Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»

МОДЕЛИ ОБЕСПЕЧЕНИЯ ДОСТОВЕРНОСТИ И ДОСТУПНОСТИ ИНФОРМАЦИИ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Монография

Владимир 2015

УДК 004.056
ББК 32.81
М74

Авторы: М. Ю. Монахов, Ю. М. Монахов, Д. А. Полянский,
И. И. Семенова

Рецензенты:
Заслуженный деятель науки России
доктор технических наук, профессор
Владимирского государственного университета
имени Александра Григорьевича и Николая Григорьевича Столетовых
О. Р. Никитин

Доктор физико-математических наук, профессор
Владимирского филиала Финансового университета
при Правительстве Российской Федерации
О. Я. Бутковский

Печатается по решению редакционно-издательского совета ВлГУ

Модели обеспечения достоверности и доступности информации
М74 в информационно-телекоммуникационных системах : монография / М. Ю. Монахов [и др.] ; Владим. гос. ун-т им. А. Г. и Н. Г. Столетовых. – Владимир : Изд-во ВлГУ, 2015. – 208 с.
ISBN 978-5-9984-0634-8

Приведены результаты теоретического и экспериментального исследований распределенных телекоммуникационных систем, находящихся под воздействием вредоносных программ.

Предназначена для научных и инженерных работников, занимающихся проблемой обеспечения информационной безопасности в автоматизированных и телекоммуникационных системах.

УДК 004.056
ББК 32.81

ISBN 978-5-9984-0634-8

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Настоящее время характеризуется значительным возрастанием социальной и экономической деятельности. Следствием и необходимым условием данных процессов становится быстрое увеличение объема социальной и экономической (производственной) информации, усложнение вычислительной и телекоммуникационной инфраструктуры обрабатываемых данных. При этом обоснованность и оперативность принимаемых решений становится все более зависимыми от качества информационного обеспечения данных процессов. Актуальной становится задача отделения ценной информации от «бесполезной», выделения из потоков данных наиболее важной части, позволяющей оперативно и точно анализировать текущее состояние объекта управления и выработать правильное решение. На данный момент наука располагает существенными возможностями в области автоматизированной обработки массовых данных статистики, основанными, например, на средствах теории распознавания образов, кластерного анализа, дискриминантного анализа и др., но проблема качества информации продолжает оставаться открытой.

Качество информации – совокупность свойств, отражающих степень пригодности конкретной информации об объектах и явлениях и их взаимосвязях для достижения целей, стоящих перед пользователем. Важнейшими взаимосвязанными характеристиками качества информации выступают ее полезность (ценность), достоверность и доступность.

Бессмысленно оценивать полезность информации, если она ошибочна, и тем более, если к ней затруднен доступ, поэтому опасно принимать решения на основании данных, достоверность которых сомнительна. С учетом того, что достоверность информации является важнейшим показателем качества принятия управленческих решений, знание о том, что этот показатель снижается в результате дезинформирования и/или под действием помех, заставляет постоянно возвращаться к поиску путей повышения «надежности» входящей и представляемой лицу, принимающему решение (ЛПР) информации.

Вопросу достоверности информации традиционно уделяется большое внимание в теории и практике управления социально-экономическими системами. В последнее время данному направлению стали уделять внимание специалисты по информационной безопасности. Данное обстоятель-

ство обусловлено тем, что в условиях конкурентной борьбы значительно участились атаки злоумышленников (конкурентов, преступных элементов) на информационные ресурсы предприятий и организаций. Следовательно, множество «помех» – дестабилизирующих факторов, оказывающих влияние на достоверность и доступность информации в современных информационных системах, не вписывается лишь в классы отказов, сбоев и ошибок аппаратно-программных средств и операторов. Кроме того, достоверность и доступность информации не может быть обеспечена простым включением в компоненты системы соответствующих элементов, процесс обеспечения достоверности и доступности должен быть непрерывным и управляемым.

Таким образом, проблема обеспечения достоверности и доступности информации выходит на новый уровень – уровень информационного противодействия. В связи с этим необходимо пересмотреть концептуальные принципы управления достоверностью и доступностью в информационно-телекоммуникационных системах, расширить их, наполнить новым содержанием.

Средой обеспечения достоверности и доступности информации и объектом исследования в данной работе является информационно-телекоммуникационная система (ИТКС) – взаимосвязанная совокупность информационных ресурсов, средств вычислительной техники, телекоммуникаций, программного обеспечения, персонала и пользователей, рассматриваемая как единое целое при реализации системных и прикладных информационных процессов, и предназначенная для того, чтобы обеспечивать потребителей надлежащим информационным обслуживанием.

В многочисленных работах отечественных и зарубежных исследователей, в работах авторов отмечаются следующие особенности управления процессом обеспечения достоверности информации в ИТКС:

– процесс обеспечения достоверности информации являются плохо формализуемым объектом управления вследствие того, что находится в условиях существенной неопределённости, источником которой служат техническая и социальная составляющие ИТКС. Здесь же отметим частую невозможность количественного измерения значения входных и выходных параметров подсистем, высоким их взаимным влиянием, а это иногда приводит к невозможности построения аналитических моделей частных процедур управления процессом;

– наличие «человеческого фактора» приводит к тому, что многие характеристики достоверности информационных ресурсов перестают быть строго определенными, связи между социальной и технической подсистемами описываются нечетко, остается открытым вопрос о количестве и составе входных данных, поскольку неизвестно, что может повлиять на поведение пользователя как элемента системы;

– если для снятия «неопределённости» при исследовании технической подсистемы применимы классические методы статистики, то для социальной подсистемы они не пригодны, поскольку неопределенность в данном случае носит субъективный характер.

Резюмируя вышеизложенное, отметим, что управление процессом обеспечения достоверности информации в ИТКС следует рассматривать как сложный интеллектуальный процесс разрешения проблем, который не может сводиться исключительно к рациональному выбору. Для поддержки этого процесса представляется целесообразным использовать когнитивный подход к моделированию и управлению, поскольку он направлен на разработку формальных моделей и методов, поддерживающих интеллектуальный процесс решения задач управления благодаря учету в этих моделях и методах когнитивных возможностей человека.

Обеспечение доступности информационных ресурсов и процессов в ИТКС, в первую очередь, связывают с задачей обнаружения и «парирования» информационных атак на доступность. Несмотря на высокую активность в исследованиях, на сегодня нет единой теории обнаружения такого типа атак. Поэтому формальные методы обнаружения атак проработаны недостаточно для широкого использования в реальных системах. Методы обнаружения, используемые сегодня в коммерческих и некоммерческих системах обнаружения вторжений (СОВ), можно определить как эвристические. Они все используют некоторые априорные предположения о том, что есть атака, какое поведение объекта в сети можно считать нормальным.

Коммерческие системы большей частью используют эвристический метод обнаружения, основанный на экспертном подходе, когда система анализирует наблюдаемое поведение объектов в сети на основе существующей у нее базы описаний известных атак. Эти описания строятся на основе знаний экспертов. Для экспериментальных систем характерно использование формализованных методов обнаружения атак, которые ис-

пользуют формальную модель атаки и пытаются приблизить процесс ее обнаружения к полной автоматизации.

Среди СОВ, декларирующих в той или иной степени обнаружение атак на доступность (DoS-атак), отметим: AAFID (Autonomous Agents for Intrusion Detection), разработанную в университете Purdue, West Lafayette, IN, USA; ASAX (Advanced Security audit trail Analyzer on uniX), разработанную в университете Namur, Belgium; NetSTAT (Network-based State Transition Analysis Tool) - развитие проекта Калифорнийского Университета в г.Санта-Барбара; Prelude; Snort; SnortNet (расширение системы Snort).

Выделим два метода обнаружения DoS-атак – анализ информационного сетевого потока и анализ журналов регистрации операционной системы или приложений. Первый подход к обнаружению атак является более эффективным по причине реагирования в реальном масштабе времени. Поэтому основные исследования в настоящий момент направлены на разработку способов и процедур обнаружения атак в сетевом трафике. Здесь основной задачей является идентификация вредоносного трафика. Большинство атак в настоящее время трудно отличить от обычных действий пользователей, в то же время, обратное утверждение так же справедливо – зачастую деятельность пользователей вызывает эффекты, идентичные эффекту от проведения распределенной атаки отказа в обслуживании.

С позиции всех способов анализа сетевого трафика атака определяется как неестественное и заметное изменение статистических свойств исследуемого трафика. Хотя авторы данных способов и декларируют успешные результаты в выявлении некоторых типов атак на доступность, имеют место существенные проблемы:

– проблема варьирования условий тестирования. Большинство способов обнаружения разработаны и исследованы не комплексно. Проведение комплексных исследований затруднено и требует больших временных затрат. Рассмотренные способы упускают из вида широкий круг условий: варьирование параметров сетей, динамики атак и т.п.;

– проблема оценки естественной сетевой активности. Выявление атак привязано к статистическим свойствам естественной сетевой активности. Модели атаки, используемые в рассмотренных подходах, составляют малую часть от общего числа возможных атак. А такие характеристики как увеличение объема трафика и распределение адресов источников трафика уже устарели, они были свойственны ранним реализациям DoS-атак, и с

тех пор стали широко известны, поэтому современные инструменты и методы проведения атак учитывают эти особенности. Кластеризация трафика позволила несколько упростить его анализ. В то же время от того, каким образом осуществляется кластеризация трафика при его анализе, существенно зависит результат обнаружения атаки. Определение способа разбиения трафика на кластеры по существу является сложной задачей, и проверить этот способ также представляется достаточно трудным;

– проблема определения параметров детектирования. Каждый из способов обнаружения атак имеет определенный набор параметров, таких как способ разбиения трафика на кластеры, значение порогов, уровень фильтра и другие. Определение значений этих параметров достаточно затруднено и зависит от конкретных условий, в которых функционирует СОВ;

– проблема раннего прогноза атаки. Для современных ИТКС мало обнаружить атаку, надо иметь механизмы ее прогнозирования (и раннего противодействия). В конце концов, не сама по себе атака вредна ИТКС, вредны ее последствия. Рассмотренные методы обнаружения атак не имеют таких возможностей. Чтобы выполнить кластеризацию трафика, надо этот трафик (с атакующими пакетами и за достаточно существенный промежуток времени) в виде файла уже иметь. Некоторым типам (неизвестных) атак не существует противодействия в существующей системе защиты, поэтому проблема раннего прогноза (даже с большой вероятностью ложной угрозы), весьма существенна для КРИВС.

– проблема самоподобия. Было замечено, что не всегда информационный поток в сети можно моделировать с использованием Пуассоновского процесса. Таким образом, множество задач, возникающих при исследовании трафика сети, пополнилось проблемой характера процесса движения пакетов по сети. На сегодняшний день существуют экспериментальные результаты того, что поведение информационного потока следует моделировать при помощи так называемого самоподобного процесса. Свойство самоподобия ассоциируется с одним из типов фрактала, то есть, при изменении шкалы корреляционная структура самоподобного процесса остается неизменной.

Теория самоподобных стохастических процессов не так хорошо развита, как теория Пуассоновских процессов. Но, учитывая то, что самоподобные модели более точно характеризуют поведение информационного

потока, чем пуассоновские модели, важной задачей стала разработка инструментальных средств для анализа и синтеза самоподобных процессов.

Целью настоящего исследования является разработка научных основ оценки уровня достоверности и доступности информационных ресурсов, синтез базовых концепций и аксиоматического фундамента теории обнаружения информационных процессов и свойств данных, влияющих на снижение достоверности и доступности информации в информационно-телекоммуникационных системах.

Монография отражает результаты теоретической и экспериментальной работы авторов по обозначенной проблеме. В многочисленных экспериментах по моделированию процессов обеспечения достоверности и доступности информации в корпоративных и социально-ориентированных ИТКС принимали участие преподаватели, аспиранты и студенты кафедры информатики и защиты информации Владимирского государственного университета.

В первой главе проводится обзор и анализ российских и зарубежных работ по вопросам обеспечения достоверности информации в современных информационно-телекоммуникационных системах, приводится методика выявления закономерностей в понятийной, категориальном, методическом аппарате в области достоверности информации на основе онтологического подхода.

Во второй главе рассматриваются особенности современных информационно-телекоммуникационных систем, влияющие на процессы обеспечения достоверности и доступности информации в системе. Предлагается концепция управления процессами обеспечения достоверности информационных ресурсов в современных ИТКС, отличающаяся учетом работы систем в условиях: дестабилизирующих факторов; активного информационного противодействия; мониторинга и динамического изменения уровня достоверности источников информации. На основе общепринятых классификаций угроз безопасности составлена карта угроз достоверности информационных ресурсов. Выявлены функции обеспечения достоверности в условиях информационного противодействия. Приведено соответствие мер и средств обеспечения достоверности выявленным функциям.

В третьей главе предлагается общая модель оценки достоверности информации в ИТКС в условиях информационных воздействий с учетом решения сопутствующих задач оценки рисков и экономической эффектив-

ности мероприятий по повышению достоверности. Приводятся результаты экспериментального исследования обеспечения достоверности информации на промышленном предприятии.

В четвертой главе рассматриваются вопросы, связанные с обеспечением доступности информационных ресурсов и процессов в ИТКС. Анализируются информационные атаки на доступность. На основе гипотезы, что самоподобность и персистентность сетевого трафика вызвана атакой на доступность, разрабатываются модели и процедуры анализа TCP-трафика на основе теории хаоса, предлагается методика раннего обнаружения информационной атаки, исследуется адекватность предложенных моделей.

В пятой главе рассматривается проблема достоверности информации в социально-ориентированных ИТКС. Представлены результаты построения моделей угрозы распространения недостоверной информации в крупномасштабных социальных сетях ИТКС, их экспериментальное исследование и практическая апробация.

Авторы выражают искреннюю благодарность рецензентам: заслуженному деятелю науки России, доктору технических наук, профессору Никитину О. Р. и доктору физико-математических наук, профессору Бутковскому О. Я. за ценные указания по существу материалов монографии.

ГЛАВА 1. **ОНТОЛОГИЯ ПОНЯТИЙНОГО АППАРАТА В ОБЛАСТИ ДОСТОВЕРНОСТИ И ДОСТУПНОСТИ ИНФОРМАЦИИ**

В главе представлен обзор и анализ российских и зарубежных работ по вопросам обеспечения достоверности и доступности информации в современных информационно-телекоммуникационных системах. Показано, что проблемы достоверности информации носят комплексный, мультидисциплинарный характер. Причем семантика понятий в разных дисциплинах несколько отличается, тем более, если рассматривать вопросы описания проблем достоверности и понимания этого понятия на разных языках. В результате было предложено разработать онтологию понятийного аппарата, методического обеспечения и признакового пространства при определении достоверности информации в информационно-телекоммуникационных системах. Так как она создана на основе открытого стандарта OWL, то ее можно развивать, модифицировать, уточнять без необходимости изменения инструментария и методов, которые ее будут использовать.

1.1. Анализ работ по проблемам обеспечения достоверности и доступности информации

Анализ работ по проблемам обеспечения достоверности и доступности информации в информационно-телекоммуникационных системах позволяет выделить следующие особенности, принципиальные для настоящего исследования.

В литературе представлен широкий спектр методов оценки и повышения достоверности информации. В иностранных публикациях [36, 16, 13, 22] под достоверностью информации, как правило, понимают качество информации, как пригодность её для использования в задачах управления принятием решений. Управление качеством данных определяет вектор развития информационного общества [19]. Эффективная работа системы планирования ресурсов предприятия во многом зависит от качества данных [31].

Значимость правильной оценки качества информации постоянно растёт в силу того, что тяжесть последствий ошибочных решений на основе недостоверных данных может быть различной: от временных, финансо-

вых и репутационных потерь до конфликтов и войн [18]. Проблема качества информации актуальна в свете критической роли, которую играет информация в экономике, основанной на знаниях и больших объёмах данных [41].

Интенсивное развитие глобальных сетей как среды для обмена информацией и отсутствие нормативных стандартов в отношении содержащейся в ней информации привело к снижению уровня достоверности информации. Пользователи глобальных сетей участвуют во взаимодействии с всё более и более разнообразной информацией, что приводит к возрастанию потребности в фильтрации информации. Существенным аспектом проблемы является неспособность поисковых технологии выделить из обширного пространстве сомнительного содержания и вернуть «качественные» результаты запроса пользователю [30, 44].

Воспринимаемое качество потребительской информации является наиболее значимым фактором для прогнозирования поведения потребителя [27]. Суждение о степени достоверности информации со стороны пользователей в целом основано на когнитивной оценке. Факторы, влияющие на суждения, определяются с точки зрения характеристик информационных ресурсов, характеристик источников, знания, ситуации и общего предположения [43]. Пользовательские критерии для определения качества данных различны и отражают объективные особенности данных и производственного процесса. Исходя из этих показателей, пользователь может оценить качество данных для конкретной области применения [55].

Достижение требуемого уровня достоверности является более сложной и широкой задачей, чем обеспечение надёжности функционирования средств обработки информации [25].

Средства поиска и исправления ошибок при обработке информации в вычислительных системах, использование надёжных и дорогостоящих хранилищ данных не решают основную проблему низкой достоверности данных [32].

Несмотря на то, что существует широкий спектр исследований по различным аспектам качества информации и её достоверности, остаётся потребность в методологии оценки достоверности информации, общей для различного вида ИС, позволяющей обеспечить базу развития информатизации общества [14, 28, 29, 8].

В последнее время количество приложений, которые используют источники данных, доступные в Интернете, сильно возросло. Одна из основных проблем связана с нахождением наиболее релевантных источников данных для данного приложения. В общем случае, источник данных считается релевантным, когда он отвечает запросам, указанным в приложении. Однако может случиться так, что конкретный источник данных отвечает на запрос, но ответ, вырабатываемый источником данных, на самом деле не соответствует требованиям пользователя [52].

Источники информации, используемые в одной информационной системе, зачастую перекрывают друг друга и формируют противоречивую информацию. Конфликты значений в противоречащих источниках часто систематические и вызваны некоторыми свойствами различных источников [23].

Доверие к информации проявляется в двух уровнях: институциональном (домен, определенный URL, тип учреждения) и индивидуальном (идентификация автора, авторская принадлежность и имя автора) [44].

Исследование информационных систем на предмет оценки и повышения уровня достоверности информации должно быть основано на анализе их в качестве подсистем, внедренных в более широкие системы управления с обратной связью. Достоверность информации не является обособленной характеристикой, а определяет факторы рисков, которые непосредственно влияют на принятие решений [38, 40].

Достоверность информации иногда понимают в очень ограниченных рамках как точность. Вместе с тем существует также и контекстное качество информации (соответствие поставленной задаче) [48, 54, 50].

Контекстные оценки столь же важны как объективные индикаторы качества, потому что они влияют на то, какая информация будет применена для принятия решений. Для человека характерен двойственный процесс познания, который позволяет одновременно оценивать как объективные, так и контекстные атрибуты качества информации [46].

Один и тот же ресурс может иметь приемлемый уровень качества для некоторых контекстов, но он может быть неприемлемым для других. Однако существующие метрики качества данных в основном получают нейтрально, без учёта специфики контекста. Это свидетельствует о необходимости пересмотра показателей качества данных и методов измерения для включения оценки контекста [6].

Качество источника данных является существенным для общего уровня достоверности. В особенности это актуально для систем, требующих изначально «неопределенных» данных [51].

Пользователи, которые нуждаются в информации для достижения целей, получают поток информации, который включает в себя дефекты. Достижение высокого уровня возможно только путём исправления самого потока, а не устранения отдельных дефектов [47]. Вместе с тем на исследование достоверности информации влияет неопределенность параметров внешних информационных систем и информационного обмена [33].

Критерии оценки достоверности информации

Качество информации является одним из наиболее важных аспектов информационной интеграции. При этом существует множество критериев оценки качества. Проблема исследования качества осложнена необходимостью количественной оценки критериев. Классификация критериев должна быть основана на оценочно-ориентированных способах, что позволяет, в частности, определить меры доверия к полученным результатам оценки, выражающие их точность [38].

Различными авторами выделены следующие критерии оценки достоверности информации:

– аутентичность (соответствие действительности), полезность (соответствие данных запросам пользователя) и содержательность (внутренняя достоверность) [7];

– полнота, однозначность, значимость и корректность [53];

– точность, полнота, последовательность и своевременность [10];

– точность, полнота, непротиворечивость и актуальность [21];

– актуальность, точность, своевременность, доступность, интерпретируемость и согласованность [55];

– непротиворечивость и точность [12];

– точность, полнота, непротиворечивость и ценность [28];

– источник, содержание, формат, презентация, актуальность, точность и скорость загрузки [44];

– полнота, достоверность, точность и актуальность [6].

Эффективный анализ качества информации требует мощных, но простых способов получения метрик достоверности [49]. Получение точ-

ных измерений и экономически целесообразных оценок качества информации затруднено из-за сложности информационных систем и субъективного характера качества информации. Непрерывное повышение качества информации возможно посредством систематической оценки и многократных измерений качества информации [45].

Подходы к оценке и обеспечению требуемого уровня достоверности

Основным ограничением существующих подходов к оценке достоверности данных является их специализация по конкретным вопросам или условиям [3]. Методы оценки достоверности информации можно разделить на две большие группы: эвристические (используемыми аудиторами) и формальные (оперирующие моделями информационных ресурсов, информационных процессов и информационных систем [41].

Подходы к оценке достоверности данных в базах данных (БД) основаны на анализе отношений и рассматривают БД в качестве графа сущность-связь, где прямые и косвенные отношения соответствуют путям в графе [29]. В то время как обычные ошибки в БД, такие как несуществующие индексы, могут быть обнаружены и исправлены с помощью традиционных инструментов очистки данных, много ошибок, обычных для производственного процесса, не могут быть решены. Решением проблемы может служить матрица качества в задачах классификации интеллектуального анализа данных [5].

Поиск недостоверных данных даже при наличии их высокой структуризации и обозначенных взаимосвязях (как, например, в реляционном представлении данных) требует использования широкой совокупности методов: профилирование данных [17], нечёткий анализ [34], интеллектуальный анализ [24] и др.

В связи с тем, что исправление недостоверных данных, которые уже были внесены в БД и использованы, вызывает значительные затруднения в связи с распространением ошибок, основной направление повышения достоверности должно быть ориентировано на процессы ввода, изменения и преобразования данных [26].

Общая модель оценки достоверности данных в системах принятия решений должна отражать исследование потоков данных с измерением ряда параметров на этапах сбора, ввода, обработки, хранения, передачи и

представления информации. Модель должна давать представление о возможных ошибках во множестве промежуточных и конечных результатов. При этом должно быть учтено распространение и изменение ошибок различных типов [9].

Управление уровнем достоверности информации невозможно без предварительного установления связи с источником информации, контекстом и проведения её структурного анализа [4]. Источники данных, содержащие последовательности событий, могут быть смоделированы на системах конечного автомата. Правила согласованности данных могут быть выражены формальными методами и автоматически проверены на данных, как до, так и после выполнения отдельных действий [15]. Формальная процедура управления достоверностью данных на всех этапах жизненного цикла информации должна преобразовывать показатели достоверности в оценки дополнительной неопределенности в связи с недостаточной достоверностью данных [55].

Комплексный подход к обеспечению достоверности данных должен объединять оценку качества данных и архитектуру данных в единую структуру с серией шагов, процедур, контрольных списков и инструментов и учитывать технологии, процессы и проблемы пользователей [20, 42].

Для оценки непротиворечивости используют условные функциональные зависимости [12].

Исследователи описали несколько подходов к работе с потерянными данными, в первую очередь, пытаясь вывести значения или оценки воздействия потерянных данных по результатам. Тем не менее, лишь немногие из этих подходов определяют скрытые структуры (смещение) в потерянных данных, то есть, определяют конкретные атрибуты, которые предсказывают утрату значений данных. Знание специфических систематических моделей смещения при потере данных может помочь аналитикам более точно оценить качество выводов из наборов данных с потерянными данными [37].

1.2. Описание методики формирования онтологии

Представленный обзор работ по проблемам обеспечения достоверности информации показывает многогранность проблемы, мультидисциплинарность и общую структурную сложность связей между элементами про-

блемы. Для систематизации сведений и последующего использования их в рамках разрабатываемой системы управления процессами обеспечения достоверности информации был проведен анализ технологий и инструментария и сделана остановка на онтологическом подходе.

Вопросам формирования и практик применения онтологий в информационной сфере на текущий момент уделяется большое внимание. При первичном знакомстве с онтологическим подходом у многих складывается впечатление, что мы имеем дело с еще одним синонимом понятия классификация, но цель подхода гораздо шире. Глубинный смысл в получении формализованной семантики тех знаний в предметных областях, которые накоплены на текущий момент, и дать возможность использовать эти знания во всемирной сети (SemanticWeb). Также ее ценность неоспорима в междисциплинарных исследованиях, когда есть потребность в согласовании понятийного, категориального аппаратов ученых и специалистов различных специальностей. Она включает машинно-интерпретируемые формулировки основных понятий предметной области и отношения между ними для совместного использования людьми и/или программными агентами, для возможности повторного использования знаний в предметной области [2].

Предлагаемая методика позволяет найти подход к систематизации данных предметной области, в которой характерны перекрестные связи и многозначность в понимании одного и того же термина/понятия, средств привычной многоуровневой классификации недостаточно для получения полной картины состояния понятийного, методического и пр. аппарата в изучаемой области.

Отбор источников для выбора понятий и связей между ними. Так как существует большое количество источников информации о вопросах достоверности, которые сами по себе обладают разными уровнями достоверности и качества в целом, то было принято решение использоваться для формирования онтологии только те понятия и их определения, которые представлены в журналах, рекомендованных ВАК, учебниках, рецензируемых монографиях, диссертациях, а также зарубежных статьях, опубликованных в журналах, входящих в международные системы цитирования.

В качестве самостоятельного класса введено понятие «источник информации», экземплярами которого становятся конкретные источники информации в онтологии по тому или иному понятию. Связь между конкрет-

ным источником информации и понятием осуществляется в соответствии со стандартом OWL через объектное свойство hasSource / имеетИсточник.

При определении основных классов онтологии следует использовать принципы квалиметрии при построении таксономии и иерархий, описанные в [1].

При описании экземпляров следует четко понимать, что относить к классу, а что есть конкретная реализация класса в виде экземпляра. При этом следует руководствоваться критерием, что класс – это есть абстракт / обобщение некоторого набора свойств, а экземпляр – воплощение с наделением этих свойств конкретными значениями.

Так как в онтологии выделятся объектные свойства и свойства данных, то следует отметить, что для нашей работы в первую очередь следует выделить объектные свойства, которые как раз и несут смысловую/семантическую нагрузку при связывании объектов онтологии.

В онтологии закладывается учет многоязычности. Для этого можно использовать стандарт SKOS и указывать для каждого понятия в метке (label) привязку к языку через префиксы у меток, например @en. Либо средствами выбранного редактора онтологий при указании метки понятия в разделе язык (language) выбирать требуемый.

Запросы по онтологии выполняются на основе языка SPARQL.

Общий порядок формирования онтологии соответствует требованиям, предъявляемым стандартом OWL, SKOS.

1.3. Формирование онтологии

В ходе выполнения работ в рамках темы «Исследование и разработка математических методов, моделей и алгоритмов обнаружения и устранения последствий снижения достоверности и доступности информации в корпоративных и распределенных информационно-телекоммуникационных системах» при анализе наработок российских и зарубежных ученых получили весьма многополярную, частично противоречивую картину в представлениях об изучаемой предметной области. Попытки пойти по пути типовых классификаций не дали системного эффекта, поэтому было принято решение прибегнуть к аппарату онтологий.

В результате проведенного системного анализа научно-производственных процессов в информационных системах предприятий и

выявления причин и условий нарушения достоверности информации была сформирована онтология, систематизирующая современное представление в области достоверности информации на понятийном, методическом уровне и показателях, определяющих уровень достоверности. Предлагаемая онтология отражает дестабилизирующие факторы и структурно-функциональные недостатки в информационных системах предприятий, оказывающие влияние на уровень достоверности научно-производственной информации, циркулирующей в системе (рис. 1.1). Разработка велась на основе стандарта OWL в программном продукте Protégé 4.2.

