ВЛАДИМИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЕКСАНДРА ГРИГОРЬЕВИЧА И НИКОЛАЯ ГРИГОРЬЕВИЧА СТОЛЕТОВЫХ
ФАКУЛЬТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ
КАФЕДРА ИНФОРМАТИКИ И ЗАЩИТЫ ИНФОРМАЦИИ

На правах рукописи

АБРАМОВ Константин Германович

МОДЕЛИ УГРОЗЫ РАСПРОСТРАНЕНИЯ ЗАПРЕЩЕННОЙ ИНФОРМАЦИИ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ

Специальность:
05.12.13 – Системы, сети и устройства телекоммуникаций

Диссертация
на соискание ученой степени кандидата технических наук

Научный руководитель -
д.т.н., проф. Монахов М.Ю.

Владимир
2014

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 4
ГЛАВА 1 Безопасность в информационно-телекомуникационных сетях. уточнение задач исследования 9
1.1 Анализ объекта исследования 9
1.2 Проблемы информационной безопасности в ИТКС 13
1.3 Моделирование ИТКС 20
1.3.1 Моделирование топологии ИТКС 21
1.3.2 Моделирование процессов информационного взаимодействия в ИТКС 23
1.3.3 Эпидемиологические модели 27
1.4 Задачи исследования 28
Выводы к первой главе 29
ГЛАВА 2 Разработка и исследование моделей угрозы распространения запрещенной информации в информационно-телекоммуникационных сетях 31
2.1 Имитационное моделирование 32
2.2 Разработка аналитической модели 42
2.3 Экспериментальное исследование аналитической модели 48
Выводы ко второй главе 50
ГЛАВА 3 Разработка методики формирования топологии крупномасштабной информационно-телекоммуникационной сети 51
3.1 Сбор данных о топологии доступной части сети 53
3.2 Формирование полного графа сети с учетом недоступной части 60
3.3 Формирование вектора топологической уязвимости полного графа сети 66
3.4 Особенности разработки программного инструментария 68
Выводы к третьей главе 71
ГЛАВА 4 Экспериментальное исследование. особенности внедрения 72
4.1 Распределенное моделирование угрозы распространения запрещенной информации в ИТКС 72
4.2 Анализ результатов экспериментальных исследований 75
4.2.1 Анализ результатов моделирования УгЗИ в ИТКС 75

4.2.2 Анализ результатов экспериментальных исследований топологии ИТКС 82
4.3 Особенности реализации автоматизированной системы противодействия угрозе распространения запрещенной информации 90
4.4 Особенности практического применения аналитической модели УгЗИ в ИТКС 94
4.5 Особенности практического внедрения 95
Выводы к четвертой главе 98
ЗАКЛЮЧЕНИЕ 100
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 103
Приложение А 118
Приложение Б 119
Приложение В 120
Приложение Г 121
Приложение Д 124

ВВЕДЕНИЕ

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, региональный аттестационный центр ООО «ИнфоЦентр». 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.

ГЛАВА 1 БЕЗОПАСНОСТЬ В ИНФОРМАЦИОННО-ТЕЛЕКОМУНИКАЦИОННЫХ СЕТЯХ. УТОЧНЕНИЕ ЗАДАЧ ИССЛЕДОВАНИЯ

1.1 Анализ объекта исследования

ИТКС обеспечивают практически полный спектр возможностей для обмена информацией между пользователями - сетевыми абонентами. ИТКС предоставляет различные сервисы для организации социальных взаимоотношений между пользователями (абонентами). На сегодняшний день наиболее популярным из них являются социальные сети.

В мире существует огромное количество различных социальных сетей, но практически в каждой стране или регионе существуют несколько наиболее популярных представителей. В США это «Facebook», «MySpace», «Twitter» и «LinkedIn»; «Nexopia» — в Канаде, «Bebo» — в Великобритании, «Facebook», «dol2day» — в Германии. В России на сегодняшний день самыми популярными являются «ВКонтакте», «Одноклассники.ru», «Мой Мир@mail.ru». На рисунке 1.1 изображена динамика роста пользователей самой крупномасштабной в России социальной сети «ВКонтакте» [137].

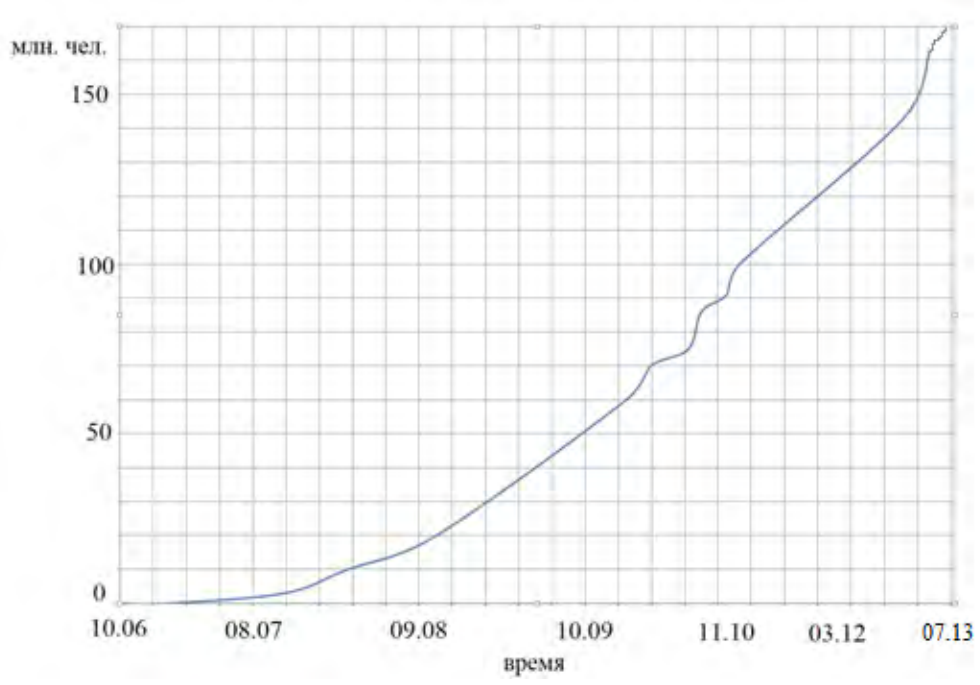


Рисунок 1.1 – Динамика роста пользователей социальной сети «ВКонтакте»

С бурным ростом числа пользователей ИТКС возникают и проблемы безопасности в них.

Обобщенная структурная схема ИТКС приведена на рисунке 1.2.

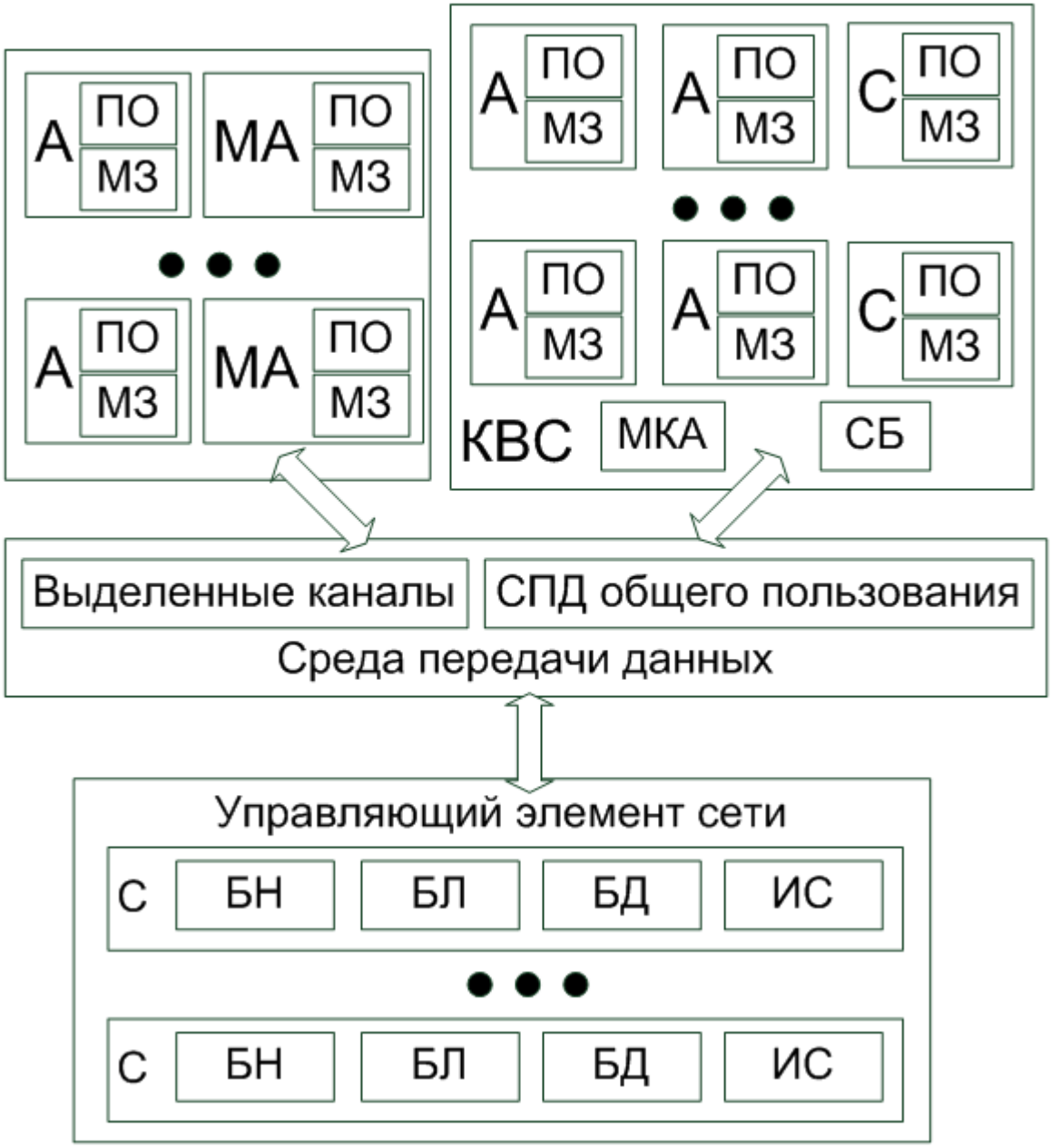


Рисунок 1.2 – Структурная схема ИТКС

Ее состав в общем случае образуют следующие функциональные элементы:

- абоненты (А). Под абонентом понимается человеко-машинная система, состоящая из устройства, через которое осуществляется доступ к сети, и непосредственно пользователя ИТКС. Абоненты могут быть отдельными узлами сети (если пользователь использует свой домашний компьютер), либо могут быть объединены в корпоративную вычислительную сеть (КВС) (если абонент

использует рабочий компьютер), 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 для всей корпоративной сети;

- средства телекоммуникации, обеспечивающие взаимодействие между собою абонентов;

- 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.).

Архитектура типового управляющего элемента представлена на рисунке 1.3.

Опишем эту многослойную архитектуру.

1. Презентационный слой.

На этом слое принимаются HTTP-запросы от абонентов, обычно веб-браузеров, и выдаются им HTTP-ответы, как правило, вместе с HTML-страницей, изображением, файлом, медиа-потоком или другими данными. Здесь же осуществляется распределение и балансировка нагрузки, ведение журнала обращений абонентов к ресурсам.

1. Слой бизнес сервисов.

Данный слой предназначен для подбора и обработки данных.

2. Персистентный слой.

Этот слой выполняет обслуживание и управление базой данных и отвечает за целостность и сохранность данных, а также обеспечивает операции ввода-вывода при доступе абонента к информации.

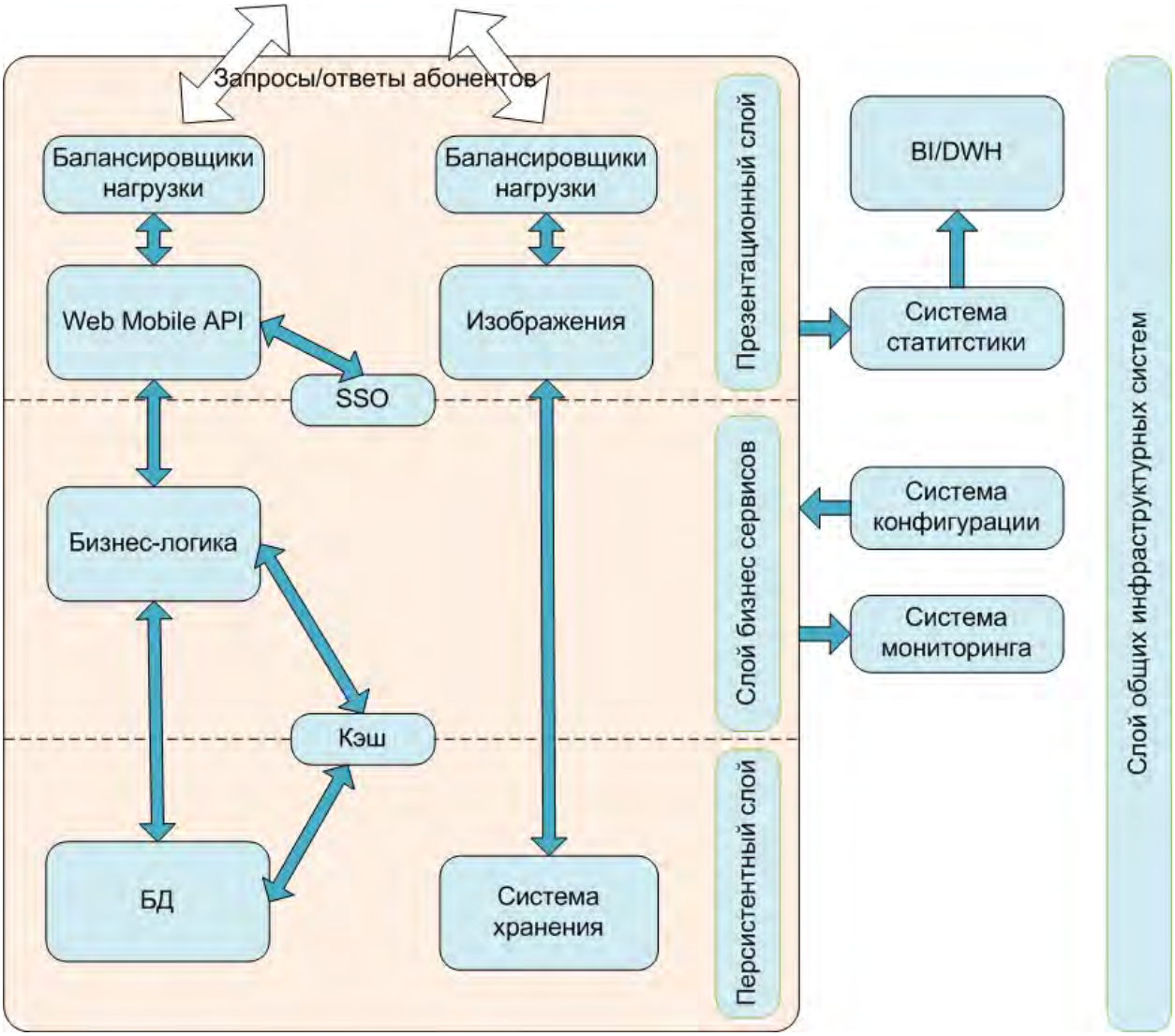


Рисунок 1.3 - Архитектура типового управляющего элемента

3. Слой общих инфраструктурных систем.

На этом слое размещаются системы протоколирования статистики, конфигурации приложений, мониторинга.

SSO (Single Sign-On, технология единого входа) - технология, при использовании которой пользователь переходит из одного раздела портала в другой без повторной аутентификации.

BI (Business intelligence, бизнес-анализ, бизнес-аналитика) - методы и инструменты для построения информативных отчётов о текущей ситуации в системе.

DWH (Data Warehouse, хранилище данных) - предметно-ориентированная информационная база данных, специально разработанная и предназначенная для подготовки отчётов и бизнес-анализа.

1.2 Проблемы информационной безопасности в ИТКС

Приведем основные проблемы информационной безопасности в ИТКС, которые актуальны для настоящего исследования.

1. Использование глобальной сети Интернет как распределенной информационно-телекоммуникационной системы.

Наиболее уязвимыми и поэтому часто атакуемыми компонентами системы являются:

1) Серверы.
2) Рабочие станции.
3) Среда передачи информации.
4) Узлы коммутации.

Типовые информационные воздействия злоумышленников:

1) Прослушивание сетевого трафика. Для прослушивания трафика (sniffing) сетевой адаптер переводится в «беспорядочный» режим. В данном режиме адаптер перехватывает все сетевые пакеты, проходящие через него, а не только предназначенные данному адресу, как в нормальном режиме функционирования - технологии – ARP Spoofing (ARP-poisoning), MAC Flooding и MAC Duplicating [19,42]. Перехват осуществляется с использованием сетевых мониторов, из которых наиболее функциональными являются Sniffer Pro от компании Sniffer Technologies [17], IRIS Network Traffic Analyzer от компании eEYE [51] и TCP Dump [88].

Последствия. Современные сетевые протоколы (TCP/IP, ARP, HTTP, FTP, SMTP, POP3 и т.д.) не имеют механизмов защиты (передаются в открытом виде).

Злоумышленник, перехватывающий трафик между сервером и любым узлом сети, может завладеть аутентификационными данными пользователя (получить пароль).

Противодействие. Известен ряд методов определения наличия запущенного сниффера в сети, например, метод пинга, метод ARP, метод DNS и метод ловушки [19, 21, 39].

2) Сканирование уязвимостей. Результатом работы сканера является информация о системе, включающая список сетевого оборудования, компьютеров с запущенными на них службами, версиями сетевого ПО (а значит и уязвимостей, присущих данному ПО), учетные записи пользователей. Сканирование уязвимостей обычно является этапом, предваряющим атаку. Именно результаты сканирования позволяют точно подобрать эксплойты для осуществления непосредственно НСД.

Обнаружение**.** Само по себе сканирование не является незаконным. Однако, если сканирование со стороны внешней, по отношению к системе, сети обыкновенное явление, то сканирование компьютеров из внутренней сети – безусловно, инцидент безопасности, требующий незамедлительной реакции со стороны сетевого администратора. Обнаружить следы сканирования можно, изучая журналы регистрации МЭ. Однако такой подход не позволяет своевременно реагировать на подобные инциденты. Поэтому современные МЭ и СОВ имеют модули (plug-in) [16, 19, 26], позволяющие обнаружить сканирование в режиме реального времени. Некоторые сканеры уязвимостей используют оригинальные методы, позволяющие производить сканирование максимально скрытно. Например, в Nmap [19] существуют возможности, позволяющие значительно затруднить обнаружение сканирования для СОВ.

Противодействие. Использование сетевых СОВ, либо периодическое изучение журналов регистрации МЭ.

3) Сетевые атаки. Сетевые атаки можно разделить на:

- атаки, основанные на переполнении буфера (overflow based attacks). Они используют уязвимость системы, заключающуюся в некорректной программной

обработке данных. При этом появляется возможность выполнения вредоносного кода с повышенными привилегиями;

- атаки, направленные на отказ в обслуживании (Denial Of Service attacks). Атаки не обязательно используют уязвимости в ПО атакуемой системы. Нарушение работоспособности системы происходит из-за того, что посылаемые ей данные приводят к значительному расходу ресурсов системы. Самым простым примером атаки этого типа является атака «Ping Of Death» [52, 84]. Сущность ее в следующем: на машину жертвы посылается сильно фрагментированный ICMP-пакет большого размера (64KB). Реакцией ОС Windows на получение такого пакета является полное зависание.

4) Атаки, основанные на использовании уязвимостей в ПО сетевых приложений - эксплойты (exploit) [31, 48, 55]. Данный класс атак основан на эксплуатации различных дефектов в ПО. Эксплойты представляют собой вредоносные программы, реализующие известную уязвимость в ОС или прикладном ПО, для получения НСД к уязвимому хосту или нарушение его работоспособности. Для эксплойтов характерно наличие функций подавления антивирусных программ и МЭ. Последствия применения эксплойтов могут быть самыми критическими. В случае получения злоумышленником удаленного доступа к системе, он имеет практически полный (системный) доступ к компьютеру. Последующие действия и ущерб от них могут быть следующими: внедрение троянской программы, внедрение набора утилит для сокрытия факта компрометации системы, несанкционированное копирование злоумышленником данных с жестким и съемных носителей информации системы, заведение на удаленном компьютере новых учетных записей с любыми правами в системе для последующего доступа как удаленно, так и локально, кража файла с хэшами паролей пользователей, уничтожение или модификация информации, осуществление действий от имени пользователя системы.

Противодействие**.** МЭ и СОВ, установленные на атакуемой системе, в ряде случаев не в состоянии отразить действие эксплойтов [38, 40, 41, 80, 82]. Для успешного отражения атак эксплойтов средства защиты необходимо обновлять,

поскольку механизм обнаружения вторжений основан на распознавании сигнатур уже известных атак. Хотя существуют разработки, способные по заверениям разработчиков отражать неизвестные атаки, практика показывает, что они все еще не эффективны.

5) Вредоносные программы (ВПр). ВПр - это компьютерная программа или переносной код, предназначенный для реализации угроз информации, хранящейся в сети, либо для скрытого нецелевого использования ресурсов или либо иного воздействия, препятствующего нормальному функционированию сети. К ВПр относятся компьютерные вирусы, троянские кони, сетевые черви и др [18, 22, 23, 28, 29, 30, 37, 57].

Противодействие. Типичным методом противодействия является использование антивирусных средств, работающих в режиме реального времени (мониторов). Для выявления троянских программ существует специализированное программное обеспечение.

2. Проблема запрещенного контента.

В зависимости от законодательства страны разные материалы могут считаться нелегальными. В большинстве стран запрещены: материалы сексуального характера с участием детей и подростков, порнографический контент, описания насилия, в том числе сексуального, экстремизм и разжигание расовой ненависти.

В российском законодательстве существует понятие запрещенной к распространению информации. Определяется такая информация постановлением правительства РФ, от 26 октября 2012 г. № 1101 – «О создании единой автоматизированной информационной системы "Единый реестр 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». В этом постановлении правительства прописаны изменения и дополнения к федеральному закону от 27 июля 2006 года № 149-ФЗ – «Об информации, информационных технологиях и защите информации», которые вступили в силу с

01 ноября 2012 года. В параграфе «Правила принятия уполномоченными Правительством Российской Федерации федеральными органами исполнительной власти решений в отношении отдельных видов информации и материалов, распространяемых посредством информационно-телекоммуникационной сети "Интернет"» постановления подробно указано, какая информация признана незаконной к распространению в России в сети Интернет, и какие правоохранительные органы ответственны за контроль над распространением незаконной информации.

Аналогично с концепцией обеспечения комплексной защиты объекта информатизации, можно сформировать полное множество функций защиты от запрещенной информации.

Под функцией защиты (ФЗ) понимается совокупность однородных в функциональном отношении мероприятий, регулярно осуществляемых в автоматизированных системах различными средствами и методами с целью создания, поддержания и обеспечения условий, объективно необходимых для надежной защиты информации.

Перечень полного множества функций защиты от запрещенной информации в социальных сетях:

1) Предупреждение условий возникновения запрещенной информации.

Функция реализуется с помощью нормативно-правовых актов. Она не может полностью исключить угрозу распространения запрещенной информации в социальных сетях, так как в целом ситуация с соблюдением законов (особенно в России) неудовлетворительна, а в интернет-пространстве обостряется из-за технических сложностей.

2) Предупреждение непосредственного проявления запрещенной информации.

Функция реализуется за счет механизмов прогнозирования распространения запрещенной информации в социальной сети. Более подробно данная функция будет рассмотрена ниже.

3) Обнаружение проявившейся запрещенной информации.

Функция связана с мониторингом ИТКС на предмет запрещенной информации на страницах абонентов. Как правило, для реализации данной защиты используется различные СОРМ. Данная ФЗ связана с проблемами контекстного поиска, а также необходимостью контроля над всей системой.

4) Предупреждение воздействия на абонентов проявившейся запрещенной информации.

Функция может быть реализована с помощью автоматической рассылки сообщений с предупреждением об ответственности за распространение запрещенной информации, вплоть до блокировки абонента. Блокировка может осуществляться легитимными средствами при наличии доступа к управлению системы и нелегитимными – при его отсутствии (взлом аккаунта). ФЗ делится на две функции ($ФЗ_{4а}$ и $ФЗ_{4б}$). Первая связана с предупреждением абонентов, на страницах которых была найдена запрещенная информация, а вторая – с рассылкой предупреждений потенциальным получателям запрещенной информации.

5) Обнаружение воздействия запрещенной информации на абонентов.

Функция связана непосредственно с фиксацией процесса распространения запрещенной информации, может быть реализована через контекстный анализ сообщений. Свойственны такие же недостатки, как и для $ФЗ_3$.

6) Локализация, ограничение воздействия запрещенной информации на абонентов.

Функция реализуется через блокировку абонентов, распространяющих запрещенную информацию ($ФЗ_{6а}$), или абонентов – потенциальных распространителей ($ФЗ_{6б}$). Данная ФЗ опирается на предыдущие функции и для ее эффективной реализации необходим контроль над системой.

7) Ликвидация последствий обнаруженного воздействия запрещенной информации на абонентов.

Функция связана с удалением запрещенной информации из системы. Для реализации данной функции также необходим контроль над системой.

На рисунке 1.4 приведены все сочетания событий, которые потенциально возможны при осуществлении всех ФЗ.

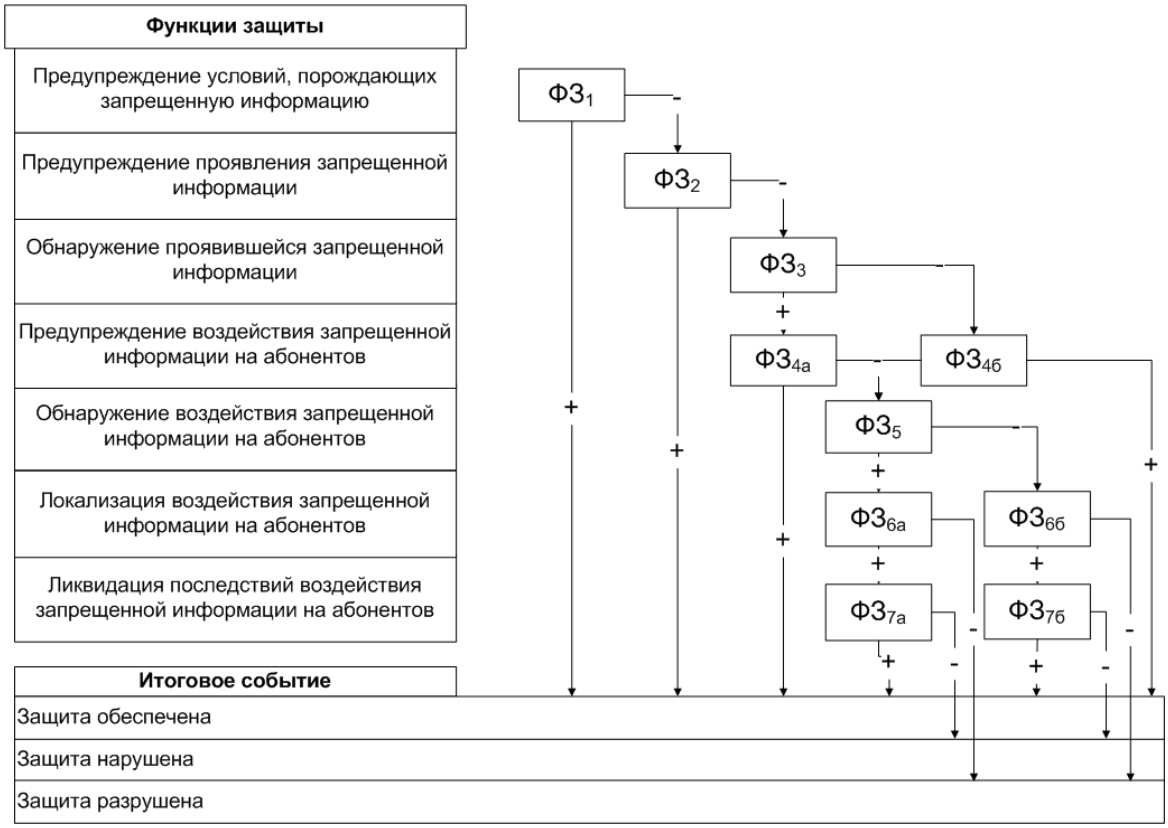


Рисунок 1.4 – Функции защиты от запрещенной информации в ИТКС

Из анализа функций защиты видно, что наиболее эффективные функции – это первые функции, так как они обеспечивают защиту на ранних этапах. Все функции имеют свои недостатки.

Наиболее перспективной ФЗ инженерно-технического направления является $ФЗ_2$. Именно ей посвящена данная работа. На данном этапе, имея информацию о топологии ИТКС и потенциальных распространителях запрещенной информации, возможно прогнозирование процесса ее распространения.

Создание моделей и алгоритмов распространения угрозы запрещенной информации – одна из ключевых задач в данном направлении. При ее решении возникают проблемы, связанные со свойствами рассматриваемой информационно-телекоммуникационной системы, а именно:

1. Отсутствие проверки подлинности данных об узле системы. Очень часто абоненты ИТКС указывают недостоверную информацию о себе.

2. Закрытость системы. Структура и информация об управлении системой является конфиденциальной информацией.

3. Проблема сбора информации. Невозможно получить полную информацию о топологии ИТКС. Существует возможность для обычного абонента сбора информации о структуре сети (функции API), но эта возможность имеет много ограничений (настройки приватности, временной интервал).

В рамках данной работы нас интересует только обмен сообщениями между абонентами, поэтому концептуальная математическая модель информационного взаимодействия представляется графом, узлами которого являются абоненты, а ребрами – связи между ними. Перечислим свойства графа, принципиальные для настоящего исследования:

1. Большая размерность. Система содержит миллионы элементов.

2. Гетерогенность. В графе, который отражает взаимосвязь элементов в системе, вершины имеют разное количество прилегающих ребер.

3. Динамика связей. В системе в течение времени происходят изменения связей.

4. Динамика узлов. В течение времени изменяется количество узлов (элементов) системы.

5. Наличие групп узлов, имеющих большое количество связей внутри и небольшое – между группами. Граф, представляющий систему, обладает определенной кластеризацией. Для таких систем характерно, что два узла, имеющие связи к какому-либо узлу, часто также имеют связь между собой.

Наиболее эффективное прогнозирование распространения угрозы запрещенной информации осуществляется с помощью моделирования данного процесса. Таким образом, мы приходим к задаче моделирования ИТКС с помощью их математической модели (графов).

1.3 Моделирование ИТКС

Один из основных способов изучения ИТКС – моделирование, которое принято рассматривать в двух аспектах. Первый касается моделирования

топологии (структуры информационных связей между узлами сети) ИТКС, а второй затрагивает проблему изучения процессов, проходящих в ней. В нашем случае это угроза распространения запрещенной информации (УгЗИ).

1.3.1 Моделирование топологии ИТКС

С точки зрения топологии ИТКС относят к сложным сетям [59, 110 и др.]. Сложные сети (комплексные сети, complex networks) — это существующие в природе сети, обладающие нетривиальными топологическими свойствами.

В [96] Jasmin Leveille приводит одну из принятых классификаций топологических моделей сетей (рис.1.5) [46]. В овалах указаны классы сетей, а в прямоугольниках конкретные модели-представители. Описываются их характеристики: распределение степеней связности узлов сети, кластерный коэффициент и средняя длина пути сети.

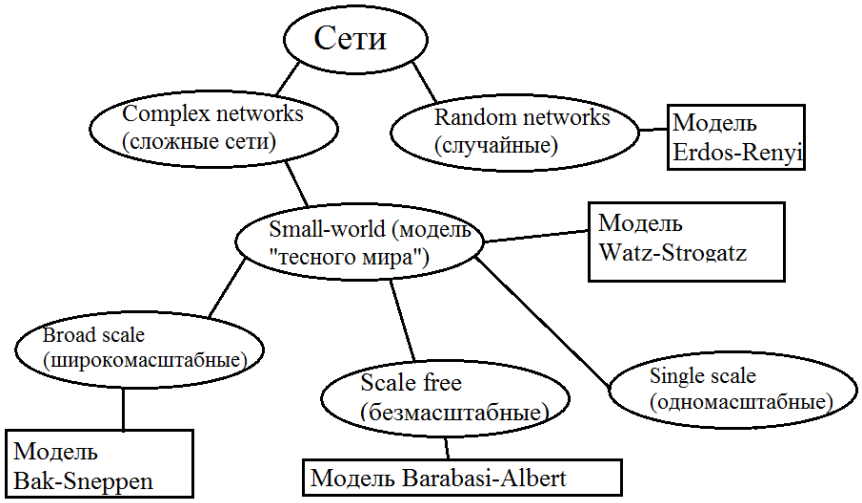


Рисунок 1.5 - Классификация сетей

Тема случайных графов (сетей) раскрыта в работах: Bela Bollobas [53], Erdos и Renyi [63]. В [66] описываются основные модели сетей и их основные характеристики. Проводится исследование топологии популярных ИТКС и осуществляется поиск наиболее адекватной топологической модели. Представлен обзор канонических работ [64, 134, 45, 44] по данной тематике. Выделены главные современные тенденции в области анализа топологии ИТКС:

1. исследование топологических характеристик ИТКС [43, 98];

2. исследование эволюции ИТКС [90, 95];

3. изучение и разработка методов для вычисления характеристик крупномасштабных ИТКС, решение проблемы получения репрезентативной выборки из ИТКС [93, 72].

ИТКС с точки зрения маркетинговых стратегий на их основе рассмотрены в [75, 97]. ИТКС часто относят к scale-free (SF) сетям [49, 133, 56, 50]. Первые работы принадлежат Barabasi и Albert и посвящены одноименной модели. Они сравнивают существующие модели и свою модель, подробно ее описывают и приводят ее сильные и слабые стороны.

В [63] Dorogovtsev и Mendes обобщают Barabasi-Albert модель и находят ее решение. Они находят распределение связности узлов и некоторые другие связанные с ним параметры сети. Также показано, что возникающая масштабируемость действительна не только для данной модели, но и для широкого класса растущих сетей. В [58] Dorogovtsev и Mendes приводят полученные универсальные отношения масштабирования, описывающие свойства развивающихся scale-free сетей и указывают пределы их действия. Дано доказательство того, что основные свойства развивающихся SF сетей могут быть описаны в рамках аналитической модели. В [60] рассматривают свойства перколяционного кластера.

Romualdo Pastor-Satorras and Alessandro Vespignani рассматривали не только топологию ИТКС как SF сеть, но анализировали процесс распространения эпидемии на таких сетях [112-126]. Получили данные по компьютерным вирусам и выявляли такие их параметры, как средняя «продолжительность жизни» вируса и устойчивость к уничтожению. Описали динамическую модель распространения инфекции в сетях. Привели метод определения наличия эпидемического порога в сети.

Еще одной точкой зрения на тип топологии ИТКС является позиция таких исследователей как M.E.J. Newman, D.J. Watts, S.H. Strogatz и ряда других ученых. Они считают, что ИТКС топологически представляет собой класс small-world сетей [99-110, 134]. В [106] рассматривается Watts-Strogatz модель, которая

относится к классу small-world сетей. Это модель имитирует структуру ИТКС. Рассмотрена проблема перколяции узлов на small-world сетях. Этот подход позволяет рассматривать простую модель распространения заболеваний (SIS) и получить аппроксимированное выражение для порога перколяции. Все аналитические результаты подтверждаются численными решениями модели.

Small-world сети рассматриваются во многих работах не только с точки зрения топологии сети [85], но и как основа для эпидемиологических моделей [91]. Более подробно эта тема раскрыта в пункте 1.3.3 «Эпидемиологические модели».

Broad Scale сети рассматриваются в [89], где анализируется Bak-Sneppen модель. Данный вид сложных сетей наименее привлекательный при моделировании топологии ИТКС.

Анализ научных трудов, в которых рассматриваются различные подходы к моделированию топологии ИТКС, показывает, что при решении данной задачи, как правило, используются small-world и scale free сети.

1.3.2 Моделирование процессов информационного взаимодействия в ИТКС

При рассмотрении вопросов, касающихся моделирования процессов, протекающих в ИТКС, основным подходом является применение моделей влияния, информационного управления и противоборства [25]. В данной работе рассматриваются модели влияния, так как они наиболее адаптивны к решаемым задачам. На рисунке 1.6 представлена обобщенная классификация моделей влияния. Коротко охарактеризуем представленные классы моделей влияния.

Пороговой моделью является любая модель, в которой есть пороговое значение или набор пороговых значений, используемых при изменении состояний. Классические модели с порогами были разработаны Schelling, Axelrod и Granovetter для моделирования коллективного поведения [78].

Модели независимых каскадов (Independent Cascade Model) принадлежит категории моделей так называемых «систем взаимодействующих частиц» (Interacting Particle Systems). Узел сети (агент) определяется аналогично

вышеописанной модели. Когда агент *i* становится активным в некоторый момент времени, он получает шанс активировать на следующем (и только на следующем) шаге каждого из своих соседей *j* с вероятностью p_{ji} (причем *j* могут пытаться независимо активировать и другие агенты) [76].

Рисунок 1.6 - Классификация моделей влияния

Модели просачивания и заражения являются популярным способом изучения распространения информации и инноваций в социальных системах. Более подробно они рассмотрены в 1.3.3 «Эпидемиологические модели».

Модель Изинга — математическая модель, описывающая возникновение намагничивания материала. В [129] предполагается, что конформность или независимость в большой социальной группе может моделироваться с помощью модели Изинга; влияние ближайших соседей является определяющим, а аналогом температуры является готовность группы мыслить творчески, готовность принять

новые идеи. Внешним полем для социальной группы является влияние «авторитета» или управление. Более сложные модели, описывающие ИТКС на термодинамических аналогиях, рассматривались в [20].

