных для фильтрации (см. шаг 5а).

Шаг 5а. Конфигурация технических средств фильтрации с использованием правил. Как правило, это антиспам системы такие как Apache Spamassassin, Yandex Spamooborona, Kaspersky Antispam, FASTBL, dnsbl и др..

Шаг 6а. Моделирование угрозы распространения запрещенной информации.

Шаг 7а. Повышение приоритета процесса фильтрации в соответствии с результатами моделирования угрозы распространения запрещенной информации.

Ликвидация последствий

Шаг 4б. Конструирование ряда поисковых запросов по формальным правилам и подстройка параметров поиска (приоритет, глубина и тд.)

Шаг 5б. Исполнение запросов и анализ результатов. На данном этапе возможно уточнение запросов.

Шаг 6б. Удаление найденных сущностей с сохранением связности БД.

Шаг 7б. Отправка сообщения о проведенных мероприятиях в контролирующие органы.

Результаты диссертационной работы находят широкое применение в учебном процессе в ВлГУ. На их основе для подготовки студентов и магистров на кафедре «Информатика и защита информации» доработан ряд курсов. Научные результаты работы использованы для написания учебных пособий, курсового и дипломного проектирований для студентов кафедры.

Разработанное ПО применяется в РОСКОМНАДЗОРе по Владимирской области. Цель применения программы – автоматизация поиска узлов ИТКС – потенциальных распространителей запрещенной информации.

Разработанная методика формирования топологии ИТКС применяется в ОАО «ВПО «Точмаш». Она помогает в создании модели внутреннего злоумышленника – создается граф социальных связей работников организации с целью выявления нежелательных контактов (связи с конкурентами, криминальными элементами, экстремистскими группами и тд.).

**В заключении** приводятся основные результаты диссертационной работы, отмечается теоретическая и практическая ценность.

1. В результате информационного обзора и проведенных экспериментов было выявлено, что на УгЗИ в ИТКС существенное влияние оказывает топология информационных связей между абонентами и модель информационного взаимодействия между ними. Выполненный анализ основных подходов к моделированию УгЗИ показал, что наиболее адекватными моделями для этой задачи являются модели влияния, просачивания и заражения.
2. Создана имитационная модель УгЗИ в ИТКС, учитывающая топологические характеристики сети, а также особенности информационного взаимодействия абонентов как человеко-машинных систем. С ее помощью проведены эксперименты, результаты которых показали зависимость реализации УгЗИ от топологической уязвимости сети.
3. Разработана аналитическая модель УгЗИ с учетом топологической уязвимости сети. Релевантность результатов аналитического решения подтверждена серией экспериментов на топологии реальной сети с использованием имитационного моделирования. При этом погрешность для процесса защиты составила не более 10%, для процесса атаки - не более 15%.
4. Разработана методика формирования топологии ИТКС, которая учитывает основные топологические характеристики доступной части сети и работает в условие недостаточной репрезентативности выборки исходных данных.

5. Разработано программное обеспечение, которое позволяет за приемлемое время получить результаты моделирования УгЗИ в ИТКС за счет использования распределенных вычислительных ресурсов. В ходе экспериментальных исследований были получены результаты, касающиеся топологии рассматриваемых ИТКС. Полученное значение средней длины пути согласуется с результатами независимых исследований и дает возможность использовать его в аналитической модели как фиксированный параметр.

Примеры эффективного апробирования механизмов прогнозирования УгЗИ в ИТКС дают основание констатировать адекватность и функциональность основных теоретических построений и разработанных на их основе алгоритмических и инструментальных средств.

**Основные результаты опубликованы в следующих работах:**

Статьи в изданиях, рекомендованных ВАК РФ

1. Монахов, Ю.М., Абрамов, К.Г. Моделирование распространения нежелательной информации в социальных медиа [Текст] / Ю.М. Монахов, К.Г. Абрамов; Вестник КГУ им. Н.А. Некрасова. - 2011. – Т.17, №3. – С. 15-18 [75%]
2. Абрамов, К.Г., Монахов, Ю.М. Алгоритмическая модель экстраполяции топологических характеристик социальных сетей [Текст] / К.Г. Абрамов, Ю.М. Монахов; Всероссийский научно-технический журнал «Проектирование и технология электронных средств», №4. – 2012. – С. 35-39. [75%]
3. Груздева, Л.М., Абрамов, К.Г., Монахов, Ю.М. Экспериментальное исследование корпоративной сети передачи данных с адаптивной системой защиты информации [Текст] / Л.М. Груздева, К.Г. Абрамов, Ю.М. Монахов; Приборостроение. – М., 2012. – Т. 55, № 8. – С. 57-59. [60%]