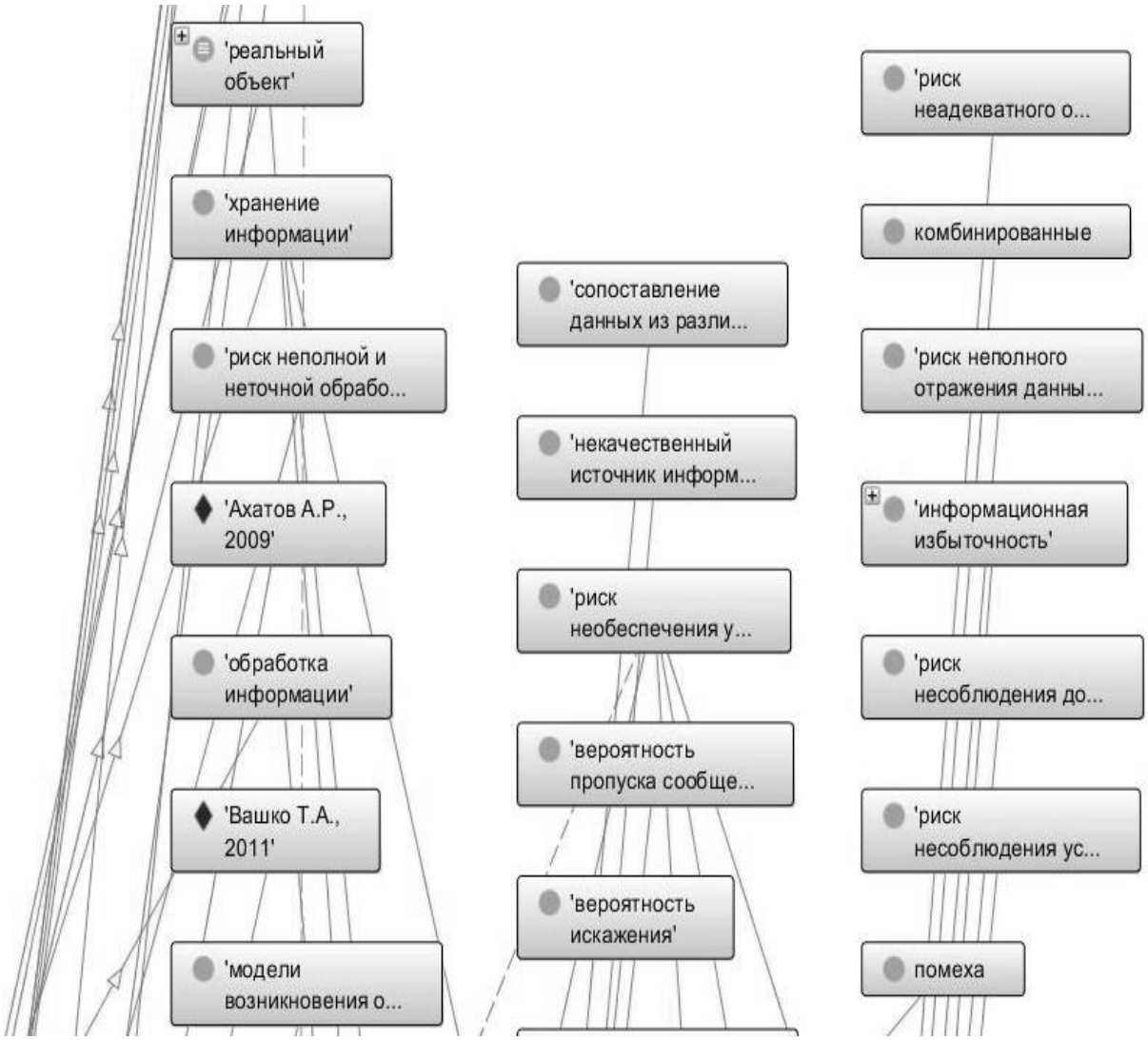


Рис. 1.1. Фрагмент онтографа (начало)

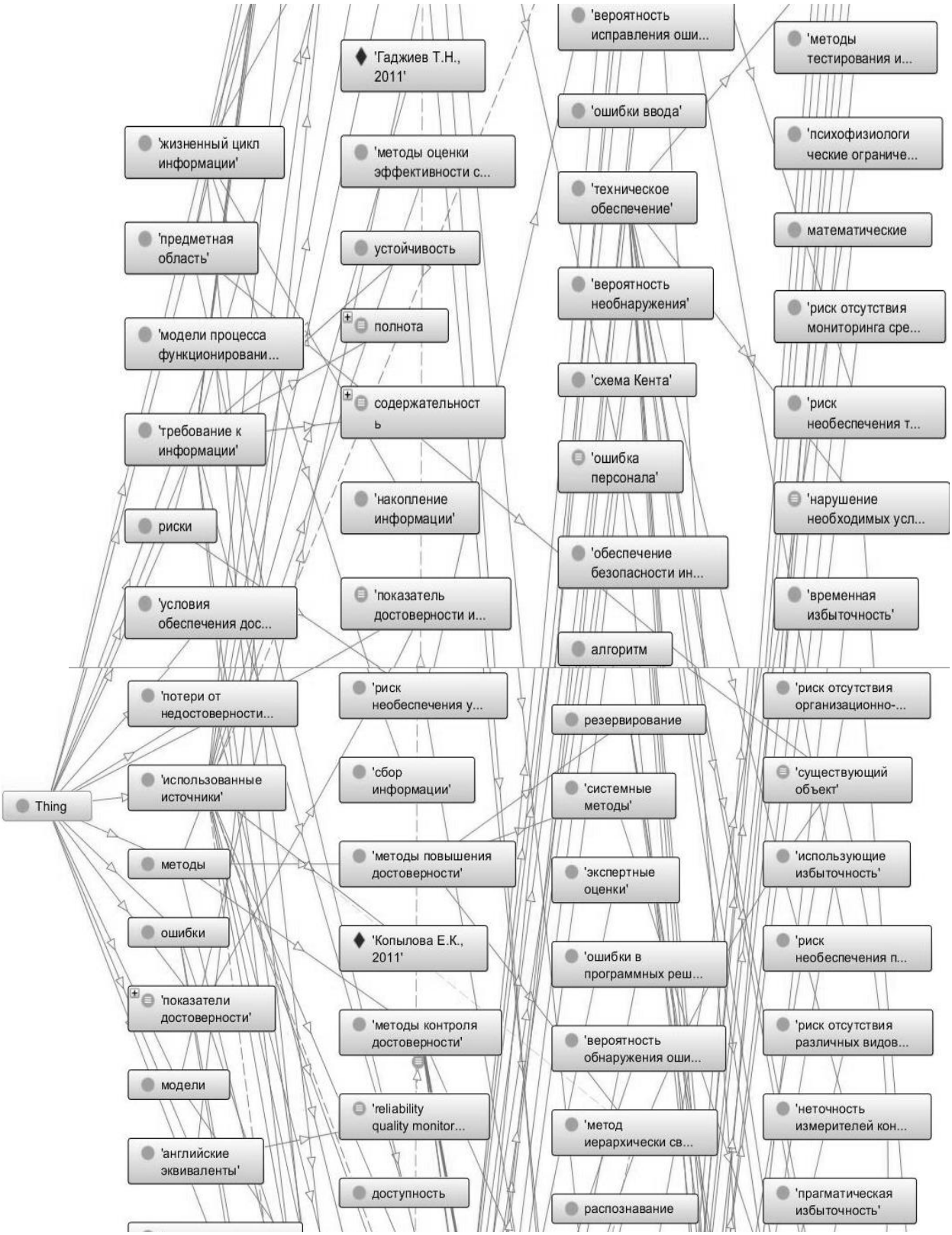


Рис. 1.1. Фрагмент онтографа (продолжение)

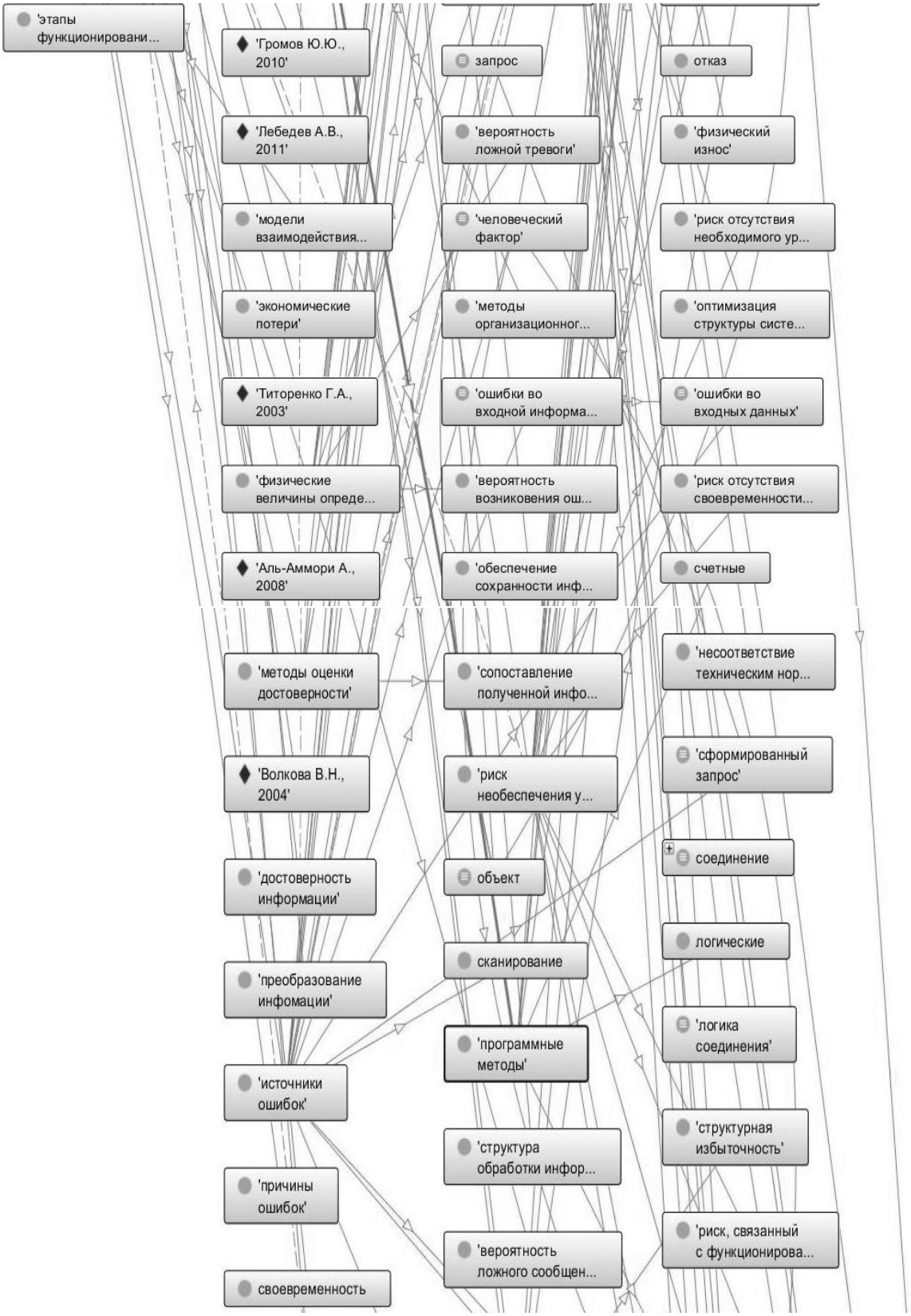


Рис. 1.1. Фрагмент онтографа (продолжение)

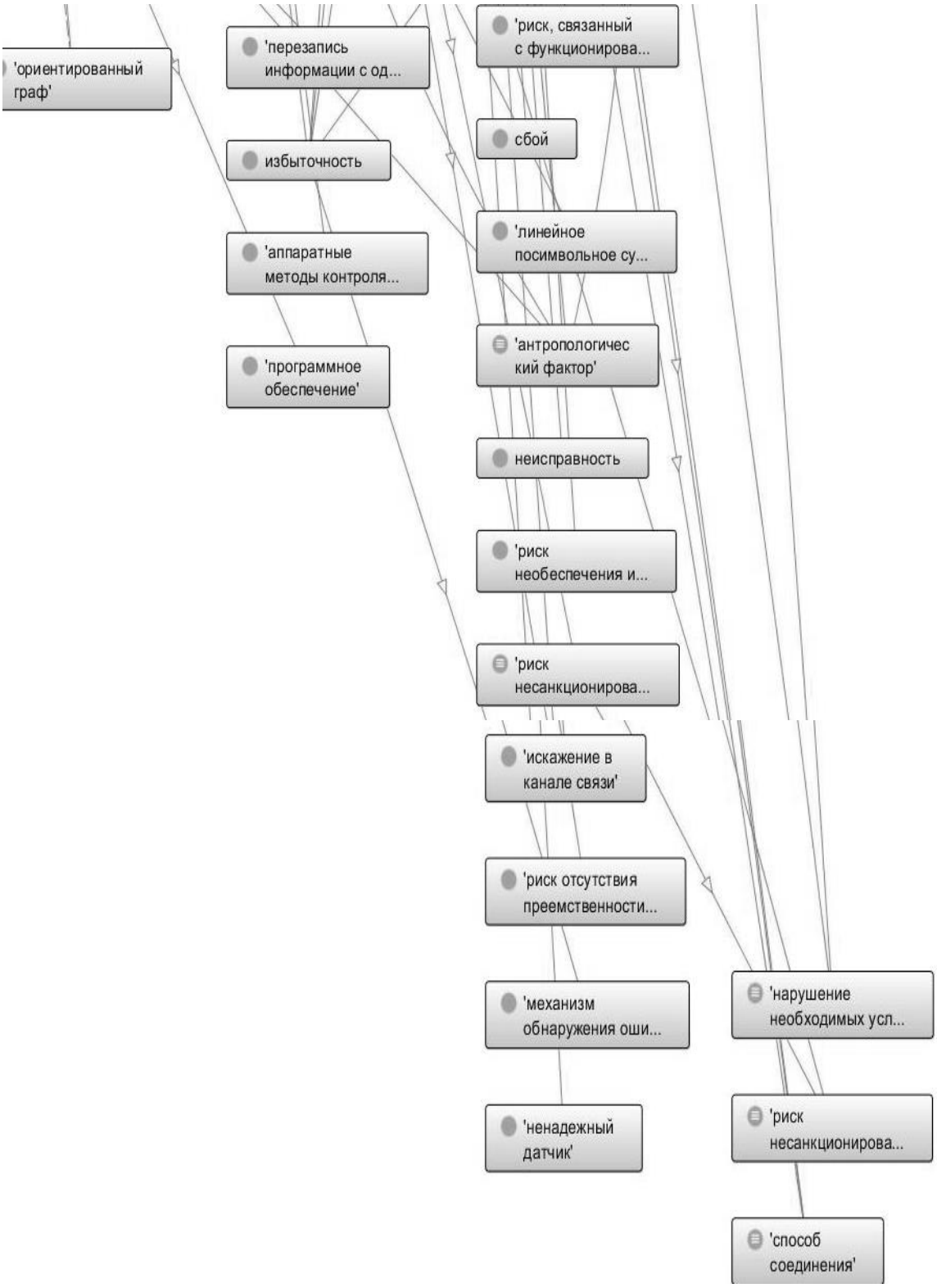


Рис. 1.1. Фрагмент онтографа (окончание)

На основе предложенной онтологии с построением системы запросов (примеры способа формирования запросов на рис. 1.2, 1.3) можно получить области, в которых наблюдаются противоречия, и перейти к разработке методов, устраняющих данные противоречия. Причем, планируется сфокусировать внимание на влиянии использования методов интеллектуального анализа данных (Data Mining, Text Mining) при автоматизированном формировании документации различных видов на уровень достоверности информации в корпоративных распределенных системах в силу все большего внедрения подобных методов в практику работы систем.

Рис. 1.2. Пример формирования запроса к онтологии средствами SPARQL

Рис. 1.3. Пример формирования запроса к онтологии средствами SPARQL

Выводы по главе

Онтология понятийного аппарата, методического обеспечения и признакового пространства при определении достоверности информации предназначена для автоматизации построения системы запросов и определения областей знаний по достоверности, в которых наблюдаются противоречия для дальнейшего перехода к разработке методов, устраняющих данные противоречия. Использование онтологии позволит повысить эффективность научных разработок в области оценки и повышения уровня достоверности информации и уточнить / согласовать существующий понятийный аппарат в данной предметной области.

С практической точки зрения онтология в комбинации с многоагентным подходом и сетями «потребности-возможности» (ПВ-сети [2]) может

стать решением оптимизационной задачи по подбору такого минимального набора конкретных средств для обеспечения достоверности информации в ИТКС, который обеспечит перекрытие всех потенциальных угроз информации при сохранении требуемого уровня ее производительности. Такое возможно при условии отражения в онтологии пространств мер и классов средств обеспечения достоверности информации, а также связанных с ними конкретных реализаций средств, существующих на ИТ-рынке, и множества потенциальных угроз. Также интересной является возможность создания web-ресурса на базе онтологии для организации поиска информации по источникам и категориям, которые нами уже систематизированы, а также подбора по онтологии аналогичных понятий на английском языке с целью проведения собственных поисковых работ.

Список библиографических ссылок

1. Азгальдов Г.Г., Костин А.В., Садовов В.В. Квалиметрия: первоначальные сведения. Справочное пособие с примером для АНО «Агентство стратегических инициатив по продвижению новых проектов». М.: Высш. шк., 2010. 143 с.
2. Онтологии и тезаурусы: модели, инструменты, приложения/ Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. М.: Бином. Лаборатория знаний, 2009. 173 с.; URL: http://www.intuit.ru/department/expert/ontoth/
3. A comprehensive data quality methodology for web and structured data / Carlo Batini, Federico Cabitza, Cinzia Cappiello, Chiara Francalanci // Int. J. Innov. Comput. Appl. 2008. V. 1. № 3. P. 205-218.
4. A framework for information quality assessment / Stvilia B., Gasser L., Twidale M. B., Smith L.C. // Journal of the American Society for Information Science and Technology. 2007. V. 58, № 12. P. 1720-1733.
5. A general approach to incorporate data quality matrices into data mining algorithms / Ian Davidson, Ashish Grover, Ashwin Satyanarayana, Giri K. Tayi // In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '04). ACM, New York, NY, USA, 2004. P. 794-798.
6. Adir Even and G. Shankaranarayanan Utility-driven assessment of data quality // SIGMIS Database. 2007. V. 38. № 2. P. 75-93.

7. Agmon N., Ahituv N. Assessing data reliability in an information system // Journal of Management Information Systems. 1987. V. 4, № 2. P. 34-44.

8. AIMQ: a methodology for information quality assessment / Yang W. Lee, Diane M. Strong, Beverly K. Kahn, Richard Y. Wang // Inf. Manage. 2002. V. 40, № 2. P. 133-146.

9. Ballou D. P., Pazer H. L. Modeling data and process quality in multi-input, multi-output information systems // Management science. 1985. V. 31. №2. P. 150-162.

10. Blake R., Mangiameli P. The Effects and Interactions of Data Quality and Problem Complexity on Classification // J. Data and Information Quality. 2011. V. 2. № 2. P. 1-28.

11. Boongoen T., Shen Q. Nearest-neighbor guided evaluation of data reliability and its applications // IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS–PART B: CYBERNETICS. 2010. V. 40, № 6. P. 1622-1633.

12. Cong G. et al. Improving data quality: Consistency and accuracy //Proceedings of the 33rd international conference on Very large data bases. VLDB Endowment. 2007. P. 315-326.

13. Coral Calero, Angelica Caro, Mario Piattini An Applicable Data Quality Model for Web Portal Data Consumers // World Wide Web. 2008. V. 11. № 4. P. 465-484.

14. Data quality / Huh Y., Keller F., Redman T., Watkins A. // Information and Software Technology. 1990. V. 32, № 8. P. 559-565.

15. Data quality through model checking techniques / Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, Fabio Mercorio // In Proceedings of the 10th international conference on Advances in intelligent data analysis X (IDA'11). Springer-Verlag, Berlin, Heidelberg, 2011. P. 270-281.

16. Data quality: concepts, methodologies and techniques. / Batini C., Scannapieco M.: Springer, 2006.

17. Data quality: the accuracy dimension (The Morgan Kaufmann Series in Data Management Systems)/ Olson J. E.: Morgan Kaufmann, 2003.

18. Data quality: the field guide / Redman T.C.: Digital Press, 2001.

19. English L.P. Information Quality Management: the Next Frontier // Quality Congress AsQs. Annual Quality Congress Proceeding. 2001. P. 529-533.

20. Enterprise Data Quality: A Pragmatic Approach / Amjad Umar, George Karabatis, Linda Ness, Bruce Horowitz, Ahmed Elmagardmid // Information Systems Frontiers. 1999. V.1. № 3. P. 279-301.

21. Fox C. et al. The notion of data and its quality dimensions // Information processing & management. 1994. V. 30. № 1. P. 9-19.

22. Francalanci C., Pernici B. Data quality assessment from the user's perspective // Proceedings of the 2004 international workshop on Information quality in information systems. ACM, 2004. P. 68-73.

23. Heiko Müller, Johann-Christoph Freytag, Ulf Leser Improving data quality by source analysis // J. Data and Information Quality. 2012. V. 2, № 4. 38 p.

24. Hipp J., Güntzer U., Grimmer U. Data Quality Mining-Making a Virtue of Necessity // DMKD. 2001.

25. Hoare C. Data reliability // ACM Sigplan Notices. 1975. V. 10. P. 528-533.

26. Huh Y. U. et al. Data quality // Information and Software Technology. 1990. V 32. № 8. P. 559-565.

27. Jeong M., Lambert C.U. Adaptation of an information quality framework to measure customers' behavioral intentions to use lodging Web sites // International Journal of Hospitality Management. 2001. V. 20. № 2. P. 129-146.

28. Kahn B. K., Strong D. M., Wang R. Y. Information quality benchmarks: product and service performance // Communications of the ACM. 2002. V. 45, № 4. P. 184-192.

29. Katerattanakul P., Siau K. Measuring information quality of web sites: development of an instrument // Proceedings of the 20th international conference on Information Systems –Association for Information Systems, 1999. P. 279-285.

30. Knight S., Burn J. Developing a Framework for Assessing Information Quality on the World Wide Web // Informing Science. 2005. V. 8. P.159-172.

31. Lan Cao, Hongwei Zhu Normal accidents: Data quality problems in ERP-enabled manufacturing // J. Data and Information Quality. 2013. V. 4. № 3. 26 p.

32. Lee Y.W. et al. Journey to Data Quality. Cambridge, MA: MIT Press. 2006.

33. Li S., Lin B. Accessing information sharing and information quality in

supply chain management // Decision support systems. 2006. V. 42. № 3. P. 1641-1656.

34. Maimon O., Kandel A., Last M. Information-theoretic fuzzy approach to data reliability and data mining // Fuzzy Sets and Systems. 2001. V. 117, № 2. P. 183-194.

35. Managing information quality: increasing the value of information in knowledge-intensive products and processes / Eppler M. J.: Springer, 2006.

36. Methodologies for data quality assessment and improvement / Batini C., Cappiello C., Francalanci C., Maurino A. // ACM Computing Surveys (CSUR). 2009. V. 41, № 3. P. 16.

37. Monica Chiarini Tremblay, Kaushik Dutta, Debra Vandermeer Using Data Mining Techniques to Discover Bias Patterns in Missing Data // J. Data and Information Quality. 2010. V.2, № 1. 19 p.

38. Nicolaou A. I., McKnight D. H. Perceived information quality in data exchanges: Effects on risk, trust, and intention to use // Information Systems Research. 2006. V. 17, № 4. P. 332-351.

39. Nuray-Turan, R. et al. Adaptive Connection Strength Models for Relationship-Based Entity Resolution // J. Data and Information Quality. 2013. V.4. № 2. P. 1-22.

40. Orr K. Data quality and systems theory // Communications of the ACM. 1998. V. 41, № 2. P. 66-71.

41. Overview and framework for data and information quality research / Madnick S. E., Wang R. Y., Lee Y. W., Zhu H. // Journal of Data and Information Quality (JDIQ). 2009. V. 1, № 1. P. 2.

42. R. Ryan Nelson, Peter A. Todd Antecedents of Information and System Quality: An Empirical Examination Within the Context of Data Warehousing // J. Manage. Inf. Syst. 2005. V. 21. № 4. P. 199-235.

43. Rieh S. Y. Judgment of information quality and cognitive authority in the Web // Journal of the American Society for Information Science and Technology. 2002. V. 53, № 2. P. 145-161.

44. Rieh S. Y., Belkin N. J. Understanding judgment of information quality and cognitive authority in the WWW // Proceedings of the 61st annual meeting of the american society for information science. 1998. V. 35. P. 279-289.

45. Sang Hyun Lee, Abrar Haider Измерение качества информации методом шести сигм // In Proceedings of the 17th international conference on Database Systems for Advanced Applications (DASFAA'12). Springer-Verlag,

Berlin, Heidelberg, 2012. P. 323-334.

46. Stephanie Watts, G. Shankaranarayanan, Adir Even Data quality assessment in context: A cognitive perspective // Decis. Support Syst. 2009. V. 48, № 1. P. 202-211.

47. Strong D. M., Lee Y. W., Wang R. Y. 10 potholes in the road to information quality // Computer. 1997. V. 30. № 8. P. 38-46.

48. Strong D. M., Lee Y. W., Wang R. Y. Data quality in context // Communications of the ACM. 1997. V. 40, № 5. P. 103-110.

49. Stvilia B. et al. Assessing Information Quality of a Community-Based Encyclopedia // In Proceedings of the International Conference on Information Quality. 2005. P. 442-454.

50. Stvilia B. et al. Information quality work organization in Wikipedia // Journal of the American society for information science and technology. 2008. V. 59. № 6. P. 983-1001.

51. Tayi G. K., Ballou D. P. Examining data quality // Communications of the ACM. 1998. V. 41, № 2. P. 54-57.

52. Using information quality for the identification of relevant web data sources: a proposal / Bernadette Farias Lóscio, Maria C. M. Batista, Damires Souza, Ana Carolina Salgado // In Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services (IIWAS '12). ACM, New York, NY, USA, 2012. P. 36-44.

53. Wand Y., Wang R. Y. Anchoring data quality dimensions in ontological foundations // Communications of the ACM. 1996. V. 39, № 11. P. 86-95.

54. Wang R. Y., Strong D. M. Beyond accuracy: What data quality means to data consumers // J. of Management Information Systems. 1996. V. 12, № 4. P. 5-33.

55. Weidema B. P., Wesnæs M.S. Data quality management for life cycle inventories– an example of using data quality indicators // Journal of Cleaner Production. 1996. V. 4. № 3-4. P. 167-174.

ГЛАВА 2. КОНЦЕПЦИЯ УПРАВЛЕНИЯ ПРОЦЕССОМ ОБЕСПЕЧЕНИЯ ДОСТОВЕРНОСТИ ИНФОРМАЦИИ В ИТКС В УСЛОВИЯХ ИНФОРМАЦИОННОГО ПРОТИВОДЕЙСТВИЯ

Современные предприятия и компании по проектированию автоматизированных систем различных классов отмечают важность обеспечения качества информации, поступающей и циркулирующей в системах. В аналитическом обзоре [20] по моделям качества данных и информации показано, что 60 % опрошенных фирм (500 средних зарубежных компаний с годовым объемом продаж более 20 млн. у.е.) испытывают проблемы с качеством данных. Одним из комплексных показателей качества информации является достоверность.

В главе рассматриваются особенности современных информационно-телекоммуникационных систем (ИТКС), влияющие на процессы обеспечения достоверности информации в системе. Предлагается концепция управления процессами обеспечения достоверности информационных ресурсов (ИР) в современных ИТКС, отличающаяся учетом работы систем в условиях: дестабилизирующих факторов, оказывающих влияние на достоверность информации в ИТКС; активного информационного противодействия; мониторинга и динамического определения уровня достоверности источников информации. Составлена карта угроз достоверности ИР. Выявлены функции обеспечения достоверности ИР в условиях информационного противодействия. Определено соответствие мер и средств обеспечения достоверности ИР выявленным функциям.