Для описания процессов распространения информации в ИТКС последнюю можно рассматривать как сложную адаптивную систему, состоящую из большого количества агентов, взаимодействие между которыми приводит к масштабному, коллективному поведению, которое трудно предсказать и анализировать. Для моделирования и анализа таких сложных систем иногда используются клеточные автоматы. Клеточный автомат состоит из набора объектов (в данном случае агентов), обычно образующих регулярную решетку. Состояние отдельно взятого агента в каждый дискретный момент времени характеризуется некоторой переменной. Состояния синхронно изменяются через дискретные интервалы времени в соответствии с неизменными локальными вероятностными правилами, которые могут зависеть от состояний ближайших соседних агентов в окрестности данного агента, а также, возможно, от состояния самого агента.

В статье [136] представлена модель цепей Маркова, в которой изучается влияние в команде (группе агентов). Предлагаемая модель является динамической байесовой сетью (Dynamic Bayesian Network — DBN) с двухуровневой структурой: уровнем индивидов (моделируются действия каждого агента) и уровнем группы (моделируются действия группы в целом).

Модели взаимной информированности [25]. Есть агент, входящий в некоторую социальную сеть. Агент информирован о текущей ситуационной обстановке (действиях и представлениях других агентов, параметрах среды — так называемом состоянии природы (state of nature) и т.п.). Ситуационная обстановка влияет на имеющийся у агента набор ценностей, установок и представлений, связанных следующим образом: ценности влияют на установки, а те, в свою очередь, приводят к предрасположенности к представлениям того или иного уровня, с предрасположенностями согласована находящаяся «в памяти» агента иерархическая система представлений о мире. Предрасположенность к тем или иным представлениям и ситуационная обстановка (например, действия других

агентов) приводят к формированию новых или модификации старых представлений. В соответствии с этими представлениями и установленной целью агент принимает решение и выполняет действие. Результаты действий приводят к изменению как самой ситуационной обстановки, так и внутренних ценностей, установок и представлений.

Модели согласованных коллективных действий. Ключевое значение здесь имеют социальные связи. С одной стороны, социальные связи могут обеспечить эффективный локальный социальный контроль для стимулирования участия в коллективном действии (в силу давления со стороны своих соседей, доверия к ним, социального одобрения, необходимости сохранения положительных отношений и соответствия ожиданиям, эмоциональной привязанности, сохранения своей репутации, отождествления себя с соседями и т.п.). Так, например, поведение соседей агента повлияет на его собственное поведение. С другой стороны, социальные связи обеспечивают агента информацией о намерениях и действиях других агентов в сети и формируют его (неполные) представления, на основе которых агент принимает свои решения. И, наконец, в пределах социальных связей агенты могут прикладывать совместные усилия по созданию локального общественного блага и совместно пользоваться им. Поэтому структура ИТКС оказывает сильное воздействие на решения агентов о принятии участия в коллективном действии.

В [54] ИТКС рассматривается как коммуникационная, посредством которой агенты сообщают друг другу о своей готовности принять участие в коллективном действии. Каждый агент информирован о готовности только своих ближайших соседей и на основе этого локального знания принимает решение об участии, используя правило принятия решений «я приму участие, если примешь участие ты» (механизм координации). То есть рассматривается координационная игра с неполной информированностью. Коммуникационная сеть способствует координации, и основной интерес представляет то, каковы свойства таких сетей, которые допускают коллективное действие. Рассматриваются минимально достаточные сети, которые выстраивают агентов в иерархию социальных ролей

/ступеней: «ведущие» (initial adopters), «последователи» (followers) и т.д. до «поздних последователей» (late adopters). Такие сети способствуют координации следующим образом:

1) информируя каждую ступень о более ранних ступенях;
2) формируя общее знание в пределах каждой ступени.

То есть обеспечивается понимание роли (локально) общего знания в коллективном действии и соотношение между структурой социальной сети и общим знанием.

Равновесие стабильной сети (stable network equilibrium) [83] - ситуация, в которой не существует агента, для которого любая комбинация изменения его действия и изменения его связей приведет к лучшему результату. Только равновесия с полным участием или полным неучастием являются равновесиями стабильной сети.

1.3.3 Эпидемиологические модели

В [81] рассматриваются модели распространения инфекционных заболеваний среди населения, проводится их математический анализ и применение к конкретным заболеваниям. Рассматривается классическая эпидемиологическая SIR модель Кермака-Маккендрика, MSEIR и SEIR эндемические модели.

В [135] рассматривается эпидемиологические модели распространения вирусов и борьбы с ними. Представлена новая модель, которая может быть использована для прогнозирования процесса распространения вредоносных программ и оценки эффективности противодействия им. Показано, как применяется модель для анализа динамики системы, инфекционных вспышек и других процессов, связанных с распространением вирусов.

В [87] Kephart и White проводят аналогию между биологическими и компьютерными вирусами и рассматривают адаптацию методов математической эпидемиологии к изучению компьютерных вирусов. Рассматриваются стандартные эпидемиологические модели на ориентированном графе,

используется моделирование для изучения распространения вирусов. Большое внимание уделяется изучению критического порога эпидемии.

В [115] Pastor-Satorras и Vespignani представляют анализ динамики развития эпидемии в сложных гетерогенных сетях, приводят аналитические и численные результаты. Рассматривается влияние начальных условий и актуальность статистических результатов исследования, касающегося гетерогенных сетей. Авторы считают, что представленные теоретические сведения представляют большой интерес и могут дать полезную информацию для разработки стратегий, направленных на адаптивное сдерживание эпидемии.

В [94] Leskovec, Adamic и Huberman вирусный маркетинг в ИТКС. Вирусный маркетинг — общее название различных методов распространения рекламы, характеризующихся распространением в прогрессии близкой к геометрической, где главным распространителем информации являются сами получатели информации. Осуществляется данный подход путем формирования содержания сообщения, таким образом, который способен привлечь новых получателей информации за счет яркой, творческой, необычной идеи. Также эффективность сообщения основывается на использовании естественных доверительных отношениях между получателем и отправителем.

В рамках решаемых задач для нас наиболее подходят оптимизационные и имитационные модели. Из них рассмотрим модели просачивания и заражения (класс эпидемиологических моделей), так как данные модели наиболее точно отражают специфику рассматриваемых нами проблем. Данный класс моделей является очень распространенным при исследованиях процессов взаимодействия в ИТКС.

1.4 Задачи исследования

После проведения анализа предметной области в рамках данной работы были поставлены следующие задачи исследования:

1. Создать имитационную модель распространения угрозы запрещенной информации в ИТКС.

– разработать алгоритм УгЗИ в ИТКС;

– на основе разработанного алгоритма создать имитационную модель УгЗИ в ИТКС;

– провести экспериментальное исследование имитационной модели УгЗИ в ИТКС.

2. Создать аналитическую модель распространения угрозы запрещенной информации в ИТКС.

– на основе экспериментальных данных по имитационной модели создать аналитическую модель УгЗИ в ИТКС;

– провести экспериментальное исследование аналитической модели, проверить адекватность модели.

3. Разработать методику формирования топологии ИТКС

- алгоритма формирования графа доступной части сети;
- алгоритма формирования полного графа.

4. Смоделировать процесс распространения угрозы запрещенной информации на реальной крупномасштабной ИТКС.

– разработать методику формирования топологии крупномасштабной ИТКС;

– реализовать методику в виде ПО;

– разработать ПО под распределенную вычислительную систему для моделирования УгЗИ на топологии крупномасштабной ИТКС;

– провести экспериментальное исследование имитационной модели УгЗИ на топологии крупномасштабной ИТКС с использованием разработанного ПО;

– провести экспериментальное исследование по полученным результатам.

Выводы к первой главе

ИТКС являются крупномасштабными сетями с постоянно растущим числом абонентов. С бурным ростом числа пользователей ИТКС возникают проблемы информационной безопасности и защиты информации в них.

Анализ проблем информационной безопасности выявил, что кроме проблем, связанных с использованием глобальной сети Интернет как распределенной информационно-телекоммуникационной системы, которые достаточно хорошо известны и решаемы, существует малоизученная проблема запрещенного контента.

Создание моделей и алгоритмов распространения угрозы запрещенной информации – один из ключевых подходов при решении данной задачи. Проведенный анализ публикаций по данной тематике показывает, что существующие решения малоэффективны. Обычно при моделировании распространения угрозы запрещенной информации не учитывается топология ИТКС (модель сети – полносвязный граф). А, если топология учитывается, то, как правило, используется простейшая SIS модель, а структура сети отражается SF сетью. При моделировании УгЗИ важно иметь топологию, отражающую структуру связей реальной сети, а также использовать адекватную модель информационного взаимодействия узлов. Еще одной важной проблемой является крупномасштабность ИТКС, которая мешает получить данные с имитационной модели за приемлемое время. Решение этой задачи состоит в создании аналитической модели УгЗИ в ИТКС.

ГЛАВА 2 РАЗРАБОТКА И ИССЛЕДОВАНИЕ МОДЕЛЕЙ УГРОЗЫ РАСПРОСТРАНЕНИЯ ЗАПРЕЩЕННОЙ ИНФОРМАЦИИ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ

По результатам обзора предметной области были поставлены задачи создания имитационной и аналитической моделей распространения угрозы запрещенной информации в ИТКС. Имитационная модель необходима для получения экспериментальных результатов для синтезирования аналитической модели. Необходимость создания аналитической модели обосновывается тем, что для имитационного моделирования на топологии существующих ИТКС (десятки миллионов узлов) необходимы большие временные затраты. Не учитывая время на сбор информации о топологии сети, которое может составлять порядка недели, непосредственно моделирование УгЗИ занимает несколько часов даже при использовании распределенных вычислительных ресурсов. Аналитическая модель может дать прогноз УгЗИ почти мгновенно. С ее помощью можно получить актуальные данные (до того момента, когда количество атакующих абонентов будет максимальным) по динамике УгЗИ.

Процесс УгЗИ характеризуется следующими особенностями [1,7,10,12,13,14,15]. В сети существуют узлы трех типов. Первый тип – атакующие узлы, это узлы, распространяющие запрещенную информацию. Второй тип – защищенные узлы, характеризующиеся тем, что не принимают участие в распространении запрещенной информации и никогда не будут этим заниматься. Третий тип – потенциально уязвимые. Узлы такого типа не участвуют в процессе распространения угрозы, но могут быть подвержены негативному влиянию со стороны атакующих узлов и могут начать распространять запрещенную информацию.

Постановка задачи

Дано: N – количество узлов, равное числу абонентов сети, I_0 – количество абонентов-злоумышленников – изначальных источников угрозы, R_0 – количество

абонентов изначально невосприимчивых к атакующим воздействиям, β – параметр, отражающий силу угрозы, вероятность осуществления атаки, γ – параметр отражающий степень противодействия угрозе, вероятность защиты абонента (β и γ в данном исследовании определены как константы, но могут быть выражены как функции, зависящие от психосемантических профилей абонентов ИТКС [33-35]), φ – коэффициент топологической уязвимости сети, отражающий внутреннее свойство ИТКС, основанное на характеристиках ее топологии, которое способствует распространению запрещенной информации (описан в главе 3.3), t – время процесса (в условных единицах времени).

Требуется разработать аналитическую модель динамики атаки $I(t)$ и защиты узлов $R(t)$

$$\begin{cases} I(t) = f(N, \beta, \gamma, \varphi, t) \\ R(t) = g(N, \beta, \gamma, \varphi, t) \end{cases}$$

Методика разработки аналитической модели включает в себя последовательность следующих действий:

1) формирование имитационной модели для исследования характера и параметров процесса УгЗИ;
2) синтез аналитических зависимостей параметров процесса;
3) проведение экспериментов с целью проверки точности (адекватности) модели.

2.1 Имитационное моделирование

Приведем алгоритм реализации УгЗИ, основываясь на описании процессов, протекающих в реальных ИТКС. Схема реализации угрозы представлена на рисунке 2.1.

Алгоритм 2.1- Алгоритм УгЗИ в ИТКС

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

Шаг 2. Абоненты-получатели (узлы 2,3,4), приняв сообщение с ЗИ, читают его и включаются в процесс атаки, распространяя ее дальше по своему списку контактов (узел 3), либо игнорируют или вообще удаляют сообщение (узел 2), 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 могут перестать воспринимать и, соответственно, распространять ЗИ (узел 5) (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.

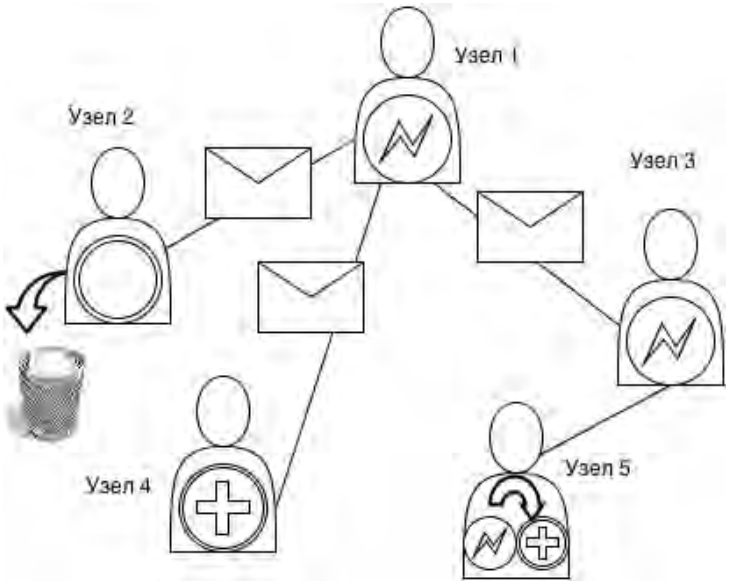


Рисунок 2.1 - Схема реализации УгЗИ

Таким образом, УгЗИ в ИТКС представляет собой сложный динамический процесс, состоящий из двух противоборствующих подпроцессов атаки и защиты узлов сети.

На основе описанного алгоритма была построена имитационная модель УгЗИ в ИТКС, которая состоит из разработанной программы ModelGraph и данных, которые могут быть сгенерированы с помощью ПО Pajek [36].

Имитационная модель УгЗИ

Входные данные: N, k - средняя степень связности узлов, α - параметр, отражающий среднюю длину пути и уровень сетевой кластеризации, β, γ (в модели считается, что β и γ одинаковы для каждого абонента), I_0, R_0.

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, α.

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 β выполнить: $V^S \backslash v_j$ и $V^I \cup v_j$; с вероятностью γ выполнить: $V^I \backslash v_i$, $V^R \cup v_i$.

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

ModelGraph - программа для имитационного моделирования УгЗИ в ИТКС [8]. Данный программный продукт является однопоточным приложением. Программа состоит из исполняемого файла ModelGraph.exe и библиотеки chartdir50.dll для построения графиков. После выбора типа сети и ввода ее

параметров происходит имитационное моделирование по алгоритму 2.1. Затем результаты отправляются в функцию построения графиков для вывода результатов в графическом виде. Программа написана в среде разработки Microsoft Visual Studio .NET 2008. Исходными данными для гетерогенной сети является файл формата .net, определенный в программе Pajek. Реализация ПО подтверждается свидетельством о государственной регистрации программ (Приложение А).

ПО Pajek представляет собой программу, для ОС MS Windows, предназначенную для анализа и визуализации больших сетей. Данная программа находится в свободном доступе и предназначена для некоммерческого использования. Pajek разработан Vladimir Batagelj и Andrej Mrvar.

Проанализируем подпроцесс атаки без защиты, проведя ряд экспериментов (эксперимент 1-3) с использованием имитационной модели (φ – коэффициент топологической уязвимости сети, см. главу 3.3).

1) Эксперимент 1. Влияние силы атаки на процесс.

Эксперименты проводились при следующих значениях параметров: N=1000, φ=20, I_0=1, β=0,1..0,9 (рисунок 2.2).

2) Эксперимент 2. Влияние значения средней степени связности узлов в сети на процесс.

Эксперименты проводились при следующих значениях параметров: N=1000, φ=0,5..60, I_0=1, β=0,5 (рисунок 2.3).

3) Эксперимент 3. Влияние количества изначально атакующих узлов на процесс.

Эксперименты проводились при следующих значениях параметров: N=1000, φ=20, I_0=1..40, β=0,5 (рисунок 2.4).

Каждый из трех типов экспериментов проводился 100 раз, брались усредненные значения.

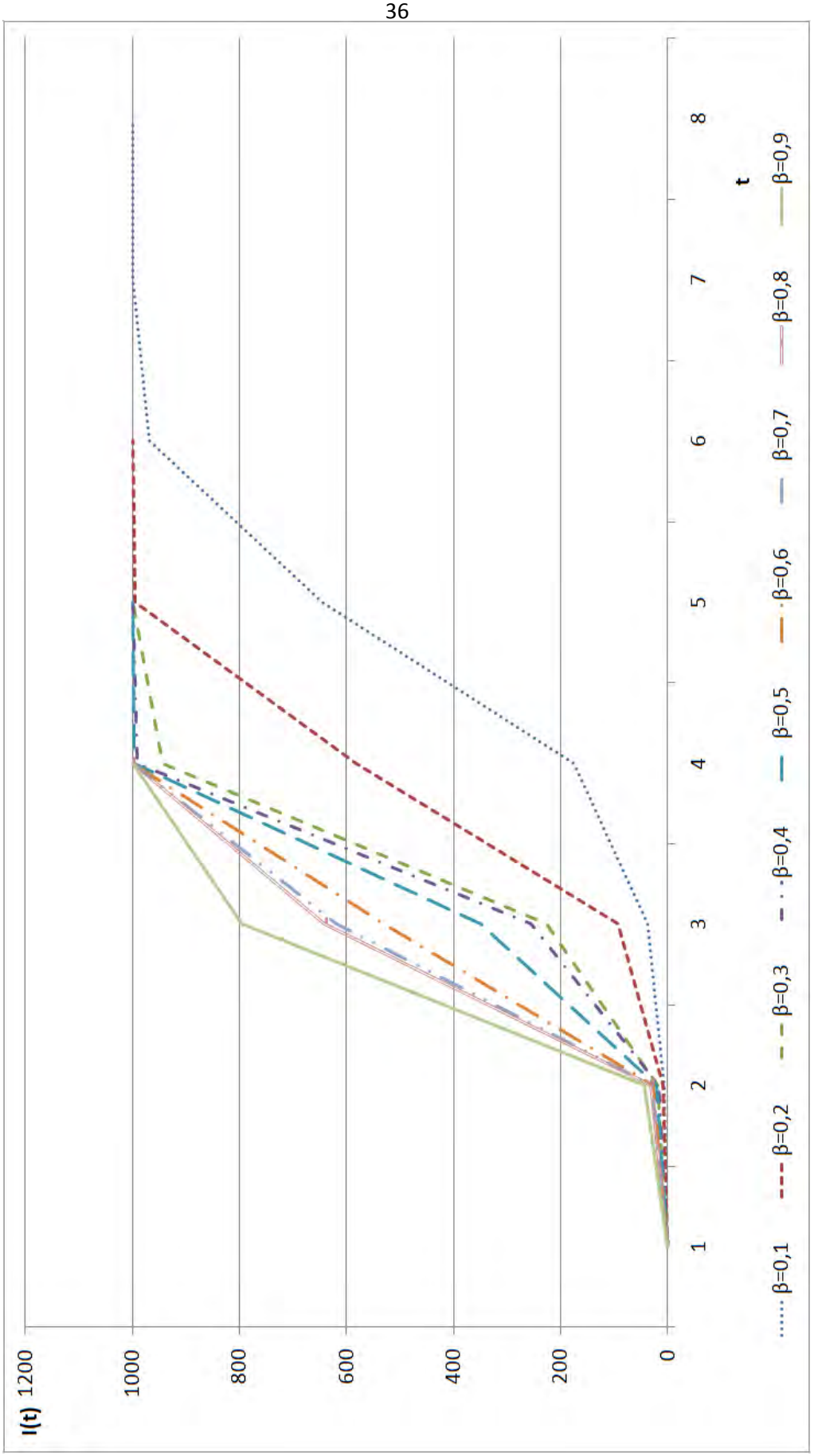


Рисунок 2.2 – Влияние β на процесс атаки

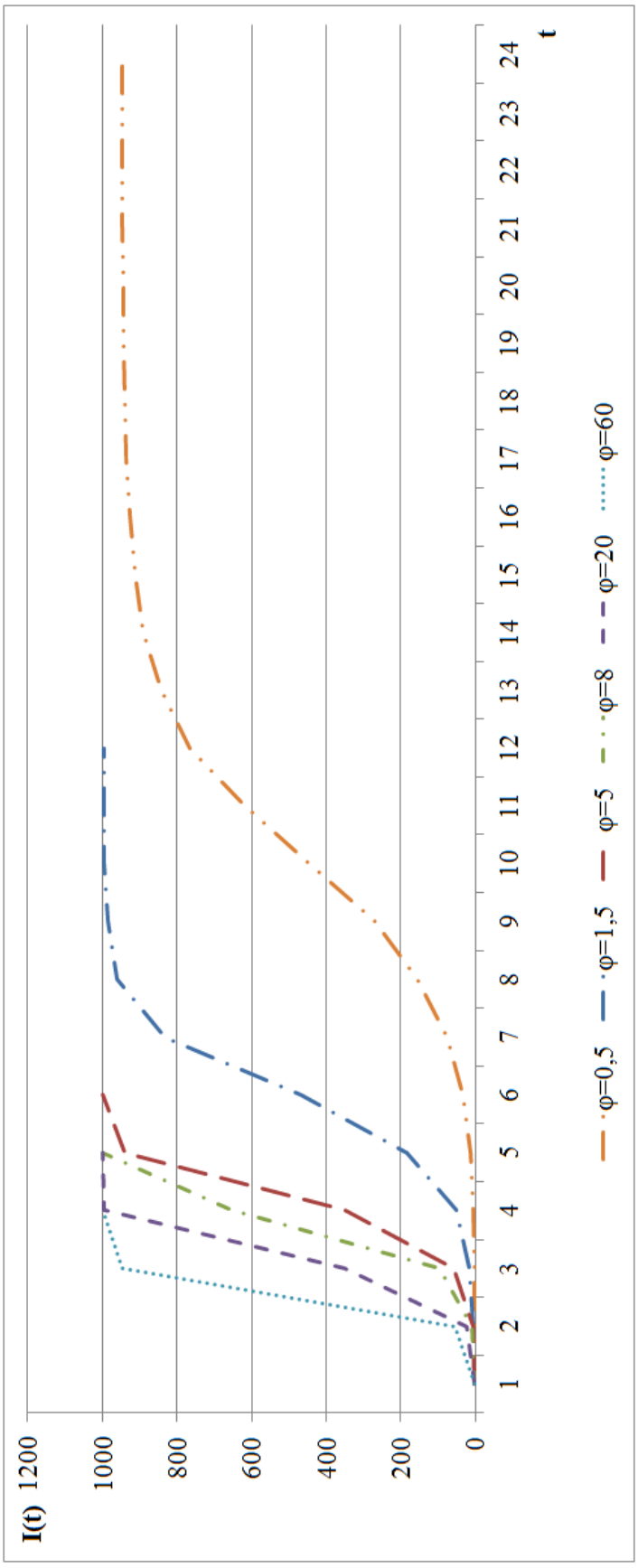


Рисунок 2.3 – Влияние φ на процесс атаки

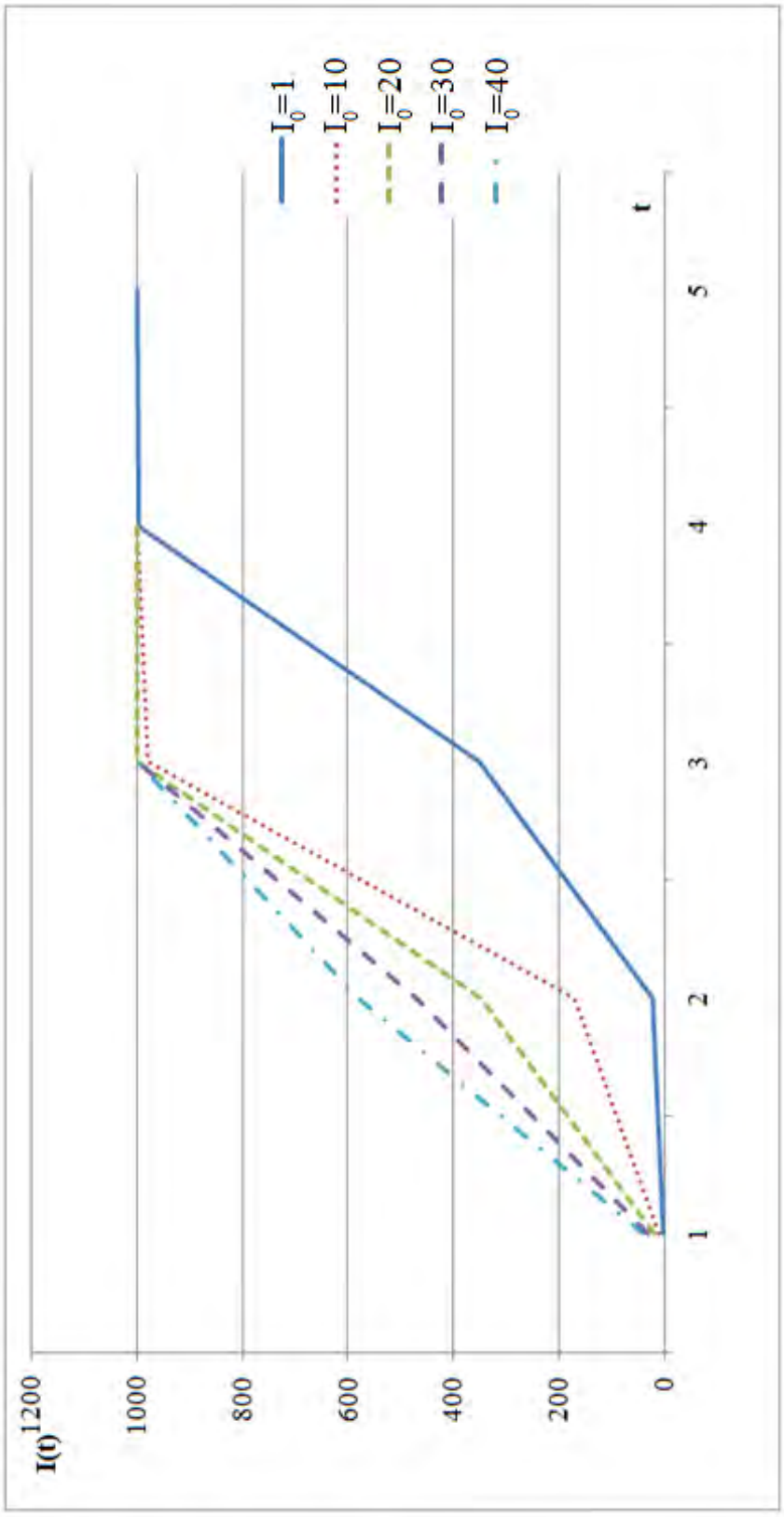


Рисунок 2.4 – Влияние I_0 на процесс атаки

По результатам экспериментов 1-3 можно сделать следующие выводы:

- процесс атаки $I(t)$ имеет экспоненциальную зависимость (эксперимент 1,2,3);
- при увеличении значений φ, I_0, β возрастает динамика заражения узлов (интенсивность атаки) (эксперимент 1,2,3);
- при росте вероятности проведения атаки β от 0,1 до 0,9, время процесса снижается в два раза (с 8 до 4 условных единиц времени) (эксперимент 1);
- коэффициент топологической уязвимости φ имеет самое большое влияние (в сравнении с I_0, β) на длительность процесса. Например, при φ = 0,5 (низкая уязвимость) атака длится 24 условные единицы времени, а при φ = 60 всего лишь 4 (эксперимент 2);
- большое количество изначально атакующих узлов I_0 снижает время, за которое происходит заражение всех узлов в сети. Например, при I_0=40 длительность процесса составляет 3 условные единицы времени (эксперимент 3).

Усложним условия экспериментов, добавив подпроцесс защиты, который зависит от начального количества защищенных узлов R_0 и вероятности защиты γ.

4) Эксперимент 4. Влияние вероятности защиты.

Эксперименты проводились при следующих значениях параметров: N=1000, φ = 20, I_0=1, β=0,5, γ =0,1..0,9, R_0 =0. (рисунок 2.5).

5) Эксперимент 5. Влияние начального количества защищенных узлов.

Эксперименты проводились при следующих значениях параметров: N=1000, φ = 20, I_0=1, β=0,5, γ =0,5, R_0 =0..200. (рисунок 2.6).

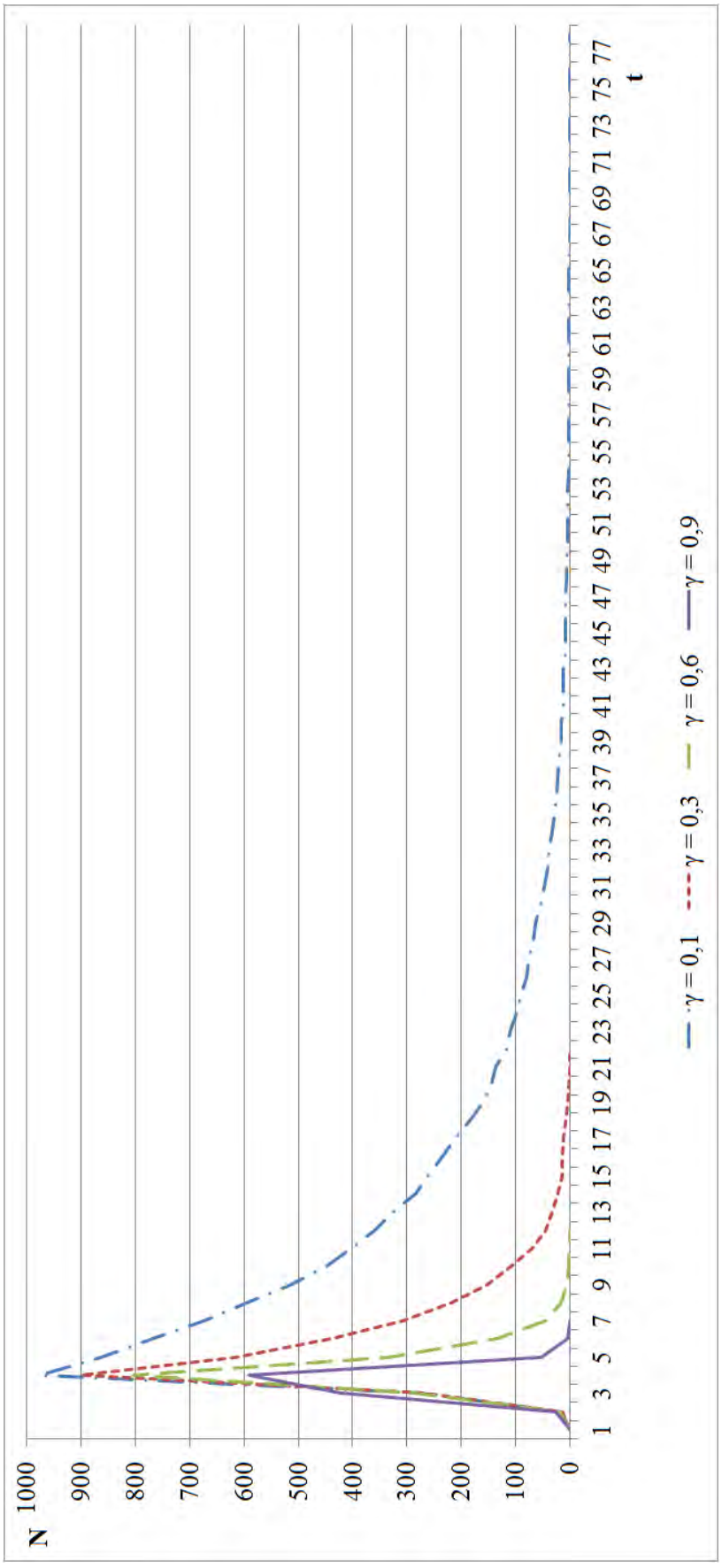


Рисунок 2.5 – Влияние γ на процесс атаки

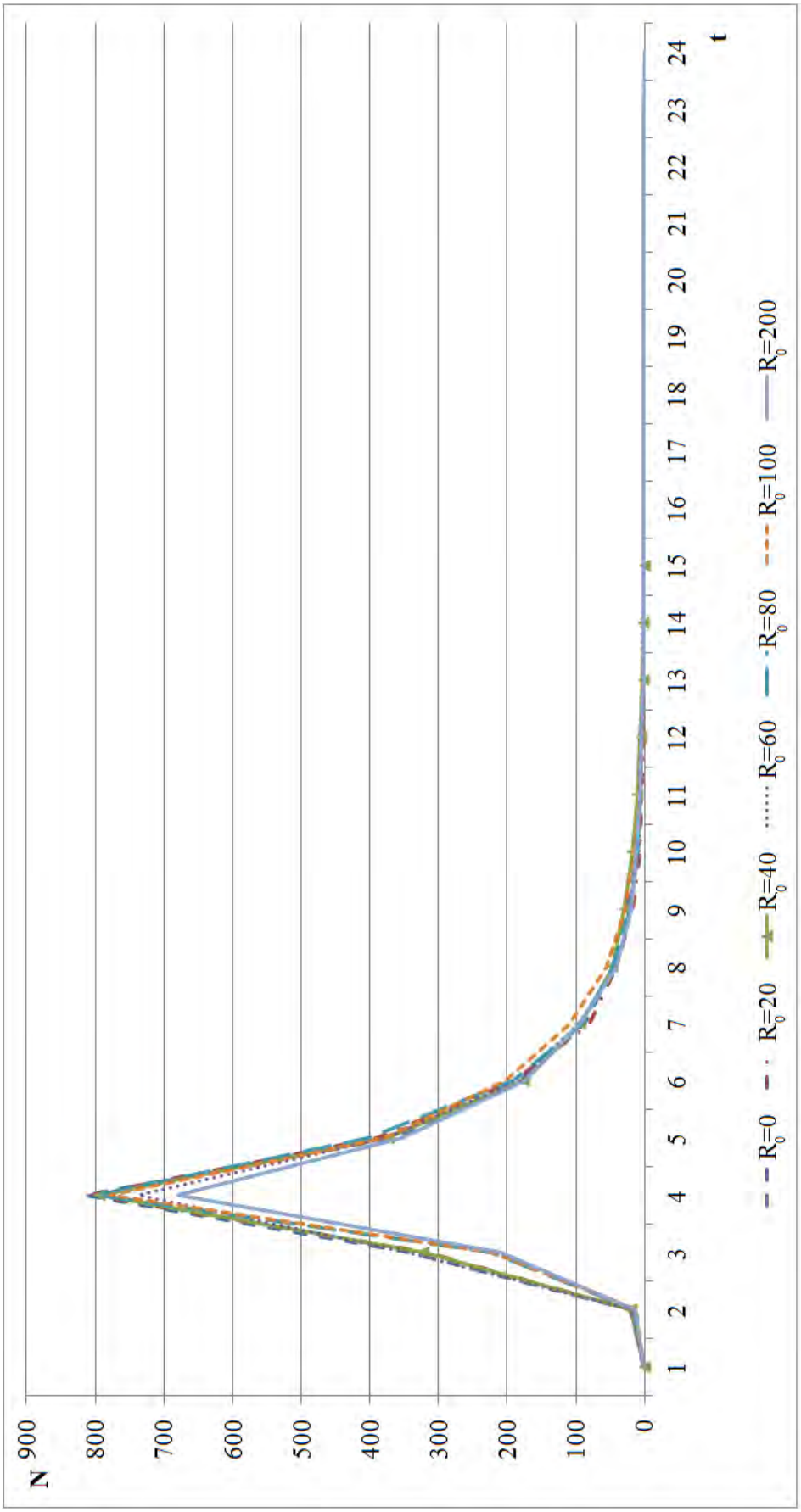


Рисунок 2.6 – Влияние R_0 на процесс атаки

По результатам экспериментов 4 и 5 можно сделать следующие выводы:

- введение подпроцесса защиты увеличивает время всего процесса УгЗИ (эксперимент 4,5);
- при небольших значениях вероятности защиты (γ<0,3) угроза реализуется практически на всех узлах в сети (эксперимент 4);
- при небольших значениях вероятности защиты (γ<0,3) время процесса составляет более 50 условных единиц времени (эксперимент 4);
- при большой вероятности защиты (≈0,9) процесс длится ≈7 условных единиц времени и максимальное количество атакующих узлов снижается в зависимости от вероятности проведения атаки (эксперимент 4);
- при случайном выборе изначально защищенных узлов картина процесса атаки практически не изменяется (эксперимент 5);
- при высокой топологической уязвимости возрастает длительность процесса УгЗИ (эксперимент 5).

2.2 Разработка аналитической модели

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 [46,87,91,96,103 и др.], в частности нашему описанию точно соответствует SIR-модель Кермака-Маккендрика [67, 126, 81, 127 и др.]. Характер графиков, полученных в результате имитационного моделирования (рисунок 2.7), схож с результатами, которая дает данная модель.

Исходя из вышесказанного, приходим к выводу, что данная модель является наиболее релевантной для настоящего исследования.

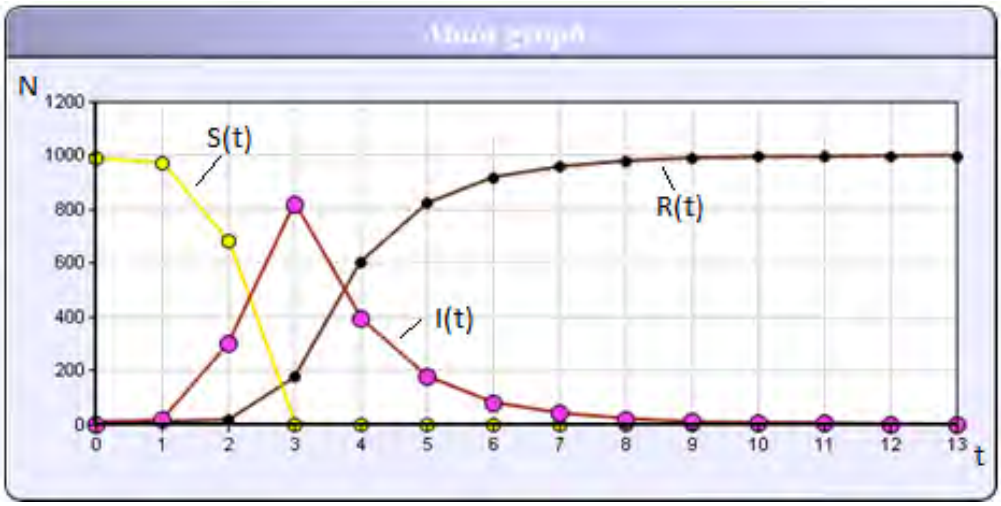


Рисунок 2.7 – Имитационное моделирование (N=1000, φ = 20, I_0=1, β=0,5, γ =0,5, R_0 =10), $S(t)$ – количество подверженных атаке узлов

SIR (от англ. Susceptibles – Infectives - Removed with immunity) – эпидемиологическая модель, упрощенно описывающая распространение заболевания, передающегося от одного индивида к другому, которая рассматривает субъектов с точки зрения трех возможных состояний: восприимчивый, инфицированный, иммунизированный.