Опубликованные доклады в зарубежных и международных НТК

1. Абрамов, К.Г., Монахов, Ю.М., Никиташенко, А.В. К вопросу об уточнении моделей распространения нежелательной информации в социальных сетях Интернета [Электронный ресурс] / К.Г. Абрамов, Ю.М. Монахов, А.В. Никиташенко; Информационные системы и технологии ИСТ-2011: материалы XVII международной научно-технической конференции (Н.Новгород, 22 апреля 2011 года) – Н. Новгород: Электронное издание, 2011. – 149 с.; – ISBN 978-5-9902087-2-8. [70%]
2. Абрамов, К.Г., Монахов, Ю.М. Моделирование распространения нежелательной информации в социальных медиа [Текст] / К.Г. Абрамов, Ю.М. Монахов; Труды XXX Всероссийской научно-технической конференции. Проблемы эффективности и безопасности функционирования сложных технических и информационных систем / Серпуховский ВИ РВ. – 2011. – ч.IV. – С. 178-182. – ISBN 978-5-91954-029-8. [75%]
3. Абрамов, К.Г., Монахов, Ю.М., Распространение нежелательной информации в социальных сетях Интернета [Текст] / К.Г. Абрамов, Ю.М. Монахов; Перспективные технологии в средствах передачи информации: Материалы 9-ой международной научно - технической конференции; редкол.: А.Г. Самойлов [и др]. – Владимир: издат. ВлГУ, 2011. – Т. 1. – 272 с.; – ISBN 978-5-905527-02-9. [75%]

4. Абрамов, К.Г. Влияние перколяционного кластера на распространение нежелательной информации в социальных медиа [Текст] / К.Г. Абрамов; Проблеми інформатики і моделювання. Тезиси одинадцятоі міжнародноі науково-технічноі конференціі. – Харків-Ялта, 2011. – С. 4-5. [100%]
5. Абрамов, К.Г., Монахов, Ю.М. Некоторые аспекты безопасности Интернета в условиях инфраструктуры web 2.0 [Текст] / К.Г. Абрамов, Ю.М. Монахов; Труды X Российской научно-технической конференции "Новые информационные технологии в системах связи и управления". (Калуга, 1-2 июня 2011г.) – Калуга: Изд. "Ноосфера", 2011. – 610 с.; – С. 593-595. – ISBN 978-5-89552-322-3. [75%]
6. Абрамов, К.Г., Монахов, Ю.М., Бодров, И.Ю. К вопросу о моделировании топологии социальных сетей [Текст] / К.Г. Абрамов [и др.]; Труды пятой всероссийской научно-практической конференции по имитационному моделированию и его применению в науке и промышленности "Имитационное моделирование. Теория и практика" ИММОД-2011. – Санкт-Петербург: ОАО "Центр технологии и судостроения", 2011. – 448 с.; – С.373-378. – ISBN 978-5-905526-02-2. [70%]
4. Абрамов, К.Г., Монахов, Ю.М., Медведникова, М.А., Трусова, А.И., Бодров, И.Ю. К вопросу о моделировании процесса пропаганды в социальных сетях [Текст] / К.Г. Абрамов и [др.]; Математика и математическое моделирование. Труды научно-практической конференции, Мордовский государственный педагогический институт имени М. Е. Евсевьева. – 2011. [60%]
5. Абрамов, К.Г., Монахов, Ю.М., Медведникова, М.А., Трусова, А.И., Бодров, И.Ю. Статистические параметры топологии социальных сетей [Текст] / К.Г. Абрамов [и др.]; Математика и математическое моделирование. Труды научно-практической конференции, Мордовский государственный педагогический институт имени М.Е. Евсевьева. – 2011. [75%]
6. Абрамов, К.Г., Малышев, Р.В., Монахов, Ю.М. К вопросу о топологических характеристиках социальной сети «В КОНТАКТЕ» [Текст] / К.Г. Абрамов, Р.В. Малышев, Ю.М. Монахов; Перспективные технологии в средствах передачи информации: Материалы 10-ой международной научно-технической конференции, Владим. гос. ун-т. – 2013. – т. 2. – С. 115-118. [75%]
7. Абрамов, К.Г., Монахов, Ю.М. Топологические характеристики социальной сети «ВКОНТАКТЕ» [Текст] / К.Г. Абрамов, Ю.М. Монахов; Труды XXXII Всероссийской научно-технической конференции. Проблемы эффективности и безопасности функционирования сложных технических и информационных систем / Серпуховский ВИ РВ. – 2013. – ч.IV. – С. 136-140. – ISBN 978-5-91954-074. [75%]

**Свидетельства о регистрации комплекса программ**

1. Программа имитационного моделирования распространения нежелательной информации в социальных сетях / Абрамов Константин Германович, Монахов Юрий Михайлович // Свидетельство о государственной регистрации программы для ЭВМ №2011617403 / Федеральная служба по интеллектуальной собственности, патентам и товарным знакам. – 23.09.2011.
2. Программа вычисления топологических характеристик социальных сетей / Абрамов Константин Германович, Бодров Иван Юрьевич, Монахов Юрий

Михайлович // Свидетельство о государственной регистрации программы для ЭВМ № 2012610825 / Федеральная служба по интеллектуальной собственности. – 18.01.2012.

3. Программный комплекс топологического анализа и моделирования распространения запрещенной информации в крупномасштабных социальных сетях / Абрамов Константин Германович, Малышев Роман Владимирович, Монахов Юрий Михайлович, // Свидетельство о государственной регистрации программы для ЭВМ №2013660757 / Федеральная служба по интеллектуальной собственности. – 18.11.2013.

Подписано в печать 20.06.2014
Формат 60×84/16. Усл. печ. л. 1,16. Тираж 100. Заказ 198
Издательство
Владимирского государственного университета имени
Александра Григорьевича и Николая Григорьевича Столетовых
600000, Владимир, ул. Горького, 87