2.1. Особенности среды обеспечения достоверности информации

Выделим принципиальные особенности (свойства) ИТКС, как среды, на базе которой формируется процесс обеспечения достоверности информации:

1) «Человеческий фактор». Исходя из современных концепций построения информационных систем, ИТКС следует рассматривать как социо-техническую систему – совокупность информационно-телекоммуникационной (технической) и социальной инфраструктур (подсистем) [13]. За счет этого расширяется базовая концепция построения ИТКС: в систему в

качестве элементов структуры добавляются пользователи, а также их (пользователей) информационные связи. Основой устойчивого функционирования ИТКС становится не только (и не столько) высокая производительность системы, живучесть структуры, надежность и защищенность ее аппаратно-программных средств и обслуживающего персонала, но и качество передаваемой и получаемой пользователями информации, в первую очередь, ее достоверность. Процессы взаимодействия пользователей, вызванные недостоверной информацией, могут приводить к дисфункциональному поведению всей ИТКС [3]. Управляемость социальной среды, профессиональные навыки и квалификация пользователей, а также общее понимание решаемых задач становятся важнейшими составляющими ИТКС, оказывающими существенное влияние на информационные процессы в системе, что предопределяет и подход к обеспечению достоверности обрабатываемой в системе информации.

2) «Конфликтная среда». Отношения между пользователями могут иметь характер противодействия (конфликта) [13]. Функционирование в конфликтной среде означает, что в ИТКС присутствуют два динамических процесса противоборства:

– процесс целенаправленного снижения достоверности информации для перевода ИТКС в функционально неустойчивое состояние. Основной причиной его возникновения являются интересы нарушителей - злоумышленников, заключающиеся в том, чтобы исказить, подменить, сделать недоступными информационные ресурсы для «легальных» пользователей. Способ реализации – информационные атаки. Возможности их удачного осуществления основаны на уязвимостях технической и социальной подсистем;

– процесс повышения достоверности информации, заключающийся в выборе «надежных» источников, в противодействии атакам злоумышленников, восстановлении пораженных информационных ресурсов, обеспечении надежного функционирования и живучести технической и социальной подсистем.

3) «Крупномасштабность». ИТКС могут быть крупномасштабными («большими») системами [7], охватывающими значительные территории, миллионы пользователей и интегрироваться в мировую систему информационного взаимодействия. ИТКС могут быть взаимно проникающими. Процессы в ИТКС, реализованные, как правило, на основе распределенных

приложений, могут проходить с различными скоростями и влиять друг на друга. Кроме того, информационные ресурсы в процессе функционирования ИТКС могут добавляться и исчезать. Все это приводит к наличию значительного количества не устраняемых (или вообще плохо локализуемых) уязвимостей и обилию векторов атак [21]. Таким образом, обеспечение достоверности информационных ресурсов выполняется в сложно контролируемой среде и требует применения адаптивных средств управления.

4) «Многосвязность». ИТКС, как правило, многосвязные: их различные элементы соединены между собой (пользователи – информационно, аппаратно-программные средства – физически) и могут иметь как прямые, так и обратные связи. Структура и топология ИТКС переменны, могут быть как управляемыми, так и неуправляемыми. Характер информационных связей в ряде социально-ориентированных ИТКС [1, 11] сильно зависит от психофизиологического, интеллектуального и др. состояний пользователей. Общая структурная надежность системы и ее компонентов совсем не означает устойчивости ИТКС, наоборот, в случае распространения дезинформации могут поменяться цели системы, и новые информационные процессы будут рассматриваться как «дисфункциональность», неустойчивость системы.

5) «Самоорганизация». ИТКС могут быть самоорганизуемыми, т.е. склонными к самостоятельному автономному (не управляемому извне) появлению и поведению. Это означает, что у ИТКС появляется способность, с одной стороны, стать «разносчиком дезинформации», с другой – вырабатывать меры к самосохранению и противодействию внешним воздействиям [9]. Кластеры узлов ИТКС с нарушенным целеполаганием, частично или полностью потерявшие санкционированное управление в результате атакующего воздействия и захватившие ресурсы, могут оказывать существенное влияние на обеспечение достоверности информации в конкретной ИТКС.

Резюмируя выделенные свойства среды, отметим следующие особенности управления процессом обеспечения достоверности информации в ИТКС:

– процесс обеспечения достоверности информации являются плохо формализуемым объектом управления вследствие того, что находится в условиях существенной неопределённости, источником которой служат техническая и социальная составляющие ИТКС. Неопределенность связана

с крупномасштабностью и слабой структурированностью ИТКС, с высокой сложностью происходящих в системе информационных процессов, их недостаточной изученностью, неточностью. Здесь же отметим частую невозможность количественного измерения значения входных и выходных параметров подсистем, высоким их взаимным влиянием, приводящим к синергетическому эффекту [4] и возникновение свойств эмерджентности [19]. Это приводит к сложностям (а иногда и невозможности) построения формальных (аналитических) моделей частных процедур управления процессом обеспечения достоверности информации, учитывающего специфику ИТКС;

– наличие «человеческого фактора» приводит к тому, что многие характеристики достоверности информационных ресурсов перестают быть строго определенными: связи между социальной и технической подсистемами описываются нечетко, остается открытым вопрос о количестве и составе входных данных, поскольку неизвестно, что может повлиять на поведение пользователя как элемента системы и т.д. Трудно предсказуем эффект влияния управляющих воздействий на человека. Поскольку цель системы формулируется ЛПР или определяется системой более высокого уровня качественно (т.е. нечетко), это приводит к размытости, появлению «диапазона допустимости» при достижении цели в управлении процессом обеспечения достоверности информации;

– если для снятия «неопределённости» при исследовании технической подсистемы применимы классические методы статистики, то для социальной подсистемы они не пригодны, поскольку неопределенность в данном случае носит субъективный характер. В отличие от объективной вероятности, которая отражает относительную частоту появления какого-либо события в общем объеме наблюдений, под субъективной вероятностью понимается мера уверенности некоторого человека или группы людей (экспертов) в том, что данное событие в действительности будет иметь место.

Таким образом, управление процессом обеспечения достоверности информации в ИТКС следует рассматривать как сложный интеллектуальный процесс разрешения проблем, который не может сводиться исключительно к рациональному выбору. Для поддержки этого процесса представляется целесообразным использовать когнитивный подход к моделированию и управлению, поскольку он направлен на разработку формальных

моделей и методов, поддерживающих интеллектуальный процесс решения задач управления благодаря учету в этих моделях и методах когнитивных возможностей человека [10].

Согласно модели качества информационной системы (Rodriguez & Casanovas, 2010), приведенной в [20] можно выделить основные классы элементов в системе, влияющие на достоверность информационных ресурсов. Объединив эти классы элементов с классом информационных источников и определив критерии, по которым можно получить количественную или качественную (в понятиях нечеткой логики) оценку взаимодействия элементов и силы влияния друг на друга, получим следующую когнитивную карту (рис. 2.1).

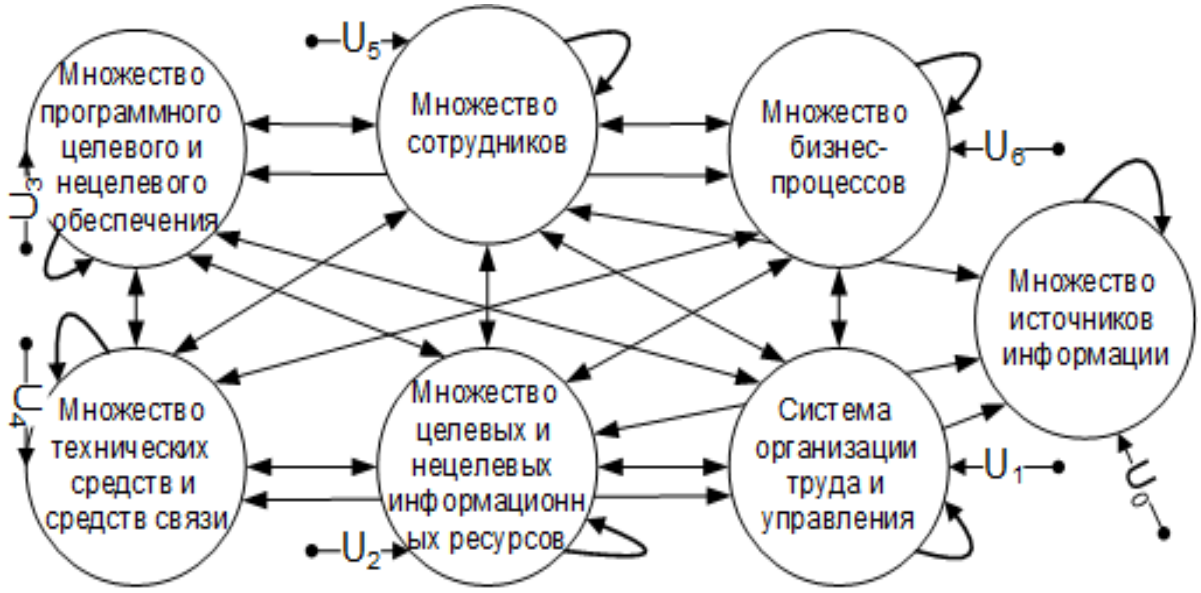


Рис. 2.1. Схема взаимного влияния классов элементов в управленческой деятельности предприятием и ИТКС

На рис. 2.1 $U_j, j = \overline{0..6}, U_j \in [-1,1]$ – это возмущающие воздействия на элементы системы со стороны внешней среды либо специальные меры, направленные на изменение ситуации в работе системы, $r_i, i = \overline{1..8}, r_i \in [-1,1]$ – это весовые коэффициенты, отражающие силу влияния одного параметра на другой, в которых знак минус указывает на обратно пропорциональную силу влияния, IS, IR, MS - критерии, описанные на рис. 2.2 и задающиеся в долях процентов от 0 по 1.

Начальные оценки критериев можно получить, имея:

– дерево с экспертными оценками уверенности в источниках информации, используемых на предприятии;

– дерево оценки уверенности в сохранности/неподверженности угрозам источников информации;

– оценки стандартизации бизнес-процессов;

– вероятностные оценки влияния внешней среды через U_j.

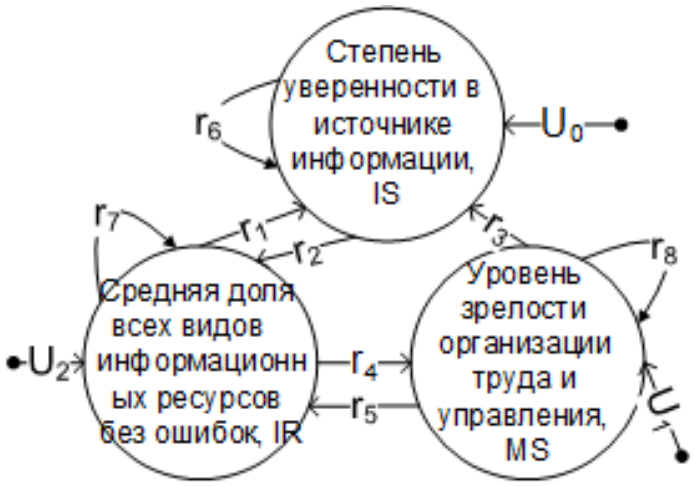


Рис. 2.2. Фрагмент когнитивной карты

Анализ сценариев развития ситуации на предприятии может быть выполнен с применением импульсного моделирования [10].

Другой подход основан на системной динамике, в частности, как показано в [15]. Для оценки рисков недостоверности источника информации возможно применить модификацию модели Вольтерра с учетом ограниченности ресурсов роста и самоограничения максимального значения. При этом покажем работу на примере фрагмента когнитивной карты (рис. 2.2). В ходе экспериментов была определена одна из форм системы дифференциальных уравнений:

$$\begin{cases} \dfrac{d(IS)}{dt} = r_6 \cdot IS + r_1 \cdot \dfrac{IS}{1+MS} + r_3 \cdot IS \cdot MS + r_9 \cdot IS^2 + U_0; \\ \dfrac{d(IR)}{dt} = r_7 \cdot IR + r_2 \cdot \dfrac{IR}{1+MS} + r_5 \cdot IR \cdot IS + r_{10} \cdot IR^2 + U_2; \\ \dfrac{d(MS)}{dt} = r_8 \cdot MS + r_4 \cdot \dfrac{MS}{1+IR} + r_{11} \cdot MS^2 + U_1. \end{cases}$$

Члены с коэффициентами $r_3, r_5, r_9, r_{10}, r_{11}$ отвечают за самоограничение значений *IR, IS, MS*. Вторые члены уравнений регулируют скорость роста значений *IR*, *IS*, *MS*. $r_9 \in [0,1], r_{10} \in [0,1], r_{11} \in [0,1]$ – коэффициенты в членах уравнений, отвечающих за срабатывание системного ферхюльстовского фактора. Начальные условия задаются, исходя из ситуации на предприятии. Представленная система была реализована в среде AnyLogic.

На рис. 2.3 представлен фрагмент эксперимента, в котором отражается варьирование параметров системы уравнений и реакции прогнозируемых значений критериев на вносимые изменения. Стоит отметить, что формализация и структурно-параметрический синтез системы уравнений, которые отражали бы реальные процессы в ИТКС достаточно трудоемкая задача. Решение путем перебора вариантов структур и параметров модели не гарантирует получения адекватной модели.

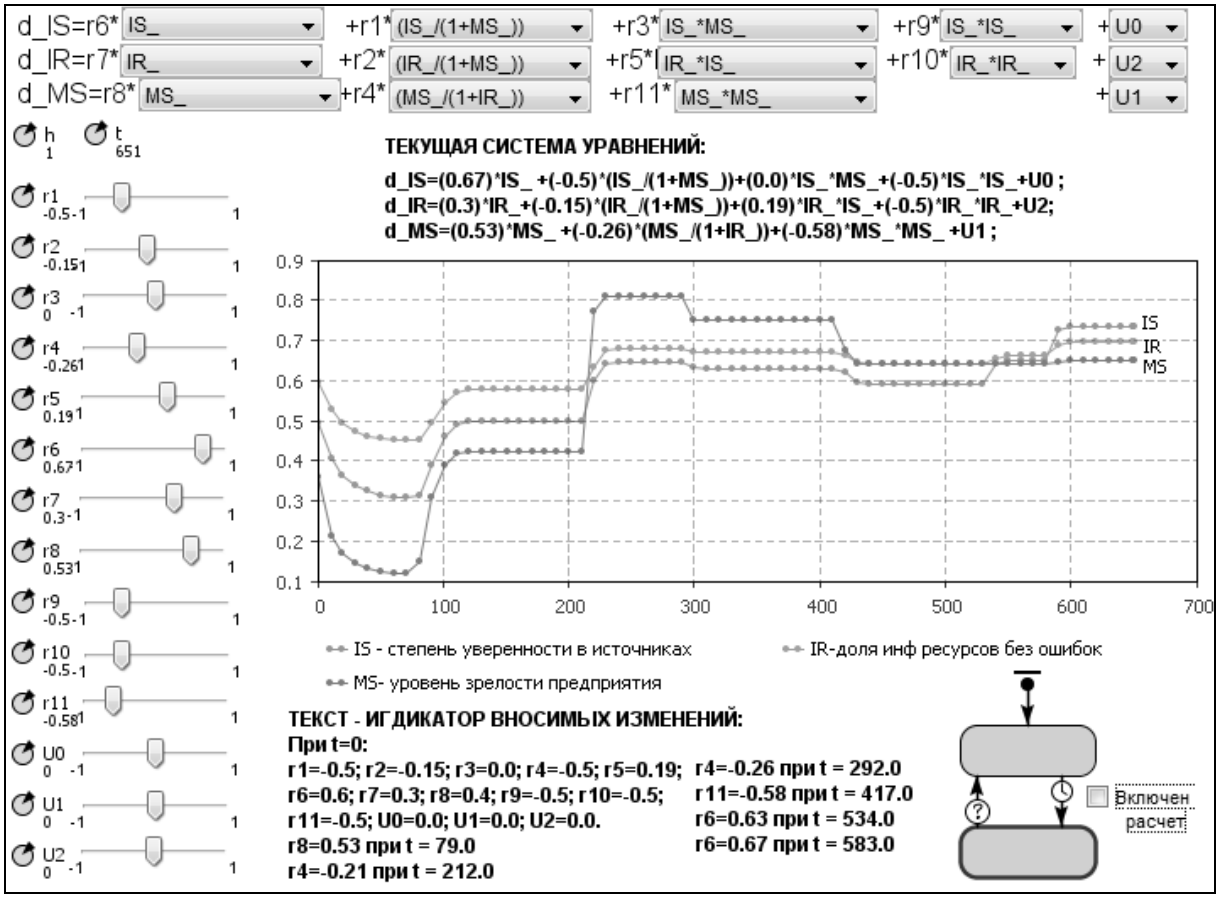


Рис. 2.3. Фрагмент эксперимента, отражающий варьирование параметров системы и реакции на изменения, по оси абсцисс – модельное время, по оси ординат – *IR, IS, MS*

2.2. Описание концепции обеспечения достоверности информации

Основные компоненты концепции обеспечения достоверности информации

Можно предложить следующие компоненты концептуальной модели процесса обеспечения достоверности информации (ОДИ) на первом (самом высоком, наиболее абстрактном) уровне декомпозиции: информационные ресурсы (ИР), источники информации (ИстИ), дестабилизирующие факторы (ДФ – угрозы достоверности), источники ДФ (угроз достоверности информации), цели злоумышленников, функции, меры и средства ОДИ, показатели достоверности.

Информационные ресурсы и их источники

Информация – объективная категория, формирующая дополнительные знания (по Шеннону) о каком-либо объекте или явлении. Информация проявляется в сообщениях ИстИ, где сообщение – выбранная порция информации, имеющая законченный смысл. Информационные сообщения, как объективная реальность, абсолютно достоверны всегда.

В ИТКС циркулируют данные – обработанные сообщения, представленные в формализованном виде (например, в виде цифрового кода), пригодном для передачи, переработки и представления в некотором информационном процессе (ИП) для решения задач пользователей. В них может присутствовать информация об объекте или явлении. Заметим, что пользователь, как правило, не является непосредственным наблюдателем объекта или явления, а должен довольствоваться данными об объектах, которые получает от некоторого ИстИ, который является либо непосредственным «наблюдателем» объекта или явления, либо транслирует данные, получаемые от других источников, в лучшем случае, от первоисточников. При этом ИстИ могут перекрывать друг друга и формировать противоречивую информацию. Конфликты значений в противоречащих источниках часто систематические и вызваны свойствами различных источников [22].

Информационные сообщения до времени «скрыты» в ИстИ, проявляются же в виде данных в момент инициирования задач пользователей

путем фиксации на физических носителях или при передаче по физическому каналу связи в ИТКС. Эти данные сохраняются, подвергаются переработке, представляются пользователям, которые посредством данных процессов обработки принимают информационные сообщения. Далее сообщения «растворяются» в потребителях (пользователях).

Данные, получаемые при кодировании сообщений, могут оказаться правдивыми (правдоподобными, неправдоподобными), полными (недостаточно полными для задач пользователей), актуальными (устаревшими для решаемой задачи) и т.п. Степень доверия к таким данным определяется их семантической и «временной» искаженностью.

Зафиксированную в ИТКС совокупность данных будем называть информационным ресурсом (ИР). Законодательно ИР – это «отдельные документы и отдельные массивы документов, документы и массивы в информационных системах» [2]. ИР – это данные различного характера, материализованные в виде документов, баз данных и баз знаний. В процессе производственной деятельности ИР рассматриваются как экономическая категория, которая является важнейшим элементом информационного менеджмента [6].

Внешние ИР формируются внешней информационной средой предприятия и отражают отношения между предприятием и экономическими и политическими субъектами, действующими за его пределами.

Внутренние ИР формируются внутренней информационной средой, т.е. совокупностью структурных подразделений предприятия и работающих специалистов, технологическими, социальными, экономическими и другими отношениями между ними. Внутренние ИР определяются внутренними бизнес-процессами [12]. При использовании ИР в предприятиях к ним предъявляются определенные требования, в том числе получение ИР в установленные сроки, полноту и неискаженность как поступающих, так и полученных ИР и т.д.

Эффективность бизнес-процессов определяется качеством ИП, реализуемых корпоративной ИТКС. Здесь важны следующие аспекты:

– решающее значение имеет реальная доступность ИР, которая на практике ограниченна;

– экономическая полезность ИР определяется фактором времени и качеством. Устаревшая или неполная информация может не только оказаться полностью обесцененной, но и привести к значительным

потерям стоимости производимых на ее основе работ.

ИП направлены на целесообразное использование ИР и снабжение ими всех элементов ИТКС. Эффективность функционирования ИП определяется наличием современных средств вычислительной техники (СВТ), распределенных БД, сетей телекоммуникаций, возможностью их модернизации и модификации, изменения структуры, включения новых компонентов и т.д., что позволяет обеспечить эффективную циркуляцию и переработку ИР. По назначению и характеру использования выделим два основных класса ИП:

– системные (обеспечивающие) ИП – представляют собой процедуры исполнения отдельных системных операций, связанных с представлением, преобразованием, хранением, обработкой или передачей данных;

– прикладные ИП – задачи пользователей. Основная цель прикладных ИП - получать посредством переработки первичных ИР информацию, на основе которой вырабатываются управленческие решения.

Будем считать, что в ИТКС циркулируют информационные ресурсы четырех типов:

– $ИР_1$ – исходные данные, полученные на хранение и обработку от ИстИ (включая потребителей и взаимодействующих ИТКС);

– $ИР_2$ – производные данные, то есть данные, полученные в ИТКС в процессе переработки исходных и производных данных;

– $ИР_3$ – программы, используемые для обработки данных, организации и обеспечения функционирования ИТКС;

– $ИР_4$ – нормативно-справочные и служебные данные.

Достоверность информации (степень доверия к данным), содержащейся в информационных ресурсах $ИР_1$, $ИР_3$, $ИР_4$ во многом определяется качеством их источника [22, 23]. Следовательно, необходимо говорить о достоверности информации как достоверности ИстИ (точнее, о степени доверия потребителя-пользователя к конкретному источнику), который искажает (неосознанно или с умыслом) формируемые им данные, делая информацию, содержащуюся в них недостоверной.

Достоверность ИстИ – апостериорная оценка, получаемая в результате наблюдения за его (источника) «информационной активностью». Источнику можно верить или нет – он субъект, может сообщать системе дезинформацию, может быть достоин доверия или недоверия.

Достоверность информации, содержащейся в $ИР_2$, $ИР_3$, $ИР_4$ в основном определяется качеством и устойчивостью процессов (функций) хранения, переработки и представления данных, происходящих в рамках технической подсистемы ИТКС при выполнении задач пользователей [14, 16]. Такая функциональная устойчивость системы достигается надежностью технических и программных средств, живучестью структурного построения системы, квалификацией и навыками в работе персонала, обеспечением безопасности ИР [17, 18]. В данном случае следует связать понятие достоверности информации с категориями целостности и доступности ИР. Целостность ИР обеспечивается, если он нелегитимно не изменяется, доступность - если легитимный процесс получает ИР за приемлемое время. Все это должно быть обеспечено при функционировании ИТКС в условиях случайных или преднамеренных информационных воздействий [5].

Таким образом, каждое звено прохождения (обработки) информации накладывает на нее свой (информационный) фильтр, вносящий свои «ослабления и запаздывания», т.е. искажения. Природа таких искажений чаще всего случайна. В итоге «достоверность информации», которую мы оцениваем, есть априорная оценка вероятности того, что сообщение для пользователя при решении определенной задачи будет содержать неискаженные данные.

Дестабилизирующие факторы (угрозы достоверности) и их источники

На основе общепринятых классификаций угроз [2,8] составлена карта ДФ (угроз достоверности ИР), представленная на рис. 2.4, и разработан их расширенный перечень. Выделены следующие угрозы достоверности:

1) Саботаж или преднамеренная угроза - умышленное нарушение информационного процесса, уклонение от работы или недобросовестное ее выполнение. Субъекты угрозы - персонал (внутренние нарушители - инсайдеры), пользователи (внешние нарушители). Даная угроза может эксплуатировать организационную уязвимость ИТКС - недовольство персонала, например, условиями труда. Относительно пользователей в качестве уязвимых могут эксплуатироваться их психофизиологические свойства.

В качестве средств осуществления угрозы могут выступать:

– вандализм - вывод из строя всех или отдельных элементов ИТКС (устройств, носителей, персонала);

– дезорганизация функционирования системы – неправомерное отключение оборудования, изменение режимов работы технических средств (ТС) или программного обеспечения (ПО);

– умышленное злоупотребление ресурсами (в том числе сетевыми);

– злоупотребление правами;

– внедрение вредоносного ПО;

– нелегитимная имперсонация – «маскарад», в том числе незаконное подключение к линиям связи;

– разглашение, передача или утрата атрибутов разграничения доступа (паролей, ключей шифрования, пропусков и т.п.);

– замена, вставка, удаление или изменение данных в информационном потоке в канале связи;

– вскрытие используемых алгоритмов шифрования;

– недобросовестное исполнение обязанностей персоналом;

– ведение агентурной работы.

Угрозы достоверности ИР

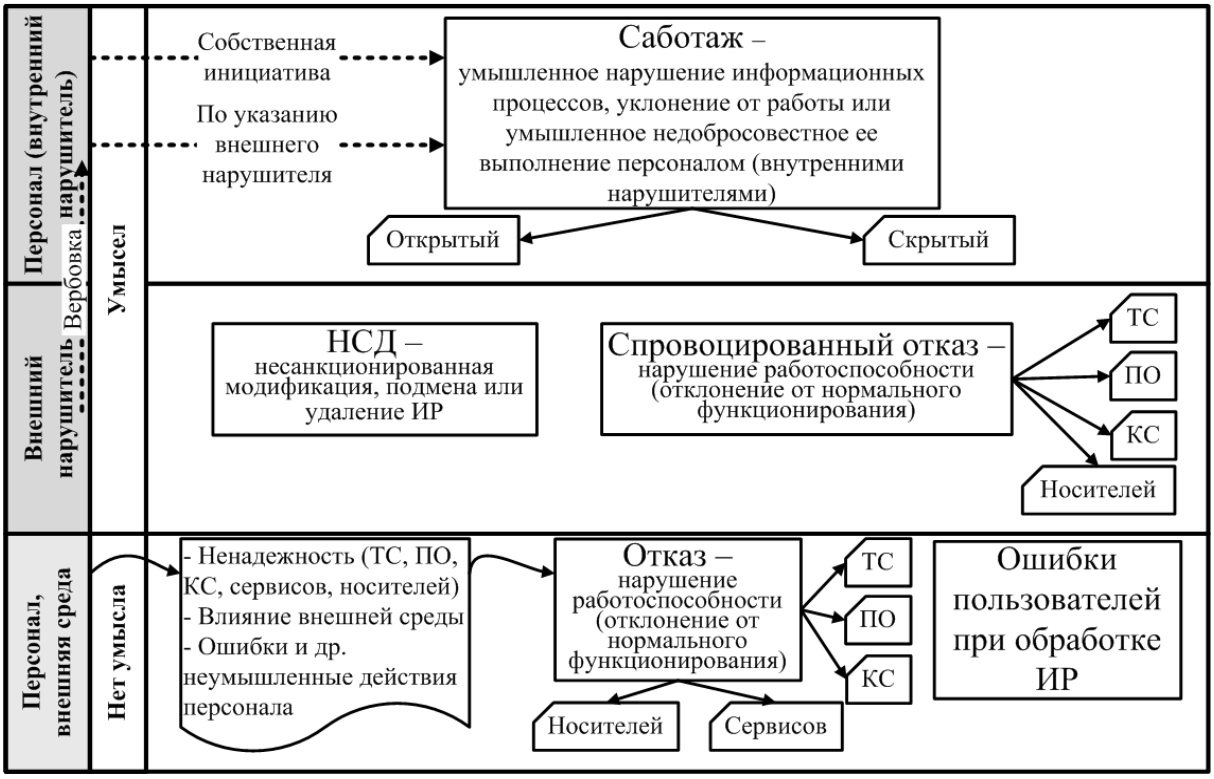


Рис. 2.4. Карта угроз достоверности ИР

2) Несанкционированный доступ (НСД) к ИР – первопричина не-

санкционированной модификации, подмены или уничтожения ИР. Это угроза нарушения целостности и доступности ИР. Субъект данной угрозы чаще всего внешний нарушитель. Даная угроза может эксплуатировать уязвимости ИТКС: недостатки ИТКС и ее компонентов (например, уязвимости ПО). В качестве средств осуществления угрозы могут выступать:

– нелегитимная имперсонация, разглашение, передача или утрата атрибутов разграничения доступа;

– внедрение вредоносного ПО;

– удаление или изменение данных в информационном потоке в канале связи;

– вскрытие используемых алгоритмов шифрования.