Система дифференциальных уравнений, описывающих SIR-модель, имеет вид [67,81]:

$$\begin{cases} \dfrac{dI}{dt} = \beta \cdot \dfrac{S(t) \cdot I(t)}{N} - \gamma \cdot I(t) \\ \dfrac{dR}{dt} = \gamma \cdot I(t) \\ \dfrac{dS}{dt} = -\beta \cdot \dfrac{S(t) \cdot I(t)}{N} \end{cases}, \qquad (2.1)$$

где $I(t)$ – количество зараженных (инфицированных) особей, $S(t)$ – количество восприимчивых особей, $R(t)$ - количество «исключенных с иммунизацией» (removed with immunity) особей, $N=I(t)+S(t)+R(t)$ – количество особей в популяции, γ – коэффициент восстановления/смерти, β – скорость заражения (инфицирования), t – время. Данная система является избыточной – любое уравнение из трех уравнений можно исключить.

При использовании системы 2.1 для анализа УгЗИ в ИТКС получаем результаты в виде графиков (рисунок 2.8), которые хотя и правильно описывают характер процесса, но не дают нужной точности прогноза.

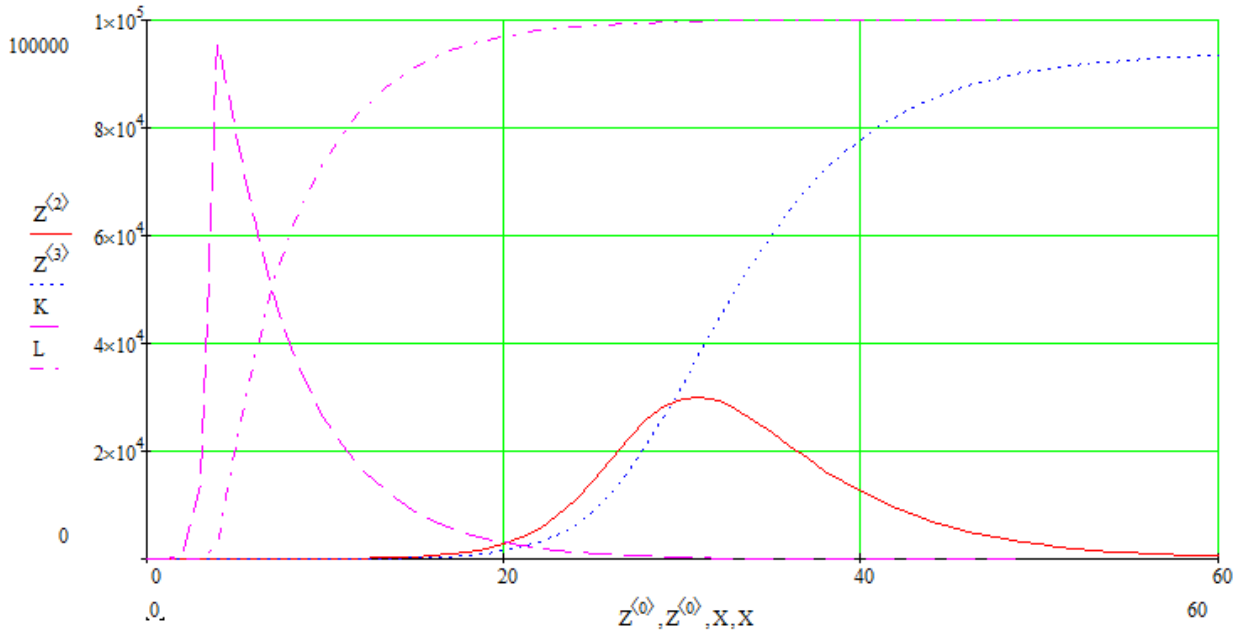


Рисунок 2.8 – Результаты имитационного моделирования (N=100000, φ = 150, I_0=1, β=0,3, γ =0,2, R_0 =0) и аналитического решения ($Z^{<2>}$, $Z^{<3>}$ - аналитическое решение для процессов атаки и защиты соответственно, K, L – результаты имитационного моделирования для процессов атаки и защиты соответственно)

Была выдвинута гипотеза о том, что система 2.1 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 2.1 под прогнозирование УгЗИ в ИТКС путем интегрирования в нее параметра топологической уязвимости сети φ.

Проанализировав графики, полученные по результатам имитационного моделирования и аналитического решения системы 2.1, и проследив физический смысл уравнений в данной системе [67, 81, 127], можно прийти к следующему выводу. Процесс защиты не зависит от топологии сети, поэтому «изменять» $R(t)$ не имеем права. А вот процесс атаки зависит от структуры связей между абонентами в сети. Параметр топологической уязвимости φ может влиять на $I(t)$ через коэффициент β. В общем виде адаптированную систему 2.1 можно представить в следующем виде:

$$\begin{cases} \dfrac{dI}{dt} = C \cdot \beta \cdot \dfrac{S(t) \cdot I(t)}{N} - \gamma \cdot I(t) \\ \dfrac{dR}{dt} = \gamma \cdot I(t) \\ \dfrac{dS}{dt} = -C \cdot \beta \cdot \dfrac{S(t) \cdot I(t)}{N} \end{cases}, \qquad (2.2)$$

где C – коэффициент, зависящий от параметра φ.

Отметим, что в [127] уже предлагался аналогичный подход, при этом отмечалось, что коэффициент C может быть выражен функцией или аппроксимирован константой.

Анализ топологий крупномасштабных ИТКС показал, что типичные значения параметра φ для них находятся в диапазоне от 100 до 600 (см. главу 3.3). Результаты серии экспериментов по имитационному моделированию УгЗИ в ИТКС (рисунки 2.9, 2.10) позволили получить зависимость параметра C от φ в виде 2*ln φ. Аппроксимация проводилась методом наименьших квадратов с использованием пакета MathCAD.

Итоговая система имеет вид:

$$\begin{cases} \dfrac{dI}{dt} = 2 \cdot \ln\varphi \cdot \beta \cdot \dfrac{S(t) \cdot I(t)}{N} - \gamma \cdot I(t) \\ \dfrac{dR}{dt} = \gamma \cdot I(t) \\ \dfrac{dS}{dt} = -2 \cdot \ln\varphi \cdot \beta \cdot \dfrac{S(t) \cdot I(t)}{N} \end{cases}, \qquad (2.3)$$

Система дифференциальных уравнений 2.3 позволяет получить прогноз УгЗИ в крупномасштабной ИТКС (N=10^5..10^8) с погрешностью до 20%.

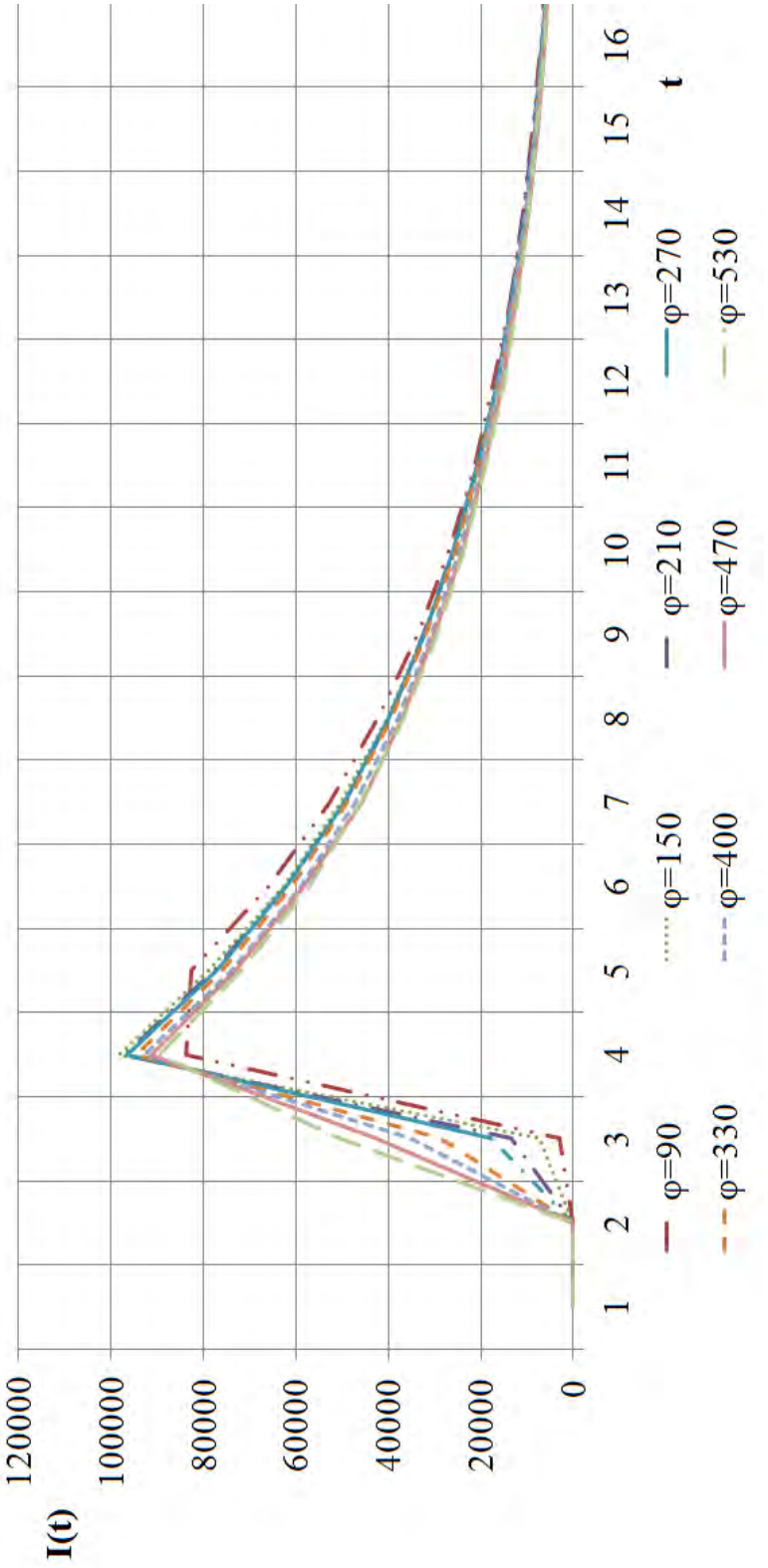


Рисунок 2.9 – Результаты имитационного моделирования (N=10^5, I_0=1, β=0,3, γ =0,2, R_0 =0)

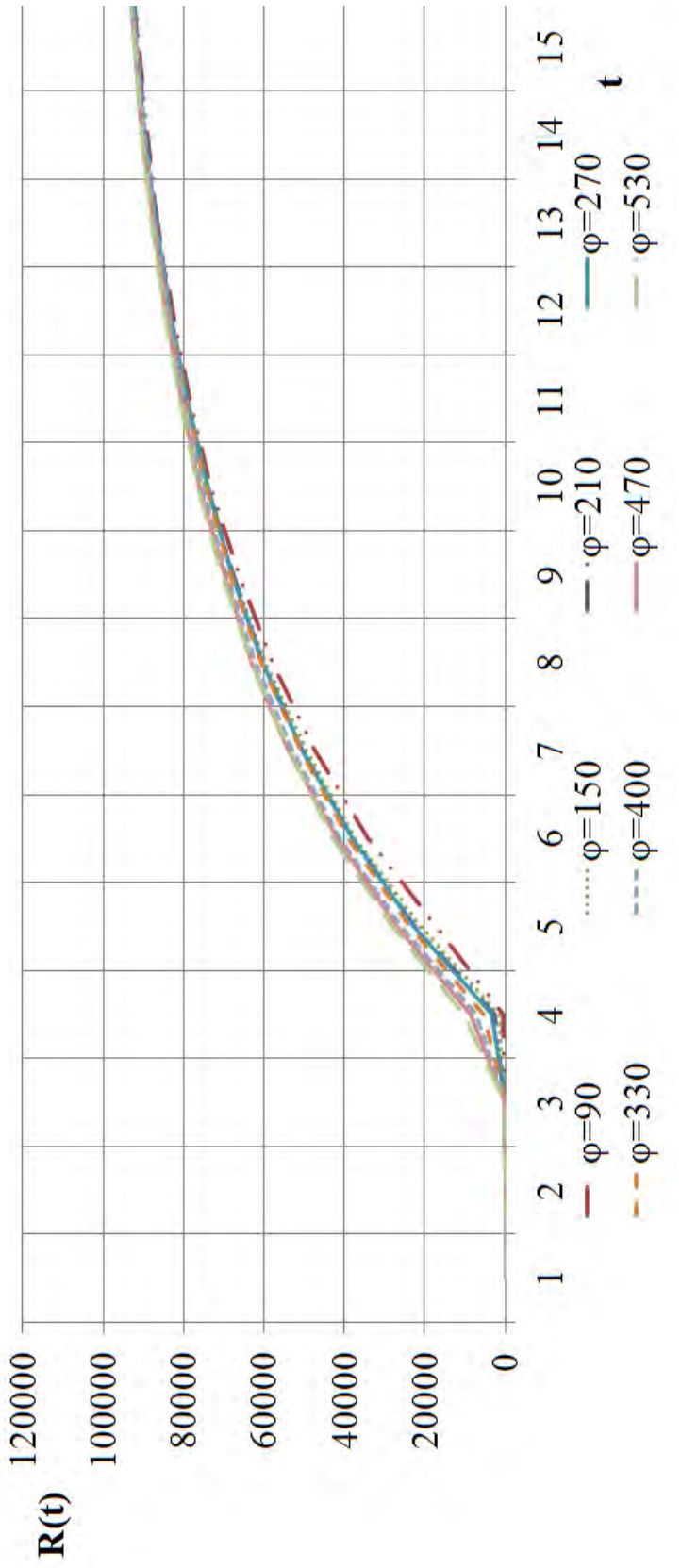


Рисунок 2.10 – Результаты имитационного моделирования (N=10^5, I_0=1, β=0,3, γ =0,2, R_0 =0)

2.3 Экспериментальное исследование аналитической модели

Результаты аналитической модели сравнивались с результатами имитационного моделирования процесса УгЗИ на топологии реальной сети. Эти данные были получены в результате работы, описанной в третьей главе («ВКонтакте»).

Эксперимент 1. На рисунке 2.13 приведены результаты имитационного моделирования и аналитического решения для β=0,5, γ=0,51, R_0=0, I_0=1.

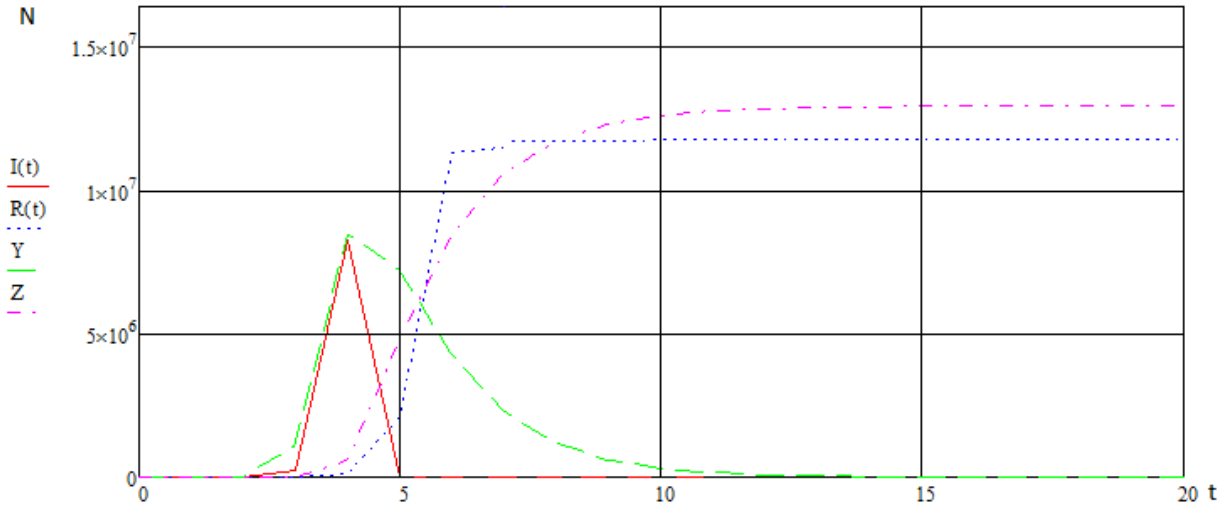


Рисунок 2.11 – Результаты аналитической и имитационной модели (I и R – аналитическое решение, Y и Z – результаты имитационной модели). β=0,5, γ=0,51, R_0=0, I_0=1.

Эксперимент 2. На рисунке 2.14 приведены результаты имитационного моделирования и аналитического решения для β=0,5, γ=0,51, R_0=0, I_0≈24000.

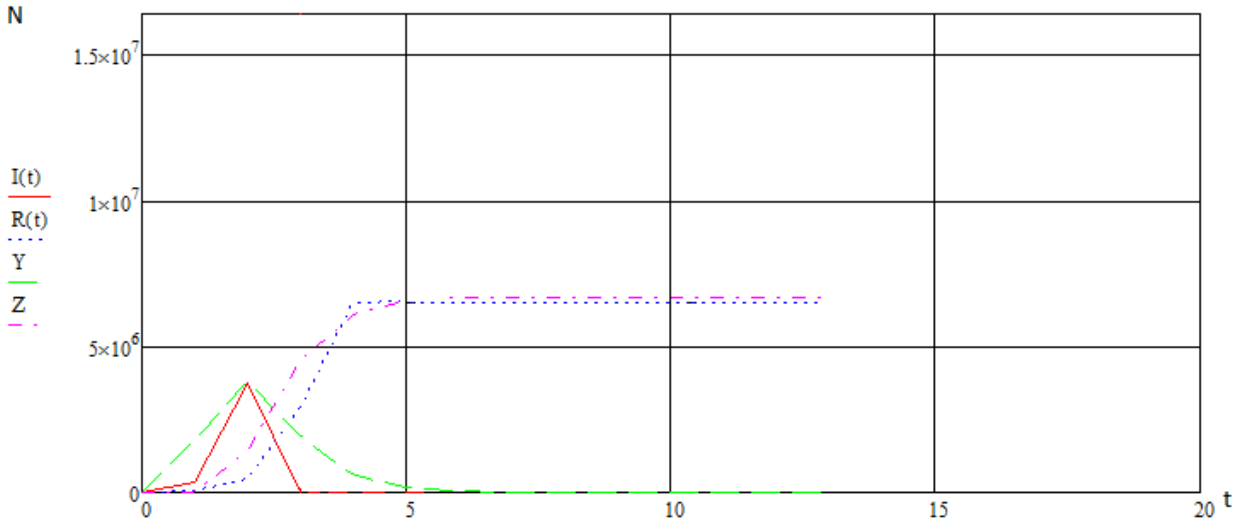


Рисунок 2.12 – Результаты аналитической и имитационной модели (I и R – аналитическое решение, Y и Z – результаты имитационной модели). β=0,5, γ=0,51, R_0=0, I_0≈24000.

Эксперимент 3. На рисунке 2.15 приведены результаты имитационного моделирования и аналитического решения для β=0,5, γ=0,51, $R_0 \approx 4*10^6$, I_0=1.

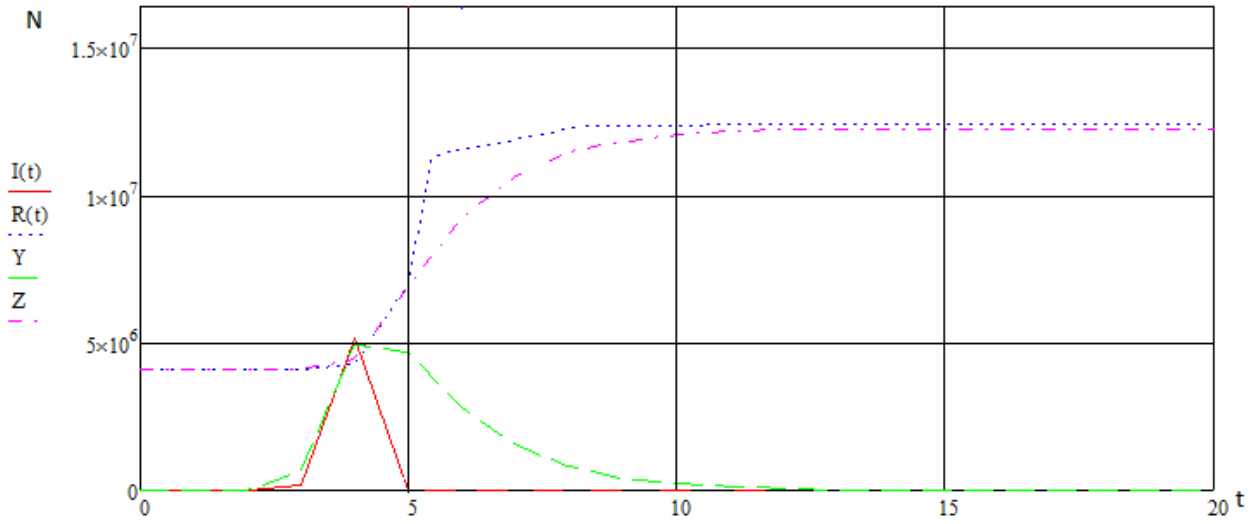


Рисунок 2.13 – Результаты аналитической и имитационной модели (I и R – аналитическое решение, Y и Z – результаты имитационной модели). β=0,5, γ=0,51, $R_0 \approx 4*10^6$, I_0=1.

По результатам экспериментов можно сделать следующие выводы:

- результаты аналитического решения подходят для аппроксимации имитационных результатов, при этом погрешность аппроксимации для процесса защиты $R(t)$ составляет не более 10%, для процесса атаки $I(t)$ - не более 15% (эксперимент 1,2,3);
- при средних значениях силы атаки и защиты ($\beta,\gamma \in [0,3;0,7]$) погрешность остается в том же диапазоне ($\Delta_{R(t)}$<10%, $\Delta_{I(t)}$<15%), при сильной атаки и слабой защите и наоборот – может составлять порядка 20%.
- при моделировании с большим количеством изначально атакующих узлов (I_0>>1) погрешность составляет: $\Delta_{R(t)}$<10%, $\Delta_{I(t)}$<15% (эксперимент 2);
- при добавлении в сеть большого количества изначально защищенных узлов ($\approx 4*10^6$) аналитическое решение также дает результат с погрешностью $\Delta_{R(t)}$<10%, $\Delta_{I(t)}$<15% (эксперимент 3);
- сравнивая данные результаты с результатом применения исходной системы дифференциальных уравнений 2.1, можно говорить о значительном увеличении точности прогнозирования процесса УгЗИ в ИТКС за счет учета влияния на процесс топологической уязвимости сети.

Выводы ко второй главе

Разработан алгоритм реализации УгЗИ в ИТКС, основанный на характерах процессов, протекающих в реальных условиях.

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.

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%.

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.

ГЛАВА 3 РАЗРАБОТКА МЕТОДИКИ ФОРМИРОВАНИЯ ТОПОЛОГИИ КРУПНОМАСШТАБНОЙ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

Под топологией будем понимать структуру информационных связей между узлами сети. Топологические характеристики (средняя степень связности узлов, распределение степеней связности узлов, кластерный коэффициент сети, средняя длина пути сети) в работе рассматриваются как основные технические уязвимости ИТКС к реализациям угроз. Другие уязвимости: использование нелицензионного ПО в узлах, некорректно настроенные межсетевые экраны и тд., отраженные в различных трудах [16, 19, 21, 31, 51, 52 и др.] в данной работе не рассматриваются.

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 [69-74, 92, 131], что закрытых узлов на порядок больше, чем открытых, поэтому при сборе данных теряется значительная часть данных.

Особенности практической реализации:

1) частота запросов абонента о связях узла ограничена администраторскими мерами (например, для сети «ВКонтакте» это значение приблизительно

составляет 10 запросов в секунду). Это ограничение приводит к тому, что, учитывая масштабность ИТКС (десятки миллионов узлов), получение информации о топологии сети превращается в длительный процесс (например, для сети «ВКонтакте» получение информации о $16*10^6$ узлов заняло около 20 суток). Учитывая, что время сессии ограничено (например, для сети «ВКонтакте» это значение равно одним суткам), данная особенность должна учитываться при практической реализации.

2) известные средства (например, Tictrac [130]) для решения задачи сбора информации о связях узлов в ИТКС не эффективны, так как напрямую не предназначены для достижения этой цели и имеют множество недостатков.

3) топология реальной ИТКС постоянно изменяется (абоненты регистрируются, добавляют связи, удаляют связи и учетные записи). В связи с этим, необходимо постоянно получать актуальную информацию о ИТКС для более точного моделирования УгЗИ.

Топология сети представляется графом $G=\{V,E\}$, где V (множество вершин графа) – множество узлов-абонентов, а E (множество ребер) - информационные связи между узлами.

Будем считать, что граф является неориентированным, то есть все связи – двунаправленные. Любые две вершины графа могут быть связаны не более чем одним ребром. Для упрощения исследований граф считается не взвешенным, т.е. сила информационных связей [77,62,65] не отображается на веса соответствующих ребер. Узел представляет собой человеко-машинную систему, на одном компьютере не может находиться несколько узлов.

В предлагаемой модели узел $v_i = \{id_i, flag_i\}$ хранит уникальный идентификатор абонента сети (id) и флаг (flag). Переменная *flag* определяет статус узла: открытый (*flag*=1) или закрытый (*flag*=0).

В главе разрабатывается методика формирования топологии ИТКС, которая состоит из последовательности шагов [6]:

- сбор данных о топологии доступной части сети;

- формирование полного графа сети с учетом добавления недоступной части на основе вычисленных прогнозируемых топологических характеристик (распределение степеней связности, средняя длина пути).
- формирование вектора топологической уязвимости узлов ИТКС.

3.1 Сбор данных о топологии доступной части сети

Введем определения.

Определение 1. Граф доступной части сети – граф, содержащий открытые и закрытые узлы и связи между ними.

Определение 2. Полный граф сети – граф, содержащий открытые узлы и закрытые узлы, перешедшие в состояние открытых, и связи между ними.

Определение 3. Соседние узлы (смежные узлы) – узлы, имеющие связи с данным узлом.

Постановка задачи: требуется составить граф доступной части сети $G(V, E)$, где

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).

Блок-схема алгоритма формирования графа доступной части сети представлена на рисунке 3.1.

Переменные, используемые в алгоритме:

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.

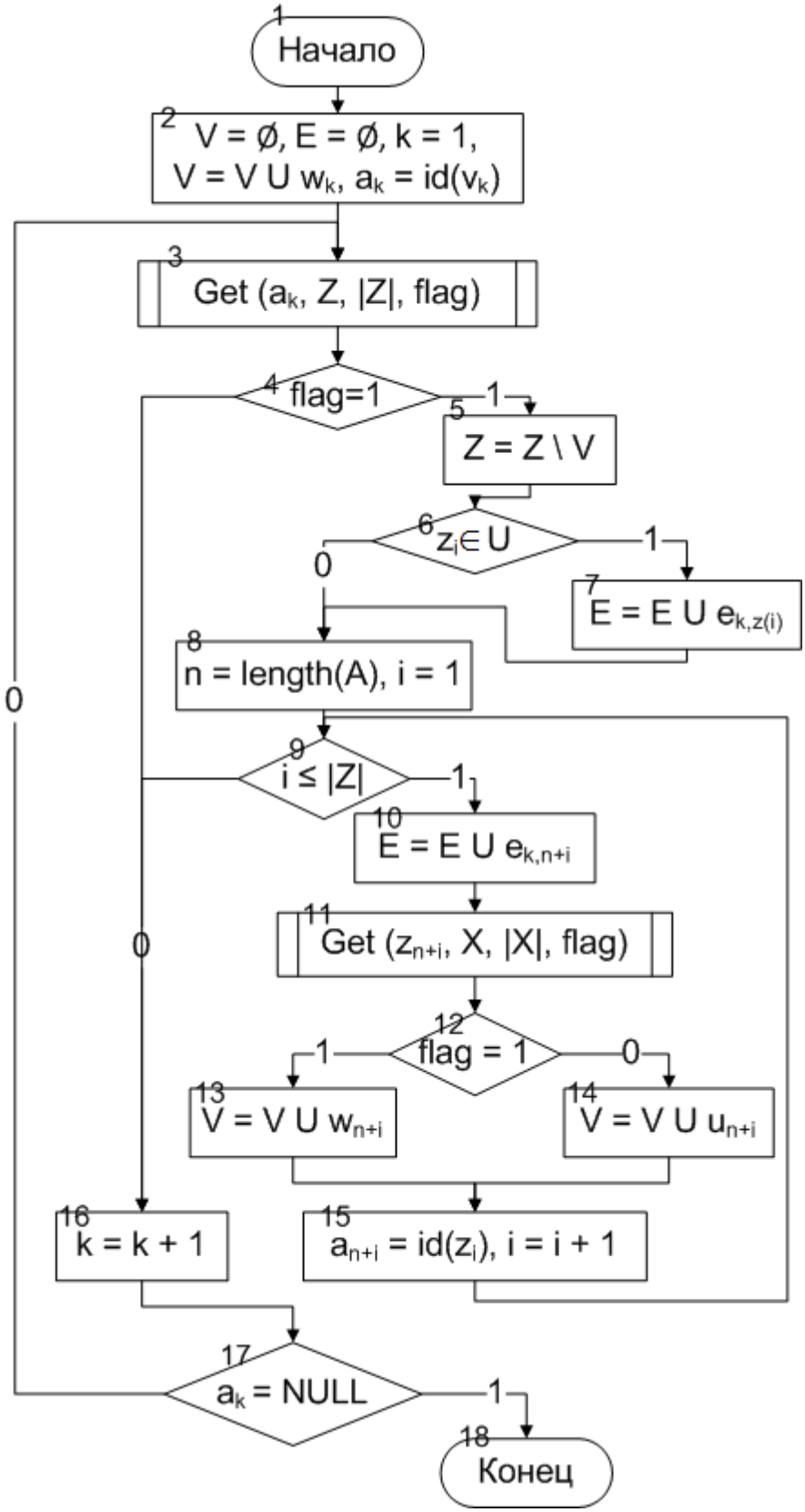


Рисунок 3.1 – Блок-схема алгоритма формирования графа доступной части сети

Алгоритм 3.1 - Алгоритм формирования графа доступной части сети

Шаг 1 (блок 2). 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 (блоки 3,4). 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 (блок 5-7). 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 (блоки 8-15). 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 (блоки 16,17). Перейти к следующему узлу $k = k + 1$. Если $a_k = NULL$, то конец алгоритма, иначе перейти к шагу 2.

Рассмотрим пример поэтапной реализации алгоритма 3.1.

Этап 1. Выполняем начальные установки согласно первому шагу алгоритма: k=1, $V=\{w_1\}$, A[12].

Этап 2. Выполняем функцию *Get*(12, Z, $|Z|$, *flag*). Получаем Z={43, 36, 39, 78}, $|Z|$=4, *flag*=1. Переходим к третьему шагу алгоритма.

Этап 3. Проверяем множество Z на наличие узлов, уже добавленных в множество V, и при наличии таковых, удаляем их. Получаем Z={43, 36, 39, 78}, $|Z|$=4.

Этап 4. Определяем длину массива A (n=1). Добавляем ребра, связывающие первую вершину с узлами из множества Z. Получаем $E=\{e_{1,2}, e_{1,3}, e_{1,4}, e_{1,5}\}$.

Выполняем функцию *Get* для всех узлов из множества Z и добавляем их в соответствующие подмножества множества V. Получаем $W=\{w_1, w_2\}$, $U=\{u_3, u_4, u_5\}$. Записываем идентификаторы узлов в массив A. Получаем A[12, 43, 36, 39, 78].

Этап 5. Увеличиваем счетчик k=1+1=2. Второй элемент массива A (a_2) существует, значит, переходим ко второму шагу алгоритма.

После выполнения первых пяти этапов получаем граф, представленный на рисунке 3.2, на котором закрытые узлы выделены серым цветом, а открытые - белым.

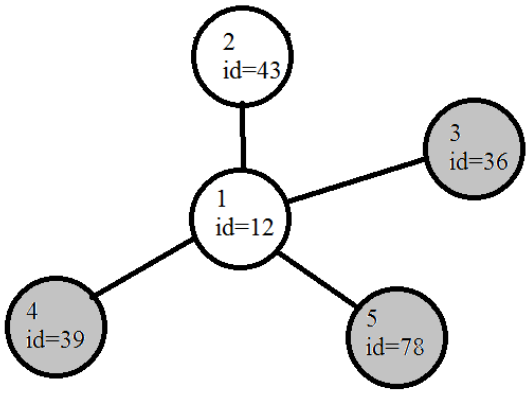


Рисунок 3.2 – Результат работы алгоритма (1-5 этапы)

Этап 6. Выполняем функцию G*e*t(43, Z, $|Z|$, *flag*). Получаем Z={12, 16, 25, 4}, $|Z|$=4, *flag*=1. Переходим к третьему шагу алгоритма.

Этап 7. Проверяем множество Z на наличие узлов, уже добавленных в множество V, и при наличии таковых, удаляем их. Получаем Z={16, 25, 4}, $|Z|$=3.

Этап 8. Определяем длину массива A (n=5). Добавляем ребра, связывающие вторую вершину с узлами из множества Z. Получаем $E=\{e_{1,2}, e_{1,3}, e_{1,4}, e_{1,5}, e_{2,6}, e_{2,7}, e_{2,8}\}$. Выполняем функцию *Get* для всех узлов из множества Z и добавляем их в соответствующие подмножества множества V. Получаем $W=\{w_1, w_2, w_8\}$, $U=\{u_3, u_4, u_5, u_6, u_7\}$. Записываем идентификаторы узлов в массив A. Получаем A[12, 43, 36, 39, 78, 16, 25, 4].

Этап 9. Увеличиваем счетчик k=2+1=3. Третий элемент массива A (a_3) существует, значит, переходим ко второму шагу алгоритма.

После выполнения этапов 6-9 получаем граф, представленный на рисунке 3.3.

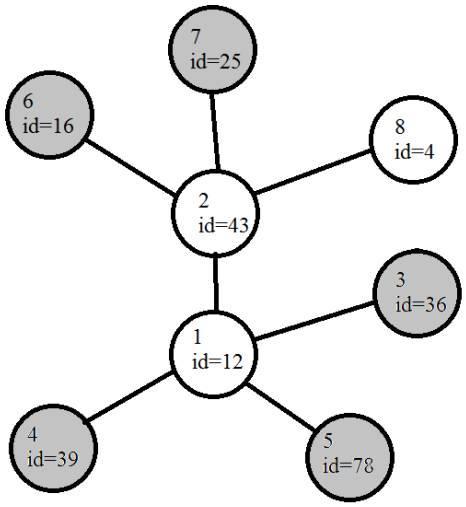


Рисунок 3.3 – Результат работы алгоритма (1-9 этапы)

Этап 10. Выполняем функцию *Get*(36, *Z*, |*Z*|, *flag*). Получаем *Z*= ∅, |*Z*|=0, *flag*=0. Переходим к пятому шагу алгоритма.

Этап 11. Увеличиваем счетчик *k*=3+1=4. Четвертый элемент массива *A* (a_4) существует, значит, переходим ко второму шагу алгоритма.

Этап 12. Выполняем функцию *Get*(39, *Z*, |*Z*|, *flag*). Получаем *Z*= ∅, |*Z*|=0, *flag*=0. Переходим к пятому шагу алгоритма.

Этап 13. Увеличиваем счетчик *k*=4+1=5. Пятый элемент массива *A* (a_5) существует, значит, переходим ко второму шагу алгоритма.

Этап 14. Выполняем функцию *Get*(78, *Z*, |*Z*|, *flag*). Получаем *Z*= ∅, |*Z*|=0, *flag*=0. Переходим к пятому шагу алгоритма.

Этап 15. Увеличиваем счетчик *k*=5+1=6. Шестой элемент массива *A* (a_6) существует, значит, переходим ко второму шагу алгоритма.

Этап 16. Выполняем функцию *Get*(16, *Z*, |*Z*|, *flag*). Получаем *Z*= ∅, |*Z*|=0, *flag*=0. Переходим к пятому шагу алгоритма.

Этап 17. Увеличиваем счетчик *k*=6+1=7. Седьмой элемент массива *A* (a_7) существует, значит, переходим ко второму шагу алгоритма.

Этап 18. Выполняем функцию *Get*(25, *Z*, |*Z*|, *flag*). Получаем *Z*= ∅, |*Z*|=0, *flag*=0. Переходим к пятому шагу алгоритма.

Этап 19. Увеличиваем счетчик *k*=7+1=8. Восьмой элемент массива *A* (a_8) существует, значит, переходим ко второму шагу алгоритма.

Этап 20. Выполняем функцию *Get*(8, *Z*, |*Z*|, *flag*). Получаем *Z*={43, 36}, |*Z*|=2, *flag*=1. Переходим к третьему шагу алгоритма.

Этап 21. Проверяем множество *Z* на наличие узлов, уже добавленных в множество *V*, и при наличии таковых, удаляем их. Получаем $Z = \emptyset$, |*Z*|=0, $E=\{e_{1,2}, e_{1,3}, e_{1,4}, e_{1,5}, e_{2,6}, e_{2,7}, e_{2,8}, e_{8,3}\}$.

Этап 22. На данном этапе ничего не изменяется, так как $Z = \emptyset$.

Этап 23. Увеличиваем счетчик *k*=8+1=9. Девятого элемента массива *A* (a_9) существует, значит, работа алгоритма завершена.

Сформированный в результате алгоритма граф доступной ИТКС представлен на рисунке 3.4.

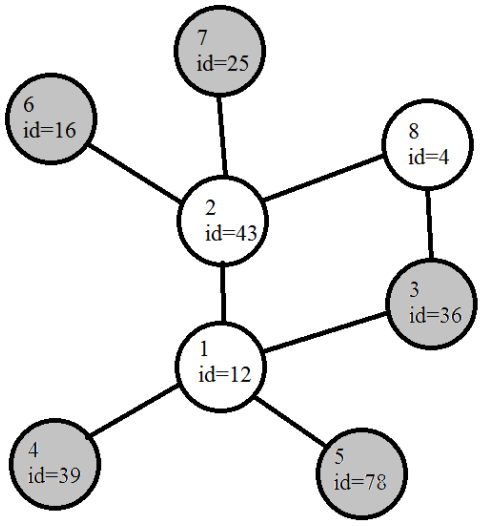


Рисунок 3.4 – Итоговый результат работы алгоритма

Результат работы после каждого этапа отражены в таблице 3.1.

Таблица 3.1 – Поэтапные результаты работы алгоритма

№	V	E	A	Z	K	N
1	w_1	Ø	12	Ø	1	0
2	w_1	Ø	12	43,36,39,78	1	0
3	w_1	Ø	12	43,36,39,78	1	0
4	w_1, w_2, u_3, u_4, u_5	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$	12,43,36,39,78	43,36,39,78	1	1
5	w_1, w_2, u_3, u_4, u_5	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$	12,43,36,39,78	43,36,39,78	2	1
6	w_1, w_2, u_3, u_4, u_5	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$	12,43,36,39,78	12,16,25,4	2	1
7	w_1, w_2, u_3, u_4, u_5	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$	12,43,36,39,78	16,25,4	2	1
8	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	16,25,4	2	5
9	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	16,25,4	3	5
10	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	3	5
11	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	4	5
12	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	4	5
13	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	5	5
14	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	5	5
15	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	6	5
16	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	6	5
17	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	7	5
18	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	7	5
19	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	Ø	8	5
20	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,16 25,4	43,36	8	5
21	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$, $e_{8,3}$	12,43,36,39,78,16 25,4	Ø	8	5
22	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$, $e_{8,3}$	12,43,36,39,78,16 25,4	Ø	8	8
23	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$, $e_{8,3}$	12,43,36,39,78,16 25,4	Ø	9	8

Рассмотренный пример показывает корректность алгоритма 3.1.

3.2 Формирование полного графа сети с учетом недоступной части

Разработан алгоритм формирования полного графа сети, который учитывает топологические характеристики доступной части сети (распределение степеней связности, средняя длина пути).

Вычисление средней степени связности сети

Степень связности узла (degree) – количество смежных с ним узлов [79].

Средняя степень связности сети (average degree) – среднее арифметическое степени связности по всей сети.

Использованный алгоритм вычисления средней степени связности основывается на вычислении степеней связности у открытых узлов с учетом их связей с закрытыми. Среднее значение берется по открытым узлам.

Получение распределения степеней связности узлов в сети

Распределение степеней связности узлов – статистическая характеристика, показывающая количество узлов с каждым значением связности в сети [59].

Учет открытых и закрытых узлов при получении распределения степеней связности ведется аналогичным образом с подходом вычисления средней степени связности.

Вычисление кластерного коэффициента сети

Кластерный коэффициент узла – характеристика, показывающая «плотность» связей вокруг узла [59]. Кластерный коэффициент узла вычисляется как отношение числа существующих связей между смежными узлами к значению общего количества возможных таких связей:

$$C_i = \frac{2n_i}{k_i \cdot (k_i - 1)},$$

где k_i – степень связности узла, n_i – количество связей между смежными узлами.

Рассмотрим пример вычисления кластерного коэффициента для узла 1 (рисунок 3.5). Сплошными линиями показаны существующие связи, пунктирными – потенциальные. Степень связности k=4. Число возможных связей между его смежными узлами равно $k(k\text{-}1)/2 = 4(4\text{-}1)/2=6$. Число существующих связей – 2. Кластерный коэффициент C=2/6=1/3.

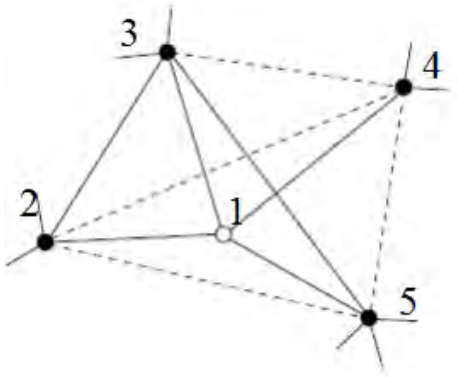


Рисунок 3.5 – Схематичный рисунок для определения кластерного коэффициента

Алгоритм вычисления коэффициента кластеризации сети заключается в подсчете кластерного коэффициента каждого узла и нахождения среднего значения. Вычисление кластерных коэффициентов осуществляется только для открытых узлов с подсчетом клик образуемых и открытыми и закрытыми узлами. Среднее значение рассчитывается по открытым узлам.

Алгоритм вычисления средней длины пути сети

Средняя длина пути узла – среднее арифметическое кратчайших путей от заданного узла до всех остальных.

Средняя длина пути сети – среднее арифметическое средних длин пути всех узлов сети.

Вычисление средней длины пути в графе осуществляется только по открытым узлам. Закрытые узлы при этом «удалялись» из сети, так как они не несут полезной информационной нагрузки для данной топологической характеристики. Данный алгоритм заключается в вычислении суммы средних длин пути для каждого открытого узла, деленной на их общее количество.

Блок-схема алгоритма формирование полного графа сети с учетом недоступной части представлена на рисунке 3.6.

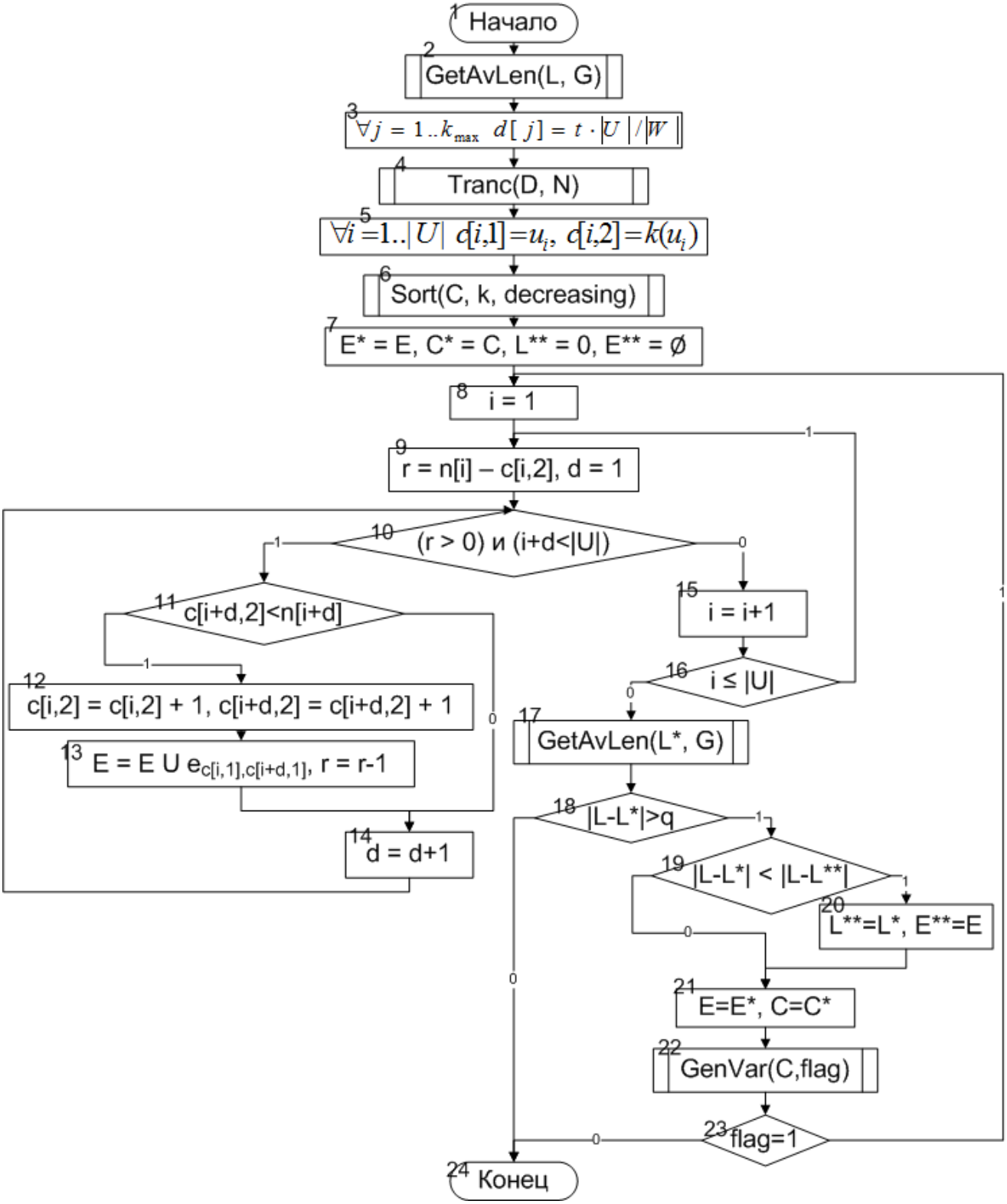


Рисунок 3.6 – Алгоритм генерации недоступной части сети

Алгоритм 3.2 – Алгоритм формирования полного графа сети

Шаг 1 (блок 2). Вычислить среднюю длину пути L в графе G.

Шаг 2 (блок 3). 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 (блок 4). 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 (блоки 5,6). 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 (блок 7). Сохранить исходную конфигурацию сети: $E^*=E$, $C^*=C$. Инициализировать переменную L^{**}=0 и множество E^{**}=Ø.

Шаг 6 (блоки 8-16). 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 $i+d \le |U|$, найти узел для связи: если он существует $c[i+d,2] < n[i+d]$, то добавить связь $c[i,2] = c[i,2]+1$, $c[i+d,2] = c[i+d,2]+1$, $E = E \cup e_{c[i,1],c[i+d,1]}$, $r=r-1$; $d=d+1$.

Шаг 7 (блок 17). Вычислить среднюю длину пути L^* для графа сети с новой конфигурацией.

Шаг 8 (блок 18). Если значение L^* удовлетворяет заданной точности q ($|L-L^*|<q$), то конец алгоритма.

Шаг 9 (блоки 19-21). 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 (блоки 22,23). Сгенерировать новый вариант расстановки узлов в массиве C. Если вариантов больше нет, то конец алгоритма, иначе перейти к шагу 6.

Рассмотрим пример поэтапной реализации алгоритма 3.2. На рисунке 3.7 представлен граф доступной части исходной сети (белым отмечены открытые, а серым – закрытые узлы). Задача – получить полный граф ИТКС с точностью средней длины пути 0,15.

Этап 1. Вычисляем среднюю длину пути – *GetAvLen(L,G)*. Получаем *L*= 2,85.

Этап 2. Получаем распределение степеней связности по закрытым узлам согласно шагу 2 алгоритма. Получаем массив *D*, который представлен в таблице 3.2.

Таблица 3.2 – Распределение степеней связности по закрытым узлам

k	1	2	3	4	5	6	7
Количество узлов	1	2	3	2	2	1	1

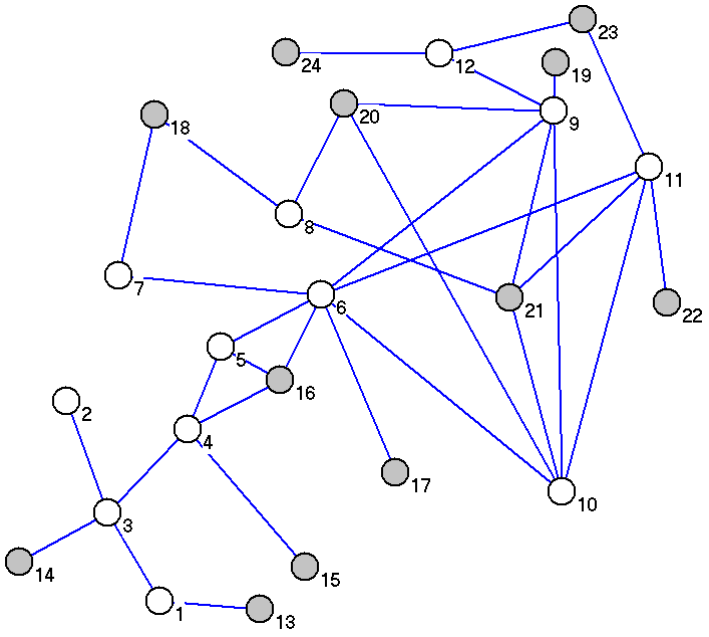


Рисунок 3.7 – Граф исходной сети

Этап 3. Формируем массив *N* и сортируем его согласно шагу 3 алгоритма. Получаем *N*={7,6,5,5,4,4,3,3,3,2,2,1}.

Этап 4. Формируем двумерный массив *C* согласно шагу 4 алгоритма. Получаем массив *C*, который представлен в таблице 3.3.

Таблица 3.3 - Двумерный массив *C*

№ узла	21	20	16	18	23	13	24	22	19	17	15	14
Текущее значение k	4	3	3	2	2	1	1	1	1	1	1	1

Этап 5. Сохраняем текущую конфигурацию согласно шагу 5 алгоритма. Получаем множество *E**, содержащее все связи исходного графа сети, и массив *C**, равный *C*. Инициализируем переменную *L***=0 и множество *E***=Ø.

Этап 6. Получаем новую конфигурацию сети согласно шагу 6 алгоритма.

Результат работы этапа представлен в таблице 3.4. На первом шаге для узла 21, чтобы получить связность 7, нужно добавить 3 связи (7-4=3). Добавляем связи к следующим узлам ($e_{21,20}$, $e_{21,16}$, $e_{21,18}$) и увеличиваем у них текущую степень связности. Шаг 2 – добавляем две связи ($e_{20,16}$, $e_{20,18}$) для узла 20. У узла 16 становится нужная степень 5, переходим к следующему узлу. Шаг 3 - добавляем одну связь ($e_{18,23}$) для узла 18. Шаг 4 - добавляем одну связь ($e_{23,13}$) для узла 23. Шаг 5 - добавляем две связи ($e_{13,24}$, $e_{13,22}$) для узла 13. Шаг 6 - добавляем одну связь ($e_{24,22}$) для узла 24. Шаг 7 - добавляем две связи ($e_{19,17}$, $e_{19,15}$) для узла 19.

Таблица 3.4 – Результаты работы 6 этапа

№ узла	21	20	16	18	23	13	24	22	19	17	15	14
Текущее значение k	4	3	3	2	2	1	1	1	1	1	1	1
Нужное значение k	**7**	**6**	**5**	**5**	**4**	**4**	**3**	**3**	**3**	**2**	**2**	**1**
Шаг 1	*7*	4	4	3	2	1	1	1	1	1	1	1
Шаг 2	*7*	*6*	*5*	4	2	1	1	1	1	1	1	1
Шаг 3	*7*	*6*	*5*	*5*	3	1	1	1	1	1	1	1
Шаг 4	*7*	*6*	*5*	*5*	*4*	2	1	1	1	1	1	1
Шаг 5	*7*	*6*	*5*	*5*	*4*	*4*	2	2	1	1	1	1
Шаг 6	*7*	*6*	*5*	*5*	*4*	*4*	*3*	*3*	1	1	1	1
Шаг 7	*7*	*6*	*5*	*5*	*4*	*4*	*3*	*3*	*3*	*2*	*2*	*1*

Этап 7. Вычисляем среднюю длину пути *L* (GetAvLen(L,G))*. Получаем *L**=2.7.

Этап 8. Сравниваем значение *L* и *L**: *q*=2,85-2,7=0,15. Полученное значение средней длины пути удовлетворяет условию задачи, следовательно, конец алгоритма.

На рисунке 3.8 представлен полный граф ИТКС, полученный в результате работы алгоритма.

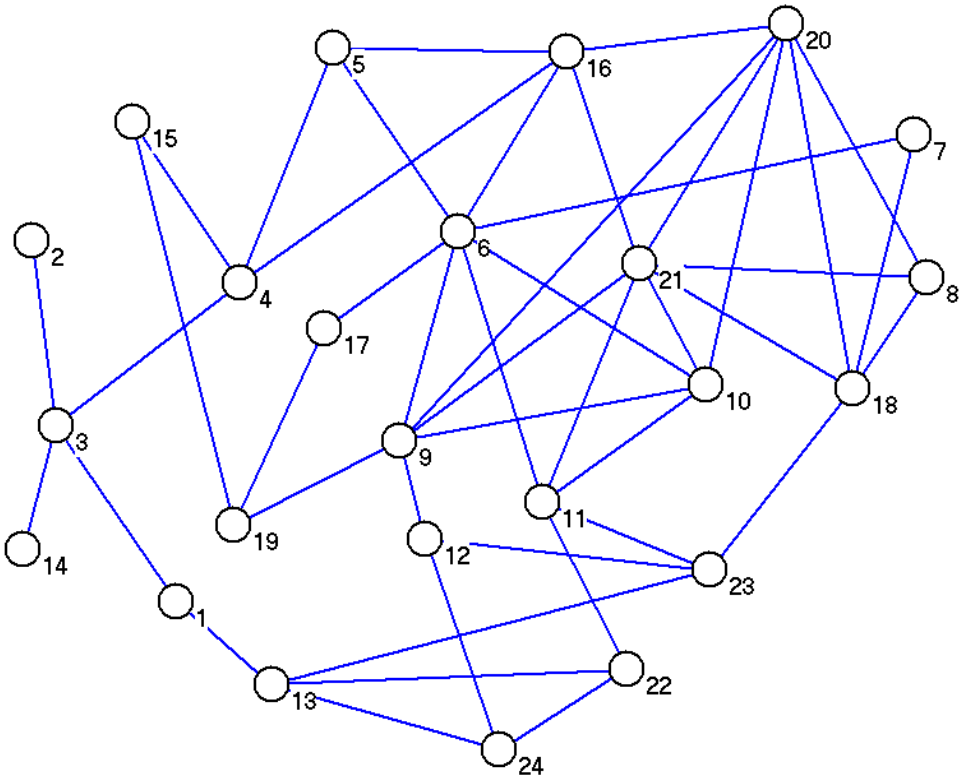


Рисунок 3.8 – Полный граф ИТКС

3.3 Формирование вектора топологической уязвимости полного графа сети

Топологическая уязвимость ИТКС – внутреннее свойство ИТКС, основанное на характеристиках ее топологии, которое способствует распространению угрозы запрещенной информации.

Топологической уязвимостью узла сети назовем показатель φ, который вычисляется по формуле:

$$\varphi_i = \frac{k_i \cdot (C_i + 1)}{L_i}, \tag{3.1}$$

где k_i – степень связности узла, C_i – кластерный коэффициент узла, L – средняя длина пути узла.

Данная характеристика показывает, насколько уязвим к атакам с точки зрения расположения в сети определенный узел.

Накладываемое условие для применения формулы 3.1 - в сети должно быть больше одного узла.

Свойства коэффициента φ:

1) $1 \leq \varphi \leq 2(N\text{-}1)$, где N – количество узлов в сети.

Крайний случай (максимальное значение) – полносвязный граф. В нем $k_i = N - 1$ и средняя длина пути равна единице L_i=1. Кластерный коэффициент имеет свойство $0 \leq C \leq 1$ и в полносвязном графе C_i=1. Следовательно, в этом случае φ_i=2(N-1).

Крайний случай (минимальное значение) – граф из двух узлов. При этом $k_i = 1$, L_i=1 C_i=0. Следовательно, в этом случае φ_i=1.

2) С увеличением φ, возрастает уязвимость узла.

Подсчет коэффициента топологической уязвимости для всей сети осуществляется по формуле:

$$\varphi = \frac{k \cdot (C+1)}{L}, \tag{3.2}$$

где k – средняя степень связности узлов в сети, C – средний кластерный коэффициент сети, L – средняя длина пути сети.

При исследовании топологий реальных крупномасштабных ИТКС (10^5-10^8), можно выделить основные значимые положения:

1) средняя степень связности узлов в таких сетях составляет 100-1000 [69,72,74,131];
2) средняя длина пути определяется теорией шести рукопожатий: в глобальных масштабах равна 6, в реальных сетях составляет значение 3-5 [4,138];
3) коэффициент кластеризации, как правило, варьируется в значениях от 0,01 до 0,2 [131].

Исходя из вышеперечисленного и полученных экспериментальных результатов, имеем типичное значение коэффициента топологической уязвимости в диапазоне от 100 до 500.

Практическое применение

1) Используя коэффициент φ, можно оценить топологическую уязвимость конкретной реальной сети по формуле 3.2.

В ходе работы были проанализированы социальные сети Facebook и ВКонтакте. Для сети Facebook $\varphi \approx 70$, ВКонтакте – $\varphi = 200$. Для сети Facebook получили не совсем типичное значение, связано это с методом выборки, примененной американскими исследователями [69-74], а также тем, что данная сеть крупнейшая и, действительно, в целом менее уязвимая, чем сеть ВКонтакте. Во второй главе диссертации топологическая уязвимость φ использовались для создания аналитической модели распространения запрещенной информации как интегральная составляющая топологических параметров сети.

2) При анализе топологических характеристик сети можно подсчитать коэффициенты уязвимости для каждого узла в сети (вектор топологической уязвимости узлов ИТКС).

Вектор топологической уязвимости узлов ИТКС – вектор вида:

№ узла	Значение φ
Узел 1	φ_1
……..	…….
Узел N	φ_N

Полученный вектор можно использовать при прогнозировании угрозы распространения запрещенной информации. С одной стороны, можно классифицировать по опасности атакующие узлы, а с другой стороны, можно выстроить наиболее эффективную стратегию противодействия угрозе.

3.4 Особенности разработки программного инструментария

Разработанная методика формирования топологии ИТКС реализована в виде программного комплекса.

Первая программа предназначена для получения доступной части сети. Хотя данное ПО ориентировано на социальную сеть «ВКонтакте», его легко

можно переработать под другую ИТКС. Работа программы основана на алгоритме обхода в ширину. Приложение написано на языке программирования Python. Для хранения топологии используется объектно-ориентированная база данных ZODB. Получение информации осуществляется при помощи API ВКонтакте. Программа собирает данные до наступления одного из следующих событий: получена информация обо всех открытых узлах в сети, сбор данных прерван пользователем.

В начале работы программы необходимо авторизоваться под аккаунтом абонента, с которого начнется сбор информации. Выходными данные приложения представляют собой текстовый файл, в котором в каждой строке записан идентификатор узла, и через пробел перечислены идентификаторы смежных с ним узлов.

В результате работы программы была получена часть топологии социальной сети ВКонтакте, содержащая 118834 открытых узла и 16270504 закрытых. Фрагмент выходного файла представлен на рисунке 3.9.

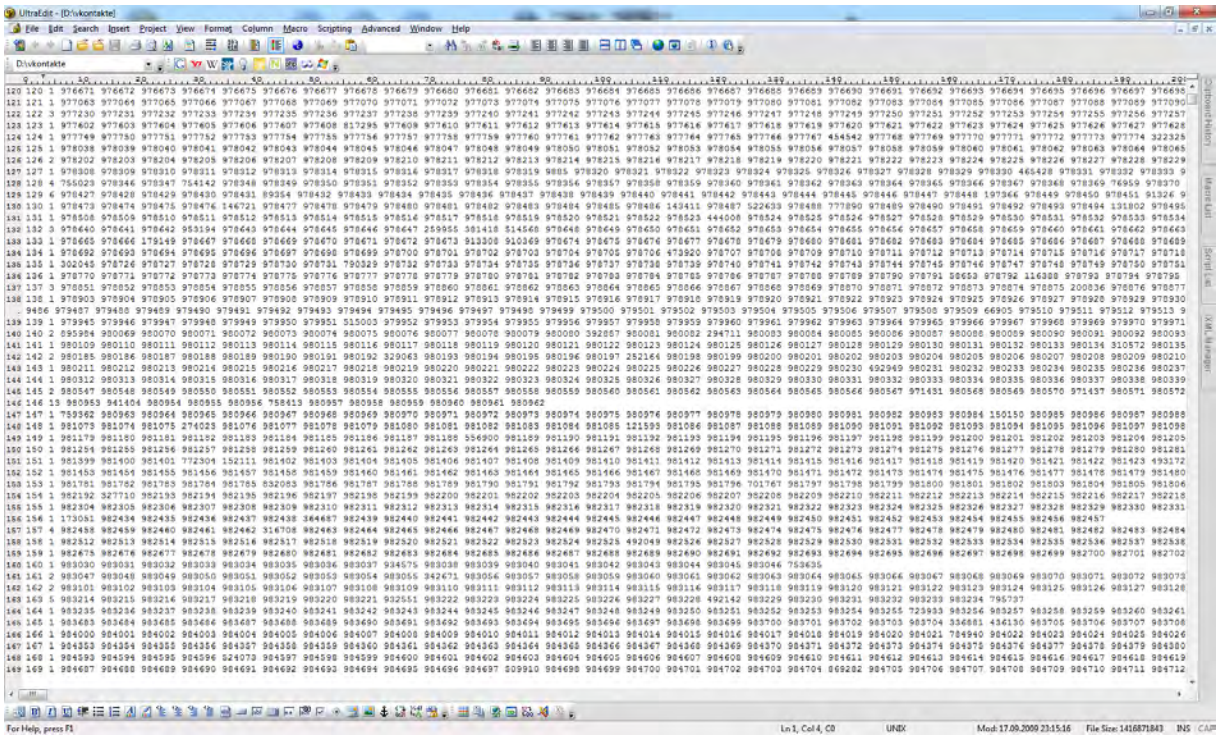


Рисунок 3.9 – Фрагмент выходного файла программы

На рисунке 3.10 показан фрагмент (1000 узлов) полученной топологии, построенный с помощью ПО Pajek.

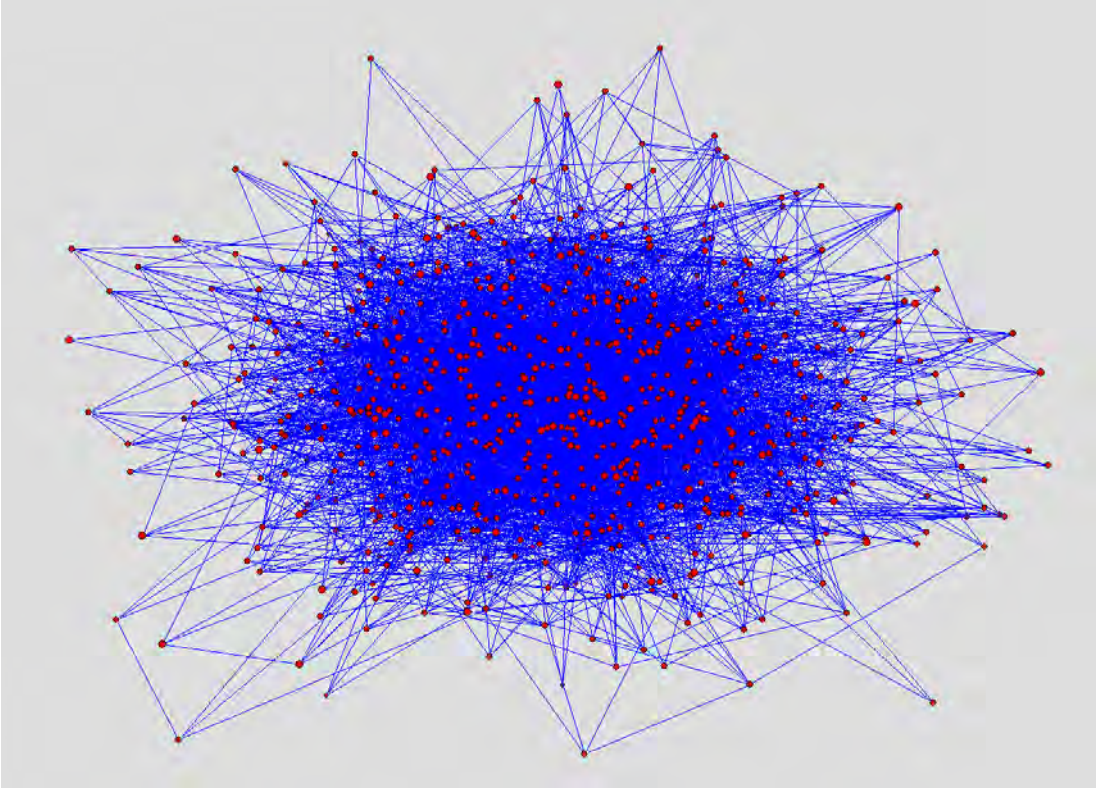


Рисунок 3.10 – Визуализированный фрагмент топологии

Вторая программа предназначена для формирования полного графа ИТКС на основе вычисленных прогнозируемых топологических характеристик и формирования его вектора топологической уязвимости. ПО создано для использования на супер-ЭВМ «Скиф-Мономах» с использованием распределенных вычислительных ресурсов. Программа написана в среде программирования Microsoft Visual Studio 2008. Интерфейсом взаимодействия между процессами в приложении является MPI. В некоторых случаях дополнительно использовалось многопоточное программирование. Для представления графа в памяти вычислительной системы использовалось два подхода: нераспределенный (локальный, использовалась библиотека Boost Graph Library) и распределенный (Parallel Boost Graph Library).

Формат выходных данных аналогичен первой программе - текстовый файл, в котором в каждой строке записан идентификатор узла, и через пробел перечислены идентификаторы смежных с ним узлов (топология полного графа сети). Второй выходной файл – файл с вектором топологической уязвимости сети.

Реализация ПО подтверждается свидетельством о государственной регистрации программ (Приложение Г).

Выводы к третьей главе

Разработана методика формирования топологии ИТКС, которая учитывает основные топологические характеристики доступной части сети и работает в условие недостаточной репрезентативности выборки исходных данных. Предлагаемая методика состоит из последовательности разработанных алгоритмов.

Создан алгоритм формирования исходных данных о топологии сети (множества вершин и связей между ними доступной части сети), который учитывает ограничения по сбору данных и реализован в виде разработанного программного обеспечения.

Разработан алгоритм формирования полного графа сети с учетом добавления недоступной части на основе вычисленных прогнозируемых топологических характеристик. Алгоритм реализован в виде разработанного программного обеспечения.

Введена оценка топологической уязвимости сети (вектор топологической уязвимости), учитывающая следующие параметры: среднюю длину пути сети, коэффициент кластеризации сети, среднюю степень связности сети и общее количество узлов в сети.

ГЛАВА 4 ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ. ОСОБЕННОСТИ ВНЕДРЕНИЯ

Моделирования 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 социальной сети «ВКонтакте») получен в рамках данной научной работы (глава 3), а второй (фрагмент из 16163521 узла социальной сети «Facebook») получен независимо американскими учеными Minas Gjoka, Maciej Kurant и др. [69-74].

4.1 Распределенное моделирование угрозы распространения запрещенной информации в ИТКС

Экспериментальное исследование УгЗИ в ИТКС осуществлялось на основе имитационной модели, подробно рассмотренной во второй главе работы.

Имитационная модель реализована в виде разработанного ПО под распределенную вычислительную систему. Для реализации параллельных вычислений на графе была использована библиотека Parallel Boost Graph Library [111]. Библиотека является свободно распространяемой и по своим функциональным возможностям не имеет альтернатив.

Parallel Boost Graph Library (PBGL) предоставляет гибкую и эффективную реализацию концепции графов. Входит в собрание библиотек boost, расширяющих функциональность C++, которые свободно распространяются по лицензии Boost Software License вместе с исходным кодом.

Библиотека позволяет выбрать представление графа, тип данных и алгоритм из большого набора алгоритмов, среди которых:

- Поиск в ширину
- Поиск в глубину
- Алгоритм Беллмана-Форда
- Алгоритм Дейкстры
- Алгоритм Прима
- Алгоритм Краскала
- Нахождение компонент связности графа
- Задача о максимальном потоке
- Обратный алгоритм Катхилла-Макки
- Алгоритм топологической сортировки и др.

Разработанная программа создана для использования на супер-ЭВМ «Скиф-Мономах».

Суперкомпьютер находится во Владимирском государственном университете (РЦНИТ) с 2008 года. Общая характеристика представлена в таблице 4.1.

Таблица 4.1 – Общая характеристика супер-ЭВМ «Скиф-Мономах»

Пиковая производительность	4771 Tflops/s
Производительность на Linpack	3756 Tflops/s (78.7 % от пиковой)
Число процессоров/ядер в системе	128/512
Модель процессора	Intel Xeon 5345 2.33 GHz
Объем оперативной памяти	512 Гбайт
Дисковая память узлов	10Тб
Число стоек/вычислительных	4/2
Число вычислительных узлов	64
Производитель	Т-Платформы

Группы вычислительных узлов: student (4 узла, 2 процессора, ОП 8 Гб, HDD 160 Гб), short (14 узлов, 2 процессора, ОП 8 Гб, HDD 160 Гб), long (14 узлов, 2 процессора, ОП 8 Гб, HDD 160 Гб), work (32 узла, 2 процессора, ОП 8 Гб, HDD 160 Гб).

Все узлы в СКИФ МОНОМАХ связаны двумя независимыми сетями: системная сеть: InfiniBand DDR (Fat Tree: 6x12 порта; латентность на уровне MPI: 1.3-1.95 мкс; скорость обмена на уровне MPI: 1540 Мбайт/с) и вспомогательная/управляющая сеть: GigabitEthernet (2x(44 портов + 4x10G)).

Программное обеспечение:

- Операционная система Suse Linux Enterprise Server v 10 sp1
- Система очереди задач Torque
- Система мониторинга узлов Ganglia
- Компиляторы GNU gcc, Intel C/C++ Compiler
- Доп. Библиотеки MPI (mpich), ANSYS, ScaLAPACK, lapack, blas

Инфраструктура суперкомпьютера. Суперкомпьютер СКИФ МОНОМАХ обладает уникальной информационно-вычислительной и инженерной инфраструктурой, необходимой для надёжной круглосуточной работы комплекса.

Разработанное ПО написано в среде программирования Microsoft Visual Studio 2008. Интерфейсом взаимодействия между процессами в приложении является MPI. В некоторых случаях дополнительно использовалось многопоточное программирование. Для представления графа в памяти вычислительной системы использовался распределенный подход с использованием библиотеки Parallel Boost Graph Library.

Формат входных данных - текстовый файл, в котором в каждой строке записан идентификатор узла, и через пробел перечислены идентификаторы смежных с ним узлов (топология полного графа сети). В выходном файле фиксируются данные о динамике УгЗИ, представленные списками атакующих и защищенных узлов в каждый квант времени.

Реализация ПО подтверждается свидетельством о государственной регистрации программ (Приложение В). В приложении Г приведена часть кода программы.

4.2 Анализ результатов экспериментальных исследований

4.2.1 Анализ результатов моделирования УгЗИ в ИТКС

Предложенный алгоритм распределенного моделирования был апробирован на двух представленных выше топологических фрагментах сетей, после применения к ним алгоритма формирования полного графа сети. Эксперименты проводились с разными начальными условиями. Сначала было проанализировано влияние параметров β и γ на характер процесса, результаты экспериментов приведены на рисунках 4.1 и 4.2 («ВКонтакте»).

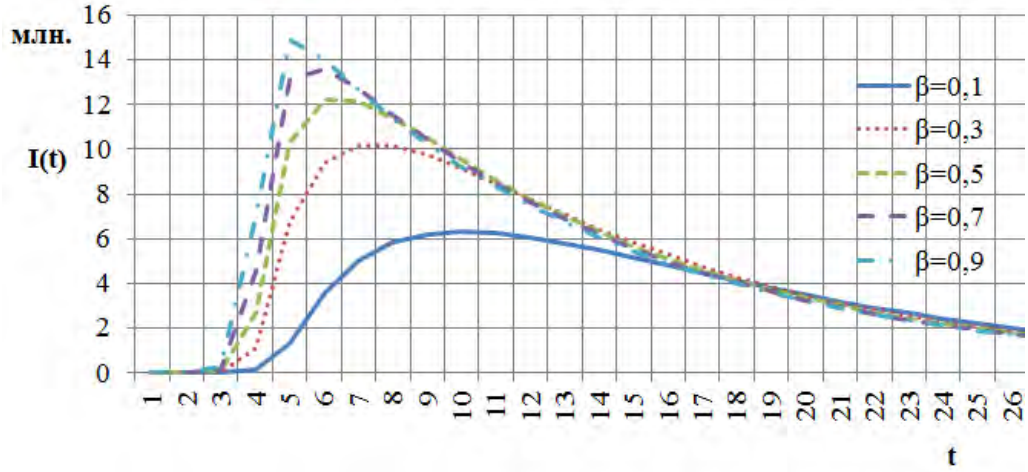


Рисунок 4.1 – Результаты моделирования с параметрами $\gamma = 0{,}1$, $I_0 = 1$, $R_0 = 0$

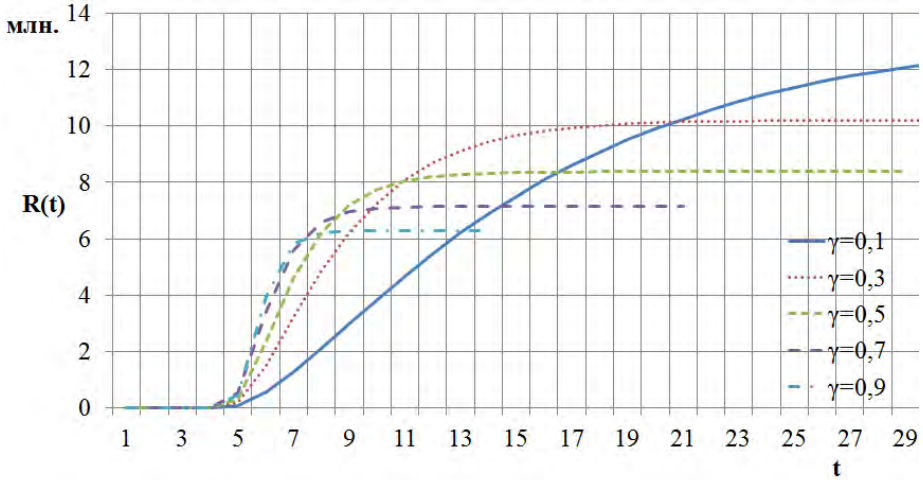


Рисунок 4.2 – Результаты моделирования с параметрами $\beta = 0{,}2$, $I_0 = 1$, $R_0 = 0$

В ходе работы во второй главе была получена аналитическая модель УгЗИ в ИТКС. В рамках данной модели предусмотрены два случая: $\beta \neq \gamma$ и $\beta = \gamma$, поэтому при моделировании использовались следующие частные случаи: β=0,2 и γ=0,8 , β=0,5 и γ=0,5. Количество изначально атакующих узлов I_0, рассматривалось

исходя из того факта, что это может быть один человек, либо несколько. В качестве нескольких распространителей выбиралось порядка 0,1% узлов случайным образом. При рассмотрении такого условия как количество изначально защищенных узлов R_0, исходим из следующих соображений. Во-первых, таких узлов может и не быть, во-вторых, их может быть достаточное количество (рассматривалось 25% от общего количества узлов в сети), и, в-третьих, такие узлы составляют основную часть сети (рассматривалось 75% от общего количества узлов в сети). Узлы, подверженные атаке (S_0), определяются: $S_0=N- I_0- R_0$, где N – общее количество узлов в сети.