3) Спровоцированный отказ – нарушение работоспособности элементов ИТКС: ТС, ПО, кабельной системы (КС). Это угроза нарушения доступности ИР. Субъект данной угрозы – внешний нарушитель. Даная угроза может эксплуатировать уязвимость ИТКС - недостатки ИТКС и ее элементов. В качестве средств осуществления угрозы могут выступать:

– вандализм;

– дезорганизация функционирования системы;

– нелегитимная имперсонация, разглашение, передача или утрата атрибутов разграничения доступа;

– умышленное злоупотребление ресурсами;

– внедрение вредоносного ПО.

4) Отказ ТС, ПО, КС, сервисов – это случайная угроза, вызывающая нарушение доступности ИР. Субъекты данной угрозы: внешняя среда, персонал, производители компонентов ИТКС, организации, предоставляющие услуги, и их работники. Даная угроза может эксплуатировать уязвимости ИТКС: агрессивная внешняя среда, низкая квалификация персонала, неопытность персонала, низкая надежность ТС, ПО, КС, носителей, недостатки структуры ИТКС, недостатки организационного обеспечения (ОО), низкая надежность организаций, предоставляющих услуги. В качестве средств осуществления угрозы могут выступать:

– неумышленное отключение оборудования или изменение режимов работы устройств и программ персоналом;

– злоупотребление ресурсами ИТКС;

– неумышленное использование несанкционированных программ и обработка данных;

– ошибки персонала (при установке, настройке оборудования и программ и т.п.).

5) Ошибки пользователей при обработке ИР – это случайная угроза, вызывающая нарушение целостности и доступности ИР. Субъекты угрозы – пользователи. Даная угроза может эксплуатировать такую уязвимость ИТКС, как низкая квалификация пользователей.

Объект, субъект или явление, которые самим своим существованием, ошибочным функционированием или целенаправленным воздействием негативно влияют на достоверность ИР, будем называть *источником угроз достоверности* (ИУД).

Функции, меры и средства обеспечения достоверности ИР

Для обеспечения требуемого уровня достоверности ИР необходимы:

– механизмы практической реализации гарантированного обеспечения требуемого уровня достоверности;

– средства рациональной реализации необходимых действий по ОДИ;

– способы оптимальной организации и проведения всех действий по ОДИ в процессе функционирования ИТКС.

С целью построения концепции, удовлетворяющей всей совокупности требований, предлагается система концептуальных решений (по аналогии с концепциями обеспечения информационной безопасности [8]):

– формирование полного множества функций обеспечения достоверности ИР (ФОДИР),

– формирование полного множества мер и средств реализации ФОДИР.

Функция обеспечения достоверности ИР – это присущий ИТКС вид деятельности, осуществляемый с целью создания и поддержания условий обеспечения достоверности ИР в условиях информационного противодействия.

Перечислим виды ФОДИР:

1. Предотвращение возникновения угроз (F_1). Угрозы достоверности информации (УгДИ) могут возникнуть случайно или намеренно и, их источник, как правило, люди. Здесь следует снижать количество источников угроз. Для этих целей необходима работа сотрудников службы безопасно-

сти с информаторами в интересах наблюдения и объективной оценки ситуации как внутри коллектива сотрудников, так и вне, среди конкурентов и преступных формирований. В таком предупреждении возникновения угроз весьма существенную роль играет информационно-аналитическая деятельность службы безопасности на основе глубокого анализа криминогенной обстановки и деятельности конкурентов и злоумышленников.

2. Сдерживание угроз (F_2). Главной целью является способствование такому архитектурному построению ИТКС, реализации ИП (включая программное и аппаратное обеспечение) и организационной структуре, чтобы свести к минимуму саму возможность появления уязвимостей в элементах ИТКС, т.е. преследуется упреждающая цель.

3. Обнаружение проявившихся угроз (F_3). Предполагается осуществление таких мер (и, соответственно, использование таких средств) обнаружения угроз, в результате которых проявившиеся УгДИ будут обнаружены еще до того, как они окажут негативное воздействие на достоверность ИР. Иными словами, это функция непрерывного слежения за характеристиками, идентифицирующими конкретные угрозы.

4. Предупреждение воздействия на ИР проявившихся угроз (F_4) – меры, осуществляемые в рамках данной функции, преследуют цель не допустить нежелательного воздействия УгДИ на ИР, если они реально проявились, т.е. данная функция является естественным продолжением предыдущей. Это предполагает использование средств, «устраняющих или ослабляющих воздействие угрозы».

5. Обнаружение воздействия (необнаруженных) угроз (F_5) – функция слежения за ИР с целью своевременного обнаружения фактов воздействия на них необнаруженных (неизвестных) УгДИ. При этом под своевременным понимается такое обнаружение, при котором сохраняются реальные возможности локализации воздействия на информацию.

6. Устранение (локализация, ограничение) обнаруженного воздействия угроз (F_6). Являясь логическим продолжением предыдущей, данная функция предусмотрена с целью недопущения распространения воздействия («недостоверности») на другие (составные) ИР (за пределы максимально допустимых размеров).

7. Ликвидация последствий реализованной атаки (F_7) – проведение таких мероприятий относительно локализованного воздействия УгДИ на информацию, в результате которых дальнейшая обработка информации

может осуществляться без учета имевшего место воздействия. Иными словами, удается восстановить то состояние информационных ресурсов, которое имело место до воздействия УгДИ.

Меры обеспечения достоверности ИР – способы и методики реализации функций обеспечения достоверности ИР за счёт противодействия внутренним и внешним угрозам, снижения воздействия УгДИ на ИР и облегчения восстановления ИР при реализации угроз.

Средства ОДИ – это действия, процедуры, механизмы и устройства, способные с той или иной степенью эффективности реализовать меры обеспечения достоверности ИР.

В табл. 3.1 приведено соответствие мер и средств обеспечения достоверности ИР функциям ОДИ. Заметим, что каждая функция обеспечивается подмножеством полного множества мер и средств.

Таблица 3.1

Связь функций с мерами и средствами обеспечения достоверности информационных ресурсов

ФОДИР / Меры и средства ОДИ	1	2	3	4	5	6	7
Организационное обеспечение	+	+		+		+	+
Физическая защита	+	+		+			
Обеспечение целостности данных			+	+	+		+
Контроль доступа				+			
Идентификация и аутентификация				+			
Аудит			+		+		
Контроль носителей данных		+		+			
Обеспечение надёжности инфраструктуры	+	+					
Сетевое администрирование		+		+		+	+
Защита от вредоносного ПО			+	+		+	+
Обнаружение вторжений			+				
Валидация данных				+	+		

Показатели достоверности информации

Оценивание достоверности информационного ресурса производится на основе:

– свойств информации, комплексно определяющих категорию «достоверность»;

– множества целей управления достоверностью, определяемой прикладными задачами ИТКС, на базе которых формулируются основные критерии обеспечения свойств;

– множества показателей (качественно-количественных представлений измеренных характеристик достоверности ИР), сопоставимых с критериями и позволяющих отнести оцениваемую достоверность ИР к тому классу, который определяется исходя из целей.

Свойства информации, предлагаемые для оценки ее достоверности:

– аутентичность – соответствие информации об объекте или явлении его действительному состоянию;

– полнота – отражение всех существенных в рамках задачи характеристик объекта;

– актуальность (своевременность) – отражение характеристик объекта или явления с задержкой, допустимой в решаемой задаче;

– целостность – неизменность в процессах хранения, передачи, переработки и представления данных в ИТКС.

Для оценки достоверности используются следующие критерии достоверности:

1) доверие к источнику информации:

– уверенность в том, что информация поступила именно из данного источника информации (ИстИ);

– уверенность в том, что данный ИстИ обладает полнотой данных для предоставления конкретной информации;

– уверенность в том, что данный ИстИ предоставил все необходимые (запрошенные) данные;

– уверенность в том, что при передаче в ИТКС данных от ИстИ не допущено искажения информации (т.е. присутствовали какие-либо ошибки или не были внесены данные с ложной информацией);

2) доверие к системе обработки данных (ИТКС):

– уверенность в том, что данная информация не искажена на любом участке технологического процесса обработки данных в направлении от ИстИ к пользователю;

– уверенность в том, что запрошенные задачей пользователя данные будут актуальны и доставлены к задаче вовремя;

– доверие к результату.

Нарушение достоверности проявляется в нарушении соответствую-

щего показателя, т.е. недостоверным ИР является неаутентичный, неполный, неактуальный или нецелостный ИР. Если ИР является структурированным, то это означает, что отдельные составляющие ИР могут быть достоверными либо нет, что обуславливает введение такого понятия как «уровень достоверности ИР».

Все показатели достоверности являются динамическими, но при этом только для показателя актуальности можно установить функциональную зависимость от времени и ряда характеристик ИТКС, т.к. он определяется в целом детерминированными процессами устаревания информации. Временная устойчивость показателя полноты ограничена снизу временной устойчивостью показателя целостности при неизменности требований к информационной модели объекта. Показатель целостности определяется стохастическими процессами и может быть спрогнозирован с определённой вероятностью. Показатель аутентичности остаётся постоянным при достаточном уровне показателя актуальности и постоянстве показателя целостности.

2.3. Общий подход к оценке достоверности информации

Рассмотрим модель функционирования ИТКС с точки зрения возникновения и реализации ДФ (угроз ДИ), а также противодействия им мер обеспечения достоверности (рис. 2.5).

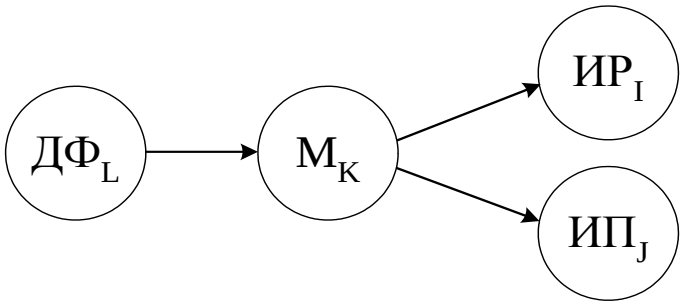


Рис. 2.5. Модель функционирования ИТКС в условиях информационных воздействий

Пусть в ИТКС определены и классифицированы имеющиеся «уязвимости», как системные структурно-функциональные недостатки (СФН), составлен полный перечень возможных ДФ и мер обеспечения достоверности. Обозначим через n общее число ДФ, s – число СФН, x – число мер

обеспечения достоверности ИР и ИП.

l-й ДФ можно охарактеризовать относительной частотой возникновения $p_l^{ДФ}$. Совокупность элементов системы обеспечения достоверности информации (СОДИ) можно охарактеризовать показателями качества защитных мер передачи, хранения и обработки *i*-го ИР в условиях воздействия *l*-го ДФ: $x_{il}^{П}$, x_{il}^{X}, x_{il}^{O}. Эти меры понижают вероятность нарушения достоверности *i*-го ИР *l*-м ДФ $p_{il}^{НД}$ путём обеспечения аутентичности (сокращение числа ошибок в процессе обработки) и целостности (недопущение искажений в процессе хранения и передачи).

Оценка полноты информации требует рассмотрения характеристик источников информации (ИстИ). Области знаний ИстИ соответствует полный тезаурус ключевых понятий T. Область знаний может быть разделена на ряд предметных областей с частными тезаурусами T_K. Способность источника к предоставлению требуемой информации из области знаний определяется соответствием предметных областей.

Пусть запрос на данные, предоставляемые ИстИ (если данные будут предоставлены по запросу), либо область интересов получателя информации (если сбор информации происходит без запроса, по наличию информационных источников) покрывает *A*-предметных областей с тезаурусами $T_a,\ a \in A$. Пусть информация от ИстИ в данном сообщении покрывает *B*-предметных областей с тезаурусами $T_b,\ b \in B$. Тогда возможны следующие состояния информированности ИстИ по запросу:

– избыточная информированность – множество предметных областей ИстИ включает множество предметных областей запроса – $A \subset B$. Тогда степень соответствия предметных областей *S=1*;

– полная информированность – множество предметных областей источника и множество предметных областей запроса совпадают – $A = B$. Степень соответствия предметных областей *S = 1*;

– неполная информированность – множество предметных областей ИстИ и множество предметных областей запроса частично совпадают, но не равны – $A \cap B \neq 0,\ A \neq B$. Степень соответствия предметных областей $S = \sum_{\forall a \in A} T_a \,/\, \sum_{\forall b \in B} T_b$;

– неинформированность ИстИ – множество предметных областей

ИстИ и множество предметных областей запроса не совпадают – $A \cap B = 0$. Степень соответствия предметных областей $S=0$.

Показатель достоверности i-го ИР, полученного из j-го ИстИ в условиях воздействия на ИР n дестабилизирующих факторов равен

$$D_{ij} = S_j \cdot \left(1 - \sum_{l=1}^{n} p_l^{ДФ} \cdot p_{il}^{НД} \right) = S_j \cdot \left(1 - \sum_{l=1}^{n} p_l^{ДФ} \cdot \left(1 - x_{il}^{П} \cdot x_{il}^{X} \cdot x_{il}^{O} \right) \right), \qquad (2.1)$$

где $p_l^{ДФ}$ – относительная частота возникновения l-го ДФ,

$p_{il}^{НД}$ – возможность нарушения достоверности (уничтожение, модификация) i-го ИР l-м ДФ, которая зависит от качества элементов СОДИ

$$p_{il}^{НД} = 1 - \left(1 - \prod_{q_1} \left(1 - \delta_{ilq_1}^{П} \cdot x_{ilq_1}^{П} \right) \right) \cdot \left(1 - \prod_{q_2} \left(1 - \delta_{ilq_2}^{X} \cdot x_{ilq_2}^{X} \right) \right) \cdot \left(1 - \prod_{q_3} \left(1 - \delta_{ilq_3}^{O} \cdot x_{ilq_3}^{O} \right) \right), \qquad (2.2)$$

где $x_{ilq_1}^{П}, x_{ilq_2}^{X}, x_{ilq_3}^{O}$ - показатели качества q-го средства защиты соответственно процесса передачи, хранения или обработки i-го ИР в условиях воздействия l-го ДФ;

$0 \le \delta_{ilq} \le 1$ - предел достаточности q-го средства защиты при условии, что оно является единственным средством по противодействию l-му ДФ, который способен нарушить достоверность i-го ИР.

2.4. Модель управления процессом обеспечения достоверности

С формальной точки зрения обеспечение достоверности информации в ИТКС рассматривается как задача дискретного управления многошаговым процессом с задаваемым желаемым (конечным) состоянием достоверности информационных ресурсов $P_{\hat{e}}$, известным начальным состоянием P_0 и набором допустимых действий D таких, что действие $d_i \in D$, реализуемое на i-м шаге, переводит достоверность информационных ресурсов из состояния P_i в состояние P_j с более высокими показателями (достоверности). Задача управления состоит в выборе оптимальной последовательности

действий $D^* = \langle d_0^*, d_1^*, ... \rangle$ и, соответственно, состояний $P^* = \langle P_0^*, P_1^*, ... \rangle$, таких, что в результате достигается желаемое (или максимально возможное – экстремальное) значение достоверности.

Формальная модель управления достоверностью ИР

Рассмотрим процесс управления достоверностью информации на примере одного ИР, а далее на этой основе предложим множественную модель. Объект управления – достоверность ИР. Состояния достоверности будем описывать идентификатором P, который представляет собой вектор частных показателей достоверности $P = \{pd_1, ...\}$.

Объект управления может подвергаться воздействиям двух видов:

а) неуправляемые. Это воздействия (влияния) внешней среды $U(t)$. Внешней средой (окружением) для ИР является ИТКС с множеством своих структурных элементов и информационных процессов (ИП), в которых «задействован» данный ИР. Элементы и ИП, инициируемые соответствующим программным обеспечением, могут быть ненадежными, уязвимыми к информационным атакам злоумышленников, сама ИТКС имеет определенный уровень функциональной устойчивости и структурной живучести. Данный вид воздействий всегда приводит к «ухудшению» показателей достоверности;

б) целенаправленные (управляющие) воздействия $X \in D$, функционально состоящие из воздействий $X^{(1)}$,обеспечивающих (может быть, даже повышающих) те или иные показатели достоверности, и воздействий $X^{(2)}$, контролирующих текущие значения данных показателей. Пусть эти воздействия вырабатывает специализированная система обеспечения достоверности информации (СОДИ) в результате обработки результатов измерения состояния объекта и внешней среды, а также выбранного алгоритма функционирования (стратегии), выделенных ресурсов и поставленной задачи по достижению определенной цели.

Процесс управления достоверностью ИР представлен на рис.2.6.

На схеме модель ИР представлена автоматом*:* преобразователями $F^{(1)}$ и $F^{(2)}$, а также модулем, позволяющим хранить идентификатор достоверности (ИД), текущие значения показателей, память состояния (*ПС)*

достоверности ИР. Преобразователь $F^{(1)}$ реализует функцию переходов к новому состоянию ИД $P(t+1)$ в зависимости от его текущего состояния $P(t)$, состояния среды $U(t)$ и воздействия $X^{(1)}(t)$:

$$P(t+1)=F^{(1)}\left\{P(t),U(t),X^{(1)}(t)\right\}. \qquad (2.3)$$

Оценку идентификатора достоверности $P^{*}(t)$ для очередного шага выработки управляющих воздействий получают по результатам «измерения» показателей $P(t)$. Данная процедура инициируется контролирующим воздействием $X^{(2)}(t)$:

$$P^{*}(t)=F^{(2)}\left\{P(t),X^{(2)}(t)\right\}. \qquad (2.4)$$

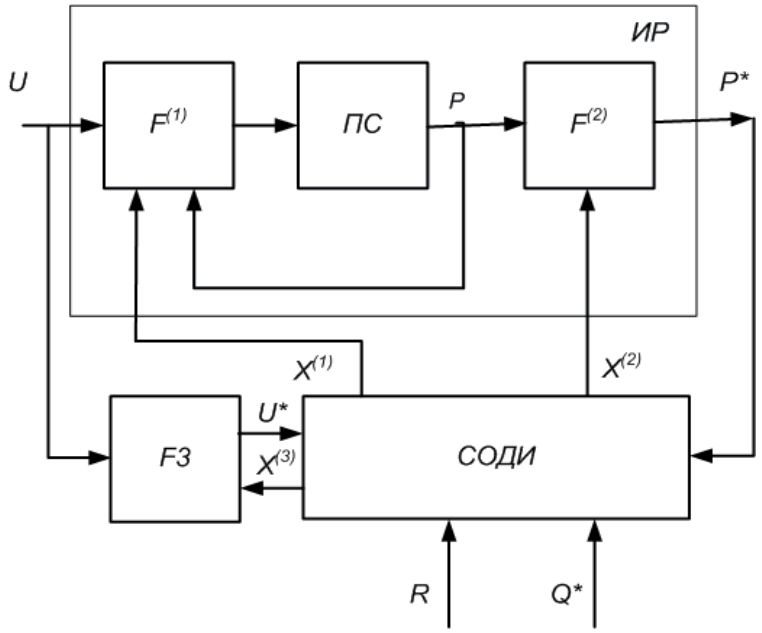


Рис. 2.6. Процесс управления достоверностью ИР

Вид функции выходов (преобразователь $F^{(2)}$) во многом определяется способом и методикой контроля (измерения) показателей ИД.

Входы автомата - управления $X \in D$ ($X^{(1)}$ и $X^{(2)}$), вырабатываемые управляющим устройством (СОДИ) согласно алгоритму управления ϕ, выбранному из множества Φ известных алгоритмов $\phi \in \Phi$ (стратегии до-

стижения конкретной цели Q^*), выделяемых ресурсов R, полученной информации о состоянии среды U^* и оценки состояния объекта P^*:

$$X = \phi\left(Q^*, P^*, U^*, R\right). \tag{2.5}$$

Информацию о состоянии внешней среды $U(t)$ для очередного шага выработки управляющих воздействий доставляют «датчики», осуществляющие функциональное преобразование $F^{(3)}$, в виде измеренных значений $U^*(t)$. Контроль осуществляется в соответствии с воздействием из $X^{(3)}$:

$$U^*(t) = F^{(3)}\left\{U(t), X^{(3)}\right\}. \tag{2.6}$$

Стратегия (алгоритм) управления ϕ, в общем плане, должна минимизировать число шагов управления (время достижения цели Q^*) для достижения $P_к$.

Модель управления достоверностью множества информационных ресурсов представлена на рис. 2.7.

Обеспечение достоверности информации в ИТКС осуществляется одновременно по N информационным ресурсам $(ИР_1, \ldots, ИР_N)$, представленных в модели кортежами типа $\left\langle F_i^{(1)}, ПС_i, F_i^{(2)} \right\rangle, i = 1, \ldots, N$. Это множественный объект управления.

«Функционирование» модели объекта описывается системой:

$$\begin{cases} P_1(t+1) = F_1^{(1)}\left\{P_1(t), U_1(t), \ldots, U_M(t), X_1^{(1)}\right\} \\ \qquad\qquad \ldots \\ P_N(t+1) = F_N^{(1)}\left\{P_N(t), U_1(t), \ldots, U_M(t), X_N^{(1)}\right\} \end{cases}, \tag{2.7}$$

где $P_1(t), \ldots, P_N(t)$ - текущие «состояния» ИР – значения показателей индикаторов достоверности;

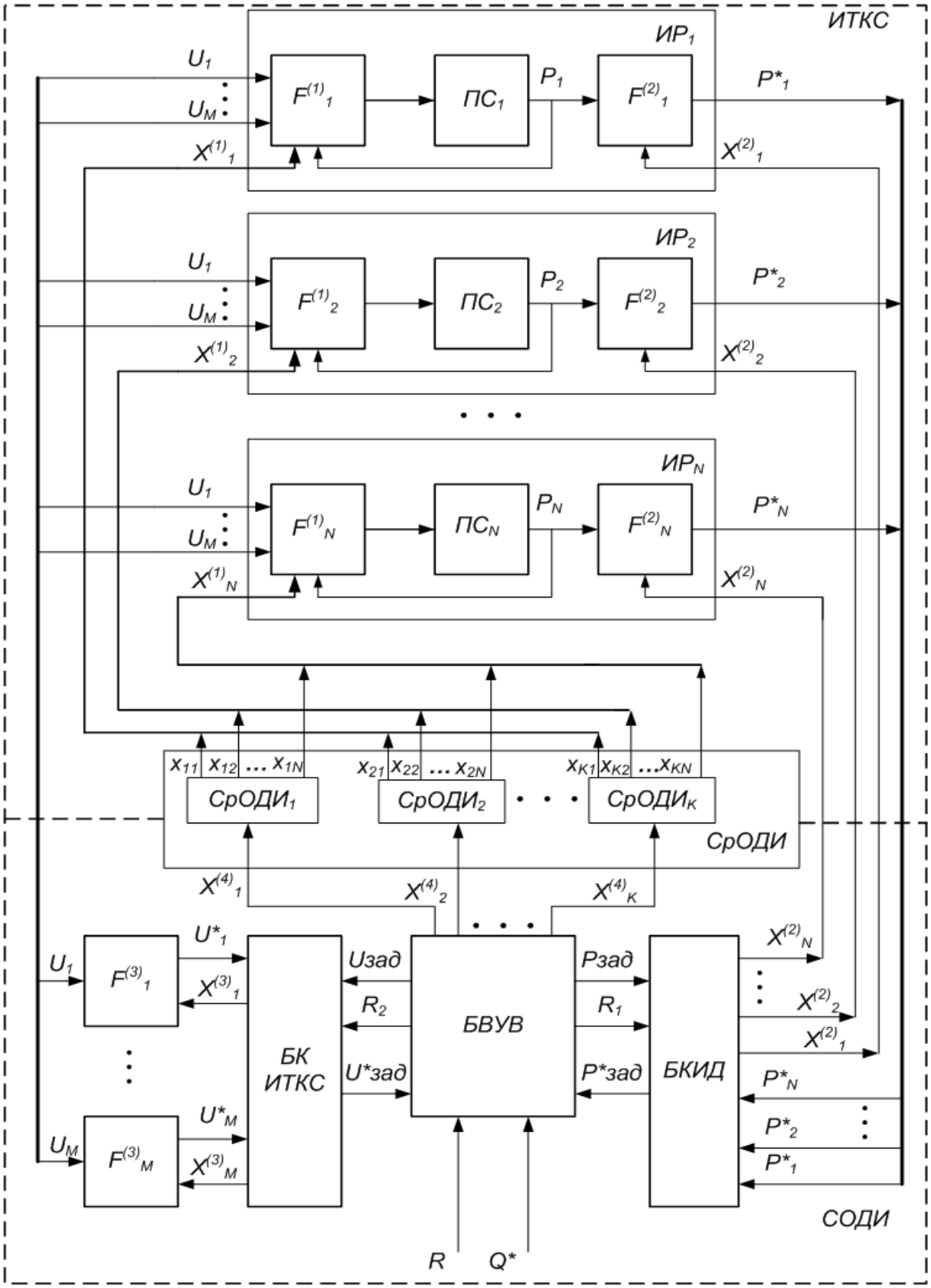


Рис. 2.7. Схема управления достоверностью множества ИР

$U_1(t),...,U_M(t)$ – множество текущих значений параметров факторов, дестабилизирующих устойчивое функционирование ИТКС и приводящих к снижению показателей достоверности. Будем полагать, что данные значения одинаковы для всех ИР;

$X_1^{(1)}(t),...,X_N^{(1)}(t)$ – управляющие воздействия.

Множество $\{x_{11}(t),...,x_{1K}(t),...,x_{N1}(t),...,x_{NK}(t)\}$ – множество текущих значений характеристик средств обеспечения ДИ. Подмножества вида $X_i^{(1)}(t)=\{x_{i1}(t),...,x_{iK}(t)\}, i=1,...,N,$ как «совокупное» участие средств обеспечения ДИ, определяют «уровень противодействия» дестабилизации ИТКС по отношению к $ИР_i$.

Средства обеспечения ДИ являются, по сути, исполнительными механизмами системы управления и позволяют повышать (или, по крайней мере, не снижать) показатели достоверности.

Система обеспечения достоверности информации включает в себя блок выработки управляющих воздействий (БВУВ), блок контроля идентификаторов достоверности (БКИД) и блок контроля информационно-телекоммуникационной системы (БК ИТКС).

Основным назначением БВУВ является формирование программы повышения достоверности (ППД) – оптимальной последовательности действий $D^*=\langle d_0^*, d_1^*, ...\rangle$ (и, соответственно, состояний $P^*=\langle P_0^*, P_1^*, ...\rangle$) таких, что в результате достигается желаемое (или максимально возможное) значение достоверности. Действия d_j^*, кроме контрольных, обеспечивают выработку сигналов управления $(X_1^{(4)},...,X_K^{(4)})$ средствами обеспечения достоверности информации, включая выбор (инициирование) определенного средства и задание режимов его функционирования.

Процесс формирования ППД D^* производится на основе:

– текущего состояния объекта (измеренные значения показателей идентификаторов достоверности $\{P_1^*(t),...,P_N^*(t)\}$) и внешней среды (измеренные значения $\{U_1^*(t),...,U_M^*(t)\}$);

– имеющихся на текущий момент времени средств обеспечения ДИ, «не задействованных» в других ППД;

– набора типовых процедур управления – функций и мер

обеспечения достоверности, которые средства обеспечения ДИ потенциально способны реализовать;

– выделенного ресурса R (например, времени) и цели Q^*. Здесь под целью в общем плане понимается задача на обеспечение (повышение) определенных показателей достоверности конкретных ИР.

Сложность процесса формирования ППД заключается в том, что средства обеспечения ДИ едины для всех ИР, а требования по достоверности в конкретной цели Q^* распространяются на каждый ИР в отдельности.

Измерения (оценки) текущих значений показателей идентификаторов достоверности осуществляются БКИД. В зависимости от выделенного БВУВ ресурса R_1 на контроль запрашиваемого подмножества показателей идентификаторов достоверности $P_{зад}$ блок вырабатывает множество $\left\{X_1^{(2)},\ldots,X_N^{(2)}\right\}$ методик измерения (оценивания) параметров. В результате преобразований $F_i^{(2)}, i=1,\ldots,N$ получаем оценки $P_1^*(t),\ldots,P_N^*(t)$.

Измерения (оценки) текущих значений характеристик ИТКС (показателей живучести, надежности, информационной активности злоумышленников и т.д.) осуществляются БК ИТКС. В зависимости от выделенного БВУВ ресурса R_2 на контроль запрашиваемого подмножества текущих характеристик ИТКС $U_{зад}$ блок вырабатывает множество $\left\{X_1^{(3)},\ldots,X_M^{(3)}\right\}$ методик измерения (оценивания) характеристик. В результате преобразований датчиков $F_i^{(3)}, i=1,\ldots,M$ получаем оценки $U_1^*(t),\ldots,U_M^*(t)$.