Графики результатов проведенного моделирования распространения запрещенной информации на топологическом фрагменте социальной сети «ВКонтакте» приведены на рисунках 4.4-4.15. На рисунке 4.3 представлена общая легенда.

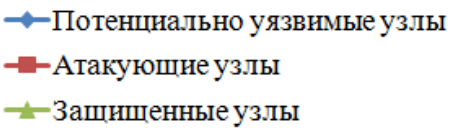


Рисунок 4.3 – Легенда

Эксперимент 1 (рис. 4.4): $\varphi = 200$, $\beta =0,2$, $\gamma =0,8$, $I_0 =1$, $R_0=0$.

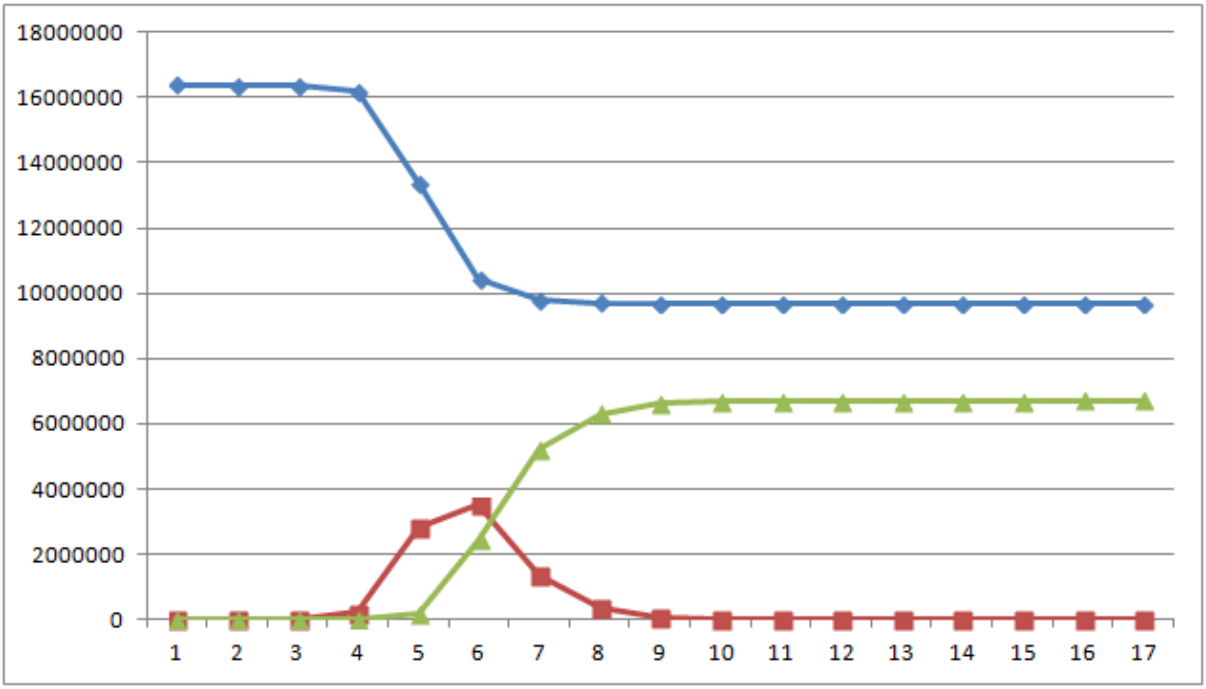


Рисунок 4.4 – Результаты эксперимента 1

Эксперимент 2 (рис. 4.5): $\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0{,}25N$.

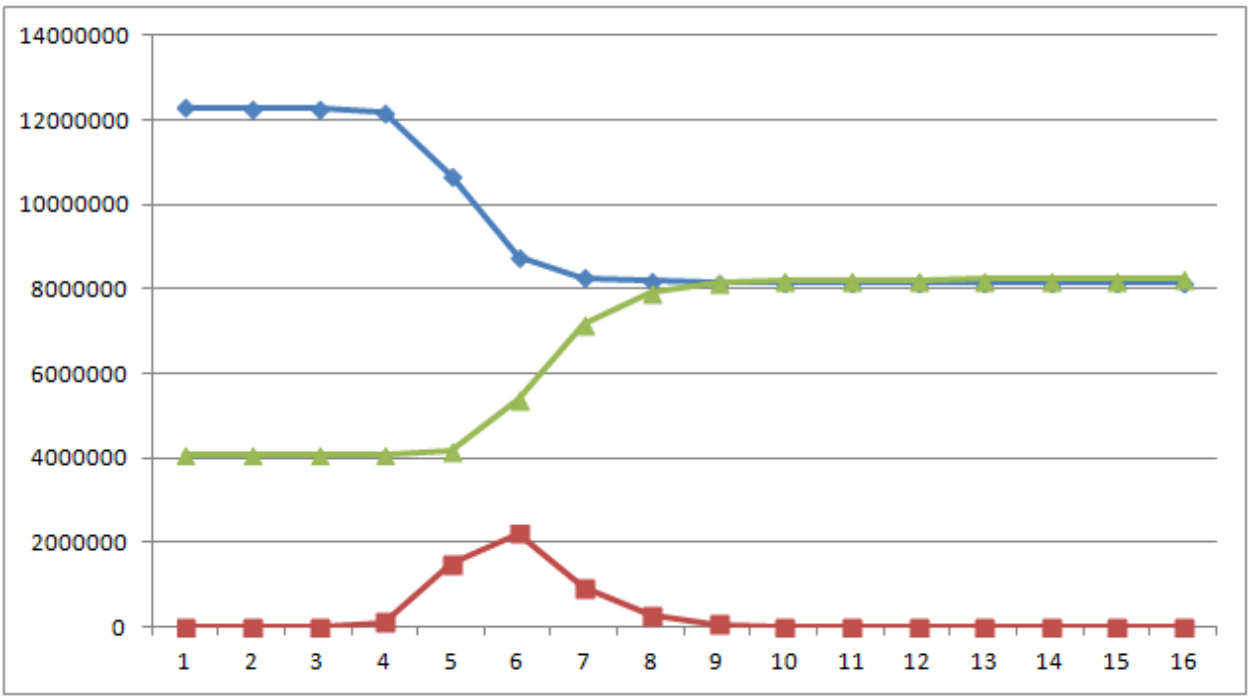


Рисунок 4.5 – Результаты эксперимента 2

Эксперимент 3 (рис. 4.6): $\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0{,}75N$.

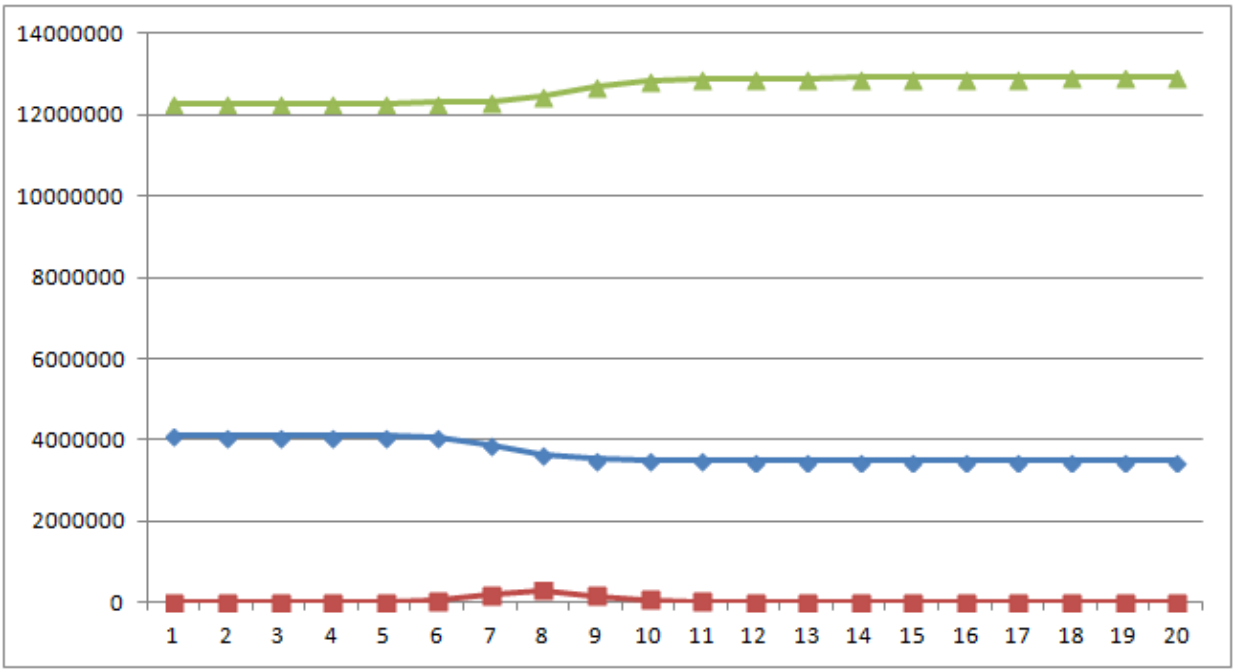


Рисунок 4.6 – Результаты эксперимента 3

По результатам первых трех экспериментов можно сделать следующие выводы:

1) Уже один атакующий узел может вызвать «вспышку» в сети, даже при большом значении вероятности защиты.

2) С ростом числа изначально защищенных узлов, максимальное число атакующих узлов падает.

Эксперимент 4 (рис. 4.7): $\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 0{,}001N$, $R_0 = 0$.

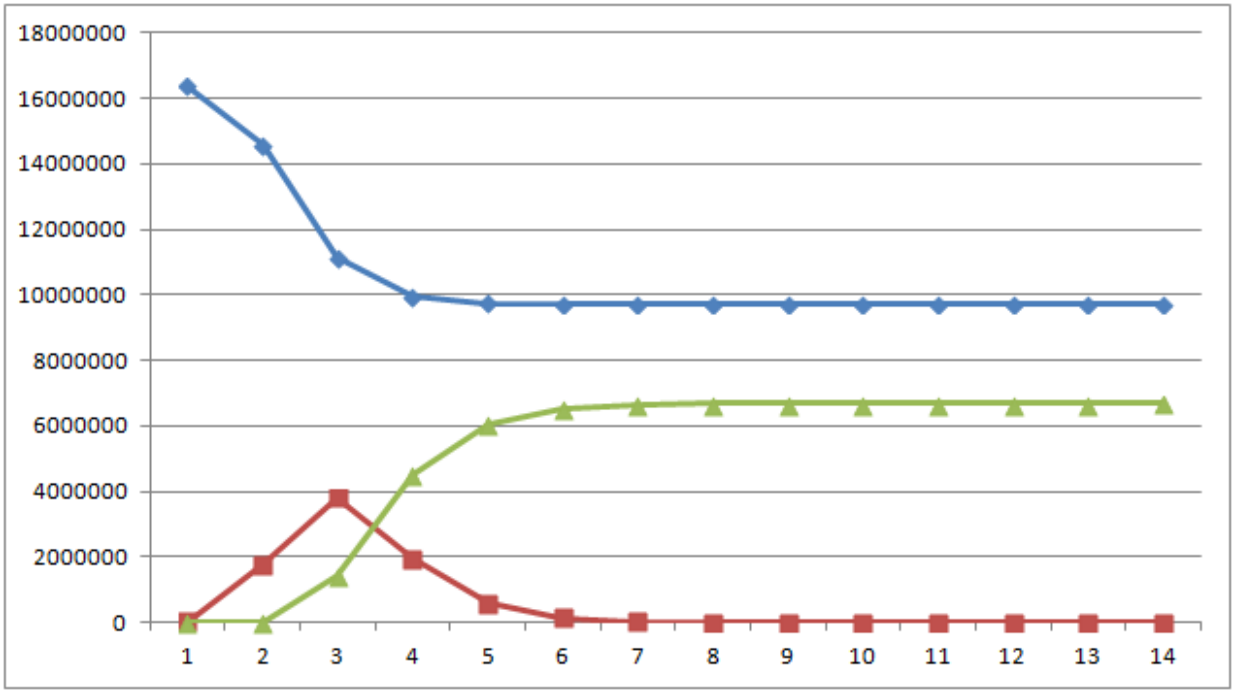


Рисунок 4.7 – Результаты эксперимента 4

Эксперимент 5 (рис. 4.8): $\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 0{,}001N$, $R_0 = 0{,}25N$.

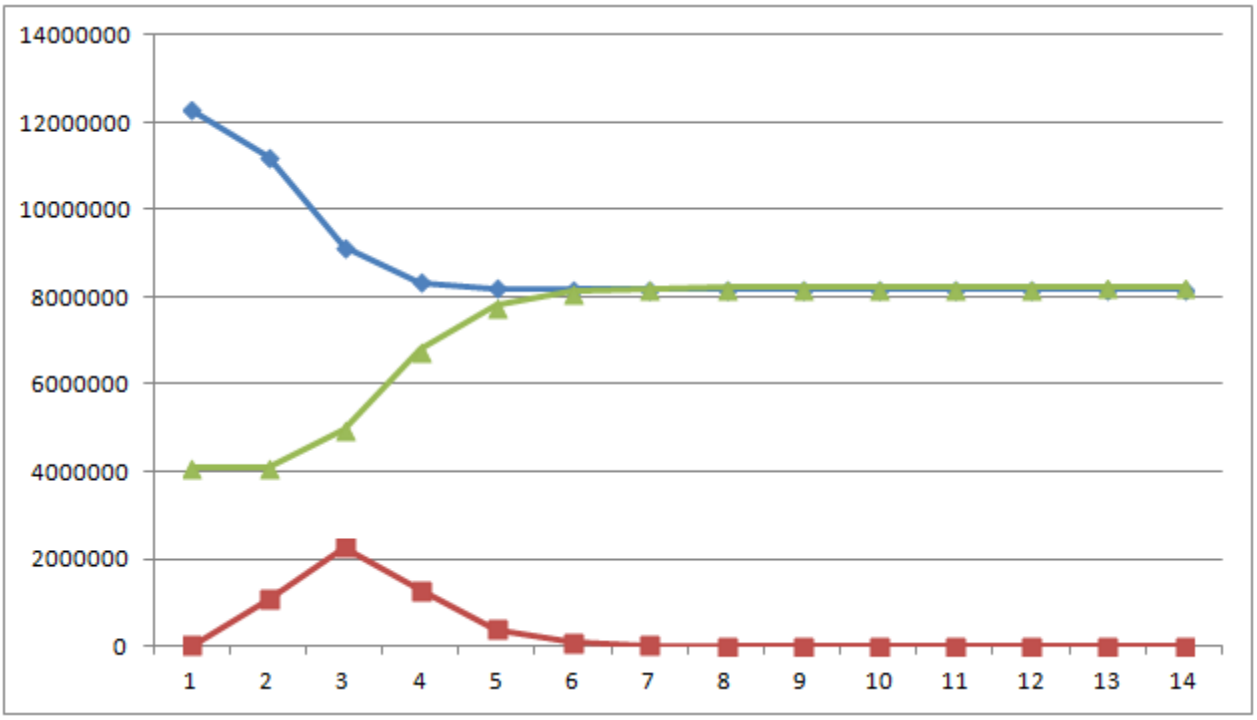


Рисунок 4.8 – Результаты эксперимента 5

Эксперимент 6 (рис. 4.9): $\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 0{,}001N$, $R_0 = 0{,}75N$.

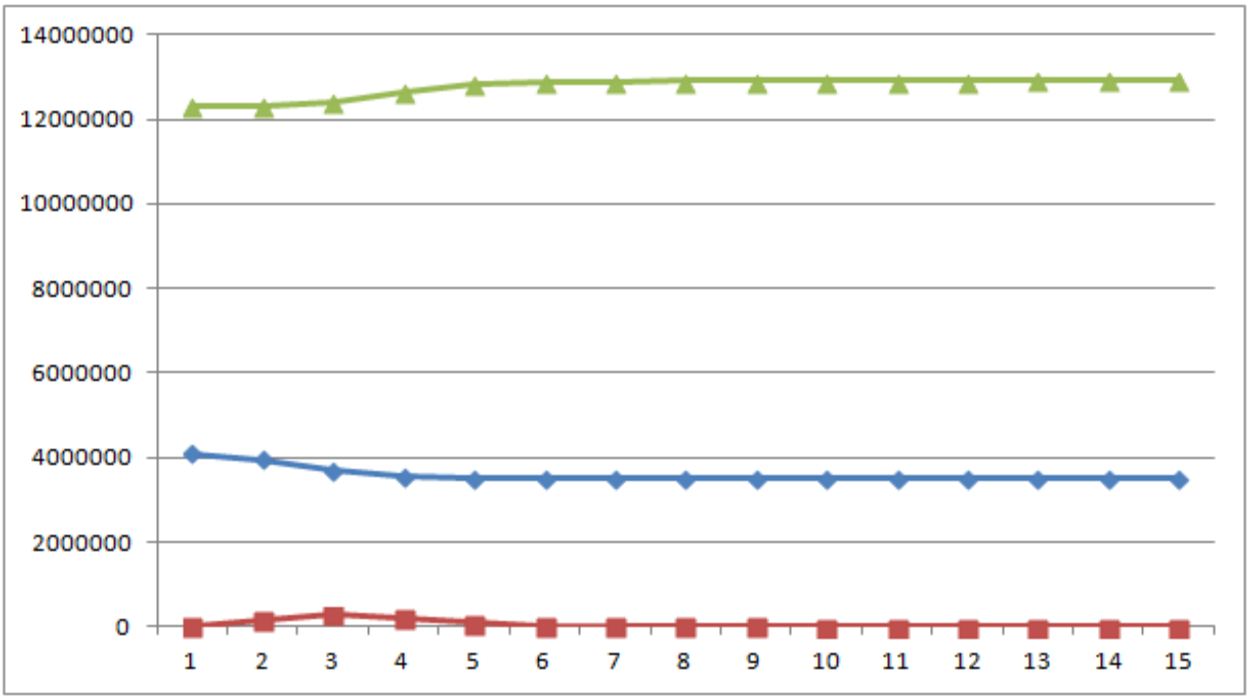


Рисунок 4.9 – Результаты эксперимента 6

По результатам экспериментов 4-6 можно сделать следующие выводы:

1) При росте числа изначально атакующих узлов наблюдается «вспышка» уже на первых этапах (1-6 тики).
2) Параметр R_0 влияет на пик также как и в 1-3 экспериментах.

Эксперимент 7 (рис. 4.10): $\varphi = 200$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 1$, $R_0 = 0$.

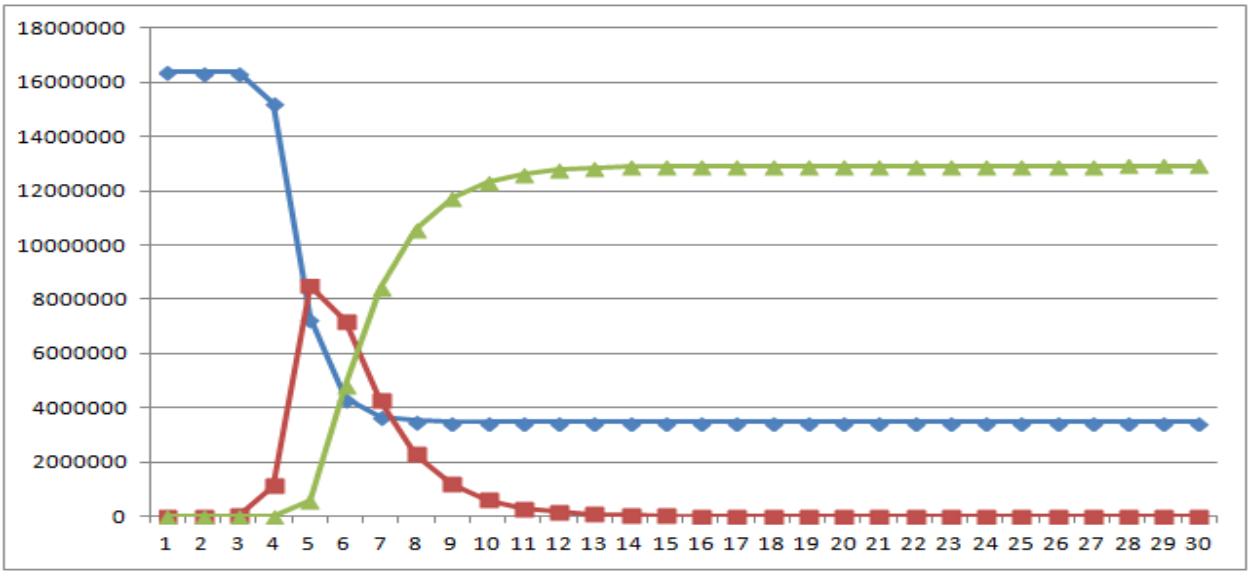


Рисунок 4.10 – Результаты эксперимента 7

Эксперимент 8 (рис. 4.11): $\varphi = 200$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 1$, $R_0 = 0{,}25N$.

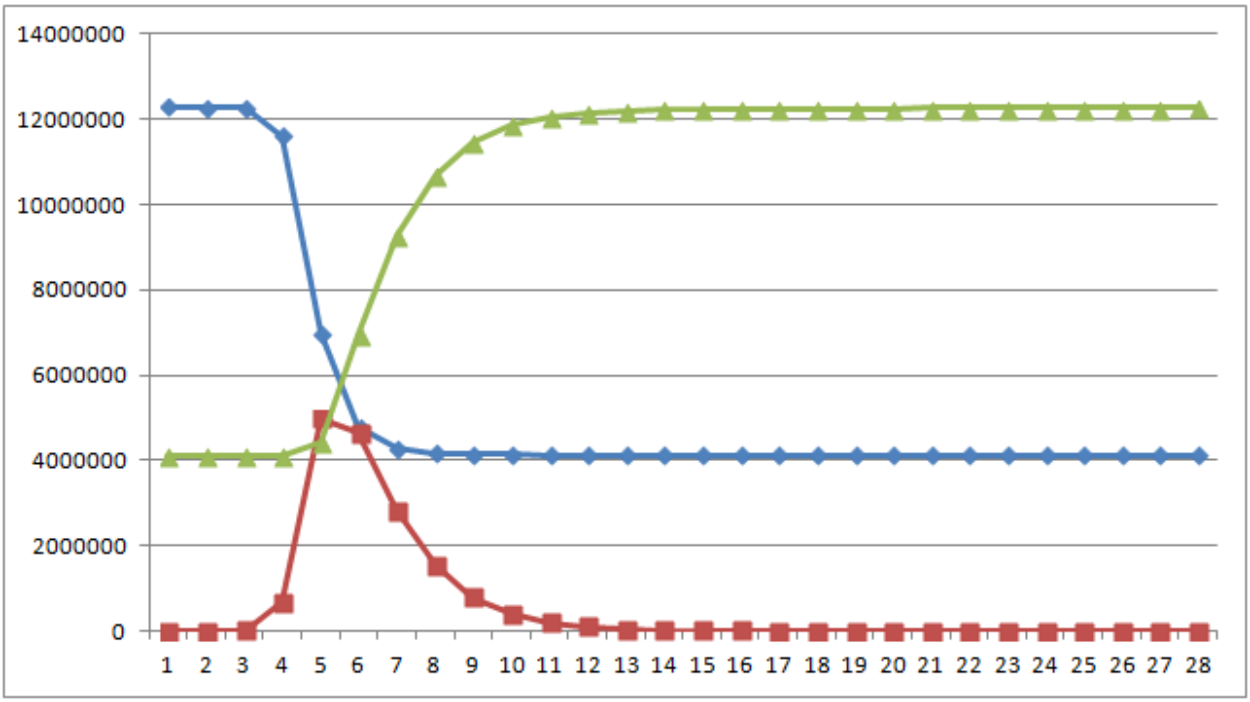


Рисунок 4.11 – Результаты эксперимента 8

Эксперимент 9 (рис. 4.12): $\varphi = 200$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 1$, $R_0 = 0{,}75N$.

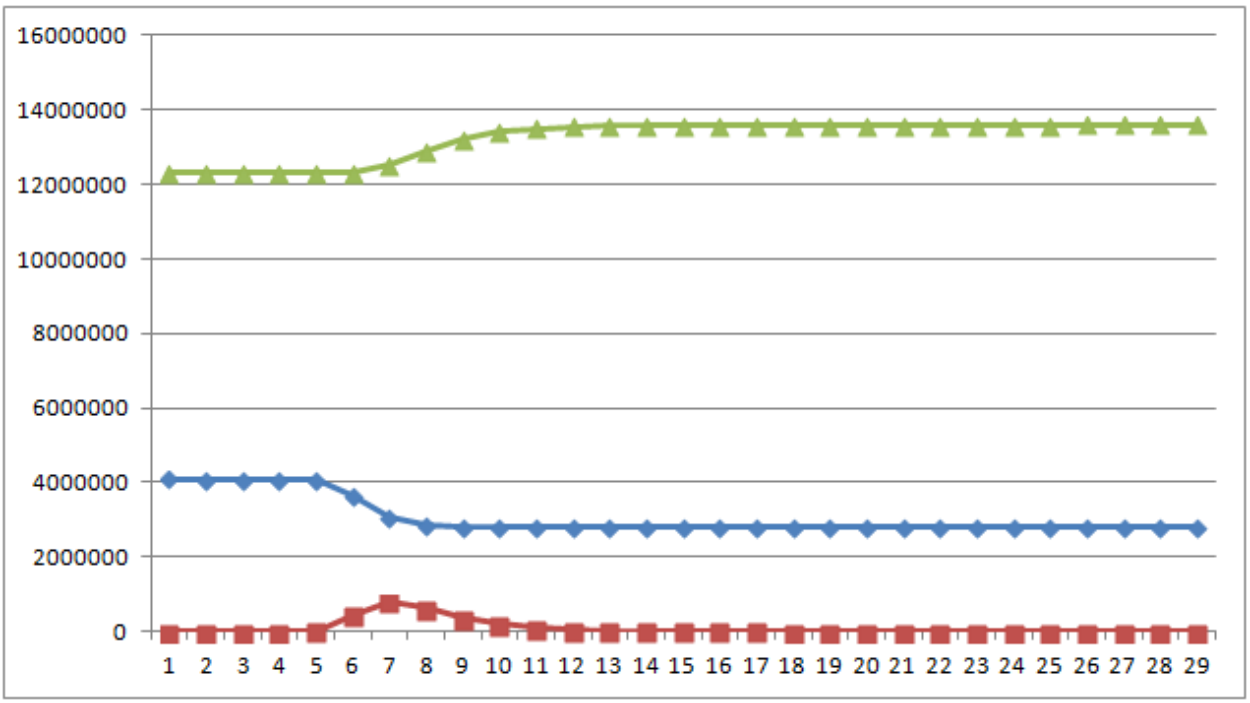


Рисунок 4.12 – Результаты эксперимента 9

По результатам экспериментов 7-9 можно сделать следующие выводы:

1) При изменении параметров β и γ резко меняется характер УгЗИ.
2) При увеличении вероятности атаки и уменьшении вероятности защиты угроза принимает глобальный характер даже при одном изначальном атакующем узле (пик атакующих узлов увеличивается более чем в два раза).

3) Процесс УгЗИ увеличивается по времени в два раза (с 15 тиков до 30).
4) Влияние числа изначально защищенных узлов остается таким же, как и в экспериментах 1-3.

Эксперимент 10 (рис. 4.13): $\varphi = 200$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 0{,}001N$, $R_0 = 0$.

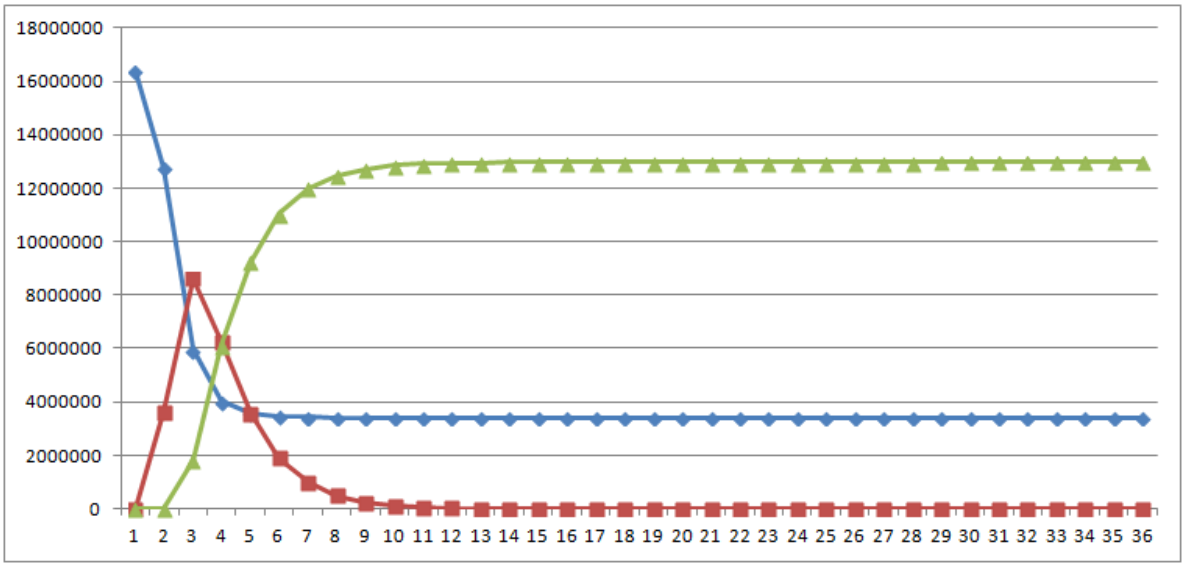


Рисунок 4.13 – Результаты эксперимента 10

Эксперимент 11 (рис. 4.14): $\varphi = 200$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 0{,}001N$, $R_0 = 0{,}25N$.

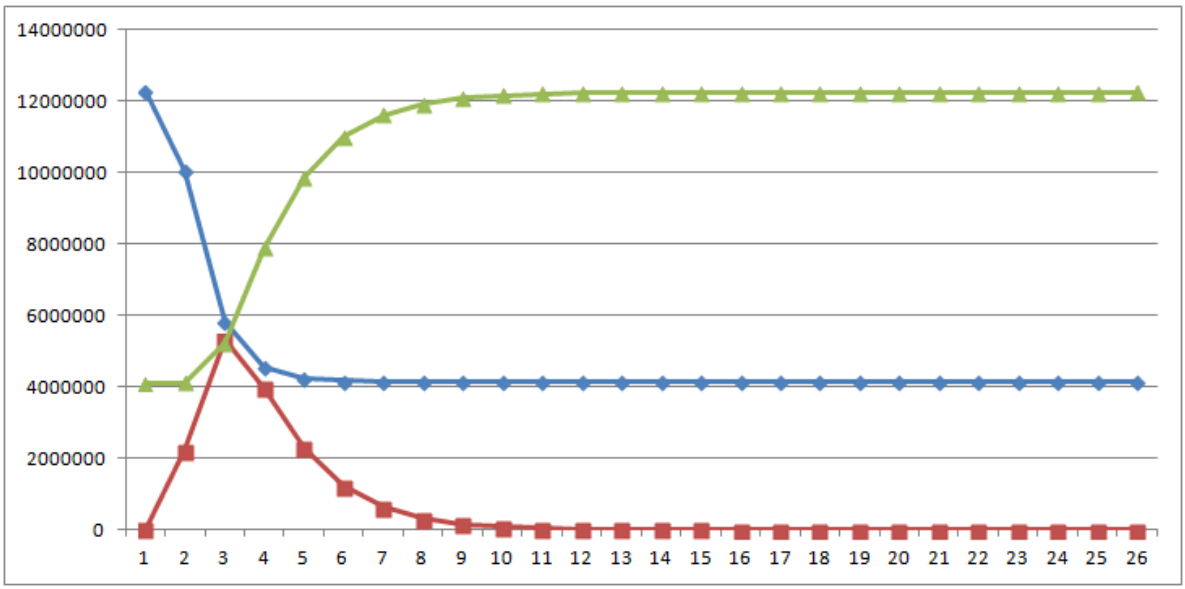


Рисунок 4.14 – Результаты эксперимента 11

Эксперимент 12 (рис. 4.15): $\varphi = 200$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 0{,}001N$, $R_0 = 0{,}75N$.

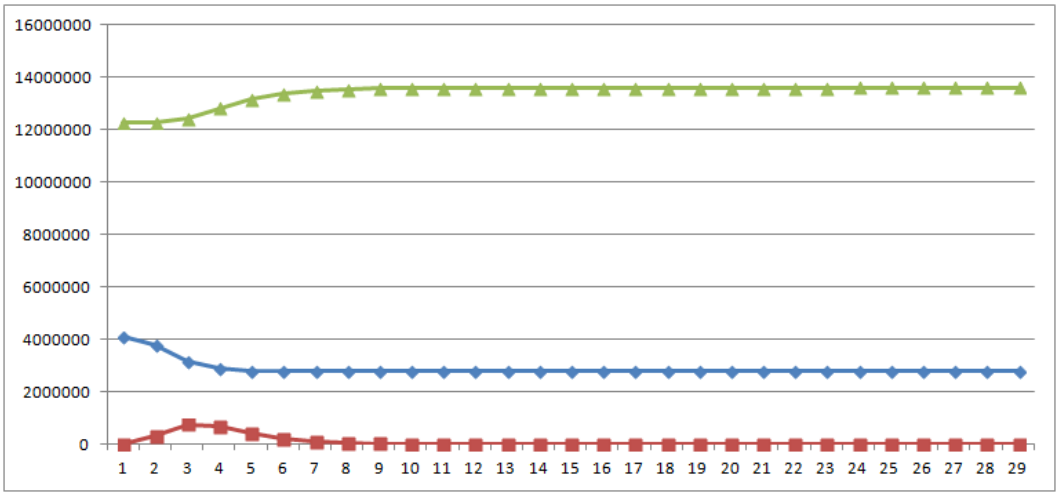


Рисунок 4.15 – Результаты эксперимента 12

По результатам экспериментов 10-12 можно сделать следующий вывод: в целом, тенденции, прослеживаемые в предыдущих экспериментах, не меняют свой характер.

Графики результатов проведенного моделирования распространения запрещенной информации на топологическом срезе социальной сети Facebook приведены в приложении Д. Характер процесса распространения запрещенной информации на этой сети такой же, как и на сети ВКонтакте. Это факт указывает на то, что разные социальные сети имеют схожую топологию.

4.2.2 Анализ результатов экспериментальных исследований топологии ИТКС

В ходе экспериментальных исследований были получены результаты, касающиеся топологии ИТКС [4,9].

В таблице 4.2 [27] представлены основные топологические характеристики для случайных графов и двух видов сложных сетей (complex networks), которые были рассмотрены в первой главе диссертации.

После проведения экспериментов можно сравнить результаты с представленными данными и сделать вывод о принадлежности социальных сетей к определенному типу, исходя из полученных топологических характеристик. Зная топологические характеристики ИТКС, можно генерировать на их основе

сети с такими же параметрами любых масштабов, что поможет изучать процессы, происходящие в них с использованием моделирования.

Таблица 4.2 - Основные топологические характеристики

Параметр	Случайные графы	Small world	Scale-Free
Средняя длина пути L	$\frac{\ln N}{\ln k}$	$\frac{\ln N}{\ln k}$	$m=1: l \sim \ln N;$ $m \geq 2: l_{BA}^{\alpha>3} \approx \ln N;$ $l_{BA}^{\alpha=3} \approx \ln N / \ln \ln N;$ $l_{BA}^{2<\alpha<3} \approx \ln \ln N.$
Кластерный коэффициент C	$\frac{k}{N}$	$C_{p \to 1} \sim k/N,$ $C_{p>0} >> C_{p \to 1}$	$5\frac{k}{N}$
Распределение степеней вершин	Закон Пуассона	Закон Пуассона	Степенной закон

Распределение степеней связности узлов

Рассмотрим результаты вычисления средней степени связности по топологическому фрагменту сети ВКонтакте. Эксперименты проводилось с использование двух подходов. При первом подходе использовался алгоритм, учитывающий все узлы. Результат представлен на рисунке 4.16, на нем показано распределение степеней связности узлов. Для информативности данные представлены на логарифмированной шкале. При этом подходе средняя степень связности по всем узлам получилась равной 8,387. При втором подходе учитывались только открытые узлы. Распределение показано на рисунке 4.17. Средняя степень связности составила 631,685.