При нормативном подходе к управлению достоверностью (например, при строго определенном критерии достижения показателей ИД и ограниченном времени) задача сводится к поиску системы с минимальной стоимостью, т.е. требуется решить задачу:

$$\begin{cases} \sum_{i=1}^{N}\sum_{j=1}^{NS} c_{ij}(p_{0_{ij}}, t_{зам_{ij}}) \to min, \\ \tilde{p}_{ij}(p_{0_{ij}}, t_{ij}) \geq p_{жел_{ij}}, \\ max_{ij}(t_{зам_{ij}}) \leq t_{доп}. \end{cases} \qquad (2.8)$$

где c_{ij} – «приведенная стоимость» повышения достоверности j-го показа-

теля ($j = 1,\ldots,NS$) идентификатора достоверности i-го ИР ($i = 1,\ldots,N$);

$p_{0_{ij}}$ – начальный уровень j-го показателя ИД i-го ИР;

$t_{зам_{ij}}$ – время, необходимое на достижение $p_{жел_{ij}}$ - желаемого уровня j-го показателя ИД i-го ИР;

$\tilde{p}_{ij}$ – достигнутый (от $p_{0_{ij}}$) за время t_{ij} уровень j-го показателя идентификатора достоверности i-го ИР;

$t_{доп}$ – допустимое время, выделяемое на повышение достоверности ИР;

N – общее количество ИР;

NS – количество показателей ИД.

Более близкой к реальной представляется задача максимального повышения показателей ИД при ограниченных ресурсах и допустимых затратах $C_{доп}$

$$\begin{cases} \sum_{i=1}^{N}\sum_{j=1}^{NS}\tilde{p}_{ij}(p_{0_{ij}}, t_{зам_{ij}}) \to max, \\ \tilde{p}_{ij}(p_{0_{ij}}, t_{ij}) \geq p_{жел_{ij}}, \\ max_{ij}(t_{зам_{ij}}) \leq t_{доп}, \\ \sum_{i=1}^{N}\sum_{j=1}^{NS} c_{ij}(p_{0_{ij}}, t_{зам_{ij}}) \leq C_{доп}. \end{cases} \qquad (2.9)$$

Выводы по главе

Предложен концептуальный подход к обеспечению достоверности информации в информационно-телекоммуникационных системах, функционирующих в условиях информационного противодействия. Описана новая модель управления процессом обеспечения достоверности, отличающаяся учетом работы ИТКС в условиях: дестабилизирующих факторов, оказывающих влияние на достоверность информации в ИТКС; активного противодействия; мониторинга и динамического определения уровня достоверности источников информации; ограничений ресурсов различных классов в ИТКС.

Показано, что структура и параметры модели закладываются подмножествами вариантов (многовариантная модель). Настройка данной системы под особенности конкретного предприятия позволит получить инструмент для прогноза развития ситуации и оценки рисков снижения достоверности информации в ИТКС. Кроме того, предложенная модель может стать частью автоматизированной системы мониторинга и управления процессами обеспечения достоверности информации в ИТКС конкретных предприятий.

Список библиографических ссылок

1. Абрамов К.Г., Монахов Ю.М. Стохастические модели распространения нежелательной информации в социальных сетях// Сборник научных трудов Sworld. 2011. Т.5(4). С. 42-45.
2. ГОСТ Р 53114-2008 Защита информации. Обеспечение информационной безопасности в организации. Основные термины и определения. М.: СандартинФорм, 2009. 20 с.
3. Груздева Л.М., Монахов Ю.М., Монахов М.Ю. Экспериментальное исследование производительности корпоративной телекоммуникационной сети // Проектирование и технология электронных средств. 2009. № 4. С. 28-31.
4. Жилин Д.М. Теория систем. М.: УРСС, 2004. 183 с.
5. Илларионов Ю.А., Монахов М.Ю. Безопасное управление ресурсами в распределенных информационных и телекоммуникационных системах. Владимир: Владим. гос. ун-т. Владимир, 2004. 204 с.
2. Костров А.В. Основы информационного менеджмента. – М.: Финансы и стат., 2001. 336 с.
6. Лернер А.Я. Начала кибернетики. М.: Наука, ГРФМЛ, 1967. 400 с.
7. Малюк А.А. Информационная безопасность: концептуальные и методологические основы защиты информации. М.: Горячая линия – Телеком, 2004. 280 с.
8. Мишурин А.О., Семенова И.И. Система управления моделями в области информационного противоборства // Вестник СГТУ. 2010. №4(49), Вып. 1. С. 150-160.
9. Монахов М.Ю., Семенова И.И. Когнитивная модель оценки

уровня достоверности информации в синтезируемой научно-производственной документации // Современные проблемы науки и образования. 2014. № 1; URL: http://www.science-education.ru/115-12147 (дата обращения: 18.09.2015)

10. Монахов Ю.М., Семенова И.И., Медведникова М.А., Костина Н.В. Методика выявления семантических дифференциалов для автоматизации оценки психосемантического профиля пользователя социальной сети // Современные проблемы науки и образования. 2013. № 5. URL: http://www.science-education.ru/111-10320 (дата обращения: 18.09.2015)

11. Монахов М.Ю., Файман О.И. Инвентаризация информационных ресурсов как основа безопасного функционирования АСУ // Известия высших учебных заведений. Приборостроение. 2012. Т. 55. № 8. С. 35-39.

12. Остапенко Г.А. Информационные операции и атаки в социотехнических системах / под ред. В.И. Борисова. М: Горячая линия-Телеком, 2006. 184 с.

13. Остринская Л.И., Семенова И.И., Дороболюк Т.Б. Теория и практика работы с современными базами и банками данных. Омск: Изд-во СибАДИ, 2005. 250 с.

14. Ризниченко Г.Ю. Математические модели в биофизике и экологии. М., Ижевск: Институт компьютерных исследований, 2003. 184 с.

15. Семенова И.И. Аспекты информационной безопасности в системах управления базами моделей // МИК-2012. Омск: Правительство Омской области, 2012. С. 246-252.

16. Сухарев М.С., Монахов Ю.М. Модель оценки функциональной устойчивости бизнес-процессов // Вестник Костромского государственного университета. 2011. №5-6. С. 4-6.

17. Сухарев М.С., Монахов Ю.М., Файман О.И. Применение системного подхода к оценке функциональной устойчивости бизнес – процессов // Сборник научных трудов Sworld. 2011. Т. 5. № 4. С. 70-73.

18. Тарасенко Ф.П. Прикладной системный анализ (наука и искусство решения проблем). Томск: Издательство Томского университета, 2004. 186 с.

19. Erwin Folmer, Jack Verhoosel State of the Art on Semantic IS Standardization, Interoperability & Quality. University of Twente, 2011. 167 p.

20. Gruzdeva L.M., Monakhov M.Yu. Early detection algorithm for attacks against information resources of automatic manufacturing control systems

// Automation and Remote Control. 2011. V. 72. № 5. P. 1075-1079.

21. Heiko Müller, Johann-Christoph Freytag, Ulf Leser Improving data quality by source analysis // J. Data and Information Quality. 2012. V. 2. № 4. 38 p.

22. Using information quality for the identification of relevant web data sources: a proposal / Bernadette Farias Lóscio, Maria C.M. Batista, Damires Souza, Ana Carolina Salgado // IIWAS '12. ACM, NY, USA, 2012. P. 36-44.

ГЛАВА 3. **МЕТОДИКА ОЦЕНКИ ПОКАЗАТЕЛЕЙ ДОСТОВЕРНОСТИ ИНФОРМАЦИИ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СИСТЕМЕ**

Обеспечение требуемого уровня достоверности информации, циркулирующей в ИТКС, невозможно без оценки текущего уровня, который определяют множество различных факторов. Методика оценки текущего уровня достоверности информации как вероятности сохранения её неизменности в информационном потоке в условиях информационных воздействий требует исследования следующих критериев [8, 9, 19, 24-27]:

– относительные частоты возникновения дестабилизирующих факторов - ДФ (угроз достоверности) $P_{ДФ}$,

– возможности нарушения достоверности информации $P_{НД}$.

– вероятности обнаружения попыток нарушения достоверности $P_{О}$.

Соответственно, повышение уровня достоверности в условиях информационных воздействий возможно работой в трёх направлениях [20]:

– изменение структуры ИТКС с целью уменьшения количества уязвимостей / структурно-функциональных недостатков (СФН), приводящих к возникновению ДФ;

– повышение качества системы обеспечения достоверности информации (СОДИ).

Необходимым условием достижения требуемого уровня достоверности информации (ДИ) является построение комплексной СОДИ. ИТКС предприятий существуют в системе товарно-денежных отношений, в основе которой лежит понятие экономической эффективности, и не могут себе позволить бесконтрольно и безосновательно тратить материальные ресурсы на проведение каких-либо мероприятий. Вследствие этого оценка уровня ДИ и принятие решений по проведению мероприятий для его повышения поднимают сопутствующую задачу оценки экономического эффекта от их проведения.

Следовательно, в общий метод оценки достоверности необходимо включить следующие дополнительные критерии:

– стоимость информационных ресурсов (ИР) $S_{ИР}$,

– стоимость средств обработки информации $S_{ОИ}$,

– стоимость системы обеспечения достоверности $S_{СОДИ}$,

– суммарный риск информации $R_{ИР}$,

– суммарный риск средств обработки информации $R_{ОИ}$,

– суммарный риск системы обеспечения достоверности $R_{СОДИ}$.

В главе предлагается общая модель оценки достоверности информации в ИТКС в условиях информационных воздействий с учетом решения сопутствующих задач оценки рисков и экономической эффективности мероприятий по повышению достоверности. Приводятся результаты экспериментального исследования обеспечения достоверности информации на промышленном предприятии.

3.1. Общая модель оценки достоверности информации в ИТКС

Вероятностные, стоимостные критерии и критерии рисков можно назвать параметрами ИТКС. Зависящие от них общий коэффициент достоверности информации $D_{ИР}$ и экономическая эффективность E_f, являются *показателями качества* СОДИ.

Можно выделить следующие зависимости между параметрами ИТКС:

$$\begin{aligned} P_{ДФ} &= f(S_{ИР}, S_{ОИ}), \\ P_{НД} &= f(P_{ДФ}, S_{ИР}, S_{ОИ}, S_{СОДИ}), \\ R_{ИР} &= f(S_{ИР}, S_{СОДИ}, P_{ДФ}, P_{НД}), \\ R_{ОИ} &= f(S_{ОИ}, S_{СОДИ}, P_{ДФ}, P_{НД}), \\ R_{СОДИ} &= f(S_{СОДИ}, P_{ДФ}, P_{НД}). \end{aligned} \qquad (3.1)$$

Оценка текущего уровня ДИ не является самоцелью, она необходима для оптимизации параметров ИТКС с целью повышения ДИ в ИТКС в условиях информационных воздействий. Общая модель контроля показателей ДИ в ИТКС представлена на рис. 3.

Зависимости (3.1) носят вероятностный характер. Математическая статистика и теория вероятностей используют экспериментальные данные, обладающие точностью и достоверностью [5, 6, 28]. В данном случае понятия точности и достоверности не всегда применимы, т.к. эти вероятности зависят от «человеческих знаний» [17]. Поэтому для достоверной ко-

личественной оценки показателей качества СОДИ необходимо использовать теорию нечётких множеств, которая оперирует не понятием «вероятность», а понятием «возможность», что более адекватно соответствует решению трудно формализуемых задач.

Вероятности возникновения, обнаружения и устранения ошибок являются числовыми характеристиками и могут быть определены статистическими методами. Но ИТКС состоит из множества разнородных компонентов, их состав может весьма существенно различаться. Т.е. получить статистику по достаточному количеству однотипных систем почти всегда невозможно. Решением здесь может быть экспертная оценка. Для показателей достоверности и эффективности необходимо получение количественного результата, что существенно осложняется рядом факторов:

– наличие сложной опосредованной взаимосвязи этих показателей с параметрами ИТКС;

– необходимость учёта и формализации ряда параметров ИТКС, многие из которых изначально заданы качественными величинами и имеют элементы неоднозначности;

– наличие существенной взаимосвязи и взаимозависимости этих параметров, имеющих противоречивый характер;

– трудность получения исходных данных, необходимых для оценки, в частности на ранних этапах проектирования СОДИ.

Проектирование, организация и применение СОДИ фактически связаны с неизвестными событиями в будущем и поэтому всегда содержат элементы неопределённости. Кроме того, присутствуют и другие причины неоднозначности, такие как недостаточно полная информация для принятия управленческих решений или социально-психологические факторы.

В процессе решения задачи оптимизации приходиться учитывать следующее [1]:

– преимущественно нечёткое описание множества исходных данных, в частности описание стандартов, используемых при построении СОДИ и задаваемых в виде требований;

– сама постановка задачи выбора обычно является нечёткой, при этом предпочтение того или иного показателя определяется экспертной информацией;

– многокритериальность задачи.

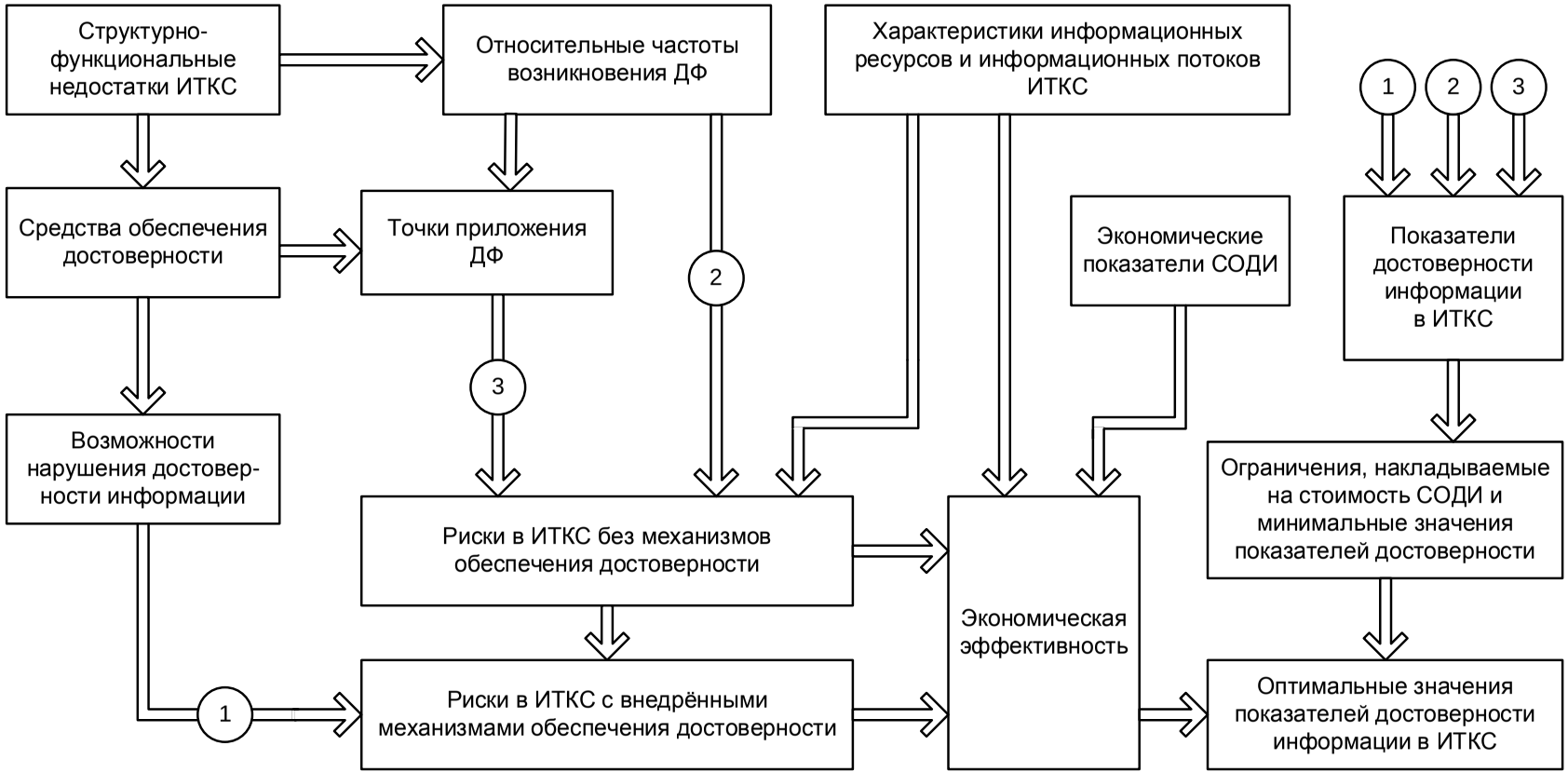


Рис. 3.1. Общая модель контроля показателей достоверности информации в ИТКС

С экономической точки зрения эффективным будет такой комплекс мероприятий по повышению достоверности информации в ИТКС, при котором выполнены следующие условия:

$$\begin{cases} S_{СОД} \le \Delta R_{ИР} + \Delta R_{ОИ} + \Delta R_{СОД} \\ S_{СОД} \le S_{ИР} + S_{ОИ} \end{cases} \quad (3.2)$$

где $\Delta R_{И} + \Delta R_{ОИ} + \Delta R_{СЗИ}$ – общее снижение рисков в ИТКС.

Если первое условие очевидно: получаемый эффект не должен быть меньше стоимости средств и мероприятий, т.е. затрат, то второе не настолько очевидно. И тем не менее, действительно большая часть затрат предприятия направлена на выполнение его основной деятельности, а не на реализацию функций СОДИ.

С функциональной точки зрения, которая определена общим коэффициентом достоверности, качество системы определяют условия, показывают соответствие некоторому минимальному порогу, например, так:

$$D_i \ge D_{i.min}, \quad \forall i = \overline{1,z}. \quad (3.3)$$

Условия (3.2) и (3.3), очевидно, противоречивы: повышение уровня ДИ в ИТКС на определённом этапе возможно только при увеличении инвестиций в СОДИ. Из этого следует постановка задачи оптимизации:

$$\begin{cases} D_{ИР} \to max, \\ D_i \ge D_{i.min}, \quad \forall i = \overline{1,z}, \\ S_{СОД} \le \Delta R_{ИР} + \Delta R_{ОИ} + \Delta R_{СОД}, \\ S_{СОД} \le S_{ИР} + S_{ОИ}. \end{cases} \quad (3.4)$$

Модель контроля показателей достоверности информации в ИТКС, представленная на рис. 3.1, является общей, отражает структуру взаимосвязей параметров ИТКС, а также порядок их оценки. Характер взаимосвязей можно описать моделью, представленной на рис. 3.2.

Область ДФ T представляет собой список всех возможных ДФ для данной ИТКС, каждый ДФ характеризуется относительной частотой возникновения.

Множество СФН *V* представляет собой перечень всех существующих уязвимостей в данной ИТКС, основными характеристиками которых являются значимость и доступность для злоумышленника. В совокупности множества ДФ и СФН и их взаимосвязи образуют перечень способов достижения цели – информационное воздействие на ИР.

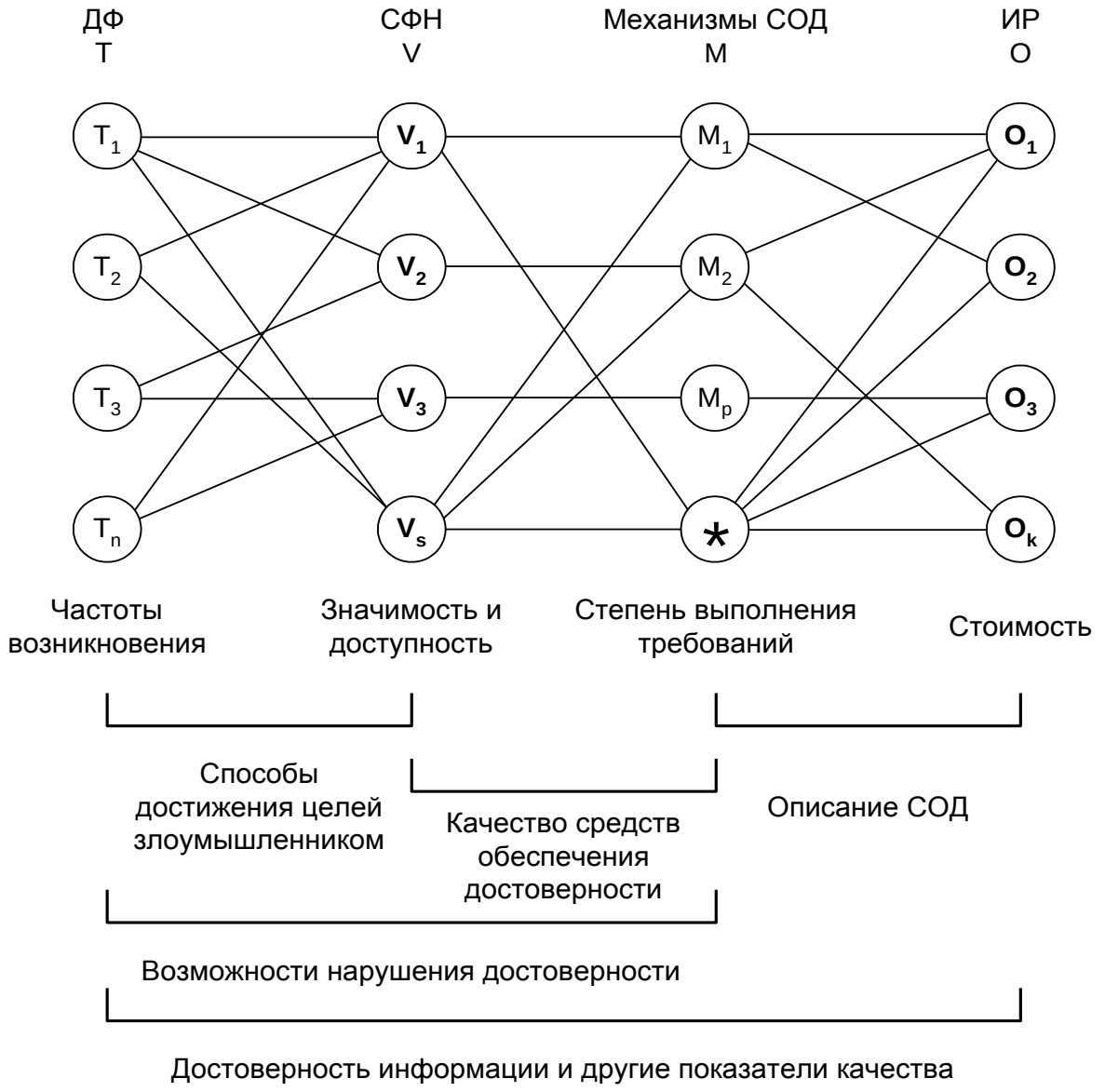


Рис. 3.2. Модель, описывающая характер взаимосвязи параметров ИТКС и показателей качества СОДИ

Система обеспечения достоверности *M* представляет собой имеющийся в организации набор средств обеспечения достоверности. Качество средств обеспечения достоверности определяется при рассмотрении мно-

жества V и множества M с их взаимосвязями, и характеризуется возможностями преодоления каждого барьера, ассоциированного с каждым СФН. Здесь искусственно добавлен элемент (*), показывающий, что ряд СФН может быть вообще не перекрыт каким-либо барьером.

Область O представляет собой совокупность ИР, характеристики которых: ценность и стоимость, определяемая исходя из финансовых потерь организации, ассоциированных с восстановлением ресурса, либо с упущенной выгодой, связанными с уничтожением, тиражированием, либо блокированием доступа к защищаемому ресурсу.

Вместе множество M, множество O и их взаимосвязи дают полное описание СОДИ организации.

Вероятности нарушения достоверности в ИТКС определяются, с одной стороны, способами достижения целей, а с другой – качеством средств обеспечения достоверности. В целом, достоверность информации и другие показатели качества будут определены при подробном рассмотрении всех областей графа, при этом эффективность характеризуется соотношением рисков от нарушения достоверности информации в ИТКС организации при отсутствии и при наличии СОДИ.

3.2. Алгоритм проведения экспертизы параметров ИТКС

В условиях разнородности элементов и параметров ИТКС и преимущественно качественного описания многих показателей для определения текущего уровня ДИ необходима процедура экспертизы, адекватная поставленной задаче. Вследствие того, что оценку текущего уровня достоверности необходимо проводить регулярно, а количество параметров ИТКС исчисляется сотнями, то такой процедурой является однотуровая анонимная процедура на основе математических методов, дающих достаточно адекватное преобразование первичных результатов со сглаживанием рассогласований оценок экспертов [19].

В общий алгоритм проведения экспертизы [12, 16] необходимо внести ряд изменений и дополнений, ориентированных на комплексный учёт различных информационных воздействий, снижающих ДИ. Ниже представлен в виде разделения на этапы доработанный таким образом алгоритм.

1) Формулирование цели экспертизы и определение её объектов.

Целью экспертизы является оценка количественных и качественных параметров ИТКС согласно модели, представленной на рисунке 3.1. При определении объектов проведения экспертизы необходимо в полной мере учитывать организационный, физический и программно-технический уровни обеспечения достоверности, но это не означает, что все уровни становятся равнозначными: в зависимости от конкретной структуры информационных процессов (ИП) в ИТКС и качества СОДИ могут выходить на первый план и оказывать большее влияние на достоверность информации СФН одного из уровней.

2) Формирование аналитической группы.

Данный этап экспертизы ИТКС не имеет каких-либо особенностей применительно к оценке ДИ и проходит известным образом [4, 22].

3) Утверждение состава экспертной группы (ЭГ).

При формировании ЭГ необходимо оценить предполагаемую степень компетентности эксперта (коэффициент авторитета). Существует ряд способов определения коэффициентов авторитета на основе статистики предыдущих экспертиз [23]. При отсутствии статистики, а также в случае участия эксперта в первой своей экспертизе коэффициенты авторитета могут быть определены на основе формальных сведений об экспертах и нормированы по условию

$$\sum_{\varepsilon=1}^{m} v_{\varepsilon}^{0} = 1 . \qquad (3.5)$$

Могут быть использованы следующие сведения об экспертах:

A. Образование;

B. Научная подготовка;

C. Стаж работы по приоритетному направлению;

D. Количество проведенных экспертиз.

Оценка может быть проведена с использованием шкалы баллов (табл. 3.1).

Количество баллов по пунктам A, B, C, D суммируем и таким образом определяем первичный балл эксперта B_{ε}^{0}.

Коэффициент авторитета с учётом нормирования вычисляем по формуле

$$\nu_{\varepsilon}^{0} = \frac{B_{\varepsilon}^{0}}{\sum_{\varepsilon=1}^{m} B_{\varepsilon}^{0}}. \qquad (3.6)$$

Таблица 3.1

Шкала оценки компетентности экспертов

Направление	Описание внутри направления	Балл
A	по приоритетному направлению	5
	по смежной специальности	4
	по направлению (неоконченное)	3
	по смежной специальности (неоконченное)	2
	не совпадает с профилем экспертизы	0
B	академик	5
	доктор наук	4
	кандидат наук	3
	аспирант, с.н.с.	2
	без степени	0
C	не менее 10 лет	5
	не менее 5 лет	4
	не менее 1 года	3
	менее 1 года	1
	отсутствует	0
D	более 20	5
	10-20	4
	4-9	3
	1-3	1
	нет	0

При проведении экспертиз обращения к экспертам сопряжены с определенными финансовыми издержками. Учитывая это обстоятельство, при формировании экспертной группы можно использовать следующий метод [22].

Пусть C_k – условная стоимость обращения к k-му эксперту, а C – граничная суммарная стоимость обращения ко всем экспертам. Пусть $x_k = 1$, если эксперт включен в группу и $x_k = 0$, если нет. Тогда задачу формирования экспертной группы, обладающей максимальной компетентностью, можно записать как задачу линейного программирования следующим образом:

$$\begin{cases} \sum_k v_k x_k \to max, \\ \sum_k c_k x_k \le c. \end{cases} \qquad (3.7)$$

Коэффициент авторитета, определённый по (3.7) является первичным. Не всегда качество оценок эксперта соответствует формальным сведениям о нём. Первичный коэффициент должен быть скорректирован на основе согласованности суждений эксперта.

4) Подготовка необходимой информации об объектах экспертизы, её анализ и систематизация.

При проведении экспертизы ИТКС наибольшее количество времени будет затрачено на изучение её характеристик, т.к. необходимо рассмотреть два основных вопроса: назначение и принципы функционирования ИТКС и области СФН, ДФ, ИР, средств СОДИ. Оба эти вопроса могут быть разрешены в ходе опросов пользователей и разработчиков ИТКС, на что требуются большие временные затраты.

Другим источником информации об объекте является проектная, рабочая и эксплуатационная документация. Иногда качество документации бывает низким или она просто отсутствует. С другой стороны, там, где она существует, её объем может исчисляться сотнями и тысячами страниц печатного текста. Документация также может содержать устаревшие сведения.

Наиболее эффективным методом сбора информации об ИТКС является комплексный метод, при котором руководство организации, разрабатывающей либо эксплуатирующей ИТКС, ставит перед её разработчиками или другим персоналом задачу подготовить такую информацию и представить её в экспертную группу.

Проведение экспертизы ИТКС для оценки текущего уровня ДИ требует следующих документов:

- документы, содержащие требования безопасности,
- описания ИП,
- описание механизмов обеспечения достоверности.

5) Предварительное ознакомление экспертов с материалами об объектах экспертизы, получение дополнительной информации.

Часто после первого ознакомления экспертов с подготовленной документацией, описывающей объект экспертизы, у них возникают различ-

ные вопросы, которые по возможности должны быть устранены подготовкой дополнительной информации. В основном это касается более детального описания ИП и механизмов обеспечения достоверности.

6) Выбор процедуры проведения экспертизы.

Существует два принципа экспертного оценивания [21]. В соответствии с первым каждому объекту экспертизы должна быть дана оценка в целом, в соответствии со вторым – многокритериальная оценка по каждому из критериев оценочной системы с последующим автоматизированным расчётом результирующей оценки. Показатели ДИ не являются такими параметрами, которые можно оценить непосредственно, т.о. первый принцип не подходит.

7) Определение оценочной системы.

Оценка количественных параметров реализуема в шкалах в соответствие с физическим смыслом параметра [10]. В задаче оценки уровня достоверности качественные параметры ИТКС можно условно разделить на два типа:

– требующие получения абсолютного результата;

– требующие получения сравнительного результата (относительной значимости) в ряде альтернатив.

Оценочной величиной первых становится нечёткое множество, функция принадлежности которого показывает распределение возможности всех предполагаемых результатов. Нормированное распределение значимостей в полном множестве альтернатив, полученное на основе парных сравнений по лингвистическим таблицам, дает искомый результат для параметров второго типа.

8) Оценка объектов экспертизы в соответствии с принятой процедурой и выбранной оценочной системой.

Данный этап алгоритма предполагает получение массива первичных оценок экспертов по всем оцениваемым параметрам. Такие оценки представлены в виде числовых величин или таблиц парных сравнений.

9) Обработка первичных результатов экспертизы.

Полученные первичные оценки могут быть качественными, тогда для получения количественного результата необходимо их преобразование в количественные значения по таблицам соответствия или с использованием баз знаний. Кроме того, оценки параметров, имеющих разный физический смысл, даны в разных шкалах, следовательно, требуется сведение

оценок в рамках единой модели и получение нескольких общих показателей.

В связи с невозможностью абсолютной формализации получения первичных оценок суждения экспертов будут расходиться. Степень такого расхождения (или обратный ей показатель – степень согласованности) показывает качество получения единой оценки и служит для коррекции коэффициентов авторитета экспертов с увеличением коэффициентов экспертов, давших более согласованные оценки, что позволяет повысить согласованность все экспертизы в целом.

10) Получение результатов – показателей качества ИТКС.

Расчёт общего коэффициента достоверности информации $D_{ИР}$ и экономической эффективности Ef проходит на основе процедур, рассмотренных далее в этой главе.

11) Принятие решения по результатам экспертизы.

Результаты экспертизы могут быть признаны адекватными, тогда решением будет выработка рекомендаций по совершенствованию СОДИ с целью повышения ДИ в ИТКС, при условии, что СОДИ экономически эффективна, либо предложения по переработке СОДИ, являющейся неэффективной, при условии, что уровень ДИ недостаточен.

3.3. Процедуры получения числовых оценок количественных параметров ИТКС

Оценка чётких количественных параметров

Подмножество параметров ИТКС определены чётким количественным значением. В качестве примеров можно привести: «в локальную сеть отдела входит 4 рабочие станции, один сервер, один принтер и один маршрутизатор»; «количество известных вирусов в БД установленного антивируса – 83468»; «период обновления антивирусных баз – 1 месяц» и др.

Простановка первичных оценок таких параметров экспертами не вызывает сложности, более того, большинство из таких параметров могут быть определены без собственно экспертизы на основе имеющейся документации. Такие оценки представляют собой точечные значения на определенной шкале. Но возникает некоторая проблема. Во-первых, в приведенных выше примерах можно отметить три вида единиц измерения: без-

размерные, пространственные и временные. Т.е. получается, что параметры нужно оценить по разным шкалам, при этом все они оказывают влияние на ДИ и должны быть сведены в единую систему. Во-вторых, для всех параметров можно указать требуемое или оптимальное значения; и если первое можно каким-то образом выделить из стандартов, то для второго без экспертизы не обойтись.

Рассмотрим подробнее решение первой стороны проблемы. В процессе проведения оценки в зависимости от физической природы оцениваемого параметра могут быть использованы различные шкалы: шкала интервалов, шкала отношений, шкала разностей, абсолютная шкала [22]. Оцениваемые параметры разнородны и для их совместного использования в одной модели необходимо нормировать количественные данные с учётом их значимости. Нормирование в частности можно выполнить сведением к единичному интервалу вещественных чисел [0, 1] с учётом весовых коэффициентов при последующих преобразованиях. Адекватность выбора именно единичного интервала обусловлена тем, что в том же интервале изменяются вероятности и функция принадлежности в теории нечётких множеств, и как конечный результат исследования – достоверность информации.

Для нормирования необходимо ввести понятия «минимальное», «максимальное», «наилучшее», «наихудшее» и «оптимальное» значения q-го параметра. Обозначим их $x_q^{MIN}, x_q^{MAX}, x_q^{НЛ}, x_q^{НХ}, x_q^{ОПТ}$. Минимальным назовем значение, наименьшее по величине и достижимое в системе. Чаще всего это 0 (кроме временных параметров). Максимальное значение определяется техническими характеристиками системы как «максимально достижимое» значение.

Очевидно, что не всегда максимальное значение является наилучшим, а минимальное – наихудшим. Обычно для всех «положительных» параметров (как то количество средств обеспечения достоверности) максимальное – есть наилучшее, а для всех «отрицательных» характеристик (время выполнения операции, количество единиц доступа и т.д.) максимальное – есть наихудшее. Кроме того, можно отметить, что не всегда экстремальное значение параметра оптимально с точки зрения организации СОДИ.

Для нормирования могут быть использованы преобразования вида:

– если максимальное значение параметра является оптимальным, то

$$x_q^{MAX} \to x_q^{НЛ} = x_q^{ОПТ}, \quad x_q^{MIN} \to x_q^{НХ};\ \bar{x}_q = \frac{x_q - x_q^{MIN}}{x_q^{MAX} - x_q^{MIN}}; \qquad (3.8)$$

– если минимальное значение параметра является оптимальным, т.е.

$$x_q^{MIN} \to x_q^{НЛ} = x_q^{ОПТ}, \quad x_q^{MAX} \to x_q^{НХ};\ \bar{x}_q = \frac{x_q^{MAX} - x_q}{x_q^{MAX} - x_q^{MIN}}; \qquad (3.9)$$

– если оптимальное значение параметра находится между минимальным и максимальным, которые оба являются наихудшими, то

$$x_q^{НЛ} = x_q^{ОПТ}, \quad x_q^{MIN} \to x_q^{НХ}, \quad x_q^{MAX} \to x_q^{НХ}, \quad x_q^{MIN} \le x_q^{ОПТ} \le x_q^{MAX},$$

$$\bar{x}_q = \begin{cases} 0, \ x_q = x_q^{МИН} \quad или \quad x_q = x_q^{MAX} \\ 1, \qquad\qquad\qquad x_q = x_q^{ОПТ} \\ \dfrac{x_q - x_q^{МИН}}{x_q^{ОПТ} - x_q^{MIN}}, \quad x_q^{MIN} < x_q < x_q^{ОПТ} \\ \dfrac{x_q^{MAX} - x_q}{x_q^{MAX} - x_q^{ОПТ}}, \quad x_q^{ОПТ} < x_q < x_q^{MAX} \end{cases}. \qquad (3.10)$$

Оценка нечётких количественных параметров

Если в качестве оптимального значения невозможно использовать статистически определенную величину, либо взятую из стандарта [11, 16], то такое значение является субъективной величиной, которая находится экспертным путем. Следовательно, она может быть представлена нечётким числом с треугольной (возможно, несимметричной) функцией принадлежности.

При определении оптимального значения количественного требования каждый эксперт представляет свои оценки в виде тройки чисел

$$\left\langle a_\varepsilon, x_\varepsilon^{ОПТ}, b_\varepsilon \right\rangle, \qquad (3.11)$$

где $x_{\varepsilon}^{ОПТ}$ – мода (чёткое значение нечёткого числа) – предполагаемое экспертом оптимальное значение параметра, a_{ε} и b_{ε} задают левую и правую границу нечёткости, т.е. величины, за границами которых значение параметра эксперт с абсолютной уверенностью считает не оптимальным.

Оценки всех экспертов можно представить графически на рис. 3.3.

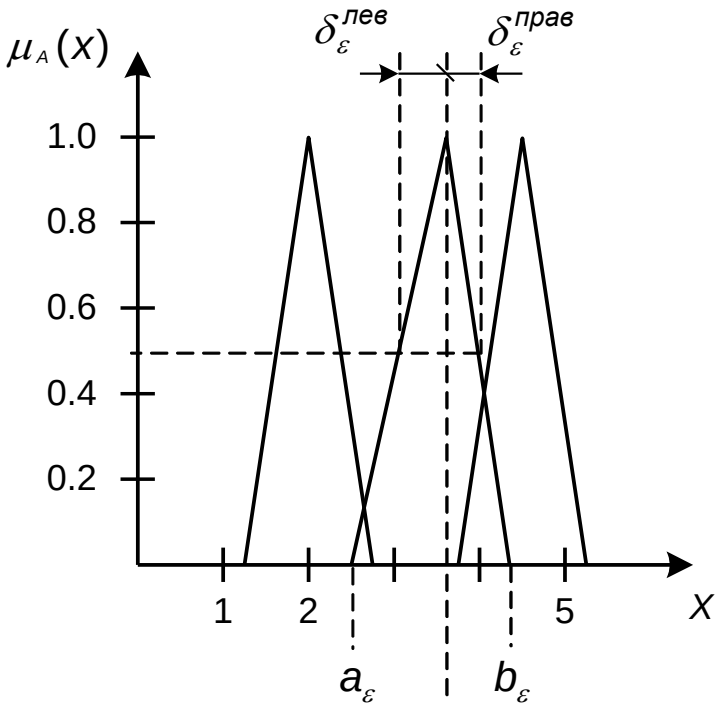


Рис. 3.3. Оценки ряда экспертов оптимального значения параметра

Величины $\delta_{\varepsilon}^{лев}$ и $\delta_{\varepsilon}^{прав}$ определяют левое и правое отклонение для α-среза 0,5, т.е. значения при которых эксперт называет значение оптимальным с уверенностью 0,5.

Очевидно, что

$$\delta_{\varepsilon}^{лев} = a_{\varepsilon} + \frac{x_{\varepsilon}^{ОПТ} - a_{\varepsilon}}{2}, \quad \delta_{\varepsilon}^{прав} = b_{\varepsilon} + \frac{x_{\varepsilon}^{ОПТ} - b_{\varepsilon}}{2}. \qquad (3.12)$$

Расчёт значений параметра $\delta^{лев}, \delta^{прав}, x^{ОПТ}$ возможен с учётом уточнённого коэффициента авторитета.

$$\delta_{лев} = \sum_{\varepsilon=1}^{m} \delta_{\varepsilon}^{лев} \cdot \nu_{\varepsilon},$$

$$\delta_{прав} = \sum_{\varepsilon=1}^{m} \delta_{\varepsilon}^{прав} \cdot \nu_{\varepsilon}, \qquad (3.13)$$

$$x^{ОПТ} = \sum_{\varepsilon=1}^{m} x_{\varepsilon}^{ОПТ} \cdot \nu_{\varepsilon},$$

где ν_{ε} – коэффициент авторитета ε-го эксперта.

Описание нечёткого количественного параметра соответствует выражению «примерно <значение z>», т.е. имеет нормальный вид функции распределения. Оценка степени соответствия значения z_q описанию может быть получена следующим образом:

$$\bar{x}_q = \begin{cases} e^{-в_{лев}\left(z_q^{ОПТ} - z_q\right)^2}, & z_q \le z_q^{ОПТ}, \\ e^{-в_{прав}\left(z_q^{ОПТ} - z_q\right)^2}, & z_q > z_q^{ОПТ}, \end{cases} \qquad (3.14)$$

где $в$ зависит от требуемой степени нечёткости и определено из

$$в_{лев} = \frac{\ln \alpha}{4\delta_{лев}^2}, \qquad в_{прав} = \frac{\ln \alpha}{4\delta_{прав}^2}, \qquad (3.15)$$

где $\delta_{лев} + \delta_{прав}$ определяет расстояние между точками перехода для функции принадлежности (3.13), т.е. точками, в которых функция принимает значение α со степенью уверенности 0.5.

Уровень α определяет степень допущения принадлежности реального значения параметра z_q требуемому $z_q^{треб}$, либо оптимальному $z_q^{ОПТ}$. Обычно $\alpha \in [0{,}5,\ 1{,}0]$.

Рис. 3.4 иллюстрирует соотношение между данными величинами.

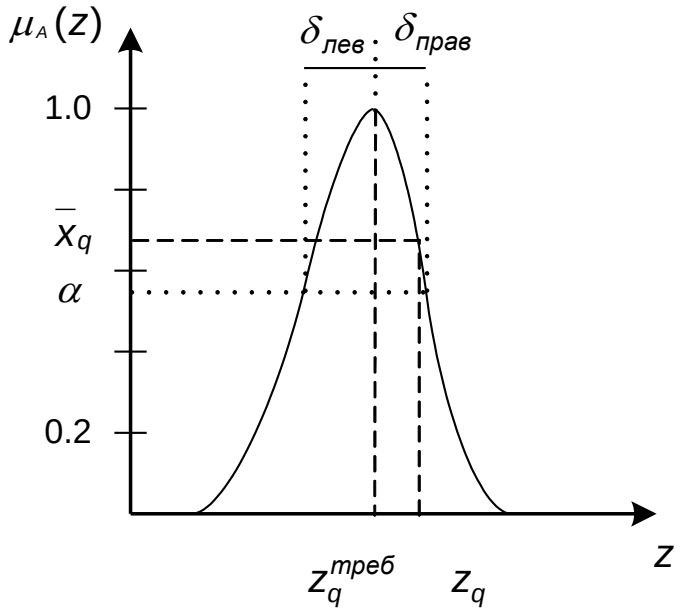


Рис. 3.4. Графическое представление функции
принадлежности качественного описания
количественного параметра

3.4. Алгоритмы получения числовых оценок качественных параметров ИТКС

В процессе проведения экспертизы эксперту приходится сталкиваться с качественными параметрами или их описаниями. В задаче оценки ДИ необходимо исследовать вероятности возникновения, обнаружения и устранения ошибок, которые определяются некоторыми числовыми значениями. В то же время ряд СФН различных элементов ИТКС характеризуется изначально качественными описаниями. Например, ошибки персонала во время эксплуатации ИТКС зависят от качества должностных инструкций, контроля со стороны руководителей, распределения обязанностей и организации труда. Это приводит к возникновению проблемы перехода от качественных описаний к количественным значениям.

Все качественные параметры можно разделить на те, которые имеют чёткое описание, т.е. существует или может быть задана последовательность описаний, упорядоченных по ранговой шкале [22], и те, которые не имеют такого ряда описаний (в частности, качественные требования различных стандартов).

Примеры параметров, которые можно описать по ранговой шкале:

– «значимость некоторого, объекта, события, условия»;

– «доступность, простота использования какого-либо метода или средства»;

– «ценность ресурса (относительная)»;

– «степень влияния некоторого условия на проявление события».

Перечисленные параметры могут быть оценены путём использования прямого преобразования по лингвистическим таблицам [34] или посредством парных сравнений альтернатив. В отличие от количественных оценка качественных параметров одним человеком обычно дает неадекватный результат. Работа экспертной группы с достаточным количеством экспертов позволяет повысить качество оценки. Алгоритм получения экспертных оценок качественных параметров с реализацией многотуровых анонимных процедур представлена на рис. 3.5.

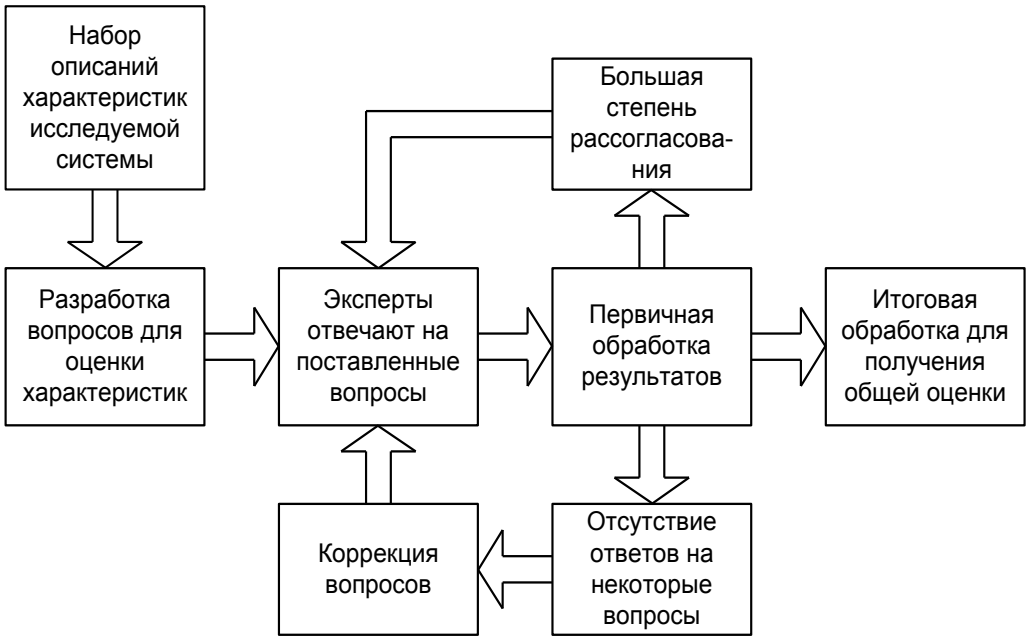


Рис. 3.5. Алгоритм получения экспертных оценок качественных параметров

В любых экспертизах качество результатов напрямую зависит от качества вопросов, поставленных перед экспертом. Чёткие вопросы усложняют задачу аналитической группы, которая их готовит и обрабатывает результаты. Более абстрактные вопросы усложняют задачу эксперта и могут привести к получению неадекватного результата или результатов с боль-

шой степенью рассогласования. Таким образом, наиболее значимой проблемой является формализация процесса подготовки вопросов и проведения экспертизы [35].

Большинство нечисловых характеристик можно описать по уровневой шкале, используя понятия «высокий», «средний» и «низкий» уровни. В данном случае вопрос перед экспертом поставлен чёткий. Но эксперт по такой шкале присвоит один и тот же, например, «средний уровень» нескольким характеристикам, имеющим на самом деле отличительные особенности, которые нельзя просто так усреднить.

Если существенно увеличить количество уровней шкалы, то тогда может возникнуть ситуация, при которой эксперту будет затруднительно выбрать наиболее подходящий. И не для всех характеристик можно придумать более трёх уровней. Человеку всегда проще выполнить сравнение двух объектов, чем дать исчерпывающую характеристику каждому в отдельности. Следовательно, адекватная оценка может быть дана путём парных сравнений альтернатив.

В задаче оценки текущего уровня ДИ в ИТКС одним из качественных параметров является значимость ряда элементов или условий для возникновения некоторого события [2].

Алгоритм оценки характеристики «значимость условия»

Одним из примеров такой характеристики является «значимости ряда СФН для возникновения ошибки, приводящей к снижению достоверности обрабатываемой информации» [32].

Алгоритм представлен на рис. 3.6.

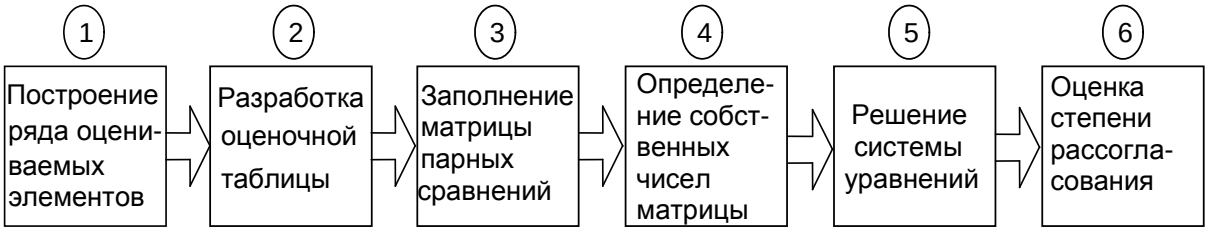


Рис. 3.6. Алгоритм определения значимости условий для возникновения события

На первом этапе формируем множество оцениваемых элементов $E = \{e_1, e_2, \ldots, e_n\}$.

Второй этап заключается в разработке лингвистических описаний, т.е. вариантов оценки элемента и присвоении им числовых значений, которые могут быть взяты, например, по методу Саати как натуральные числа $\{1,2,...,9\}$ [17, 18]. Но типовая таблица лингвистических описаний как правило, мало информативна. Поэтому для различных параметров на её базе должны быть разработаны более конкретные описания. Они даны в табл. 3.2-3.6.

Для тех параметров, определение значимости которых возможно непосредственно (без парных сравнений), используем тот же диапазон значений, но уже не в шкале относительной значимости, а в ранговой шкале.

Для программно-аппаратного и организационного типа СФН можно уточнить формулировки табл. 3.3. Один из вариантов представлен в табл. 3.4 и 3.5.

Таблица 3.2

Значимость СФН (по шкале относительной значимости)

Лингвистическая оценка сравнения 1-го и 2-го СФН	Значение
При наличии 1-го СФН наличие 2-го можно не учитывать	9
Существенное превосходство значимости 1-го СФН над 2-м	7
Использование 1-го СФН предпочтительнее, чем 2-го	5
Чуть более высокая значимость 1-го СФН против 2-го	3
Одинаковая значимость сравниваемых СФН	1

Таблица 3.3

Доступность СФН (по ранговой шкале)

Лингвистическая оценка	Значение
СФН является общеизвестным, для его использования не требуется спецсредств и особых способностей	9
СФН является общеизвестным, но для его использования требуются относительно доступные технические средства	7
СФН распространенный, для его использования требуются дорогостоящие или широко недоступные спецсредства	5
СФН является малоизвестным и/или для его использования требуются дорогостоящие или широко недоступные спецсредства, ресурсы, мероприятия и т.д.	3
Использование СФН требует таких ресурсов и средств, применить которые скрытно невозможно или их применение требует огромных затрат времени	1

Таблица 3.4

Доступность СФН программно-аппаратного уровня

Лингвистическая оценка	Значение
СФН известен, для его использования не требуется специальных вычислительных мощностей, программных средств, большого ресурса времени и денежных затрат	9
Требуются типовые программно-аппаратные средства, возможно с небольшими затратами времени	7
СФН малоизвестен (для его использования требуется специальные знания и навыки), существенные денежные затраты, специальное оборудование и/или ресурсы времени,	5
СФН является малоизвестным, для его использования требуются специальные знания и оборудование, большие вычислительные мощности, существенные денежные затраты	3
Для использования СФН требуется оборудование спецслужб и/или вычислительные мощности супер-компьютеров, а также большие затраты и/или ресурсы времени	1

Таблица 3.5

Доступность СФН организационного плана

Лингвистическая оценка	Значение
СФН известен широкому кругу лиц, не требует спецсредств, подготовки и проникновения в АСУП.	9
СФН известен широкому кругу лиц, не требует проникновения в АСУП, но требует относительно доступных специальных средств	8
СФН известен широкому кругу лиц, для его использования требуются относительно доступные специальные средства, проникновение в АСУП или дополнительная подготовка.	7
СФН известен широкому кругу лиц, для его использования требуются относительно доступные специальные средства, проникновение в АСУП и дополнительная подготовка в течение длительного времени.	5
СФН малоизвестен, для его использования требуются дорогие или малодоступные спецсредства, проникновение в АСУП и дополнительная подготовка в течение длительного времени.	3
СФН малоизвестен, для его использования требуются значительные затраты ресурсов и длительная подготовка, проникновение в АСУП сопряжено со значительными сложностями.	1

Таблица 3.6

Ценность ИР (по ранговой шкале)

Лингвистическая оценка ценности ИР на основе расчёта затрат на восстановление	Значение
Данный ИР является важнейшим для организации. Его потеря нанесет непоправимые последствия для организации.	9
Затраты на ликвидацию последствий из-за потери ресурса сопоставимы с годовыми экономическими показателями	8
Затраты на восстановление из-за потери ресурса существенны для организации	6
Затраты на восстановление несущественны, но требуется дополнительное время	2
Восстановление из-за потери ресурса будет проведено в штатном режиме	1

На третьем этапе эксперт заполняет матрицу парных сравнений для каждой пары элементов

$$M_{\varepsilon}^{A}=\begin{vmatrix} a_{11}^{\varepsilon} & a_{12}^{\varepsilon} & \dots & a_{1n}^{\varepsilon} \\ a_{21}^{\varepsilon} & a_{22}^{\varepsilon} & \dots & a_{2n}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ a_{n1}^{\varepsilon} & a_{n2}^{\varepsilon} & \dots & a_{nn}^{\varepsilon} \end{vmatrix} \tag{3.16}$$

где $A=\{a_1, a_2, \dots, a_n\}$ – множество параметров, $a_{\alpha\beta}$ – описание отношения, взятое из таблицы 3.2, $\varepsilon \in \{1, 2, \dots, m\}$ – номер эксперта.

Матрица (3.16) является обратно симметричной, т.е. для каждой пары элементов выполняется условие:

$$a_{\alpha\beta}^{\varepsilon}=\frac{1}{a_{\beta\alpha}^{\varepsilon}}, \quad \forall \alpha,\beta=\overline{1,n}, \ \forall \varepsilon=\overline{1,m}. \tag{3.17}$$

Таким образом, эксперт определяет только значения элементов матрицы выше (или наоборот ниже) главной диагонали. При этом количество сравниваемых элементов определяем по формуле

$$KC=\frac{n(n-1)}{2}. \tag{3.18}$$

Элементарный анализ данной зависимости количества сравнений от размерности матрицы показывает, что при n от 20-30 и выше количество сравнений становится слишком большим (от 200 и более – и это только по одному параметру), что ставит перед экспертом практически не выполнимую задачу.

В процессе сравнения возникает ещё одна парадоксальная ситуация: те же СФН (и другие характеристики) разделены на категории по типу (например, технические, программные, аппаратные и организационные СФН). Адекватно сравнить разнородные характеристики просто невозможно.