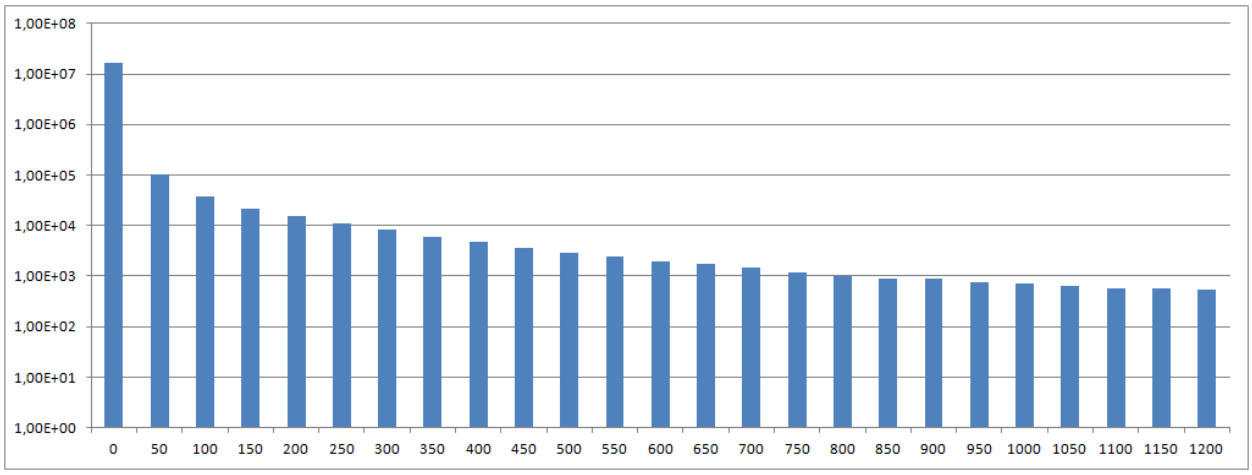


Рисунок 4.16 – Распределение степеней связности по всем узлам (ВКонтакте)

Подход, учитывающий только открытые узлы, является более корректным. Обосновывается это тем, что по закрытым узлам у нас нет полной информации и, следовательно, степень связности у них маленькая, также их на два порядка больше, чем открытых, поэтому они значительно меняют характер распределения.

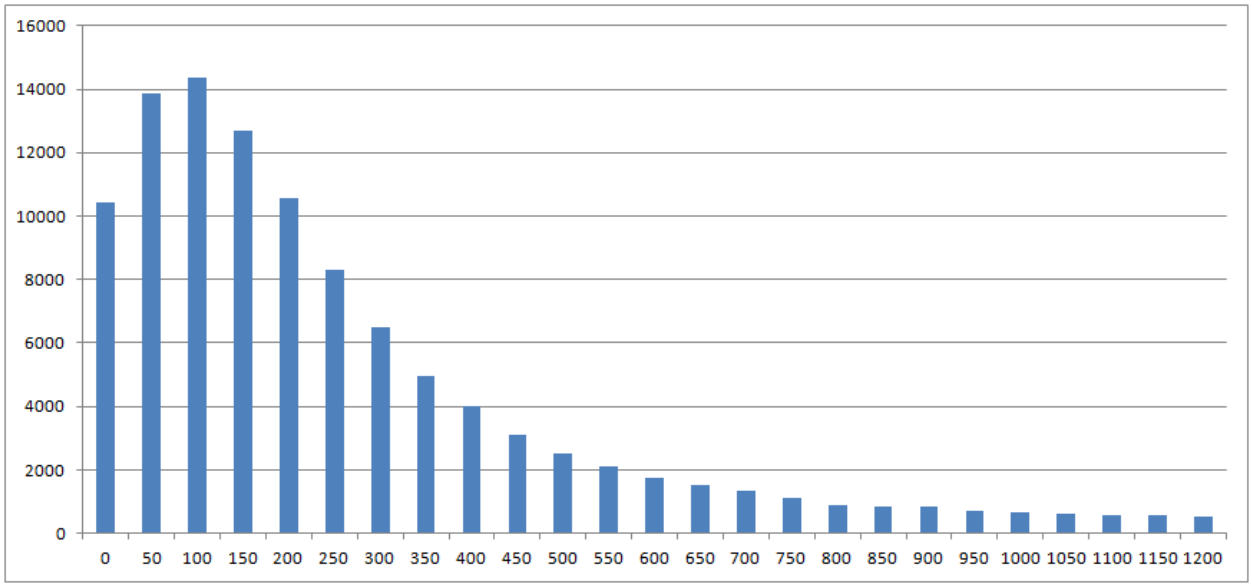


Рисунок 4.17 - Распределение степеней связности по открытым узлам (ВКонтакте)

В ходе анализа данных по распределению средней степени связности были получены следующие результаты [2,11]. Можно видеть, что представленное распределение, показанное на рисунке 4.4, нельзя аппроксимировать ни пуассоновским, ни степенным распределением. Хорошо подходит гамма-распределение, плотность вероятности которого имеет вид:

$$f(x)=\begin{cases} x^{k-1}\dfrac{e^{-x/\theta}}{\Gamma(k)\theta}, & \text{x} \geq 0 \\ 0, & \text{x} < 0 \end{cases}, \qquad (4.1)$$

где $\Gamma(k)$ – гамма-функция Эйлера.

Для полученных экспериментальных данных с помощью программного продукта Microsoft Office Excel были подсчитаны математическое ожидание и мода. Далее, используя формулы математического ожидания и моды (формулы 4.2 и 4.3 соответственно) для гамма-распределения, были вычислены значения параметров k и θ. После подстановки этих значений в формулу 4.1, было получено аппроксимирующее распределение, представленное на рисунке 4.18.

Погрешность аппроксимации равна $2{,}2 \cdot 10^{-4}$, что составляет приблизительно 10%. Следует отметить, что к выводу о том, что распределение степеней связности узлов социальной сети отлично от пуассоновского и степенного, приходили и другие исследователи [128].

$$M = k\theta \tag{4.2}$$

$$Mo = (k-1)\theta \tag{4.3}$$

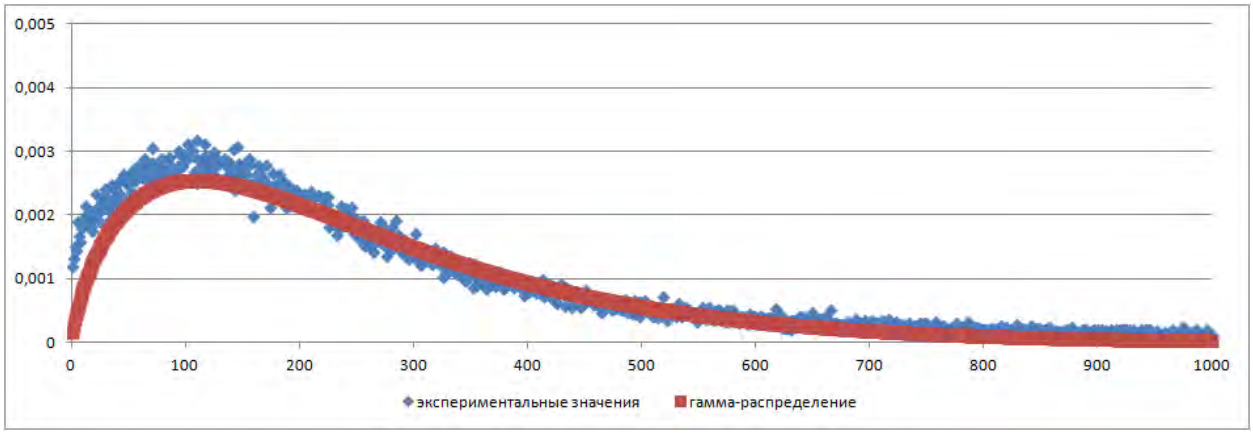


Рисунок 4.18 – Экспериментальное и аппроксимирующее распределения средней степени связности (ВКонтакте)

Рассмотрим результаты вычисления средней степени связности по топологическому срезу социальной сети Facebook. Эксперименты проводилось также с использование двух подходов. При первом подходе средняя степень связности по всем узлам получилась равной 4,15377. Результат представлен на рисунке 4.19. При втором средняя степень связности составила 295,677. Распределение показано на рисунке 4.20.

Аналогично с сетью ВКонтакте получено аппроксимирующее распределение для среза Facebook, которое представлено на рисунке 4.21. Погрешность аппроксимации получилась равной $2 \cdot 10^{-4}$ (≈10%).

Анализируя результаты моделирования распределения степеней связности по двум сетям, можно сделать следующие выводы. В целом характер распределения одинаков для обоих срезов. А средние значения у сети Facebook в два раза меньше, чем у сети ВКонтакте. Этот факт указывает на то, что русская аудитория более коммуникабельная, чем пользователи социальных сетей всей планеты в целом.

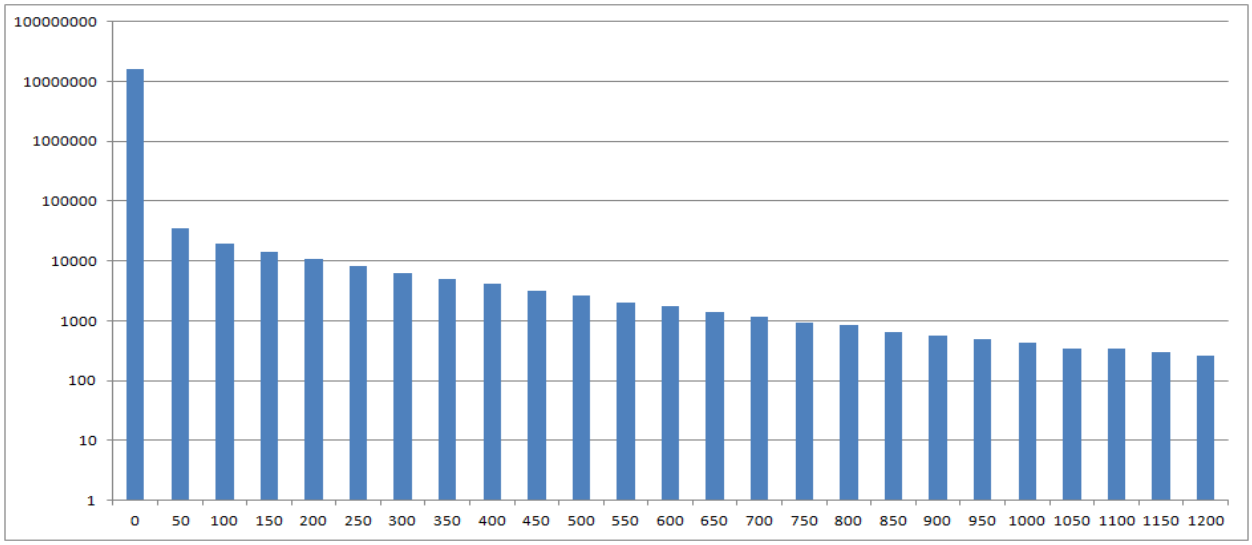


Рисунок 4.19 – Распределение степеней связности по всем узлам (Facebook)

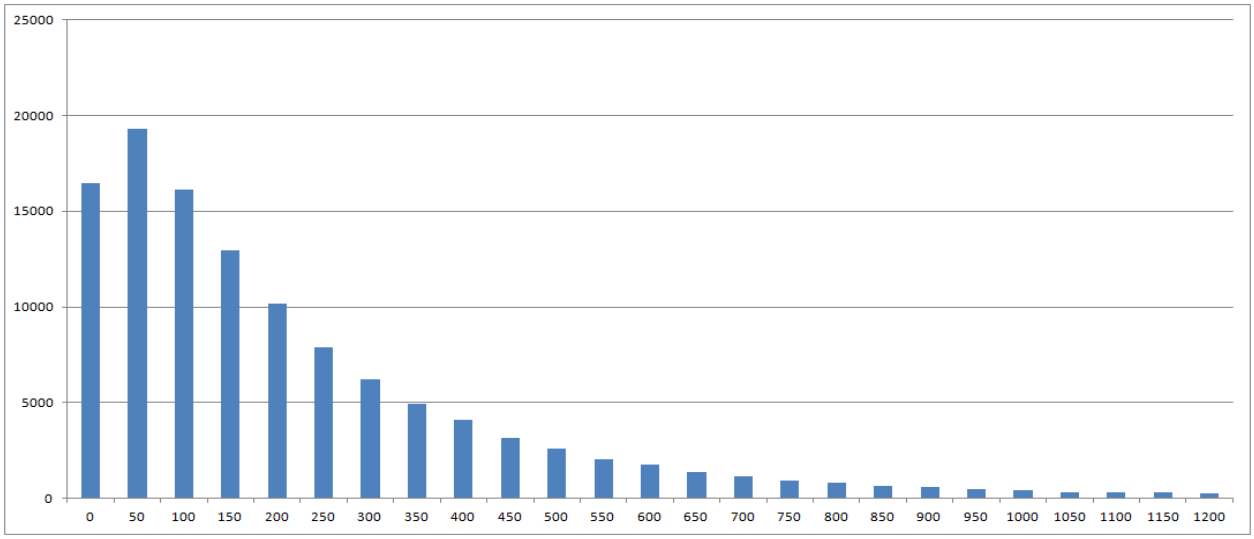


Рисунок 4.20 - Распределение степеней связности по открытым узлам (Facebook)

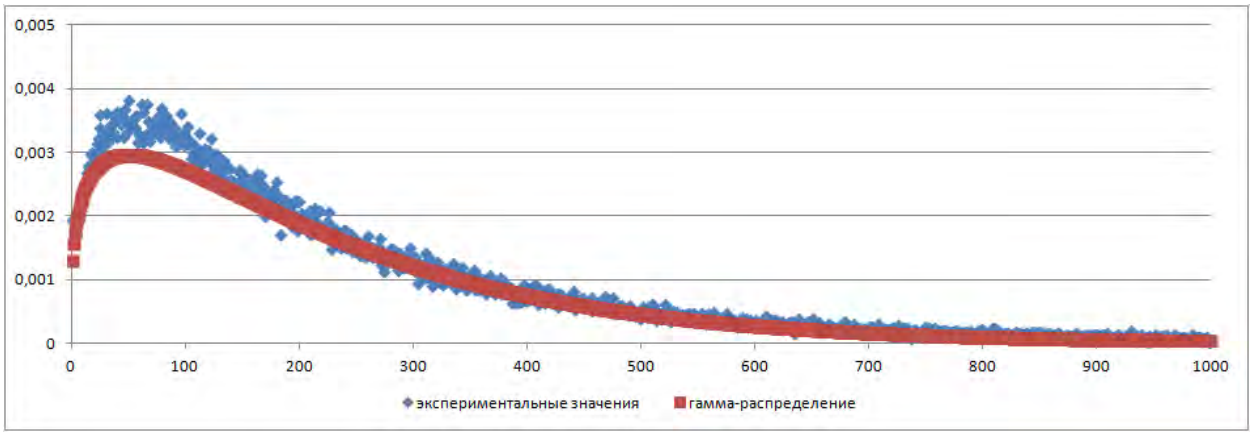


Рисунок 4.21 – Экспериментальное и аппроксимирующее распределения средней степени связности (Facebook)

Кластерный коэффициент сети

Рассмотрим результаты вычисления кластерного коэффициента по топологическому фрагменту сети ВКонтакте. Значение среднего кластерного коэффициента сети получилось равным 0,048087. Подробные данные в виде диапазонов значения коэффициента и количества узлов, попадающих в соответствующие интервалы, приведены в таблице 4.3. На рисунке 4.22 представлено распределение значений кластерного коэффициента (на логарифмической шкале).

Таблица 4.3 - Кластерный коэффициент (ВКонтакте)

Кластерный коэффициент (интервал)	Количество узлов
[0;0,1)	104810
[0,1;0,2)	9094
[0,2;0,3)	2423
[0,3;0,4)	1198
[0,4;0,5)	587
[0,5;0,6)	332
[0,6;0,7)	188
[0,7;0,8)	67
[0,8;0,9)	39
[0,9;1)	8
1	88

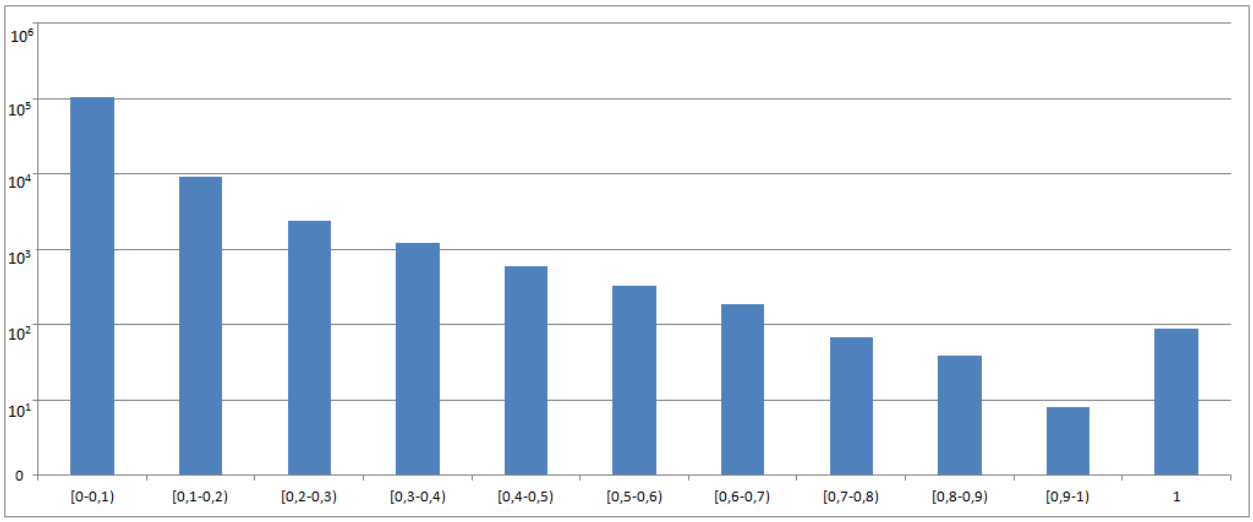


Рисунок 4.22 – Распределение кластерного коэффициента (ВКонтакте)

Анализируя полученные данные, можно сказать следующее. Большинство узлов имеют кластерный коэффициент в интервале [0;0,1), что свидетельствует о низкой степень кластеризации рассматриваемого фрагмента сети. Тем не менее, присутствует группа узлов с коэффициентом равным единице, которая выбивается из зависимости – чем больше значение кластерного коэффициента, тем меньше узлов. Физический смысл этого явления можно объяснить следующим образом. В нашей выборке мы захватили группы пользователей, которые поддерживают тесные связи между собой, например, в связи с родом деятельности. Захват, в свою очередь, таких групп определяется использованным методом выборки – обходом в ширину.

Рассмотрим результаты вычисления кластерного коэффициента по топологическому фрагменту сети Facebook. Значение среднего кластерного коэффициента сети получилось равным 0,040362. Это немного меньше, чем по срезу сети ВКонтакте (0,048087). Подробные данные в виде диапазонов значения коэффициента и количества узлов, попадающих в соответствующие интервалы, приведены в таблице 4.4. На рисунке 4.23 представлено распределение значений кластерного коэффициента (на логарифмической шкале).

Таблица 4.4 - Кластерный коэффициент (Facebook)

Кластерный коэффициент (интервал)	Количество узлов
[0;0,1)	105481
[0,1;0,2)	8006
[0,2;0,3)	2738
[0,3;0,4)	1155
[0,4;0,5)	518
[0,5;0,6)	306
[0,6;0,7)	178
[0,7;0,8)	49
[0,8;0,9)	44
[0,9;1)	17
1	122

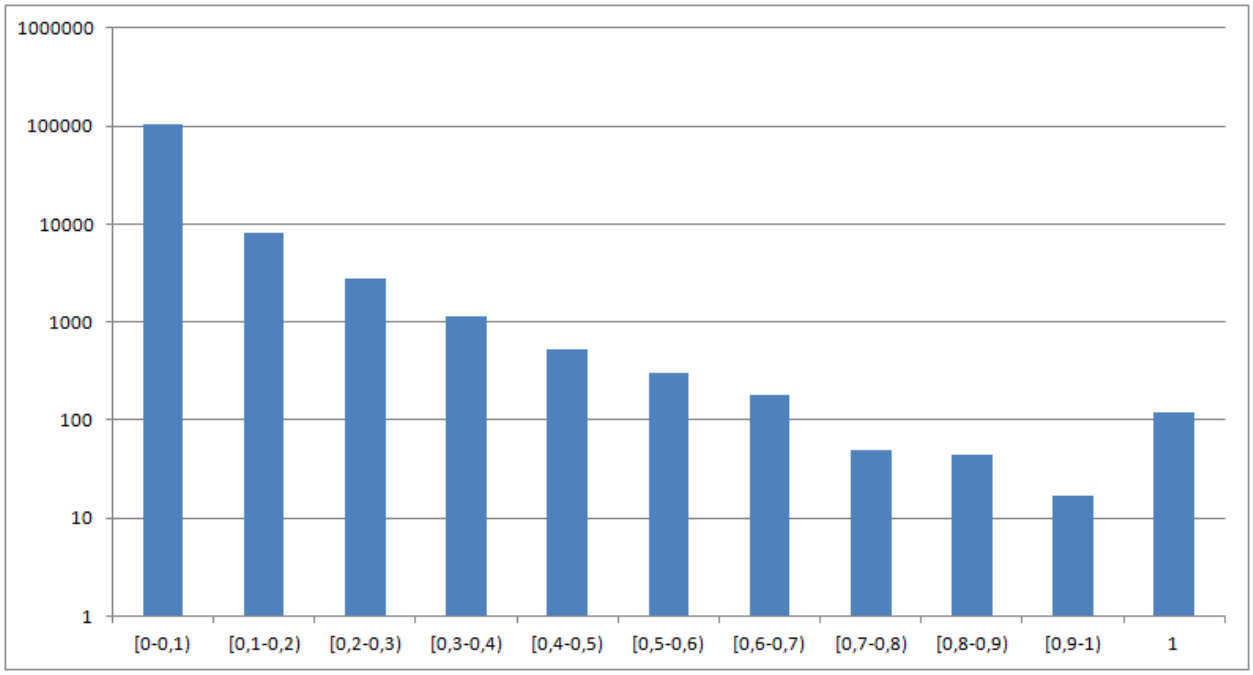


Рисунок 4.23 – Распределение кластерного коэффициента (Facebook)

Результаты моделирования на срезе Facebook аналогичны результатам по сети ВКонтакте, следовательно, характер кластеризации для рассматриваемых сетей одинаков. Если мы применим формулы для вычисления кластерного коэффициента для разных типов сетей из таблицы 4.2, то увидим, что ни одна из них не дает правильного результата для наших случаев.

Средняя длина пути сети

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 [138]. 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.

Сравнивая полученные результаты по топологическим характеристиками и данные в таблице 4.2, можно утверждать, что представленные срезы не относятся ни к одному из типов сложных сетей. В данном случае можно говорить о новом типе, который обладает полученными свойствами и является представителями топологии информационных связей абонентов ИТКС.

4.3 Особенности реализации автоматизированной системы противодействия угрозе распространения запрещенной информации

При наличии административного ресурса можно реализовать автоматизированную систему противодействия угрозе распространения запрещенной информации. Обобщенный алгоритм работы такой системы представлен на рисунке 4.24. 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 (рисунок 4.25) 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 4.26).

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.

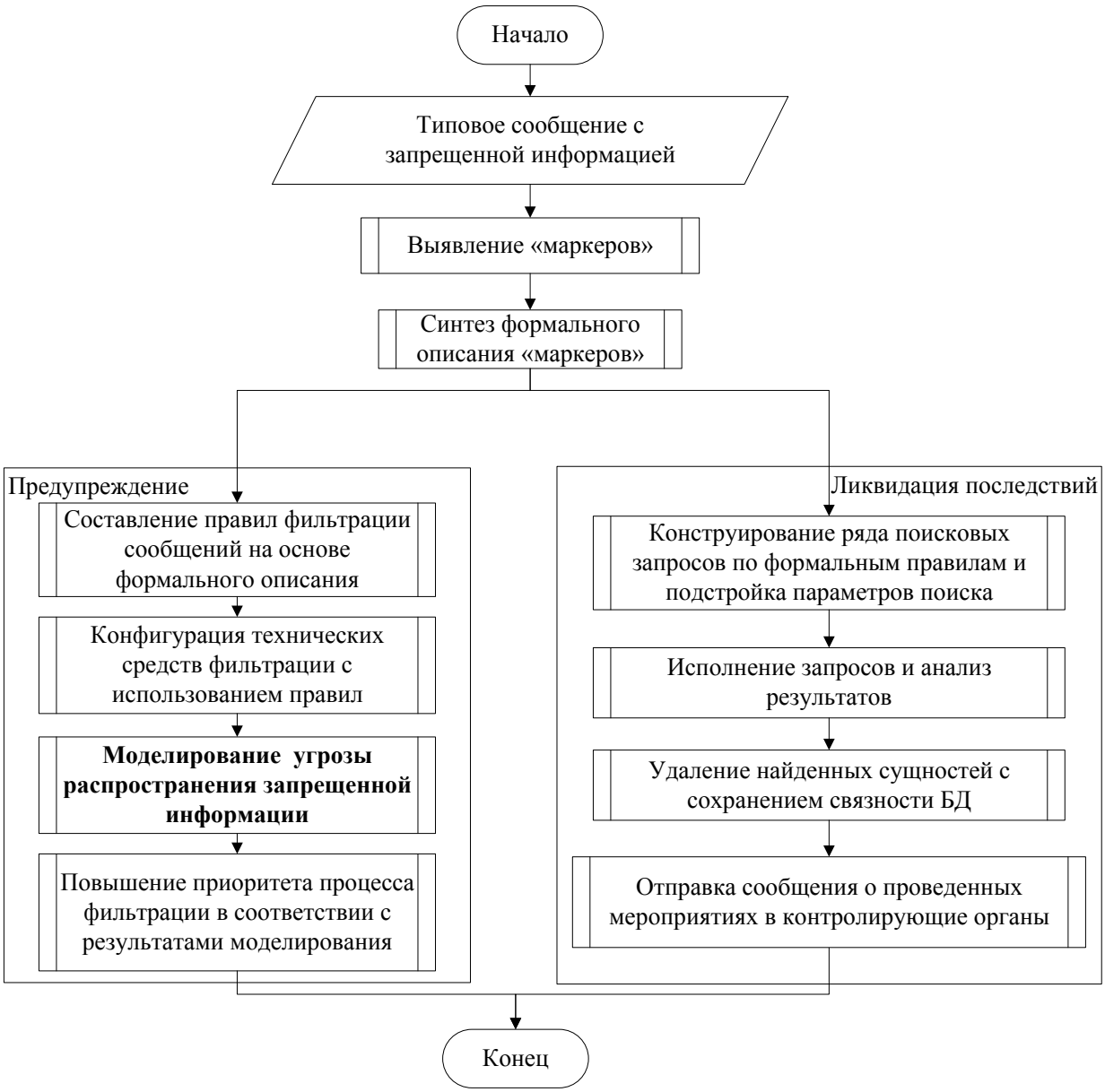


Рисунок 4.24 – Алгоритм противодействия распространению угрозы запрещенной информации

Рисунок 4.25а - Федеральный список экстремистских материалов

416	Номер 42 газеты «Евпатий коловрат», сентябрь- октябрь 2005 года (решение Смольнинского районного суда г. Санкт-Петербурга от 20.05.2008 и определение Смольнинского федерального суда Центрального района г. Санкт-Петербурга от 25.11.2008).
417	Номер 43 газеты «Евпатий Коловрат», ноябрь 2005 года (решение Смольнинского районного суда г. Санкт-Петербурга от 20.05.2008 и определение Смольнинского федерального суда Центрального района г. Санкт-Петербурга от 25.11.2008).
418	Номер 46 газеты «Евпатий Коловрат», март-апрель 2006 года (решение Смольнинского районного суда г. Санкт-Петербурга от 20.05.2008 и определение Смольнинского федерального суда Центрального района г. Санкт-Петербурга от 25.11.2008).
419	Брошюра «Аль-Ваъй» №№ 246-247 - Двадцать первый год издания - август - сентябрь 2007 г. (решение Демского районного суда г. Уфы Республики Башкортостан от 10.07.2009).
420	Брошюра «Аль-Ваъй» № 238 - Двадцать первый год издания - декабрь 2006 г. (решение Демского районного суда г. Уфы Республики Башкортостан от 10.07.2009).
421	Брошюра «Аль-Ваъй» № 238 - Двадцать первый год издания - май 2007 г. (решение Демского районного суда г. Уфы Республики Башкортостан от 10.07.2009).
422	Брошюра «Аль-Ваъй» № 248 - Двадцать второй год издания - октябрь 2007 г. (решение Демского районного суда г. Уфы Республики Башкортостан от 10.07.2009).
423	Книга Такыйюддина Набханий «Экономическая система в исламе» (решение Демского районного суда г. Уфы Республики Башкортостан от 10.07.2009).
424	Брошюра «Политические воззрения Хизб-ут-Тахрир» (решение Демского районного суда г. Уфы Республики Башкортостан от 10.07.2009).
425	Статья, начинающаяся и заканчивающаяся словами: «Если внимательно прочитать комментарии к этой новости, то видно, что мнения разделились примерно поровну….. пока не прекратится тотальное крышевание и взяточничество в правоохранительной Системе - ситуация будет только ухудшаться», комментирующая статью «Нападение скинов во Владивостоке» с Интернет сайта с электронным адресом http://www.ssdv.ru/ (решение Ленинского районного суда г. Владивостока от 18.06.2009).
426	Материал с сайта WWW.SSDV.RU, начинающийся и заканчивающийся словами: «Никогда и нигде еще суть и структура скин движения не разъяснялось так подробно и достоверно…Уничтожение государства, осуществляющего геноцид Русского народа - единственный путь освобождения бойцов Белого Сопротивления, путь восстановления справедливости, путь выживания Русского народа и Белой расы. Слава Руси!», комментирующий книгу «Библия Скинхеда» (решение Ленинского районного суда г. Владивостока от 18.06.2009).
427	Листовки, начинающиеся и заканчивающиеся словами: «Все Мы - Русский уже давно не живем в Своей стране. Мы имеем меньше прав, чем иммигранты… Вставай в наши ряды, участвуй в программах «Славянский Банк», «Земля - Народу» (решение Ленинского районного суда г. Владивостока от 18.06.2009).

Рисунок 4.25б - Федеральный список экстремистских материалов

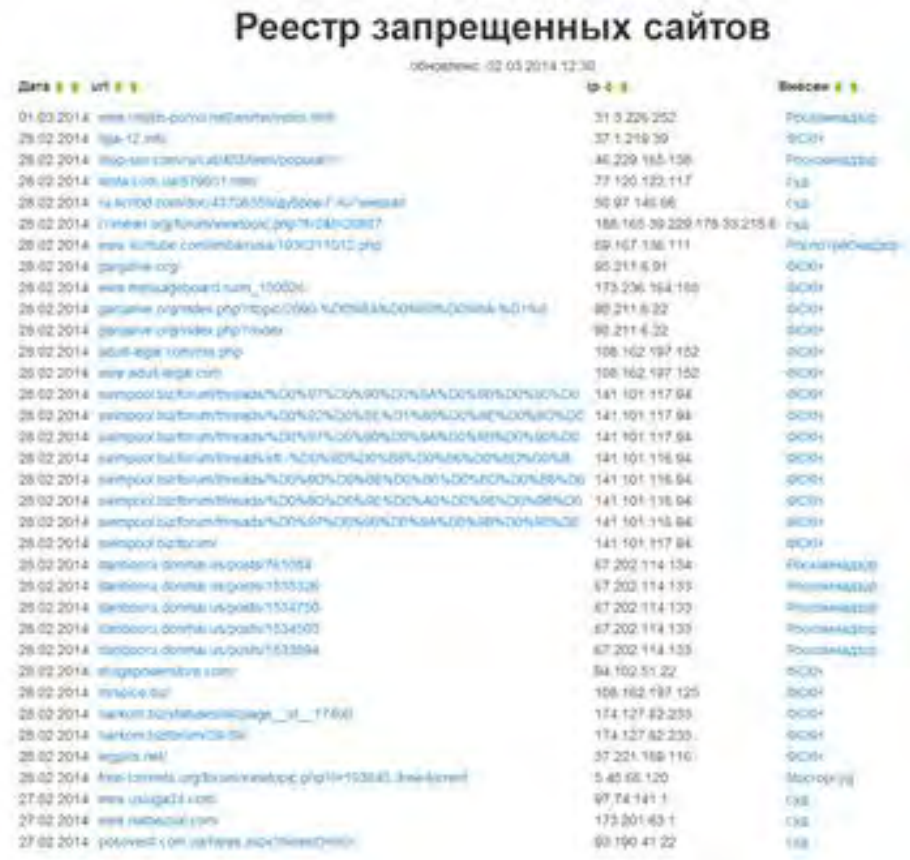


Рисунок 4.26 – Реестр запрещенных сайтов

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. (рисунок 4.27).

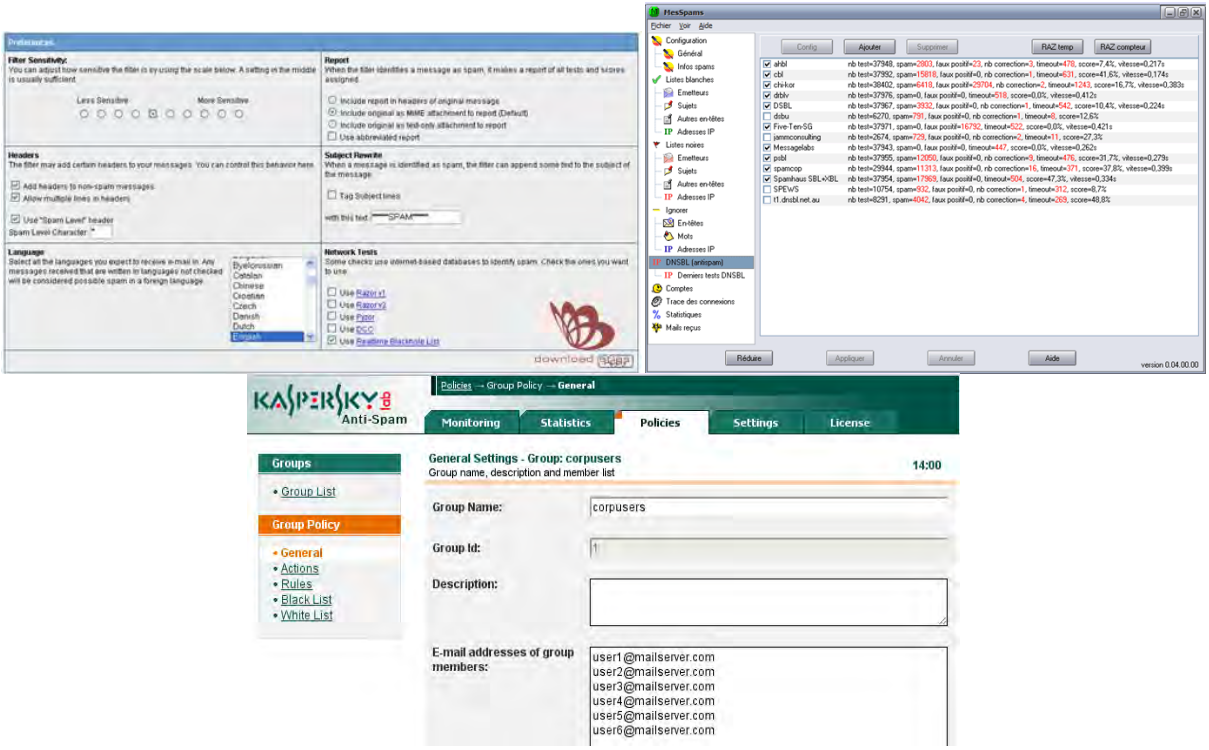


Рисунок 4.27 – Скриншоты программ

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.

4.4 Особенности практического применения аналитической модели УгЗИ в ИТКС

Используя разработанную аналитическую модель, можно получить прогноз по динамике УгЗИ в ИТКС за приемлемое время. Алгоритм получения прогноза состоит из последовательности следующих шагов.

Шаг 1. Определить коэффициент топологической уязвимости рассматриваемой ИТКС. Необходимо постоянно проводить мониторинг значения данного параметра для самых крупномасштабных и популярных сетей для использования его актуального значения.

Шаг 2. При появлении первых сообщений с запрещенной информацией собрать статистику таких сообщений. Данный шаг необходимо выполнить на ранних стадиях возникновения угрозы. С одной стороны, чем больше данных удастся собрать, тем точнее будет прогноз, с другой стороны, при задержке выполнения данного шага, актуальность прогноза может быть потеряна.

Шаг 3. Аппроксимировать собранные данные при помощи системы дифференциальных уравнений, описывающих модель, подобрав нужные значения β и γ (вероятности атаки и защиты).

В результате получаем прогноз на весь период распространения угрозы запрещенной информации.

Аналитическая модель 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) [139]. Результаты распространения данной дезинформации в сети «ВКонтакте» были аппроксимированы с помощью аналитической модели (рисунок 4.28). Погрешность аппроксимации составила приблизительно 13%.

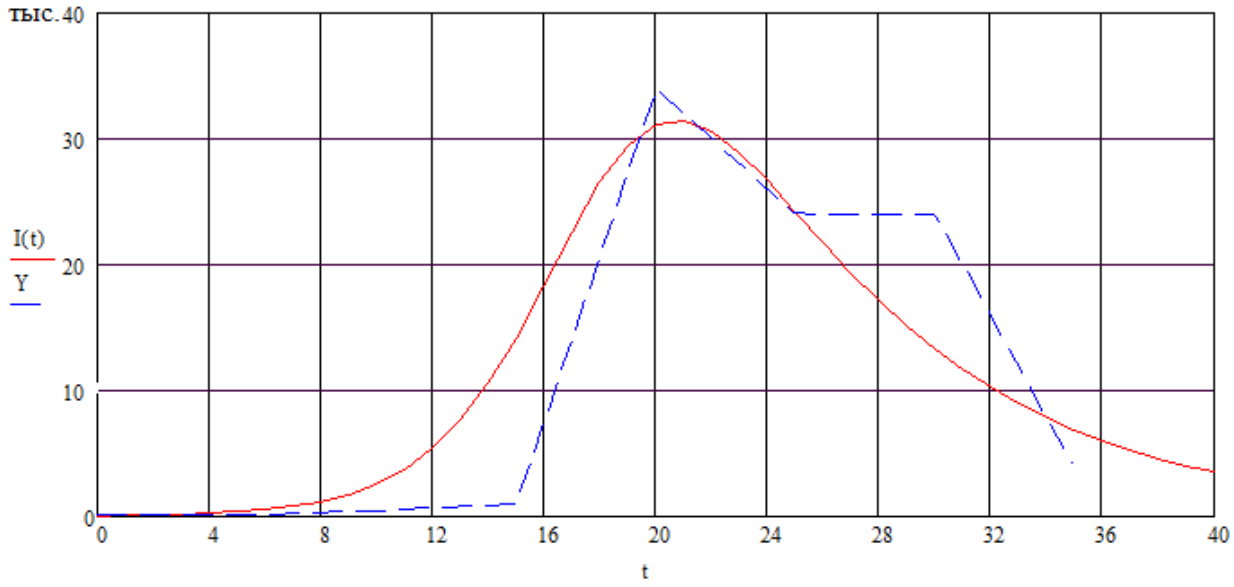


Рисунок 4.28 – Исходные данные (Y) и результат аналитической модели (I), условная единица времени равна 5 часам

4.5 Особенности практического внедрения

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.