Предложен следующий метод оценки в условиях рассмотренной ситуации. Предположим, что выделены две группы сравниваемых альтернатив (данный метод может быть легко распространён на любое количество групп). Для каждой группы необходимо построить частные матрицы типа (3.16), но с изменённым условием (3.17)

$$M_{\varepsilon}^{Br} = \begin{vmatrix} br_{11}^{\varepsilon} & br_{12}^{\varepsilon} & \dots & br_{1k_r}^{\varepsilon} \\ br_{21}^{\varepsilon} & br_{22}^{\varepsilon} & \dots & br_{2k_r}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ br_{k_r 1}^{\varepsilon} & br_{k_r 2}^{\varepsilon} & \dots & br_{k_r k_r}^{\varepsilon} \end{vmatrix} \tag{3.19}$$

где $Br = \{br_1, br_r, \dots, br_{k_l}\}$ – множество параметров r-группы из k_r-параметров;

всего l групп, $br_{\alpha\beta} = a_{\alpha\beta} - 1, \ \forall r$ – ($a_{\alpha\beta} \geq 1$ взято из таблицы 3.2), $\varepsilon \in \{1, 2, \dots, m\}$ – номер эксперта;

новое условие следующее:

$$br_{\alpha\beta}^{\varepsilon} = -br_{\beta\alpha}^{\varepsilon}, \quad \forall \alpha, \beta = \overline{1, k_r}, \ \forall \varepsilon = \overline{1, m}, \quad \forall r == \overline{1, l}. \tag{3.20}$$

Далее эксперт дает парное сравнение значимости каждой группы по условию аналогично (3.20)

$$M_{\varepsilon}^{\Gamma}=\begin{vmatrix} \Gamma_{11}^{\varepsilon} & \Gamma_{12}^{\varepsilon} & \dots & \Gamma_{1l}^{\varepsilon} \\ \Gamma_{21}^{j} & \Gamma_{22}^{\varepsilon} & \dots & \Gamma_{2l}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ \Gamma_{l1}^{\varepsilon} & \Gamma_{l2}^{\varepsilon} & \dots & \Gamma_{ll}^{\varepsilon} \end{vmatrix} \tag{3.21}$$

Затем в общую матрицу вносим частные оценки каждой группы. Для двух групп она будет выглядеть следующим образом:

$$M_{\varepsilon}^{B}=\left|\begin{array}{ccc|ccc} b1_{11}^{\varepsilon} & \dots & b1_{1k_1}^{\varepsilon} & & & \\ \dots & \dots & \dots & & & \\ b1_{k_1 1}^{\varepsilon} & \dots & b1_{k_1 k_1}^{\varepsilon} & & & \\ \hline & & & b2_{11}^{\varepsilon} & \dots & b2_{1k_2}^{\varepsilon} \\ & & & \dots & \dots & \dots \\ & & & b2_{k_2 1}^{\varepsilon} & \dots & b2_{k_2 k_2}^{\varepsilon} \end{array}\right|. \tag{3.22}$$

Остаётся заполнить позиции сравнения элементов из разных групп на основе (3.21), что может быть сделано в автоматизированном режиме без участия эксперта.

Общая картина сравнения всех элементов, скажем, 1-й и 2-й группы определяется величиной $\Gamma_{12}^{\varepsilon}$, которая не отражает отношения каждой отдельной пары элементов.

Рассмотрим любую строку матрицы (3.19). Величина $br_{\alpha\beta}^{\varepsilon}$ отражает сравнительную характеристику α- и β-элементов, которое при условии (3.20) можно считать смещением по шкале рангов.

Сумма $Sbr_{\alpha}^{\varepsilon}=\sum_{\beta=1}^{k_r} br_{\alpha\beta}^{\varepsilon}, \quad \left(\forall\alpha=\overline{1,k_r},\ \forall r=\overline{1,l},\ \forall\varepsilon=\overline{1,m}\right)$ показывает сумму смещений α-элемента относительно всех $\beta=\overline{1,k_r}$, т.е. по существу его превосходство, если сумма положительна, подавление, если она отрицательна и равную значимость, если равна нулю.

Очевидно, что из условия (3.20)

$$\sum_{\alpha=1}^{k_r} Sbr_\alpha^\varepsilon = 0 \quad \left(\forall r = \overline{1,l},\ \forall \varepsilon = \overline{1,m}\right). \tag{3.23}$$

Таким образом, в пока незаполненные места матрицы (3.22) необходимо внести величины, равные: сумма рангов из строки 1-й сравниваемой группы + сумма рангов из столбца 2-й группы + ранг сравнения групп в целом.

Для двух групп параметров заполненная матрица будет иметь вид

$$M_\varepsilon^B = \left|\begin{array}{ccc|ccc} b1_{11}^\varepsilon & \dots & b1_{1k_1}^\varepsilon & Sb1_{1\_}^\varepsilon + Sb2^\varepsilon_{\_1} + \Gamma_{12}^\varepsilon & \dots & Sb1_{1\_}^\varepsilon + Sb2^\varepsilon_{\_k_2} + \Gamma_{12}^\varepsilon \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b1_{k_1 1}^\varepsilon & \dots & b1_{k_1 k_1}^\varepsilon & Sb1_{k_1\_}^\varepsilon + Sb2^\varepsilon_{\_1} + \Gamma_{12}^\varepsilon & \dots & Sb1_{k_1\_}^\varepsilon + Sb2^\varepsilon_{\_k_2} + \Gamma_{12}^\varepsilon \\ \hline Sb1^\varepsilon_{\_1} + Sb2_{1\_}^\varepsilon + \Gamma_{21}^j & \dots & Sb1^\varepsilon_{\_k_1} + Sb2_{1\_}^\varepsilon + \Gamma_{21}^j & b2_{11}^\varepsilon & \dots & b2_{1k_2}^\varepsilon \\ \dots & \dots & \dots & \dots & \dots & \dots \\ Sb1^\varepsilon_{\_1} + Sb2_{k_2\_}^\varepsilon + \Gamma_{21}^\varepsilon & \dots & Sb1^\varepsilon_{\_k_1} + Sb2_{k_2\_}^\varepsilon + \Gamma_{21}^\varepsilon & b2_{k_2 1}^\varepsilon & \dots & b2_{k_2 k_2}^\varepsilon \end{array}\right| \tag{3.24}$$

Для дополнительного пояснения на рис. 3.7 показано формирование отдельных слагаемых в матрице (3.24).

Матрица (3.24) может содержать не нормализованные значения, а именно возможно такое, что $Sbr_\alpha^\varepsilon + Sbr_\beta^\varepsilon + \Gamma_{ab}^\varepsilon > 8$ или $Sbr_\alpha^\varepsilon + Sbr_\beta^\varepsilon + \Gamma_{ab}^\varepsilon < -8$, т.е. значения после обратного преобразования выйдут за пределы шкалы Саати.

Нормализация задаётся формулой

$$\bar{b}_{\alpha\beta} = \begin{cases} b_{\alpha\beta}, & -8 \le b_{\alpha\beta} \le 8 \\ 8, & b_{\alpha\beta} > 8 \\ -8, & b_{\alpha\beta} < -8 \end{cases}, \quad \left(\forall \alpha, \beta = \overline{1,n}\right). \tag{3.25}$$

Нормализованная по (3.25) матрица (3.24) примет вид

$$\overline{M}_{\varepsilon}^{B} = \begin{vmatrix} \bar{b}_{11}^{\varepsilon} & \bar{b}_{12}^{\varepsilon} & \dots & \bar{b}_{1n}^{\varepsilon} \\ \bar{b}_{21}^{\varepsilon} & \bar{b}_{22}^{\varepsilon} & \dots & \bar{b}_{2n}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ \bar{b}_{n1}^{\varepsilon} & \bar{b}_{n2}^{\varepsilon} & \dots & \bar{b}_{nn}^{\varepsilon} \end{vmatrix}. \qquad (3.26)$$

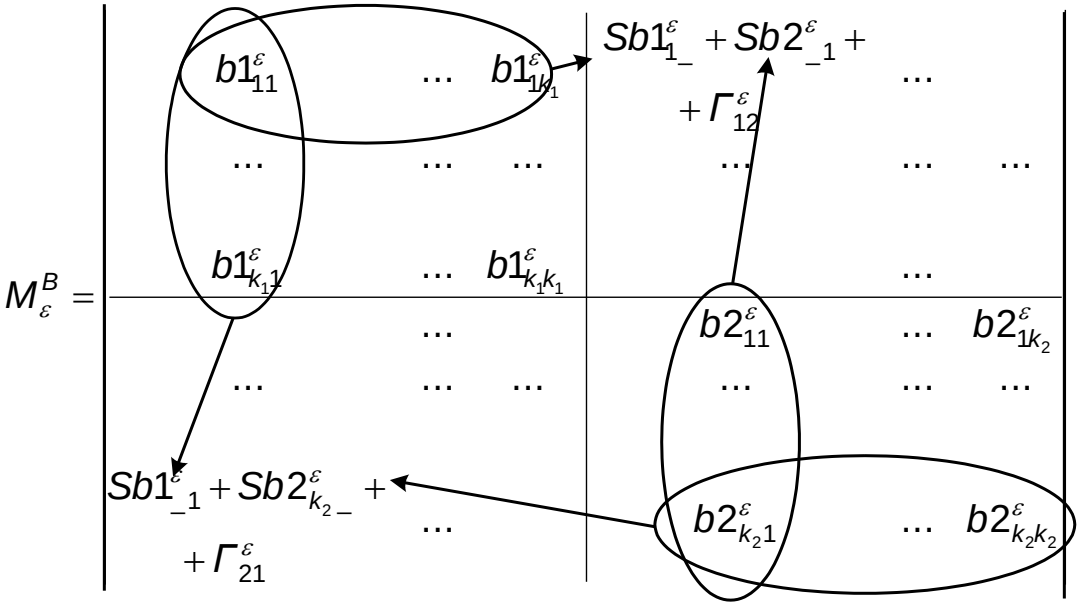


Рис. 3.7. Пояснение формирования матрицы (3.24)

Затем из матрицы (3.26) необходимо получить матрицу типа 3.16, используя преобразование вида

$$a_{\alpha\beta} = \begin{cases} \bar{b}_{\alpha\beta} + 1, & \bar{b}_{\alpha\beta} \geq 0 \\ \dfrac{1}{1 - \bar{b}_{\alpha\beta}}, & \bar{b}_{\alpha\beta} < 0 \end{cases}. \qquad (3.27)$$

Все преобразования (3.22) - (3.27) могут быть выполнены программно. Несомненным плюсом предложенного подхода является значительное снижение количества сравнений, которое максимально при равном числе параметров в каждой группе и равно

$$KC = l\frac{\frac{n}{l}\cdot\frac{n-1}{l}}{2} + \frac{l(l-1)}{2} \approx \frac{n(n-1)}{2l}, \qquad (3.28)$$

т.е. может быть уменьшено почти в такое число раз, сколько выделено групп.

Несущественным недостатком подхода является необходимость строить матрицы сравнительной значимости отдельных групп (3.21). Если это делать один раз, и применять для оценки ко всем ДФ, то такой подход немного не достоверный (например, для проявления одних ДФ могут быть более значимы программные СФН, а для других – физические).

Для улучшения качества оценки возможно строить матрицы (3.21) (а их размерность равна числу групп, т.е. относительно невелика) отдельно для каждого типа ДФ.

На <u>четвёртом этапе</u> необходимо определить собственные значения матрицы (3.16) и собственный вектор

$$A_\varepsilon = \left\{x_i^\varepsilon\right\}, \quad i = \overline{1,n}, \quad \varepsilon = \overline{1,m}, \qquad (3.29)$$

соответствующий максимальному собственному значению $\lambda_{max}^{\varepsilon}$ оценок ε-го эксперта.

Данная матрица является положительно определённой, обратно симметричной и определение собственного вектора на <u>пятом этапе</u> возможно методом Гаусса.

На <u>шестом этапе</u> проводится исследование качества оценки. При проведении сравнений в реальной ситуации вычисленное максимальное собственное число $\lambda_{max}^{\varepsilon}$ будет отличаться от соответствующего собственного числа λ_{max} для идеальной матрицы, вследствие нарушения её транзитивности (эксперту чрезвычайно сложно проводить все новые парные сравнения с учетом предыдущих). Найденные значения тем точнее, чем ближе $\lambda_{max}^{\varepsilon}$ к n. Причем всегда $\lambda_{max}^{\varepsilon} \geq n$. Разница $\lambda_{max}^{\varepsilon} - n$ даёт абсолютную меру несогласованности оценок. Относительная мера (коэффициент рассогласования)

$$K_P^{\varepsilon} = \frac{\lambda_{max}^{\varepsilon} - n}{n-1} \tag{3.30}$$

может быть использована для коррекции коэффициента авторитета ε-го эксперта.

Для такой коррекции вычисляем средний коэффициент рассогласования

$$K_{CP}^{\varepsilon} = \frac{1}{k}\sum_{i=1}^{k} K_{Pi}^{\varepsilon}, \tag{3.31}$$

где k – количество матриц парных сравнений, построенных экспертом;

K_{Pi}^{ε} – коэффициент рассогласования i-й матрицы, вычисленный по (3.30).

Вычисляем уточнение коэффициента авторитета:

$$\Delta\nu_{\varepsilon} = \frac{1}{2}\left(\nu_{\varepsilon}^{0} + \frac{\nu_{\varepsilon}^{0}}{\sqrt{K_{CP}^{\varepsilon}}}\right). \tag{3.32}$$

Уточнение предполагает, что исправленный коэффициент будет средним арифметическим первичного коэффициента и его снижения из-за рассогласования оценок. Такое снижение будет различным у всех экспертов.

Скорректированный коэффициент, удовлетворяющий правилу нормирования, определяем по формуле

$$\nu_{\varepsilon} = \frac{\Delta\nu_{\varepsilon}}{\sum_{\varepsilon=1}^{m}\Delta\nu_{\varepsilon}}. \tag{3.33}$$

3.5. Общая модель оценки рисков

Пусть существует конечное множество ДФ, которые характеризуются относительными частотами возникновения $p_i^{ДФ}$ и ущербом, наносимым

предприятию u_l, где $l=\overline{1,n}$. СОДИ выполняет функцию противодействия нарушению достоверности ИР. Основной характеристикой средства обеспечения достоверности является возможность сохранения достоверности ИР каждого *i*-го ИР при воздействии на него *l*-го ДФ p_{il}^{C}, которая связана с возможностью нарушения достоверности $p_{il}^{НД}$ через соотношение

$$p_{il}^{C}=1-p_{il}^{НД}. \tag{3.34}$$

Ущерб ИТКС при отсутствии СОДИ может быть представлен как суммарный по каждому ДФ [7, 25-33]

$$U=\sum_{l=1}^{n}u_l. \tag{3.35}$$

В свою очередь остаточный ущерб (т.е. ущерб, который все равно будет нанесен ИТКС даже при наличии СОДИ – злоумышленником могут быть использованы новые способы и средства нарушения достоверности) также представляет собой сумму потерь от всех ДФ

$$W=\sum_{l=1}^{n}w_l. \tag{3.36}$$

Риск для ИТКС при отсутствии СОДИ представляет собой функцию относительных частот возникновения ДФ и ущерба в случае нарушения достоверности

$$R^{НЗ}=f\left(p_l^{ДФ},u_i\right)=\sum_{l=1}^{n}p_l^{ДФ}\cdot u_i, \tag{3.37}$$

а риск для ИТКС при наличии СОДИ зависит также и от возможностей сохранения достоверности ИР

$$R^{ЗАЩ}=f\left(p_l^{ДФ},u_l,p_l^{C}\right)=\sum_{l=1}^{n}p_l^{ДФ}\cdot u_l\cdot\left(1-p_l^{C}\right). \tag{3.38}$$

Учитывая вероятностный характер ДФ, можно заменить предотвращённый ущерб $\Delta W = U - W$ на устранённый риск $\Delta R = R^{НЗ} - R^{ЗАЩ}$ [31]. Тогда экономическая эффективность функционирования СОДИ может быть определена как

$$Ef = \frac{R^{НЗ} - R^{ЗАЩ}}{S_{СОД}} = \frac{\sum_{l=1}^{n} p_l^{ДФ} \cdot u_l - \sum_{l=1}^{n} p_l^{ДФ} \cdot u_l \cdot \left(1 - p_l^{С}\right)}{S_{СОД}}. \qquad (3.39)$$

Для расчёта экономической эффективности необходимо определить перечень ИР и их стоимость, а также провести экспертизу таких параметров ИТКС, как значимости и доступности СФН и степень воздействия каждого ДФ на все ИР ИТКС. Достоверная экспертиза параметров ИТКС возможна только на основе определения полного списка актуальных ДФ и СФН, при условии адекватной оценки степени выполнения количественных и качественных требований к СОДИ.

Ущерб от нарушения достоверности ИР

Рассмотрим методы, применяемые для оценки ущерба от нарушения достоверности ИР.

Ущерб u_i, наносимый ИТКС l-м ДФ при отсутствии СОДИ, может быть определён в абсолютных единицах: экономических потерях, временных затратах, объёме уничтоженной или “испорченной” информации и т.д. [35].

Ущерб от нарушения достоверности ИР – это функция двух параметров: степень воздействия l-го ДФ на i-й ИР и стоимость ресурса $S_{ИР\,i}$.

Степень воздействия должна быть определена экспертным путем по лингвистической таблице и внесена в матрицу

$$M_{\varepsilon}^{H} = \begin{vmatrix} h_{11}^{\varepsilon} & h_{12}^{\varepsilon} & \dots & h_{1z}^{\varepsilon} \\ h_{21}^{\varepsilon} & h_{22}^{\varepsilon} & \dots & h_{2z}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ h_{n1}^{\varepsilon} & h_{n2}^{\varepsilon} & \dots & h_{nz}^{\varepsilon} \end{vmatrix}, \qquad (3.40)$$

где $0 \le h_{ik}^{\varepsilon} \le 1$ показывает степень воздействия l-го ДФ на i-й ИР.

$Sh_i^{\varepsilon} = \sum_{l=1}^{n} h_{li}^{\varepsilon}, \ \forall i = \overline{1,z}, \ \forall \varepsilon = \overline{1,m}$ показывает суммарный ущерб для i-го ИР и не должна быть больше $S_{ИР.i}$. Но при достаточно большом количестве ДФ и их влиянии на множество ресурсов такое условие может быть нарушено. Тогда столбцы матрицы (3.37) необходимо масштабировать:

$$\bar{h}_{li} = \begin{cases} h_{li}, & Sh_i \le S_{ИР.i} \\ h_{li} \cdot \dfrac{S_{ИР.i}}{Sh_i}, & Sh_i > S_{ИР.i} \end{cases}, \quad \forall i = \overline{1,z} \tag{3.41}$$

Ущерб, наносимый l-м ДФ незащищённой ИТКС равен:

$$u_l = \sum_{i=1}^{z} \left(\bar{h}_{li} \cdot S_{ИР.i} \right). \tag{3.42}$$

Алгоритмы определения стоимости ИР

Стоимость ресурсов ИТКС определяем на расчётный период, который для промышленных предприятий равен 1 году [13]. Для всех ресурсов, стоимость которых (исходя из общих затрат [15]) можно непосредственно определить, проводим именно такой расчёт стоимости ресурса на 1 год.

Наиболее сложным вопросом является оценка стоимости ИР. Пусть они представлены в виде конечного множества элементов и необходимо оценить суммарную их стоимость в денежных единицах.

Предложен следующий алгоритм оценки стоимости ИР.

1) Разделим ИР на две категории:

– ресурсы, стоимость которых можно определить, исходя из затрат на их приобретение и обслуживание;

– ресурсы, ценность которых является концептуальной, которые не приносят непосредственной прибыли, но потеря или порча которых нанесет предприятию ущерб.

2) Стоимость ресурса из 1-й категории

$$S_{ИР.i}^{1кат} = \frac{S_i^{приоб}}{t_{ИР.i}} + S_i^{обсл}, \quad i = \overline{1, z_{1кат}}, \tag{3.43}$$

где $S_i^{приоб}$ – стоимость приобретенного ИР (если ресурс не приобретен, а создан в самой организации, то сюда входит сумма затрат на его разработку), $t_{ИР.i}$ – срок эксплуатации ресурса в годах, $S_i^{обсл}$ – стоимость обслуживания ИР за год.

3) Оценка стоимости ресурсов 2-й категории начинаем с определения ценности ИР и 1-й и 2-й категории на основе ранговой шкалы (по табл. 3.6).

Вектор ценности по оценке ε-го эксперта:

$$C_\varepsilon = \left\{ c_1^\varepsilon, c_2^\varepsilon, \ldots, c_z^\varepsilon \right\}. \tag{3.44}$$

Для каждого *i*-го ресурса *j*-м экспертом определен ранг c_i^j.

4) Сгруппируем оценки ресурсов 1-й категории так, чтобы в каждой из максимум 9 групп (9 градаций в таблице 3.6) были ресурсы с одинаковым значением ранга.

Рассмотрим группу со значением ранга R. Пусть она состоит из z_R элементов. Вычислим среднюю стоимость ресурса в данной группе:

$$E_R = \frac{1}{z_R} \sum_{i=1}^{z_R} S_{ИР.i}^{1кат}. \tag{3.45}$$

Полученную величину можно считать стоимостью ресурса, имеющего ранга *R.* Ряд полученных величин должен быть упорядочен по возрастанию, и должно быть выполнено условие существенного различия оценок при существенном отличии рангов (в соответствии с лингвистическими описаниями табл. 3.6):

$$\begin{cases} E_R < E_{R+1}, & \left(\forall R = \overline{1, 8}\right) \\ E_{R_1} << E_{R_2}, & \left(R_2 \geq R_1 + 2\right) \end{cases}. \tag{3.46}$$

5) Если условие (3.46) не выполнено, то эксперт должен скорректировать результаты, полученные по (3.45) в соответствии с (3.46).

6) Далее всем ИР из 2-й категории, имеющим ранг R, присваиваем значение стоимости равное E_R:

$$E_R \underset{C_i^{\varepsilon}=R}{\rightarrow} S_{ИР.i}^{2кат}, \qquad \left(\forall i = \overline{1, z_{2кат}}\right). \tag{3.47}$$

7) Общая стоимость ИР по оценкам ε-го эксперта определяем суммированием:

$$S_{ИР}^{\varepsilon} = \sum_{i=1}^{z_И} S_{ИР.i}^{\varepsilon}. \tag{3.48}$$

Стоимость элементов ИТКС, подверженных воздействию ДФ $S_{ОИ}$, будем определять суммированием стоимости всех устройств в расчёте на 1 год:

$$S_{ОИ} = \sum_{k=1}^{z_{ОИ}} \frac{S_{ОИ.k}}{t_k}, \tag{3.49}$$

где $S_{ОИ.k}$ – стоимость элемента, t_k – срок службы устройства (в годах).

Стоимость элементов СОДИ $S_{СОД}$ также определяем в расчёте на 1 год путем суммирования затрат на обеспечение достоверности по всем позициям [15, 24], а также стоимости технических средств обеспечения достоверности с учётом их срока службы:

$$S_{СОД} = \sum_{k=1}^{z_{мероп}} S_{мероп.k} + \sum_{k=1}^{z_{ТС}} \frac{S_{ТС.k}}{t_k}. \tag{3.50}$$

3.6. Методы оценки показателей достоверности ИР

Одним из методов формального определения вероятностей является статистическая оценка. Она может быть адекватно применена в тех случа-

ях, когда есть статистика для аналогичной ИТКС, эксплуатируемой со сходной спецификой деятельности, т.е. когда можно говорить о наличии аналогичного набора СФН [25]. Но в связи с многообразием частных характеристик ИТКС найти аналог исследуемой системы очень не просто. Кроме того, надо учесть тот факт, что получить достоверную всеобъемлющую статистику по какому-либо предприятию от него бывает очень затруднительно.

Когда такая статистика недоступна, или в том случае, когда близких аналогов исследуемой ИТКС нет, адекватная оценка относительных частот возникновения ДФ и возможностей сохранения достоверности может быть проведена только методами экспертных оценок на основе первичной оценки не самих вероятностей, а множества частных показателей, от которых они зависят. Тогда некоторый процент ошибок в общем массиве данных не на много понизит адекватность экспертизы.

Относительные частоты возникновения ДФ

Состояние ИТКС характеризуют не только возможности нарушения достоверности информации, но и относительные частоты возникновения ДФ, приводящих к такому нарушению.

Возникновение любого ДФ преднамеренного действия обусловлено, с одной стороны, наличием в ИТКС СФН и, с другой стороны, заинтересованностью злоумышленника в выполнении такого действия. Взаимодействие двух систем: ИТКС предприятия и информационной системы злоумышленника происходит в едином информационном пространстве. С точки зрения обеспечения достоверности следует рассмотреть множество событий в этом информационном пространстве $E=\{e_0, e_1, ..., e_n\}$, где событие e_0 заключается в том, что в данный момент времени не возник ни один из полного множества ДФ, а события $e_1,...,e_n$ заключаются в возникновении l-го ДФ, $l=\overline{1,n}$.

Множество $E=\{e_0, e_1, ..., e_n\}$ является полным множеством независимых и несовместных событий, т.к. во-первых, возникновение одного ДФ никак не обусловлено возникновением любого другого ДФ, во-вторых, рассматриваемое событие “возникновение l-го ДФ” происходит в бесконечно малый промежуток времени и не может совпасть с другим событием “возникновение k-го ДФ”.

Вероятность возникновения ДФ зависит от заинтересованности злоумышленника в выполнении несанкционированных действий, естественно, при наличии некоторого СФН в ИТКС. Заинтересованность злоумышленника в использовании СФН определённо зависит от того, насколько значим данный СФН для нормального функционирования ИТКС – чем больше значимость СФН, тем более заинтересован злоумышленник в его использовании. Но на относительную частоту возникновения ДФ также оказывает влияние ещё и осведомленность злоумышленника о наличии СФН.

Относительную частоту возникновения ДФ $p_l^{ДФ}$ можно представить как функцию значений критериев "значимость СФН" и "доступность СФН" для всех СФН (т.к. один СФН может служить причиной появления нескольких ДФ, и, наоборот, множество СФН может вызвать появление одного ДФ):

$$p_l^{ДФ} = f\left(\gamma_{\alpha\beta}, d_\alpha, \forall\alpha,\beta = \overline{1,s}\right). \qquad (3.51)$$

Для вычисления количественного значения $p_l^{ДФ}$ необходимо:

– определить полное конечное множество СФН, характерных для данной ИТКС;

– получить лингвистические экспертные оценки $\gamma_{\alpha\beta}, d_\alpha, \forall\alpha,\beta = \overline{1,s}$ от каждого ε-го эксперта;

– определить причинно-следственные связи между k-м СФН и l-м ДФ;

– вычислить значения $p_{l\varepsilon}^{ДФ}$ по оценкам каждого ε-эксперта;

– агрегировать полученные значения $p_{l\varepsilon}^{ДФ}$, т.е. преобразовать каждое нечёткое множество $A\left(p_i^{ДФ}\right) = \left\{p_{i\varepsilon}^{ДФ} / \varepsilon\right\}$ в чёткое число $p_l^{ДФ}$.

На первом этапе данного алгоритма экспертиза должна быть проведена в виде круглого стола. Её результатом будет составление полного списка всех СФН, характерных для данной ИТКС.

На втором этапе эксперты представляют свои индивидуальные оценки. Для коэффициентов значимости они представлены в виде матрицы парных сравнений Саати в соответствии с табл. 3.2

$$M_{\varepsilon}^{\Gamma}=\begin{vmatrix}\gamma_{11}^{\varepsilon} & \gamma_{12}^{\varepsilon} & \dots & \gamma_{1s}^{\varepsilon}\\ \gamma_{21}^{\varepsilon} & \gamma_{22}^{\varepsilon} & \dots & \gamma_{2s}^{\varepsilon}\\ \dots & \dots & \dots & \dots\\ \gamma_{s1}^{\varepsilon} & \gamma_{s2}^{\varepsilon} & \dots & \gamma_{ss}^{\varepsilon}\end{vmatrix}. \qquad (3.52)$$

Матрица (3.52) удовлетворяет выражению (3.17).

Т.к. количество СФН типовой ИТКС исчисляется десятками, а то и сотнями, при чём их всегда можно чётко разделить на группы, то для упрощения работы по их сравнению рационально использовать метод преобразований по (3.22)-(3.27).

Вектор показателей доступности определяем по табл. 3.3-3.5 как

$$D_{\varepsilon}=\left(d_1^{\varepsilon},\, d_2^{\varepsilon},\, \dots,\, d_s^{\varepsilon}\right). \qquad (3.53)$$

На <u>третьем этапе</u> нахождения $p_l^{ДФ}$ необходимо построить матрицы причинно-следственных связей между k-м СФН и l-м ДФ

$$M_{\varepsilon}^{ПСС}=\begin{vmatrix}\rho_{11}^{\varepsilon} & \rho_{12}^{\varepsilon} & \dots & \rho_{1s}^{\varepsilon}\\ \rho_{21}^{\varepsilon} & \rho_{22}^{\varepsilon} & \dots & \rho_{2s}^{\varepsilon}\\ \dots & \dots & \dots & \dots\\ \rho_{n1}^{\varepsilon} & \rho_{n2}^{\varepsilon} & \dots & \rho_{ns}^{\varepsilon}\end{vmatrix}, \qquad (3.54)$$

где $\rho_{ik}^{\varepsilon}=1$ указывает на то, что k-й СФН может быть причиной появления l-го ДФ по мнению ε-го эксперта, $\rho_{ik}^{\varepsilon}=0$ – соответственно, не может.