На основании научных результатов, приведенных в настоящей работе, разработаны следующие программные продукты:

1. Программа имитационного моделирования распространения запрещенной информации в социальных сетях (свидетельство о государственной регистрации программы для ЭВМ №2011617403 - 23.09.2011).

2. Программа вычисления топологических характеристик социальных сетей (свидетельство о государственной регистрации программы для ЭВМ № 2012610825 - 18.01.2012).

3. Программный комплекс топологического анализа и моделирования распространения запрещенной информации в крупномасштабных социальных сетях (свидетельство о государственной регистрации программы для ЭВМ №2013660757 - 18.11.2013).

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» доработаны курсы: «Математические методы в информационной безопасности», «Экспертные системы комплексной оценки безопасности автоматизированных информационных и телекоммуникационных систем». Научные результаты работы использованы для написания учебных пособий, курсового и дипломного проектирований для студентов кафедры.

Разработанное программное обеспечение было внедрено во владимирском РОСКОМНАДЗОРе с целью автоматизации задачи поиска распространителей запрещенной информации.

В рамках работы был проведены эксперименты по оценке эффективности внедрения разработанного программного продукта. Рассматриваемый период – 3 месяца.

Сотрудник, занимающийся поиском запрещенной информацией, работает по обращениям граждан. Для того, чтобы проанализировать поступившую информацию по одному заявлению, он тратит в среднем 62 минуты. При этом он проверяет корректность текущего обращения и проводит анализ по возможным распространителям запрещенной информации, исходя из контактов злоумышленника. Учитывая среднее количество обращений граждан в месяц, сотрудник тратит в месяц 5*62=310 мин., при этом в среднем за один месяц он находит 60 распространителей запрещенной информации.

При использовании разработанного ПО также начинаем с известного распространителя (по обращению граждан). Пример - владимирский пользователь социальной сети ВКонтакте, распространяющий экстремистские идеи (рисунок 4.29).

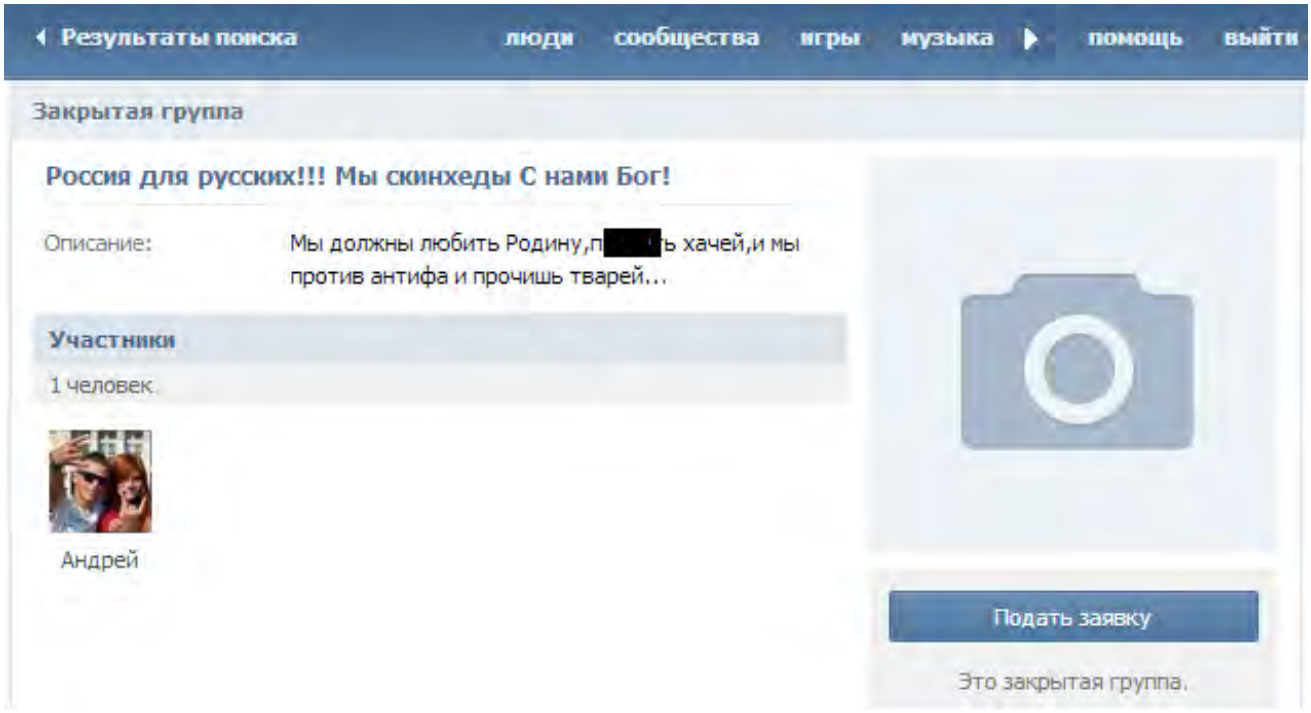


Рисунок 4.29 – Узел распространяющий запрещенную информацию

С помощью созданной программы от данного распространителя получаем топологический срез размером около 100 узлов. На этом фрагменте производим моделирование процесса распространения запрещенной информации. При этом изначальным распространителем отмечаем исходного пользователя, изначально «иммунизированные» узлы не указываются, вероятность «заражения» выставляется равной 0,2, а «иммунизации» 0,3 (исходя из эмпирических данных). Результатом моделирования, нужным для специалиста, является список узлов (id номера пользователей социальной сети), которые могут быть распространителями запрещенной информации. Для просмотра страниц по списку сотрудник РОСКОМНАДЗОРа тратит в среднем 4 минуты. В среднем по каждому инциденту обнаруживается 12 потенциальных распространителей из 100. Использую программное обеспечение тратим 5*4*12=240 минуты. В среднем

получаем выигрыш по времени в 1,3 раз. Использую этот автоматизированный подход, мы имеем дело с ошибками первого и второго рода. Ошибка первого рода заключается в том, что узел, содержащийся в списке потенциальных распространителей, на поверку не являлся таковым. В среднем имеем: из 60 человек 7 человек оказывается не причастным к запрещенной информации. Таким образом, ошибка первого рода составляет ≈12%. Ошибка второго рода заключается в том, что из среза в список не попали распространители. В среднем при анализе 100 пользователей находим 5 человек, которые не попадают в список, но при этом они распространяли запрещенную информацию. Ошибка второго рода составляет 5%. Результаты внедрения показывают, что разработанное программное обеспечение повышает эффективность по времени.

В ОАО «ВПО «Точмаш» внедрена методика формирования топологии информационно-телекоммуникационной сети, предназначенная для реконструирования связей между абонентами сетей в условиях недостатка данных. Внедрение разработки позволяет снизить трудозатраты специалистов при составлении модели внутреннего нарушителя, а также существенно повысить субъективную точность анализа связей персонала объекта с вероятными злоумышленниками.

Выводы к четвертой главе

Разработано программное обеспечение, которое позволяет за приемлемое время получить результаты моделирования УгЗИ в ИТКС за счет использования распределенных вычислительных ресурсов.

С помощью разработанного ПО были проведены экспериментальные исследования, результаты которых показали, что крупномасштабные информационно-телекоммуникационные сети нельзя отнести ни к одному из существующих классов сложных сетей. Эксперименты на топологиях реальных ИТКС показали, что распределение степеней связности узлов сети аппроксимируется гамма-распределением, а не степенным и не пуассоновским распределением, как принято считать. Также результаты экспериментов указывают, что коэффициент кластеризации ИТКС значительно ниже, чем

принято считать (0,04-0,05 против 0,16). Результаты, полученные по значению средней длины пути (подтверждение теории шести рукопожатий), позволяют нам при исследовании крупномасштабных ИТКС использовать фиксированное значение средней длины пути.

В качестве рекомендаций предложен алгоритм работы автоматизированной системы противодействия распространению угрозы запрещенной информации.

Практическая ценность работы заключается в созданном программном обеспечении, задача которого - автоматизация поиска узлов социальной сети, которые являются потенциальными распространителями запрещенной информации. Результаты внедрения показывают, что разработанное программное обеспечение повышает эффективность по времени.

ЗАКЛЮЧЕНИЕ

Информационно-телекоммуникационные сети являются крупномасштабными сетями с постоянно растущим числом абонентов. С бурным ростом числа пользователей ИТКС возникают проблемы информационной безопасности и защиты информации в них.

Анализ проблем информационной безопасности выявил, что кроме проблем, связанных с использованием глобальной сети Интернет как распределенной информационно-телекоммуникационной системы, которые достаточно хорошо известны и решаемы, существует малоизученная проблема запрещенного контента.

Создание моделей и алгоритмов распространения угрозы запрещенной информации – один из ключевых подходов при решении данной задачи. Проведенный анализ публикаций по данной тематике показывает, что существующие решения малоэффективны. Обычно при моделировании распространения угрозы запрещенной информации не учитывается топология ИТКС (модель сети – полносвязный граф). А, если топология учитывается, то, как правило, используется простейшая SIS модель, а структура сети отражается SF сетью. При моделировании УгЗИ важно иметь топологию, отражающую структуру связей реальной сети, а также использовать адекватную модель информационного взаимодействия узлов. Еще одной важной проблемой является крупномасштабность ИТКС, которая мешает получить данные с имитационной модели за приемлемое время. Решение этой задачи состоит в создании аналитической модели УгЗИ в ИТКС.

Разработан алгоритм реализации УгЗИ в ИТКС, основанный на характерах процессов, протекающих в реальных условиях.

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.

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%.

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.

Разработана методика формирования топологии ИТКС, которая учитывает основные топологические характеристики доступной части сети и работает в условие недостаточной репрезентативности выборки исходных данных. Предлагаемая методика состоит из последовательности разработанных алгоритмов.

Создан алгоритм формирования исходных данных о топологии сети (множества вершин и связей между ними доступной части сети), который учитывает ограничения по сбору данных и реализован в виде разработанного программного обеспечения.

Разработан алгоритм формирования полного графа сети с учетом добавления недоступной части на основе вычисленных прогнозируемых топологических характеристик. Алгоритм реализован в виде разработанного программного обеспечения.

Введена оценка топологической уязвимости сети (вектор топологической уязвимости), учитывающая следующие параметры: среднюю длину пути сети, коэффициент кластеризации сети, среднюю степень связности сети и общее количество узлов в сети.

Разработано программное обеспечение, которое позволяет за приемлемое время получить результаты моделирования УгЗИ в ИТКС за счет использования распределенных вычислительных ресурсов.

С помощью разработанного ПО были проведены экспериментальные исследования, результаты которых показали, что крупномасштабные информационно-телекоммуникационные сети нельзя отнести ни к одному из существующих классов сложных сетей. Эксперименты на топологиях реальных ИТКС показали, что распределение степеней связности узлов сети аппроксимируется гамма-распределением, а не степенным и не пуассоновским распределением, как принято считать. Также результаты экспериментов указывают, что коэффициент кластеризации ИТКС значительно ниже, чем принято считать (0,04-0,05 против 0,16). Результаты, полученные по значению средней длины пути (подтверждение теории шести рукопожатий), позволяют нам при исследовании крупномасштабных ИТКС использовать фиксированное значение средней длины пути.

В качестве рекомендаций предложен алгоритм работы автоматизированной системы противодействия распространению угрозы запрещенной информации.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. 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.

2. 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.

3. 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.

4. 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.

5. Абрамов, К.Г., Монахов, Ю.М. Модель распространения спама в социальных сетях [Текст] / К.Г. Абрамов, Ю.М. Монахов; Современные информационные технологии в образовательном процессе и научных исследованиях: Материалы III Международной научно-практической конференции - Шуя-Иваново-Владимир: Изд. ГОУ ВПО "ШГПУ", 2010. - 136 с. - ISBN 978-5-86229-219-0.

6. 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.

7. 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.

8. Абрамов, К.Г., Монахов, Ю.М. Программа для моделирования распространения нежелательной информации в социальных сетях ModelGraph [Электронный ресурс] / К.Г. Абрамов, Ю.М. Монахов; Материалы выставки основных результатов научных исследований, разработок, технического творчества студентов, аспирантов и молодых специалистов - Дни науки студентов ВлГУ - Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых (6-8 апреля 2011 г.). – Режим доступа: http://izi.vlsu.ru/HTC/5.pptx

9. 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

10. Абрамов, К.Г., Монахов, Ю.М. Стохастические модели распространения нежелательной информации в социальных сетях [Текст] / К.Г. Абрамов, Ю.М. Монахов; Сборник научных трудов SWorld. Материалы международной научно-практической конференции «Современные проблемы и пути их решения в науке, транспорте, производстве и образовании '2011», №4. –Одесса: Черноморье, 2011. – 411-0886 – С. 42-46

11. 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.

12. Абрамов, К.Г., Монахов, Ю.М. Модели распространения вредоносных программ в топологически гетерогенных социальных сетях [Электронный ресурс] / К.Г. Абрамов, Ю.М. Монахов; Труды НТС. Комитет по информатизации, связи и телекоммуникациям Администрации Владимирской области. – 2010. – Режим доступа: http://ksi.avo.ru/

13. 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.

14. 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.

15. 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.

16. Алешин, Л.И. Защита информации и информационная безопасность [Текст] / Л.И. Алешин; - М.: МГУК, 1999. - 176 с.

17. Анализатор Sniffer Pro LAN [Электронный ресурс] / Sniffer Technologies. - Режим доступа: http://www.securitylab.ru/software/233623.php

18. Безруков, Н.Н. Компьютерная вирусология [Текст] / Н.Н. Безруков; - Киев:

Укр. сов. энцик., 1991. - 416 с.

19. Биячуев, Т.А. Безопасность корпоративных сетей [Текст]: учеб. пособие / Т.А. Биячуев; под ред. Осовецкого Л.Г. - СПб.: СПбГУ ИТМО, 2004. - 161 с.

20. Бреер, В.В. Стохастические модели социальных сетей [Текст] / В.В. Бреер; Управление большими системами, № 27. – 2009. - С. 169-204.

21. Брэгг, Р., Родс-Оусли, М., Страссберг, К. Безопасность сетей. Полное руководство [Текст] / Р. Брэгг, М. Родс-Оусли, К. Страссберг; - М.: Эком, 2006. - 912 с.

22. Гошко, С.В. Энциклопедия по защите от вирусов [Текст] / С.В. Гошко; - М.: СОЛОН-Р, 2005. - 352 с.

23. Груздева, Л.М, Монахов, Ю.М., Монахов, М.Ю. Оценка сетевых характеристики компьютерных сетей в условиях информационного вредоносного воздействия [Текст]: учеб. пособие (с грифом УМО) / Л.М. Груздева, Ю.М. Монахов, М.Ю. Монахов; Владим. Гос.ун-т. – Владимир: Изд-во Владим. Гос. ун-та, 2010. – 86 с.

24. 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

25. Губанов, Д.А., Новиков, Д.А., Чхартишвили, А.Г. Социальные сети: модели информационного влияния, управления и противоборства [Текст] / Д.А. Губанов, Д.А. Новиков, А.Г. Чхартишвили; под ред. чл.-корр. РАН Д.А. Новикова — М.: Издат. физико-математической литературы, 2010. - 228 с.; - ISBN 9785-94052-194-5.

26. Гусева, А.И. Технология межсетевых взаимодействий [Текст] / А.И. Гусева; - М.: Бином, 1997. - 238 с.

27. Жаринов, И.В., Крылов, В.В. Конструирование графов с минимальной средней длиной пути [Текст] / И.В. Жаринов, В.В. Крылов; Вестник ИжГТУ, №4. – 2008. - С. 164-169. - ISSN 1813-7903.

28. Касперски, К. Жизненный цикл червей [Электронный ресурс] /

К.Касперски; Режим доступа: http://daily.sec.ru/dailypblshow.cfm?rid=9&pid=11697

29. Касперски, К. Компьютерные вирусы: изнутри и снаружи [Текст] / К. Касперски; - Спб: "Питер", 2005. - 528 с.

30. Качалин, А.И. Моделирование процесса распространения сетевых червей для оптимизации защиты корпоративной сети [Текст] / А.И. Качалин; Искусственный интеллект, № 2. – 2006. - С. 84-88.

31. Лукацкий, А. Обнаружение атак [Текст] / А. Лукацкий; - СПб.: БХВ-Петербург, 2001. - 624 с.

32. 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

33. Монахов, Ю.М., Медведникова, М.А. Аналитическая модель дезинформированности узла социальной сети [Текст] / Ю.М. Монахов, М.А. Медведникова; Всероссийский конкурс научно-исследовательских работ студентов и аспирантов в области информатики и информационных технологий в рамках всероссийского фестиваля науки (Белгород, 7 сентября – 9 сентября 2011) - Белгород, 2011. –Т. 1. - 602 с. - С. 595-597

34. Монахов, Ю.М., Медведникова, М.А., Аналитическая модель дезинформированности узла социальной сети [Текст] / Ю.М. Монахов, М.А. Медведникова; ИММОД-2011. - Санкт-Петербург, 2011. – Т. II. – 400 с., - С. 178-180; – ISBN 978-5-905526-02-2.

35. Монахов, Ю.М., Медведникова, М.А., Абрамов, К.Г., Бодров, И.Ю. Аналитическая модель дезинформированности узла социальной сети [Электронный ресурс] / Ю.М. Монахов, М.А. Медведникова, К.Г. Абрамов, И.Ю. Бодров; Комплексная защита объектов информатизации: Труды НТС. - Владимирский государственный университет, 2012. – Режим доступа: http://sntk.vlsu.ru

36. Программное обеспечение Pajek [Электронный ресурс] / Vladimir Batagelj, Andrej Mrvar; - Режим доступа: http://pajek.imfm.si/doku.php

37. Собейкис, В.Г. Азбука хакера 3. Компьютерная вирусология [Текст] / В.Г.

Собейкис; - М.: Майор, 2006. - 512 с.

38. Статистические системы обнаружения вторжений [Электронный ресурс]. - Режим доступа: http://stra.teg.ru/lenta/security/2081

39. Столлингс, В. Основы защиты сетей. Приложения и стандарты [Текст] / В. Столлингс; - М.: Издательский дом "Вильямс", 2002. - 432 с.

40. Чипига, А.Ф., Пелешенко, В. С. Математическая модель процессов связи узлов в сети при обнаружении и предотвращении несанкционированного доступа к информации [Электронный ресурс] / А.Ф. Чипига, В.С. Пелешенко; Режим доступа: http://science.ncstu.ru/articles/ns/002/elen/29.pdf/file_download

41. Чипига, А.Ф., Пелешенко В.С. Формализация процедур обнаружения и предотвращения сетевых атак [Электронный ресурс] / А.Ф. Чипига, В.С. Пелешенко; Режим доступа: http://www.contrterror.tsure.ru/site/magazine8/05-17-Chipiga.htm

42. Чубин, И. ARP-spoofing [Электронный ресурс] / И. Чубин; Режим доступа: http://xgu.ru/wiki/ARP-spoofing

43. Ahn, Y., Han, S., Kwak, H., Moon, S., Jeong, H., Analysis of topological characteristics of huge online social networking services [Text] / Y. Ahn, S. Han, H. Kwak, S. Moon, H. Jeong; 16th International Conference on the World Wide Web. – 2007. - P. 835-844.

44. Albert, R., Barabasi, A., Statistical mechanics of complex networks [Text] / R. Albert, A. Barabasi; Reviews of Modern Physics. – 2002. - Vol. 74, no. 1. - P. 47-97.

45. Albert, R., Jeong, H., Barabasi, A., Diameter World Wide Web [Text] / R. Albert, H. Jeong, A. Barabasi; Nature. - 1999. - Vol. 401, no. 6749. –130 p.

46. Amaral, LAN, Scala, A., Barthelemy, M., Stanley HE (2000) Classes of small-world networks [Text] / Amaral LAN, A. Scala, M. Barthelemy, Stanley HE; Proceedings of the National Academy of Sciences of the United States of America. - 97: 11149

47. Andersson, H., Britton, T. Stochastic Epidemic Models and Their Statistical Analysis [Text] / H. Andersson, T. Britton; Lecture Notes in Statistics. - Springer-Verlag, 2000.

48. Bace R., Mell P. Special Publication on Intrusion Detection Systems. [Text] / R. Bace, P. Mell; Tech. Report SP 800-31; National Institute of Standards and Technology. – 2001.

49. Barabasi, R. Albert Emergence of scaling in random networks [Text] / Albert R. Barabasi; Science. - 1999. – P. 509-512.

50. Barabasi, R.Albert, H.Jeong Physica [Text] / Barabasi, R.Albert, H.Jeong; A 272. - 1999. – P.173.

51. Barillère, R., Baggiolini, V., Beharell, M., Chmielewski, D., Gras, P., Mil-cent, H., Kostro, K., Khomoutnikov, V. Results of the OPC evaluation done within JCOP for the control of the LHC experiments [Text] / R. Barillère [et al.]; International Conference on Accelerator and Large Experimental Physics Control Systems. - Trieste, Italy, 1999. - P. 511-513.

52. Blazek, R.B. A Novel Approach to Detection of «Denial-of-Service» Attacks via Adaptive Sequential and Batch-Sequential Change-Point Detection Methods [Text] / R.B. Blazek; Proc. IEEE Workshop Information Assurance and Security. - IEEE CS Press, 2001. - P. 220–226.

53. Bollobás, B. Random Graphs [Text] / B. Bollobás; Cambridge University Press. - 2001. – 520 p. - ISBN 0521809207.

54. Chwe, M.S. Communication and Coordination in Social Network [Text] / M.S. Chwe; Review of Economic Studies, № 67. - 2000. - P.1-16.

55. Cohen, F. Simulating Cyber Attacks, Defenses, and Consequences [Text] / F. Cohen; IEEE Symposium on Security and Privacy. - Berkeley, 1999.

56. Cohen, R., Havlin, S. Scale-free networks are ultrasmall [Text] / R. Cohen, S. Havlin; Phys. Rev. Lett., 90. - 2003.

57. Deszo, Z., Barabasi, A.L. Halting viruses in scale free networks [Electronic resource] / Z. Deszo, A.L. Barabasi.Access mode: http://www.arxiv.org/PS_cache/cond-mat/pdf/0107/0107420.pdf.

58. Dorogovtsev, S.N., Mendes, J.F.F. Scaling properties of scale-free evolving networks: continuous approach [Text] / S.N. Dorogovtsev, J.F.F. Mendes; Phys. Rev., E 63. - 2001.

59. Dorogovtsev, S.N., Mendes, J.F.F. Evolution of Networks: From Biological Networks to the Internet and WWW [Text] / S.N. Dorogovtsev, J.F.F. Mendes; - Oxford, USA: Oxford University Press, 2003. — 280 p. - ISBN 978-0198515906.

60. Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A. N. Giant strongly connected component of directed networks [Text] / S.N. Dorogovtsev, J.F.F. Mendes, A. N. Samukhin; Phys. Rev., E 64. - 2001.

61. Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A.N. Structure of Growing Networks: Exact Solution of the Barabasi Albert's Model [Text] / S.N. Dorogovtse, J.F.F . Mendes, A. N. Samukhin; Phys. Rev. Lett. 85. - 2000.

62. Easley, D., Kleinberg, J., Networks, Crowds, and Markets Reasoning About a Highly Connected World [Text] / D. Easley, J. Kleinberg; - 2010.

63. Erdös, P., Rényi, A. On the evolution of random graphs [Text] / P. Erdös, A. Rényi; Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5. - 1960. – P. 17-61.

64. Erdös, P., Rényi, A. On random graphs [Text] / P. Erdös, A. Rényi; Publicationes Mathematicae. - 1959. -Vol. 6, no. 26. - P. 290-297.

65. Espinoza, Vicente, Social Networks Among the Urban Poor, in Networks in the Global Village [Text] / Espinoza, Vicente; - 1999.

66. Ferrara, E., Fiumara, G., Topological features of Online Social Networks. Communications on Applied and Industrial Mathematics [Text] / E. Ferrara, G. Fiumara; - 2011.

67. Frauenthal, J.C. [Text] / J.C. Frauenthal; Mathematical Models in Epidemiology. – New York: Springer-Verlag, 1980. – 335 p.

68. Garetto, M., Gong, W., Towsley, D. Modeling Malware Spreading Dynamics [Text] / M. Garetto, W. Gong, D. Towsley; Proc. of 22nd Annual Joint Conference of the IEEE Computer, Communications societies (INFOCOM) (March-April 2003) – 2003.

69. Gjoka, M., Kurant, M., Butts, C. T., Markopoulou, A. A Walk in Facebook: Uniform Sampling of Users in Online Social Networks [Text] / M. Gjoka [et al.]; IEEE INFOCOM '10. IEEE Journal on Selected Areas in Communications - 2010.

70. Gjoka, M., Kurant, M., Butts, C.T., Markopoulou, A. Multigraph Sampling of Online Social Networks [Text] / M. Gjoka [et al.]; IEEE J. Sel. Areas Commun. on Measurement of Internet Topologies - 2011.

71. Gjoka, M., Kurant, M., Butts, C.T., Markopoulou, A. Walking on a Graph with a Magnifying Glass: Stratified Sampling via Weighted Random Walks [Text] / M. Gjoka [et al.]; - Sigmetrics, 2011.

72. Gjoka, M., Kurant, M., Butts, C. T., Markopoulou, A. Walking in Facebook: A Case Study of Unbiased Sampling of OSNs [Text] / M. Gjoka, M. Kurant, C.T. Butts, A. Markopoulou; IEEE INFOCOM (San Diego, CA, 2010) – 2010.

73. Gjoka, M., Kurant, M., Wang, Y., Almquist, *Z*.W., Butts, C.T., Markopoulou, A. Coarse-Grained Topology Estimation via Graph Sampling [Electronic resource]: Arxiv preprint / M. Gjoka [et al.]; - 2011. -arXiv:1105.5488

74. Gjoka, M., Sirivianos, M., Markopoulou, A., Yang, X. Poking facebook: characterization of osn applications [Text] / M. Gjoka [et al.]; Proc. of WOSN - 2008.

75. Golbeck, J., Hendler, J. Inferring binary trust relationships in web-based social networks [Text] / J. Golbeck, J. Hendler; Transactions on Internet Technology - 2006. - Vol. 6, no. 4. - P. 497-529.

76. Goldenberg, J., Libai, B., Muller, E. Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth [Text] / J. Goldenberg, B. Libai, E. Muller; Marketing Letters - 2001. - № 2. - P. 11-34.

77. Granovetter, M. The strength of weak ties [Text] / M. Granovetter; American Journal of Sociology - 1973. – Vol. 78. – P. 1360–1380.

78. Granovetter, M. Threshold Models of Collective Behavior [Text] / M. Granovetter; American Journal of Sociology - 1978. – Vol. 83, no. 6. – P. 1420–1443.

79. Grimaldi, R. P. Discrete and Combinatorial Mathematics [Text] / R.P. Grimaldi; an applied introduction. - 4th edition. - New York, 1998.

80. Heberlein, L.T., Dias, G.V., Levitt, K.N., Mukherjee, B., Wood, J., Wolber, D.A.

Network security monitor [Text] / L.T. Heberlein [et al.]; Proc. of IEEE Symposium on Re-search in Security and Privacy. – Los Alamitos, CA, USA: IEEE Computer Society, 1990. - P. 296–304.

81. Hethcote, H.W. The Mathematics of Infectious Diseases [Text] / H.W. Hethcote; - 2000. - P. 599-653,

82. Hofmeyr, S.A., Forrest, S., Somayaji, A. Intrusion detection using se-quences of system calls [Text] / S.A. Hofmeyr, S. Forrest, A. Somayaji; Journal of Computer Security. - Amsterdam: IOS Press, 1998. – Vol. 6, no 3. - P. 151-180.

83. Janky, B., Takacs, K. Social Control, Participation in Collective Action and Network Stability [Text] / B. Janky, K. Takacs; HUNNET Working Paper. - 2002.

84. Jung, J., Krishnamurthy, B., Rabinovich, M. Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites [Electronic resource] / J. Jung, B. Krishnamurthy, M. Rabinovich; WWW2002 (May 7-11, 2002) – Honolulu, Hawaii, USA, 2002. - Access mode: http://www.research.att.com/~bala/papers/www02-fc.html

85. Kasturirangan, R. Multiple Scales in Small-World Networks [Text] / R. Kasturirangan; Brain and Cognitive Science Department, MIT. - 1999.

86. Kenah, E., Robins, J. M. Network-based analysis of stochastic SIR epidemic models with random and proportionate mixing [Text] / E. Kenah, J. M. Robins; Departments of Epidemiology and Biostatistics Harvard School of Public Health. - 2007.

87. Kephart, J.O., White, S.R. Directed-Graph Epidemiological Models of Computer Viruses [Text] / J.O. Kephart, S.R. White; Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy. -1991. P. 343 - 359.

88. Kolotov, A. Мониторинг сети с помощью tcpdump [Electronic resource] / A. Kolotov; - Access mode: http://www.linuxshare.ru/docs/net/tcpdump.html

89. Kulkarni, R.V., Almaas, E., Stroud, D. Evolutionary dynamics in the Bak-Sneppen model on small-world networks [Text] / R.V. Kulkarni, E. Almaas, D. Stroud; - 2008.

90. Kumar, R., Novak, J., Tomkins, A. Structure and evolution of online social networks [Text] / R. Kumar, J. Novak, A. Tomkins; Link Mining: Models, Algorithms, and Applications. - 2010. - P. 337-357.

91. Kuperman, M., Abramson, G. Small world effect in and epidemiological model [Text] / Kuperman M., Abramson G.; Physical Review Letters. - 2001. –Vol.86, no 13.

92. Kurant, M., Markopoulou, A., Thiran, P. On the bias of BFS (Breadth First Search) [Text] / M. Kurant, A. Markopoulou, P. Thiran; in Proc. 22nd Int. Teletraffic Congr.; - 2010. - arXiv:1004.1729.

93. Leskovec, J., Faloutsos, C., Sampling from large graphs [Text] / J. Leskovec, C. Faloutsos; 12th International Conference on Knowledge Discovery and Data Mining. - 2006. - P. 631-636.

94. Leskovec, J., Adamic, L.A., Huberman, B.A. The Dynamics of Viral Marketing [Text] / J. Leskovec, L.A. Adamic, B.A. Huberman; HP Labs Palo Alto, CA 94304. - 2008.

95. Leskovec, J., Kleinberg, J., Faloutsos, C. Graphs over time: densification laws, shrinking diameters and possible explanations [Text] / J. Leskovec, J. Kleinberg, C. Faloutsos; in Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining. - ACM, 2005. - P. 177-187.

96. Leveille, J. Epidemic Spreading in Technological Networks [Text] / J. Leveille; Information Infrastructure Laboratory HP Laboratories Bristol. - 2002. – P. 65-76.

97. Liben-Nowell D., Kleinberg J. The link-prediction problem for social networks [Text] / D. Liben-Nowell, J. Kleinberg; J. American Society for Information Science and Technology. - 2007. - Vol. 58, no. 7. - P. 1019-1031.

98. Mislove, A., Marcon, M., Gummadi, K., Druschel, P., Bhattacharjee, B. Measurement and analysis of online social networks [Text] / A. Mislove [et al.]; 7th ACM conference on Internet measurement. - 2007. - P. 29-42.

99. Newman, M.E.J., Barkema, G. T. New Monte Carlo algorithms for classical spin systems [Text] / M.E.J. Newman, G.T. Barkema; - New York, 1999.

100. Newman, M.E.J., Strogatz, S.H., Watts, D.J. Random graphs with arbitrary degree distributions, their applications [Text] / M.E.J. Newman, S.H. Strogatz, D.J. Watts; Phys. Rev. - 2001. – Vol. E 64.

101. Newman, M.E.J., Barkema, G.T. Monte Carlo Methods in Statistical Physics [Text] / M.E.J. Newman, G.T. Barkema; Oxford University Press. - Oxford, 1999.

102. Newman, M.E.J. Models of the small world [Text] / M.E.J. Newman; J. Stat. Phys - 2000. – P. 819-841

103. Newman, M.E.J. The spread of epidemic disease on networks [Text] / M.E.J. Newman; Physical Review E. - 2002. – P. 16-128.

104. Newman, M.E.J., Watts, D.J., Moore, C. Meaneld solution of the small-world network model [Text] / M.E.J. Newman [et al.]; Phys. Rev. Lett. - 2000.

105. Newman, M.E.J., Watts, D.J. Renormalization group analysis of the small-world network model [Text] / M.E.J. Newman, D.J. Watts; Phys.Lett. A 263. - 1999. – P. 341-346.

106. Newman, M.E.J., Watts, D.J. Scaling and percolation in the small-world network model [Text] / M.E.J. Newman, D.J. Watts; Physical Review E. - 1999. – Vol. 60.

107. Newman, M.E.J., Jensen, I., *Z*iff, R.M. Percolation, epidemics in a two-dimensional small world [Text] / M.E.J. Newman, I. Jensen, R.M. *Z*iff; Phys.Rev. E 65. - 2002.

108. Newman, M.E.J. The Structure and Function of Complex Networks [Text] / M.E.J. Newman; SIAM REVIEW. - 2003. - Vol. 45, No. 2. - P. 167-256.

109. Newman, M.E.J., *Z*iff, R.M. Efficient Monte Carlo algorithm high-precision results for percolation [Text] / M.E.J. Newman, R.M. *Z*iff; Phys. Rev. Lett. 85. – 2000.

110. Newman, Mark, Barabasi, Albert-La*sz*lo, Duncan, Watts, J. The Structure and Dynamics of Networks: (Princeton Studies in Complexity) [Text] / Mark Newman, Albert-La*sz*lo Barabasi, Duncan, J. Watts; — Princeton, USA: Princeton University Press, 2006. — 624 p. — ISBN 978-0691113579.

111. Parallel Boost Graph Library [Electronic resource] / - Access mode: http://www.boost.org/doc/libs/1_55_0/libs/graph_parallel/doc/html/index.html

112. Pastor-Satorras, R., Vespignani, A. Epidemic Spreading in Scale-Free Networks [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. Lett., 86. – 2001.

113. Pastor-Satorras, R., Vespignani, A., Absence of epidemic threshold in scale-free networks with connectivity correlations [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. Lett. – Pub.: American Physical Society, 2002. – Vol. 90, Iss. 2. – P. 1-4.

114. Pastor-Satorras, R., Vespignani, A. Critical load, congestion instabilities in scale-free networks [Text] / R. Pastor-Satorras, A. Vespignani; Europhys. Lett. - 2002. - Vol. 62. - P. 292.

115. Pastor-Satorras, R., Vespignani, A. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks [Text] / R. Pastor-Satorras, A. Vespignani; Journal of Theoretical Biology. - 2005. – P. 275-288.

116. Pastor-Satorras, R., Vespignani, A. Emergence of clustering correlations communities in a social network model [Text] / R. Pastor-Satorras, A. Vespignani; - 2003.

117. Pastor-Satorras, R., Vespignani, A. Epidemic dynamics, endemic states in complex networks [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. E. - 2001.

118. Pastor-Satorras, R., Vespignani, A. Epidemic spreading in complex networks with degree correlations [Text] / R. Pastor-Satorras, A. Vespignani; Contribution to the Proceedings of the XVIII Sitges Conference "Statistical Mechanics of Complex Networks". - Berlin, 2003.

119. Pastor-Satorras, R., Vespignani A. Epidemic spreading in correlated complex networks [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys. - 2002.

120. Pastor-Satorras, R., Vespignani, A. Immunization of complex networks [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. E. - 2002.

121. Pastor-Satorras, R., Vespignani, A. Large-scale topological, dynamical properties of the Internet [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. E. – 2002. - Vol. 65.

122. Pastor-Satorras, R., Vespignani, A. Reaction-diffusion processes, meta-population models in heterogeneous networks [Text] / R. Pastor-Satorras, A. Vespignani; Nature Physics 3. - 2007. – P. 276-282.

123. Pastor-Satorras, R., Vespignani, A. Velocity, hierarchical spread of epidemic outbreaks in scale-free networks [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. Lett. – 2004. - Vol. 92. - P. 178-701.

124. Pastor-Satorras, R., Vespignani, A. Dynamical, Correlation Properties of the Internet [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. Lett. – 2001. - Vol. 87, No. 258701.

125. Pastor-Satorras, R., Vespignani, A. Epidemic dynamics in finite size scale-free networks [Text] / R. Pastor-Satorras, A. Vespignani; Phys. Rev. E. - 2002.

126. Pastor-Satorras, R., Vespignani, A. Topology, Hierarchy, Correlations in Internet Graphs [Text] / R. Pastor-Satorras, A. Vespignani; Lecture Notes in Physics. - Berlin – Heidelberg: Springer, 2004. – P. 425-440.

127. Roberts, M.G., Heesterbeek, JAP Mathematical models in epidemiology [Text] / M.G. Roberts, Heesterbeek JAP; In JA. Filar (Ed.) Mathematical Models. Oxford: EOLSS Publishers Ltd, 2004.

128. Sala, A., Zheng, H., Zhao, Gaito, S., Rossi Brief announcement: revisiting the power-law degree distribution for social graph analysis [Text] / A. Sala, H. Zheng, Zhao, S. Gaito, Rossi; PODC. – 2010. – P. 400-401.

129. Tarnow, E. Like Water and Vapor, Conformity and Independence in the Large Group [Electronic resource] / E. Tarnow; Access mode: URL: http:/cogprints.org/4274/l/LargeGroupOrderTarnow.pdf.

130. Tictrac [Electronic resource] / Access mode: https://www.tictrac.com

131. Ugander, J., Karrer, B., Backstrom, L., Marlow, K. The Anatomy of the Facebook Social Graph [Text] / J. Ugander, B. Karrer, L. Backstrom, K. Marlow; CoRR . - 2011.

132. Volz, E. SIR dynamics in random networks with heterogeneous connectivity [Text] / E. Volz; Journal of Mathematical Biology manuscript. - 2007.

133. Wang, H., Guo, Y. Consensus on scale-free network [Text] / H. Wang, Y. Guo; American Control Conference. - 2008. - P.748 – 752.

134. Watts, D., Strogatz, S. Collective dynamics of small-world networks [Text] / D. Watts, S. Strogatz; Nature. – 1998. - Vol. 393, No. 6684. - P. 440-442.

135. Williamson, M.M., Léveillé, J. An epidemiological model of virus spread and cleanup [Text] / M.M. Williamson, J. Léveillé; Information Infrastructure Laboratory HP Laboratories Bristol HPL. - 2003.

136. Zhang D., Gatica-Perez D., Bengio S., Roy D. Learning Influence among Interacting Markov Chains [Text] / D. Zhang, D. Gatica-Perez, S. Bengio, D. Roy; Neural Information Processing Systems (NIPS). -2005. - P. 132-141.

137. [Электронный ресурс] / Режим доступа: http://vk.com/about

138. [Электронный ресурс] / Режим доступа: http://ria.ru/science/20111122/495222495-print.html

139. [Электронный ресурс] / Режим доступа: http://www.sostav.ru/articles/rus/2011/11.08/news/expressmonitoringnestlebananarus-110809042619-phpapp01.pdf

ПРИЛОЖЕНИЕ А

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2011617403

Программа имитационного моделирования распространения нежелательной информации в социальных сетях

Правообладатель(ли): ***Государственное образовательное учреждение высшего профессионального образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (RU)***

Автор(ы): ***Абрамов Константин Германович,***
Монахов Юрий Михайлович (RU)

Заявка № **2011615699**
Дата поступления **26 июля 2011 г.**
Зарегистрировано в Реестре программ для ЭВМ
23 сентября 2011 г.

Руководитель Федеральной службы по интеллектуальной собственности, патентам и товарным знакам

Б.П. Симонов

ПРИЛОЖЕНИЕ Б

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2012610825

Программа вычисления топологических характеристик социальных сетей

Правообладатель(ли): ***Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (RU)***

Автор(ы): ***Бодров Иван Юрьевич, Монахов Юрий Михайлович, Абрамов Константин Германович (RU)***

Заявка № **2011618861**

Дата поступления **22 ноября 2011 г.**

Зарегистрировано в Реестре программ для ЭВМ

18 января 2012 г.

Руководитель Федеральной службы по интеллектуальной собственности

Б.П. Симонов

ПРИЛОЖЕНИЕ В

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2013660757

Программный комплекс топологического анализа и моделирования распространения запрещенной информации в крупномасштабных социальных сетях

Правообладатель: ***Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (RU)***

Авторы: ***Малышев Роман Владимирович (RU), Монахов Юрий Михайлович (RU), Абрамов Константин Германович (RU)***

Заявка № **2013618490**

Дата поступления **20 сентября 2013 г.**

Дата государственной регистрации в Реестре программ для ЭВМ ***18 ноября 2013 г.***

Руководитель Федеральной службы по интеллектуальной собственности

Б.П. Симонов

ПРИЛОЖЕНИЕ Г

Фрагмент кода из файла DistributedGraph_main.cpp

```
void SIR_modelling(DistributedGraph& g,
		vector<Node::id_type> preinf,
		vector<Node::id_type> prerec,
		double inf_prob,
		double rec_prob,
		string outFile)
	{
		mpi::communicator world;

		if(main_process(g))
			log_mpi() << "Preparing preinfected and prerecovered nodes..." << endl;
		auto sir = get(&Node::sir, g);
		if(main_process(g)) {
			BOOST_FOREACH(Node::id_type v, preinf) {
				put(sir, add_vertex(v, g), Node::SIR::I);
			}
			BOOST_FOREACH(Node::id_type v, prerec) {
				put(sir, add_vertex(v, g), Node::SIR::R);
			}
		}
		synchronize(g.process_group());
		if(main_process(g))
			log_mpi() << "Finished." << endl;

		SIRHistory sirHistory;
		SIRManager sirManager(g, inf_prob, rec_prob);
		for(unsigned iter = 0;; ++iter) {
			log_mpi() << "Iteration: " << iter << endl;
			SIRVector condition(3);
			vector<DistributedGraph::vertex_descriptor> infected;
			vector<DistributedGraph::vertex_descriptor> potentially_susceptible;
			BGL_FORALL_VERTICES(v, g, DistributedGraph) {
				if(g[v].sir == Node::SIR::I) {
					infected.push_back(v);
					BGL_FORALL_ADJ(v, av, g, DistributedGraph) {
						// Собираем всех соседей зараженных узлов, в том числе
и со
						// статусами отличными от 'S', поскольку получение
инфо о статусе узла
						// из другого процесса занимает большое количество
времени
						if(av.owner != g.processor() || av.owner ==
g.processor() && g[av].sir == Node::SIR::S)
							potentially_susceptible.push_back(av);
					}
					condition[Node::SIR::I]++;
				} else if(g[v].sir == Node::SIR::S) {
					condition[Node::SIR::S]++;
				} else if(g[v].sir == Node::SIR::R) {
					condition[Node::SIR::R]++;
				}
			}

			sirHistory.push_back(condition);
			log_mpi() << "SIR condition: " << condition << endl;

			// Принимаем решение о прекращении моделирования на основании отсутствия
зараженных узлов
			bool infected_exists = !infected.empty();
```

```
                bool cont;
                if(main_process(g)) {
                        mpi::reduce(world, infected_exists, cont, logical_or<bool>(), 0);
                } else {
                        mpi::reduce(world, infected_exists, logical_or<bool>(), 0);
                }
                mpi::broadcast(world, cont, 0);
                // Прекращаем моделирование
                if(!cont) break;

                // Пробуем заразить
                BOOST_FOREACH(auto v, potentially_susceptible) {
                        sirManager.infect(v);
                }
                // Пробуем вылечить
                BOOST_FOREACH(auto v, infected) {
                        sirManager.recover(v);
                }

                // End of BSP superstep
                synchronize(g.process_group());
        }

        log_mpi() << "finishing..." << endl;
        if(main_process(g)) {
                SIRHistory res;
                mpi::reduce(world, sirHistory, res, SIRHistoryReducer(), 0);
                ofstream os(outFile);
                os << res;
        } else {
                mpi::reduce(world, sirHistory, SIRHistoryReducer(), 0);
        }
    }
```

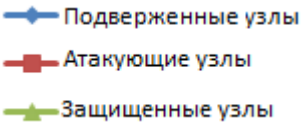
Фрагмент кода из файла aux_types.h

```
struct SIRManager {
    private:
        typedef Node::id_type id_type;

    public:
        SIRManager(DistributedGraph& g, double inf_prob, double rec_prob)
            : g(g),
            process_group(g.process_group(),
graph::parallel::attach_distributed_object()),
            sir(get(&Node::sir, g)),
            inf_gen(std::time(0)),
            rec_gen(std::time(0))
        {
            graph::parallel::simple_trigger(process_group, infect_tag, this,
&SIRManager::trigger);
            double inf_probs[] = {1.0 - inf_prob, inf_prob};
            double rec_probs[] = {1.0 - rec_prob, rec_prob};
            inf = random::discrete_distribution<>(inf_probs);
            rec = random::discrete_distribution<>(rec_probs);
        }

        void infect(const DistributedGraph::vertex_descriptor& v){
            using namespace std;
            if(v.owner == g.processor()) {
                infect_local(v);
```

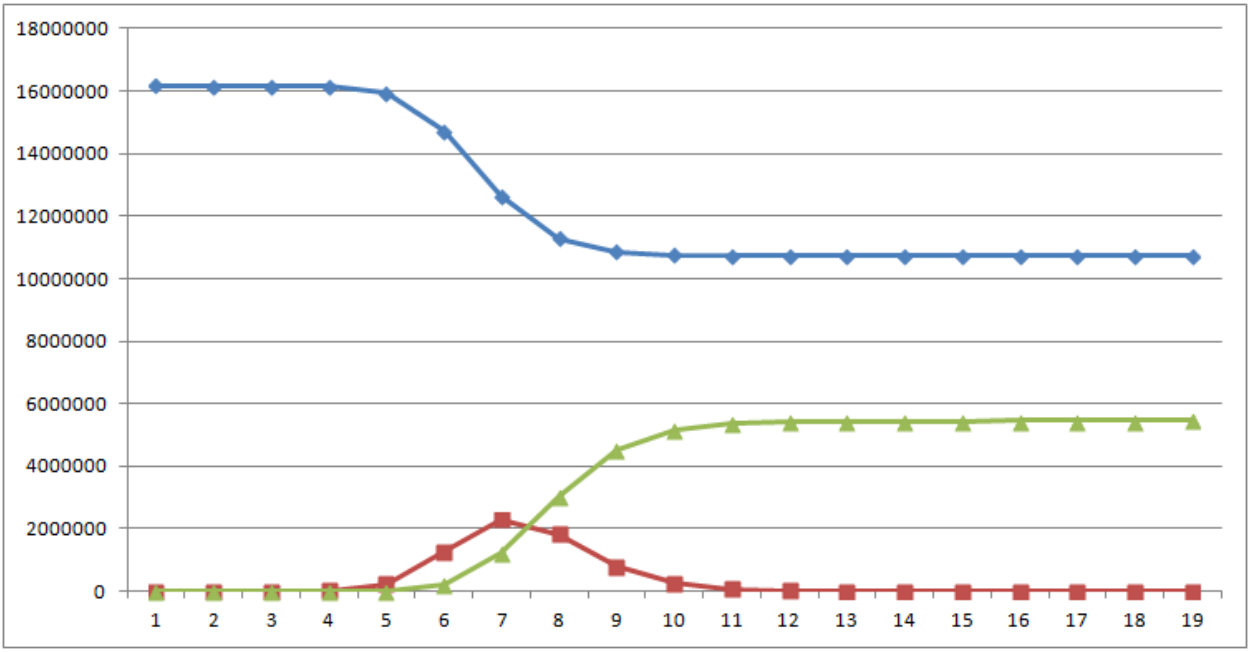
```
                } else {
                        graph::distributed::send(process_group, v.owner, infect_tag, v);
                }
        }

        void recover(const DistributedGraph::vertex_descriptor& v) {
                // На текущий момент оздоравляются только локальные узлы
                recover_local(v);
        }

private:
        void trigger(int source, int tag, const DistributedGraph::vertex_descriptor& v,
                graph::parallel::trigger_receive_context context)
        {
                using namespace std;
                infect_local(v);
        }

        void infect_local(const DistributedGraph::vertex_descriptor& v){
                if(g[v].sir == Node::SIR::S && inf(inf_gen)) {
                        put(sir, v, Node::SIR::I);
                }
        }

        void recover_local(const DistributedGraph::vertex_descriptor& v){
                if(g[v].sir == Node::SIR::I && rec(rec_gen)) {
                        put(sir, v, Node::SIR::R);
                }
        }

        DistributedGraph& g;
        DistributedGraph::process_group_type process_group;

        property_map<DistributedGraph, Node::SIR Node::*>::type sir;

        random::mt19937 inf_gen;
        random::mt19937 rec_gen;
        random::discrete_distribution<> inf;
        random::discrete_distribution<> rec;

        enum Tags {
                infect_tag
        };
};
```

ПРИЛОЖЕНИЕ Д

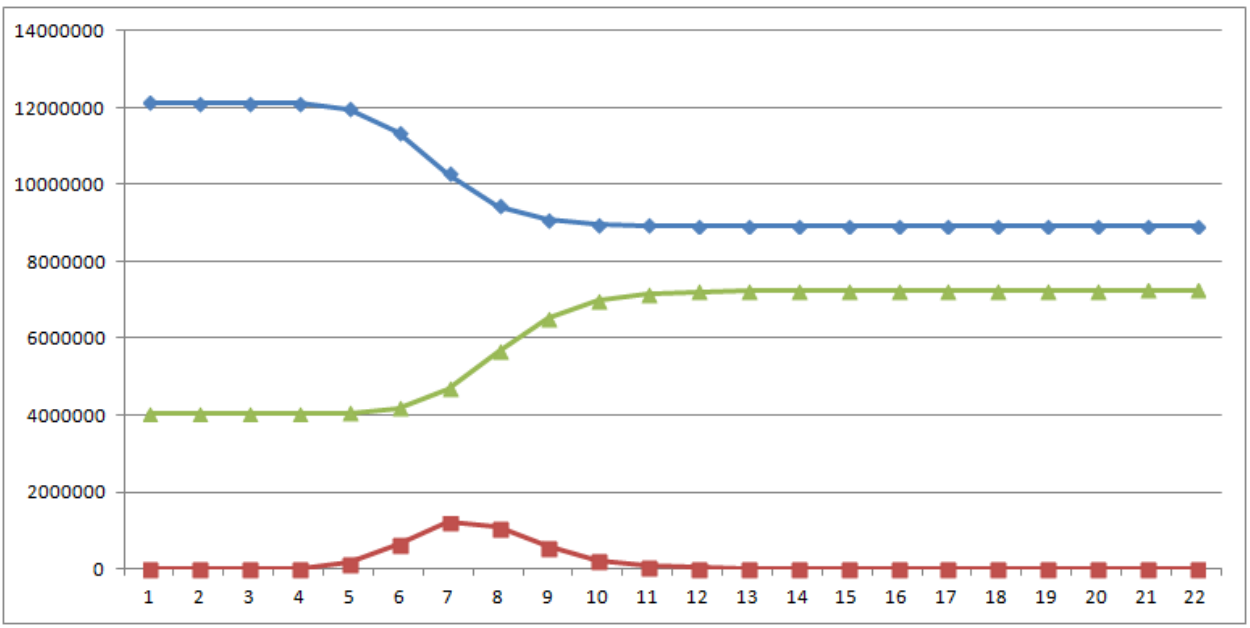
Легенда



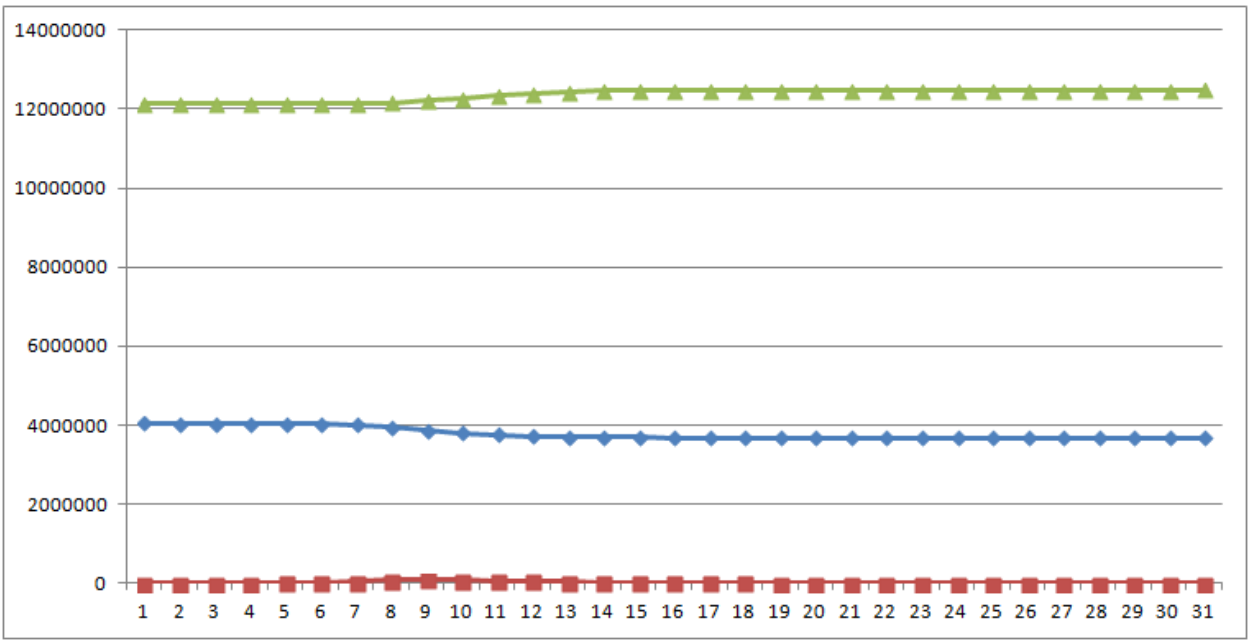
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 0{,}001N$, $R_0 = 0{,}75N$.



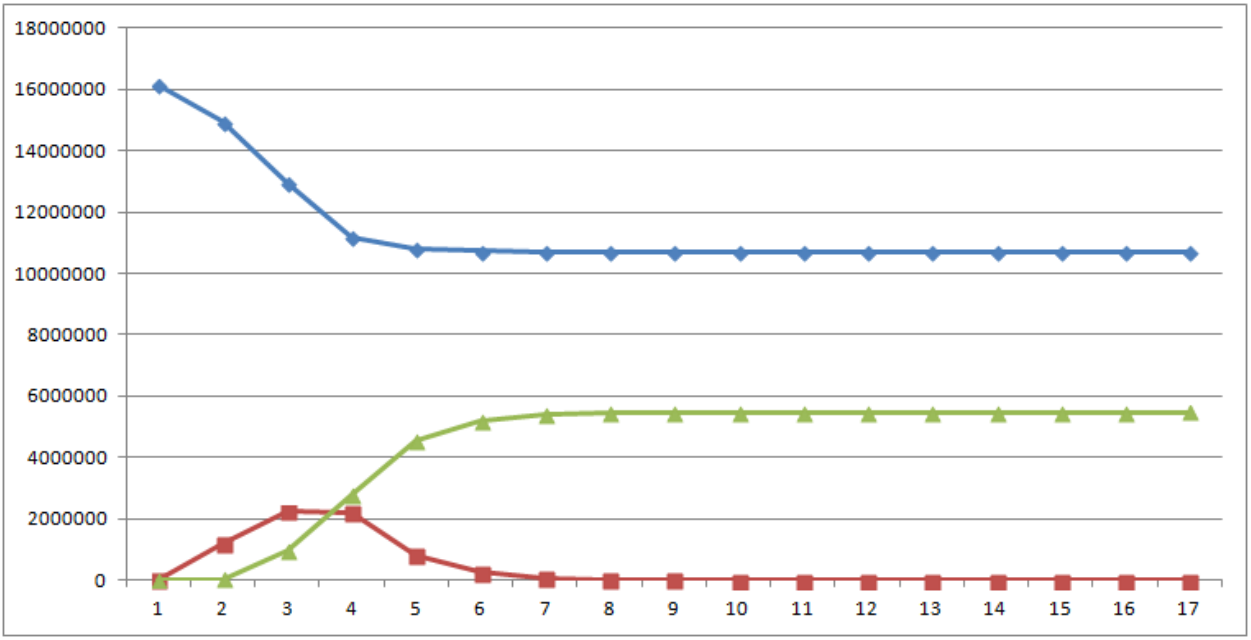
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 0{,}001N$, $R_0 = 0{,}25N$



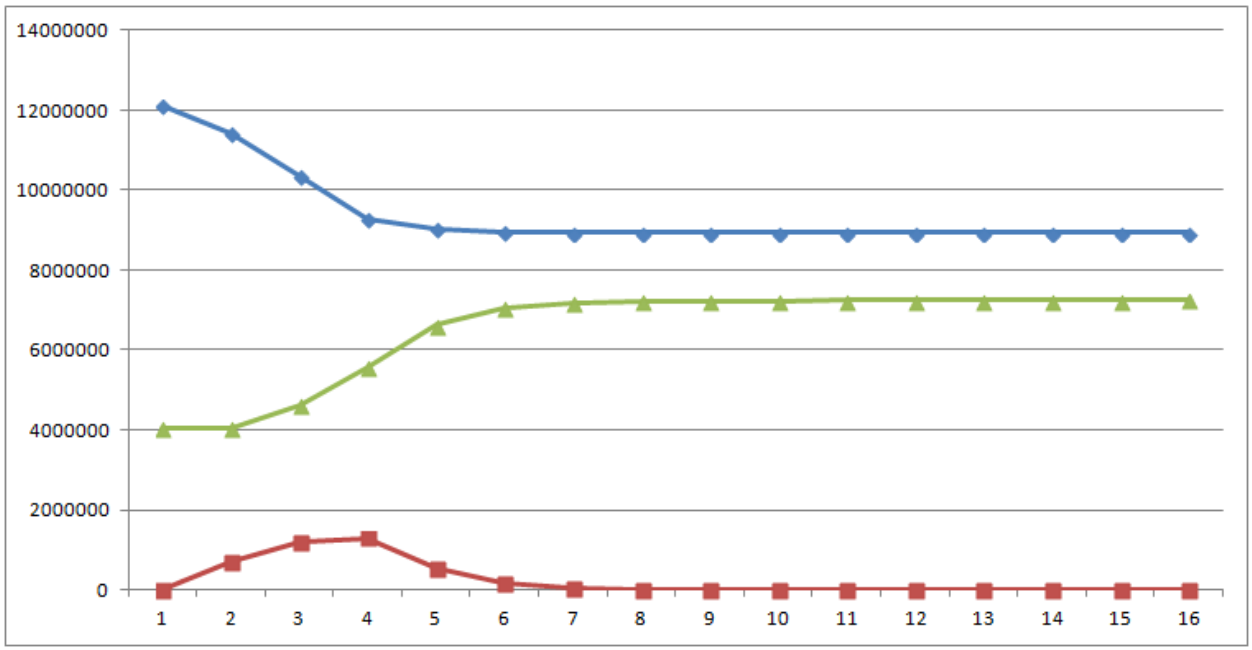
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 0{,}001N$, $R_0 = 0$.



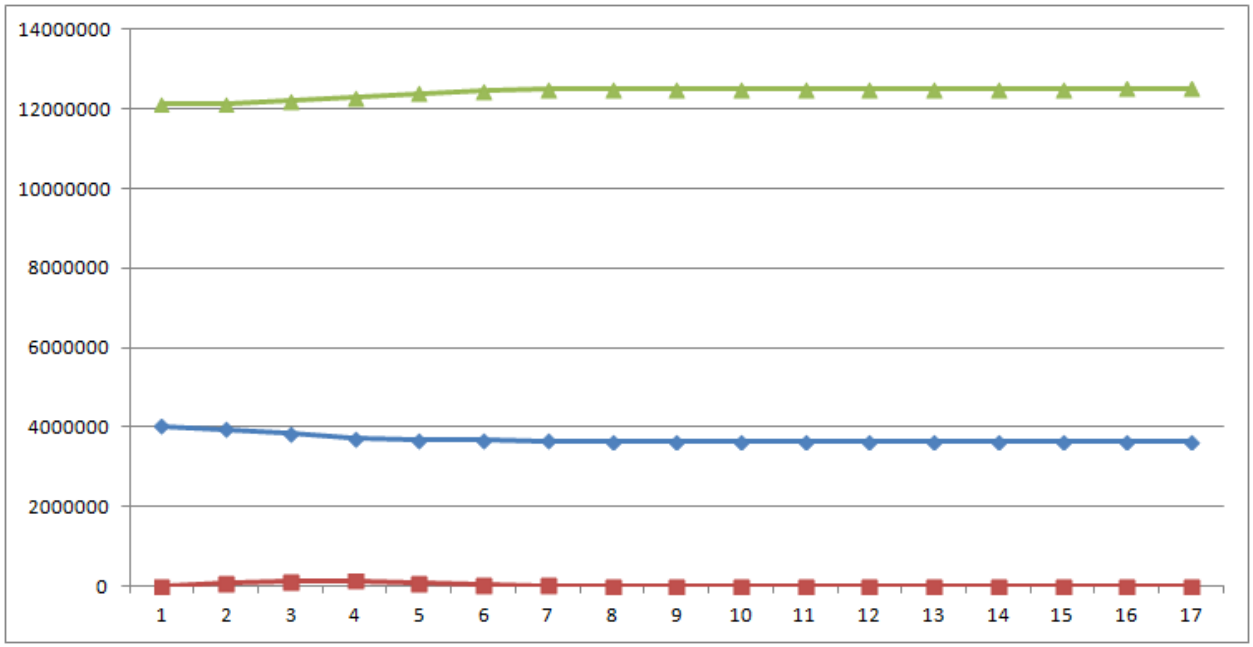
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 1$, $R_0 = 0{,}75N$.



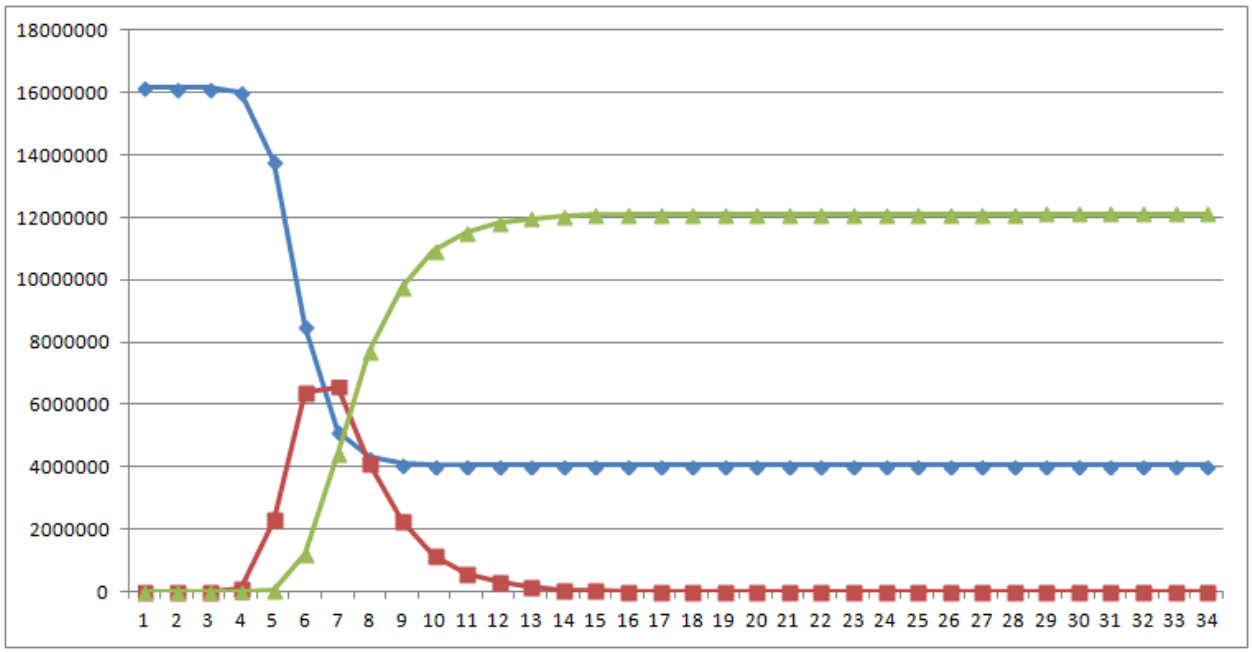
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 1$, $R_0 = 0{,}25N$.



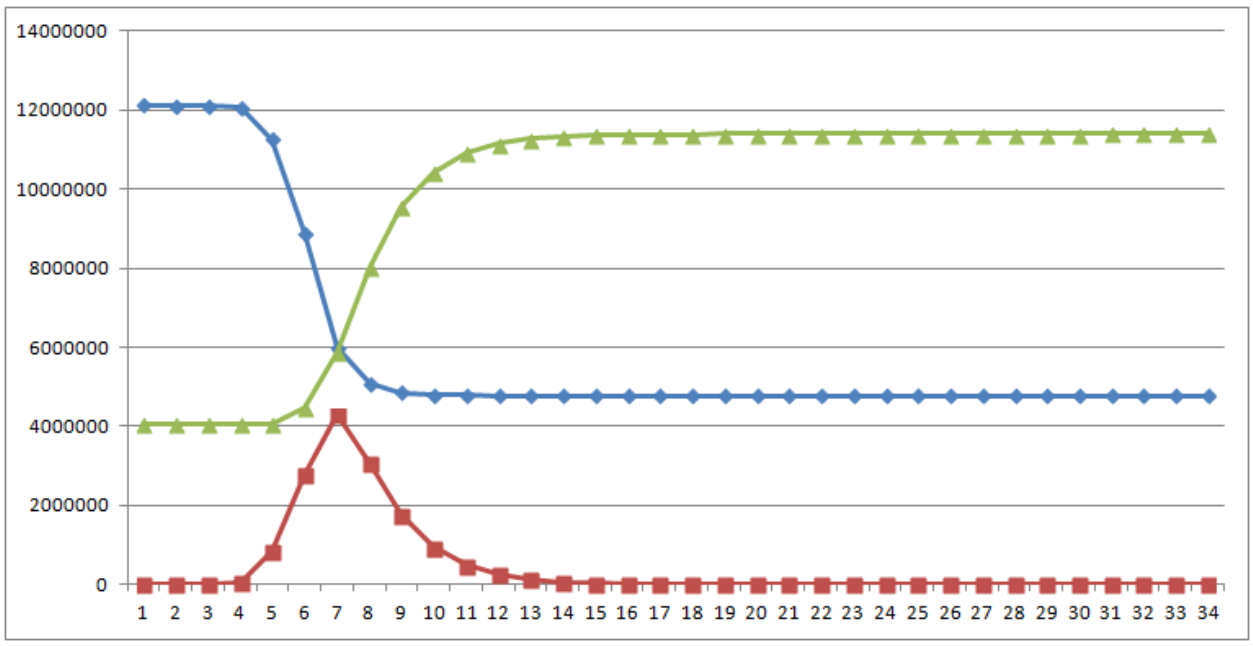
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $I_0 = 1$, $R_0 = 0$.



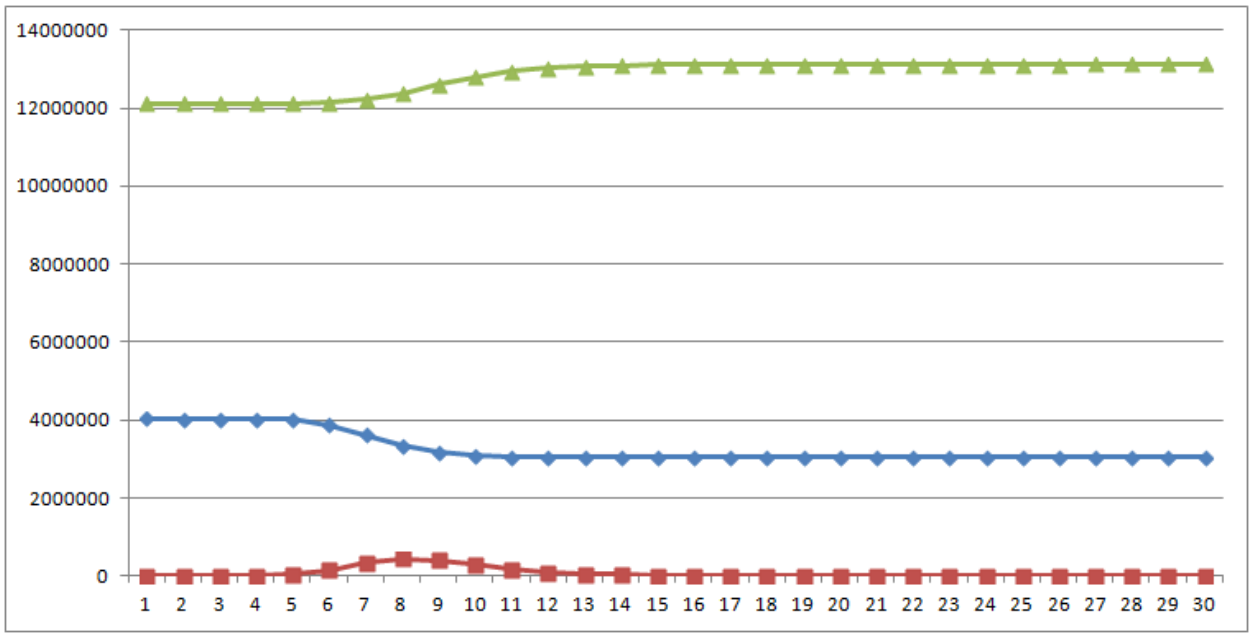
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 0{,}001N$, $R_0 = 0{,}75N$.



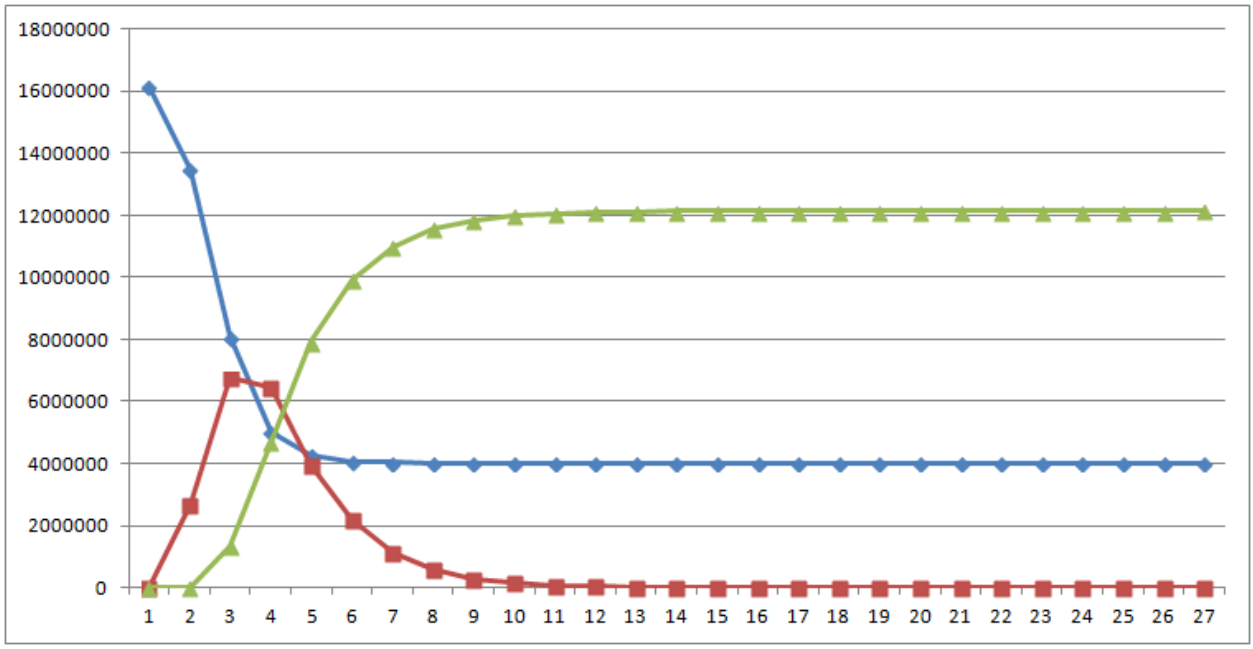
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 0{,}001N$, $R_0 = 0{,}25N$.



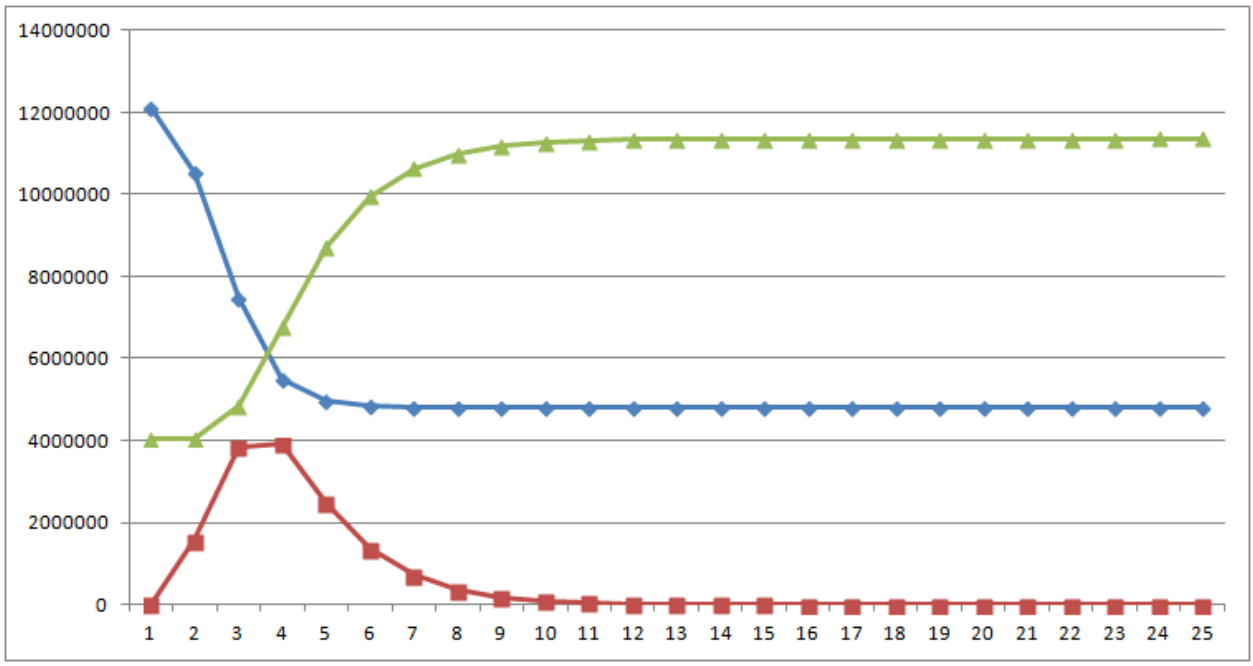
Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 0{,}001N$, $R_0 = 0$.



Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0{,}75N$.



Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0{,}25N$.



Результаты моделирования с параметрами $\varphi = 70$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0$.