Умножая матрицу (3.54) размерностью $n\times s$ на матрицу (3.52) размерностью $s\times s$, получим матрицу показателей критичности СФН с размерностью $n\times s$

$$M_j^{ПЗ}=M_{\varepsilon}^{ПСС}\times M_{\varepsilon}^{\Gamma}=\begin{vmatrix}\omega_{11}^{\varepsilon} & \omega_{12}^{\varepsilon} & \dots & \omega_{1s}^{\varepsilon}\\ \omega_{21}^{\varepsilon} & \omega_{22}^{\varepsilon} & \dots & \omega_{2s}^{\varepsilon}\\ \dots & \dots & \dots & \dots\\ \omega_{n1}^{\varepsilon} & \omega_{n2}^{\varepsilon} & \dots & \omega_{ns}^{\varepsilon}\end{vmatrix}. \qquad (3.55)$$

Величина $\omega_{ik}^{\varepsilon}$ показывает степень влияния k-го СФН на появление l-го ДФ. Величины $\omega_{ik}^{\varepsilon},\ \forall i=\overline{1,n},\ k=\overline{1,s}$ не нормированы, но при этом максимальное значение такой величины показывает максимальную степень влияния.

Нормализация (приведение к шкале Саати) должна быть проведена с учётом диапазона значений в матрице (3.50) по

$$\begin{cases} \omega_{ik}^{\varepsilon} \to \bar{\omega}_{ik}^{\varepsilon} \\ \max \gamma_{\alpha\beta}^{\varepsilon} \sim \max \bar{\omega}_{ik}^{\varepsilon}, \quad \left(\forall i=\overline{1,n},\ k,\alpha,\beta=\overline{1,s}\right). \\ \min \gamma_{\alpha\beta}^{\varepsilon} \sim \min \bar{\omega}_{ik}^{\varepsilon} \end{cases} \quad (3.56)$$

Практическая реализация такого нормирования может быть задана как

$$\bar{\omega}_{lk}^{\varepsilon} = \left(\max_{\alpha,\beta=1,\ldots,s} \gamma_{\alpha\beta}^{\varepsilon} - \min_{\alpha,\beta=1,\ldots,s} \gamma_{\alpha\beta}^{\varepsilon} \right) \cdot \frac{\omega_{lk}^{\varepsilon} - \min\limits_{\substack{l=1,\ldots,n \\ k=1,\ldots s}} \omega_{lk}^{\varepsilon}}{\max\limits_{\substack{l=1,\ldots,n \\ k=1,\ldots s}} \omega_{lk}^{\varepsilon} - \min\limits_{\substack{l=1,\ldots,n \\ k=1,\ldots s}} \omega_{lk}^{\varepsilon}} + \min_{\alpha,\beta=1,\ldots,s} \gamma_{\alpha\beta}^{\varepsilon}. \quad (3.57)$$

Преобразование (3.57) приведет к тому, что будет верно

$$\begin{cases} \max\limits_{\substack{i=1,\ldots,n \\ k=1,\ldots s}} \bar{\omega}_{ik}^{\varepsilon} = \max\limits_{\alpha,\beta=1,\ldots,s} \gamma_{\alpha\beta}^{\varepsilon} \\ \min\limits_{\substack{i=1,\ldots,n \\ k=1,\ldots s}} \bar{\omega}_{ik}^{\varepsilon} = \min\limits_{\alpha,\beta=1,\ldots,s} \gamma_{\alpha\beta}^{\varepsilon} \end{cases}. \quad (3.58)$$

при сохранении пропорций между всеми значениями из матрицы (3.55).

Таким образом, получим матрицу с нормированными значениями

$$\overline{M}_{\varepsilon}^{ПК} = \begin{vmatrix} \overline{\omega}_{11}^{\varepsilon} & \overline{\omega}_{12}^{\varepsilon} & \dots & \overline{\omega}_{1s}^{\varepsilon} \\ \overline{\omega}_{21}^{\varepsilon} & \overline{\omega}_{22}^{\varepsilon} & \dots & \overline{\omega}_{2s}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ \overline{\omega}_{n1}^{\varepsilon} & \overline{\omega}_{n2}^{\varepsilon} & \dots & \overline{\omega}_{ns}^{\varepsilon} \end{vmatrix}. \quad (3.59)$$

Матрицу (3.59) необходимо дополнить вектором (3.53) и получим

$$M_{\varepsilon}^{ДФ} = \begin{vmatrix} 10 - d_1^{\varepsilon} & 10 - d_2^{\varepsilon} & \dots & 10 - d_s^{\varepsilon} \\ \overline{\omega}_{11}^{\varepsilon} & \overline{\omega}_{12}^{\varepsilon} & \dots & \overline{\omega}_{1s}^{\varepsilon} \\ \overline{\omega}_{21}^{\varepsilon} & \overline{\omega}_{22}^{\varepsilon} & \dots & \overline{\omega}_{2s}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ \overline{\omega}_{n1}^{\varepsilon} & \overline{\omega}_{n2}^{\varepsilon} & \dots & \overline{\omega}_{ns}^{\varepsilon} \end{vmatrix}. \quad (3.60)$$

Нулевая строка данной матрицы показывает влияние показателя «доступность СФН» на возникновение в любой момент времени ситуации, когда нет никаких ДФ.

Далее необходимо определить суммарный показатель влияния всех СФН на возникновения *l*-го ДФ как

$$\omega_l^{\varepsilon} = \sum_{k=1}^{s} \overline{\omega}_{lk}^{\varepsilon}, \ \forall l = \overline{0,n}. \quad (3.61)$$

Исходя из утверждений, приведённых в начале главы, распределение суммарных показателей влияния соответствует распределению частот возникновения ДФ, которое можно получить так:

$$p_{l\varepsilon}^{ДФ} = \omega_l^{\varepsilon} \left(\sum_{l=0}^{n} \omega_l^{\varepsilon} \right)^{-1}. \quad (3.62)$$

Агрегирование полученных значений $p_{l\varepsilon}^{ДФ}$ с учётом коэффициентов авторитета экспертов v_{ε} производим следующим образом:

$$p_l^{ДФ} = \frac{2\cdot(m-2)!}{m!}\sum_{\alpha=1}^{m-1}\sum_{\beta=\alpha}^{m}\overline{H}_{\alpha\beta}\,, \qquad (3.63)$$

где

$$H_{\alpha\beta} = \frac{v_{MAX}-v_\alpha}{v_\alpha - v_\beta}(p_{l\alpha}^{ДФ} - p_\beta^{ДФ}) + p_{l\alpha}^{ДФ}, \quad \overline{H}_{\alpha\beta} = \begin{cases} 0, & H_{\alpha\beta} < 0 \\ 0 \le H_{\alpha\beta} \le 1. \\ 1, & H_{\alpha\beta} > 1 \end{cases} \qquad (3.64)$$

Агрегирование по (3.63) и (3.64) применимо к любым параметрам, определённым экспертами и должно быть использовано для всех экспертных оценок.

Возможности сохранения достоверности информации

Возможности сохранения достоверности *i*-го ИР под воздействием *l*-го ДФ по мнению **ε**-го эксперта $p_{il\varepsilon}^{C}$ определяется тем, насколько полно учтены качественные и количественные требования к передаче, хранению и обработке ИР:

$$p_{il\varepsilon}^{C} = f\left(x_{ilq_1}^{П\varepsilon}, x_{ilq_2}^{Х\varepsilon}, x_{ilq_3}^{О\varepsilon}, \quad \forall q \in \overline{1,t}\right), \qquad (3.65)$$

где $x_{ilq_1}^{П\varepsilon}$, $x_{ilq_2}^{Х\varepsilon}$, $x_{ilq_3}^{О\varepsilon}$ – степени выполнения *q*-го требования в отношении функционирования соответственно механизмов передачи, хранения или обработки *i*-го ИР в условиях воздействия *l*-го ДФ.

Пусть первые *h*-требований будут количественными ($q = \overline{1,h}$) остальные *t*–*h* – качественными ($q = \overline{h+1,t}$).

Степень выполнения количественного требования определяется его отношением к требуемому (оптимальному) количественному значению параметра СОДИ. Процедуры получения оценок количественных параметров рассмотрены в п.3.3.

Для оценки степени выполнения качественных требований (а таких требований в различных стандартах несравнимо больше, чем количественных) необходимо использовать представление требования в виде нечёткого множества.

Степень выполнения каждого качественного требования определяется функцией принадлежности ряда характеристик СОДИ, от которых зависит выполнение этого требования, к их оптимальным значениям.

Пусть $G = \{g_1, g_2, ..., g_p\}$ – универсальное множество характеристик СОДИ. На множестве G задано нечёткое множество A_q, отражающее степень принадлежности СОДИ к оптимальной по q-му требованию.

Нечёткое множество A_q определяется:

1) Множеством степеней соответствия каждой характеристики СОДИ выполнению q-го качественного требования $Y_q = \{y_1, y_2, ..., y_p\}$.

2) Множеством степеней влияния характеристик на выполнение требования в целом $\Sigma_q = \{\sigma_1, \sigma_2, ..., \sigma_p\}$.

Если некоторая характеристика полностью удовлетворяет требованию, то её степень соответствия равна 1, если полностью не удовлетворяет, то 0.

В целом степень соответствия определяется отношением к требуемому качественному описанию характеристики СОДИ, т.е. функцией принадлежности. Базовыми отношениями для уровня соответствия реальной характеристики системы требованию стандарта являются: «низкий», «ниже среднего», «средний», «выше среднего», «высокий» [22]. На множестве упорядоченных описаний характеристики в порядке следования от полного несоответствия к полному соответствию стандарту будут отношения, представленные на рис. 3.8.

Пример взят для десяти возможных описаний.

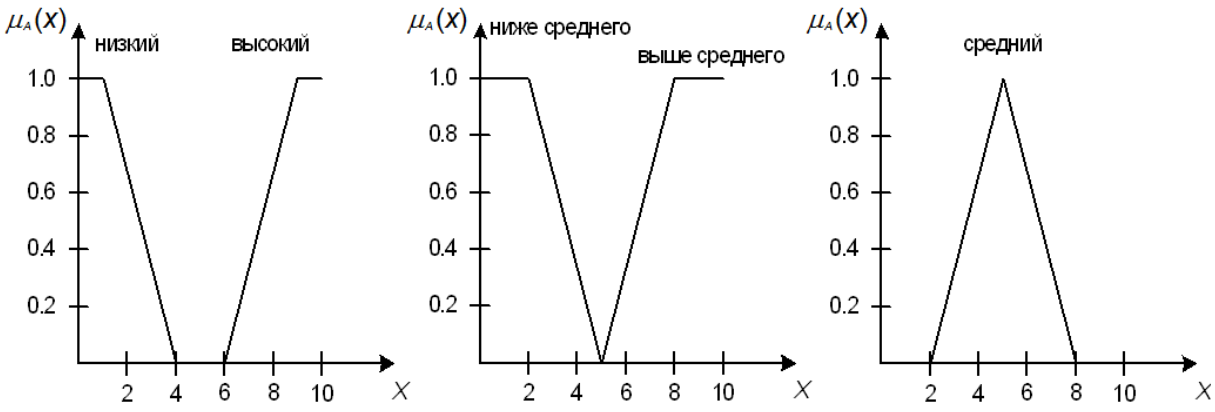


Рис. 3.8. Базовые отношения уровня соответствия характеристики требованию

Используя такие операции над нечёткими множествами, как дополнение, пересечение, объединение, концентрация и размытие можно получить функцию принадлежности для любого качественного экспертного описания, построенного на основе базовых [3, 14]. Значение функции принадлежности в точке X, соответствующей исследуемой характеристике СОДИ, будет искомым числовым значением.

Степень влияния характеристики СОДИ на выполнение требования можно найти из матрицы парных сравнений

$$M_{\varepsilon}^{\Sigma} = \begin{vmatrix} \sigma_{11}^{\varepsilon} & \sigma_{12}^{\varepsilon} & \dots & \sigma_{1p}^{\varepsilon} \\ \sigma_{21}^{\varepsilon} & \sigma_{22}^{\varepsilon} & \dots & \sigma_{2p}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ \sigma_{p1}^{\varepsilon} & \sigma_{p2}^{\varepsilon} & \dots & \sigma_{pp}^{\varepsilon} \end{vmatrix}, \qquad (3.66)$$

где $\sigma_{\alpha\beta}^{\varepsilon}$, $\forall \alpha, \beta = \overline{1,p}$ показывает, насколько влияние α-й характеристики существеннее влияния β-й.

Матрица (3.66) обычно имеет малую размерность (количество характеристик, от которых зависит требование и составляет единицы для каждого требования.

Распределение степеней влияния характеристик задано собственным вектором $\Sigma_{\varepsilon} = \left\{ \sigma_{\alpha}^{\varepsilon} \right\}$ $\alpha = \overline{1,p}$, вычисленным для максимального собственного числа матрицы (3.66).

Степень выполнения q-го качественного требования по оценкам ε-го эксперта может быть найдена по формуле

$$x_{ilq}^{\varepsilon} = \sum_{\alpha=1}^{p} y_{\alpha}^{\varepsilon} \sigma_{\alpha}^{\varepsilon}. \qquad (3.67)$$

При этом степени выполнения

x_{ilq}^{ε} $\left(\forall q = \overline{h+1,t},\ \forall i = \overline{1,z},\ \forall l = \overline{1,n},\ \forall \varepsilon = \overline{1,m} \right)$ удовлетворяют условию $0 \le x_{ilq}^{\varepsilon} \le 1$.

Возможность сохранения достоверности *i*-го ИР под воздействием *l*-го ДФ по мнению ε-го эксперта определяется как

$$p_{il\varepsilon}^{C} = \sum_{q_1}\left(v_{iq_1}^{\varepsilon} \cdot x_{ilq_1}^{\Pi\varepsilon}\right) \cdot \sum_{q_2}\left(v_{iq_2}^{\varepsilon} \cdot x_{ilq_2}^{X\varepsilon}\right) \cdot \sum_{q_3}\left(v_{iq_3}^{\varepsilon} \cdot x_{ilq_3}^{O\varepsilon}\right), \tag{3.68}$$

где $0 \le v_{iq}^{\varepsilon} \le 1$ – весовые коэффициенты важности требований по противодействию ДФ по оценке ε-го эксперта при условии, что

$$\sum_{q=1}^{t} v_{iq}^{\varepsilon} = 1, \quad \left(\forall i = \overline{1,z}, \quad \forall \varepsilon = \overline{1,m}\right). \tag{3.69}$$

Конечно, этот коэффициент, определяющий относительную значимость некоторого условия, можно определять по методам, изложенным в п. 3.4, но обычно количество требований исчисляется десятками-сотнями (количество сравнений соответственно – сотни-тысячи), а количество ДФ – также десятками, то общее число сравнений получается порядка десятков тысяч; упрощенный вариант сравнений с разделением, скажем, на десять групп даст в десять раз меньше действий, т.е. несколько тысяч.

Изложенное невозможно провести за приемлемое время. Да и количество операций в несколько тысяч, конечно, не в сотню раз повысит адекватность оценки по сравнению с количеством в несколько десятков.

Предложен следующий алгоритм проведения оценки. Эксперт заполняет предварительную матрицу

$$M_{\varepsilon}^{v^l} = \begin{vmatrix} v^{l\varepsilon}_{11} & v^{l\varepsilon}_{12} & \dots & v^{l\varepsilon}_{1t} \\ v^{l\varepsilon}_{21} & v^{l\varepsilon}_{22} & \dots & v^{l\varepsilon}_{2t} \\ \dots & \dots & \dots & \dots \\ v^{l\varepsilon}_{n1} & v^{l\varepsilon}_{n2} & \dots & v^{l\varepsilon}_{nt} \end{vmatrix}, \tag{3.70}$$

где $v^{l\varepsilon}_{iq}, \ \left(\forall q = \overline{1,t}, \ \forall i = \overline{1,n}, \ \forall \varepsilon = \overline{1,m}\right)$ показывает частное влияние *q*-го требования для устранения *i*-го ДФ по оценке ε-го эксперта. Эта величина может быть задана в пределах шкалы баллов (0 .. 5) по табл. 3.7.

Далее возможно преобразование вида

$$v_{iq}^{\varepsilon} = \frac{v'^{\varepsilon}_{iq}}{\sum_{q=1}^{t} v'^{\varepsilon}_{iq}} \quad \left(\forall i = \overline{1,n},\ \forall \varepsilon = \overline{1,m}\right), \tag{3.71}$$

удовлетворяющее условию (3.69).

Таким образом, получаем матрицу

$$M_{\varepsilon}^{V} = \begin{vmatrix} v_{11}^{\varepsilon} & v_{12}^{\varepsilon} & \dots & v_{1t}^{\varepsilon} \\ v_{21}^{\varepsilon} & v_{22}^{\varepsilon} & \dots & v_{2t}^{\varepsilon} \\ \dots & \dots & \dots & \dots \\ v_{t1}^{\varepsilon} & v_{t2}^{\varepsilon} & \dots & v_{tt}^{\varepsilon} \end{vmatrix}. \tag{2.72}$$

Таблица 3.7

Шкала оценки влияния требований на сохранение достоверности

Описание вида влияния	Балл
Абсолютное и постоянное влияние в любых условиях	5
Оказывает влияние практически во всех условиях	3
Оказывает влияние только в отдельных случаях (например, при использовании злоумышленником спецсредств)	1
Не оказывает влияния	0

Значения 2 и 4 могут быть взяты как промежуточные.

3.7. Оценка общего показателя достоверности информации

Из (2.1) с использованием (3.63) и (3.68) находим частные показатели достоверности отдельных ИР. Т.к. различные ресурсы оказывают разное влияние на качество функционирования ИТКС, то необходимо ввести некоторую величину η_i^{ε} $\left(\forall i = \overline{1, z_И}\right)$, показывающую относительную ценность *i*-го информационного ресурса. Для коммерческих предприятий ценность ресурса всегда сводится к его денежному эквиваленту, следовательно, она может быть определена из (3.48) как

$$\eta_i^{\varepsilon} = \frac{S_{ИР.i}^{\varepsilon}}{S_{ИР}^{\varepsilon}}, \quad \left(\forall i = \overline{1, z_И}\right). \tag{3.73}$$

Тогда общий показатель достоверности достоверность информации в ИТКС по оценке ε-го эксперта определим следующим образом:

$$D_{\varepsilon} = \sum_{i=1}^{z_И} \frac{S_{ИР.i}^{\varepsilon}}{S_{ИР}^{\varepsilon}} \cdot D_i^{\varepsilon} . \tag{3.74}$$

3.8. Оценка показателей достоверности информации на примере ИТКС предприятия

Корпоративная ИТКС ОАО «ВЗ «Электроприбор»

Структурная схема корпоративной ИТКС ОАО «Владимирский завод «Электроприбор» представлена на рис. 3.9.

В состав ИТКС входят: контроллер домена, сервер для работы бухгалтерской программы «Инфин», дополнительный контроллер домена, сервер АСУ, сервер-терминал, сервер резервного копирования, интернет-сервер, внутренний web-сервер, АРМы пользователей.

Контроллер домена – сервер, который управляет всей политикой безопасности в сети. Все пользовательские компьютеры регистрируясь на этом сервере, получают доступы к тем ИР, которые ему разрешены на сервере. Средства обеспечения достоверности: стандартные средства разграничения доступа ОС Windows 2000 Advanced Server, антивирус Касперского. Сервер находится в специальном помещении, оборудованном техническими средствами охраны. Доступ к системным ресурсам сервера имеют только администратор ЛВС, администратор безопасности.

Сервер для работы бухгалтерской программы «Инфин». Комплекс «Инфин» автоматизирует бухгалтерский учёт на предприятии, обеспечивает ведение журналов и книги проводок, настройку плана счетов и регистров аналитического учета. Автоматизированы все процедуры, связанные с накоплением, расчётом и анализом заработной платы сотрудников. Работает с MS SQL Server. Средства обеспечения достоверности те же. Сервер находится в специальном помещении, оборудованном техническими сред-

ствами охраны. Доступ к системным ресурсам имеют администратор ЛВС, администратор безопасности, к ИР – сотрудники бухгалтерии, центра по управлению финансами, отдела организации и оплаты труда.

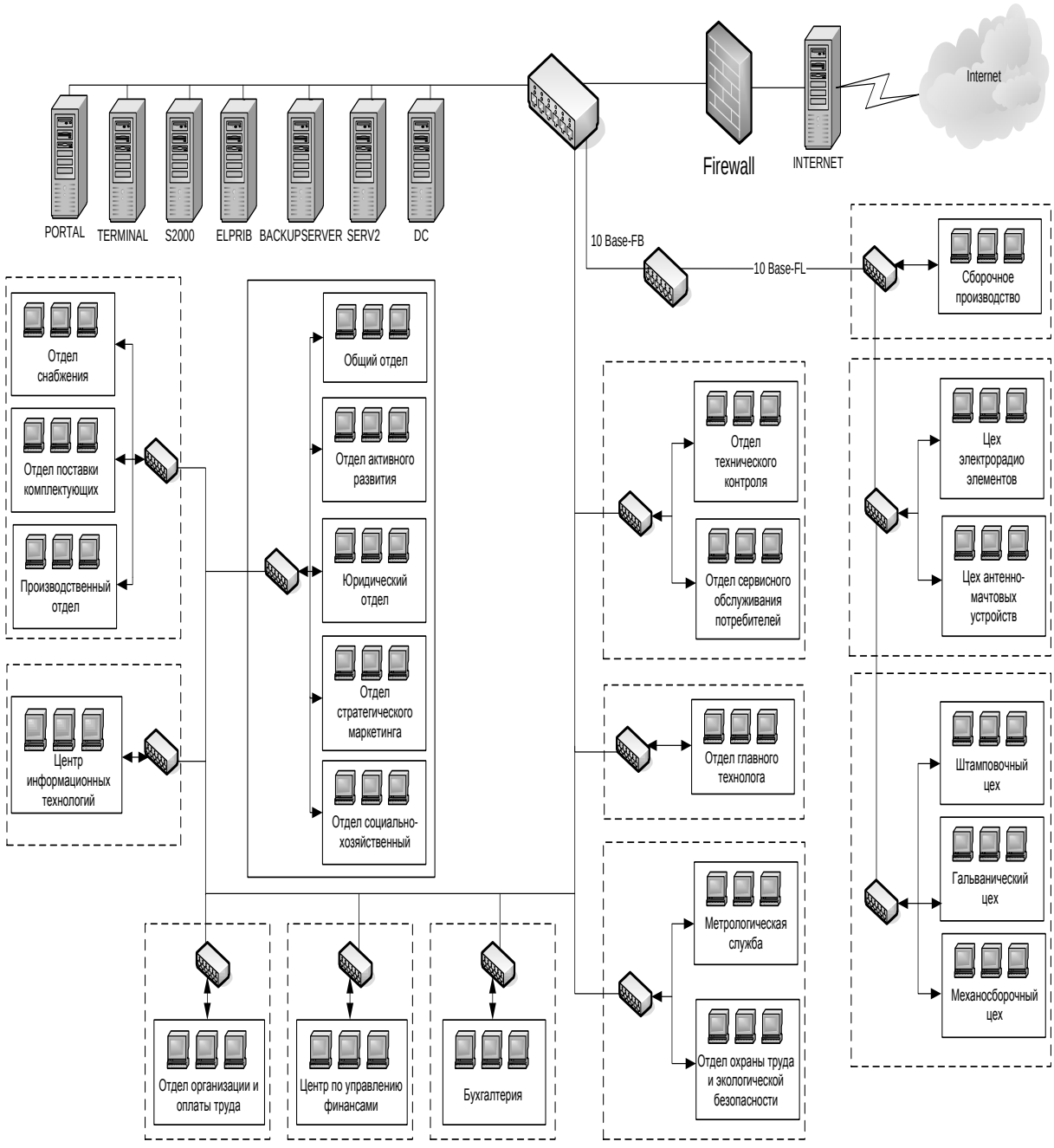


Рис. 3.9. Структурная схема корпоративной ИТКС ОАО «Владимирский завод «Электроприбор»

Дополнительный контроллер домена хранит полную копию записей главного контроллера домена и периодически синхронизирует с основным

по выбранным протоколам. Средства обеспечения достоверности те же. Сервер находится в специальном помещении, оборудованном техническими средствами охраны. Доступ к системным ресурсам сервера имеют только администратор ЛВС, администратор безопасности.

Сервер АСУ. АСУ – программный комплекс собственной разработки, предназначенный для управления технологическими и производственными процессами предприятия. Работает с СУБД FoxPro. Средства обеспечения достоверности те же. Сервер находится в специальном помещении, оборудованном техническими средствами охраны. Доступ к системным ресурсам сервера имеют только администратор ЛВС, администратор безопасности. Доступ к ИР сервера имеют сотрудники всех подразделений.

Сервер-терминал обеспечивает удаленный доступ к рабочему столу на сервере при помощи ПО клиента, функционирующего в качестве эмулятора терминала. Средства обеспечения достоверности те же. Сервер находится в специальном помещении, оборудованном техническими средствами охраны. Доступ к системным ресурсам сервера имеют только администратор ЛВС, администратор безопасности. Доступ к ИР сервера имеют сотрудники всех подразделений предприятия.

Сервер резервного копирования выполняет в системе функцию резервного хранения на носителях информации. Средства обеспечения достоверности те же. Данный сервер находится в специальном помещении – серверной, которая оборудована техническими средствами охраны. Доступ к системным ресурсам сервера имеют только сотрудники ЦИТ (администратор ЛВС, администратор безопасности).

Внутренний web-сервер. Внутренний портал объединяет имеющиеся у организации ИР и предоставляет пользователям единый защищённый доступ к служебной информации и сервисам. Средства обеспечения достоверности те же. Данный сервер находится в специальном помещении – серверной, которая оборудована техническими средствами охраны. Доступ к системным ресурсам сервера имеют только сотрудники ЦИТ (администратор ЛВС, администратор безопасности). Доступ к ИР сервера имеют сотрудники всех подразделений ОАО ВЗ «Электроприбор».

Интернет-сервер обеспечивает доступ пользователей в Интернет, а также предоставляет сервис корпоративной электронной почты. Средства обеспечения достоверности: стандартные средства разграничения доступа ОС Windows server 2003, антивирус Касперского, анти-спамовое ПО

MailShell Spam Catcher, прокси-сервер WinGate 6.2. Находится в серверной.

Информационные ресурсы:

1) схема инфраструктуры сети;
2) сведения о подготовке, принятии и исполнении отдельных решений руководства по коммерческим, организационным и иным вопросам;
3) планы развития предприятия;
4) разрабатываемые проекты;
5) базы данных клиентов, круг поставщиков;
6) сведения о подготовке и результатах проведения переговоров;
7) сведения о получаемых и прорабатываемых заказах и предложениях;
8) данные о контрактах с партнерами;
9) данные о заказах;
10) банковские операции, состояние банковских счетов предприятия и производимых операций;
11) отчёты о кредитах поставщиков;
12) сведения о контрактах;
13) сведения, составляющие коммерческую тайну предприятий-партнеров и переданные на доверительной основе;
14) ресурсы внутреннего web-сервера;
15) технологическая информация.

Элементы и средства обработки и передачи информации:

1) рабочие станции специалистов отдела снабжения;
2) рабочие станции специалистов отдела поставки комплектующих;
3) рабочие станции специалистов производственного отдела;
4) рабочие станции специалистов юридического отдела;
5) рабочие станции специалистов отдела стратегического маркетинга;
6) рабочие станции специалистов центра по управлению финансами;
7) рабочие станции бухгалтерии;
8) серверы;
9) съёмные носители информации.

Множество СФН:

1) слабая техническая укрепленность дверей;
2) незапертые двери;
3) свободный доступ в помещения отделов;
4) устаревшее антивирусное программное обеспечение;

5) ошибки в конфигурации (подготовка к работе вручную, приводящая к несогласованным конфигурациям);
6) нерегулярная смена паролей пользователей;
7) отсутствие контроля целостности файлов прикладных программ.
8) недоработанные инструкции пользователей по работе с ИР;
9) неопределенность в отношении ответственности в случае нарушений;
10) отсутствие политики обучения персонала правилам ИБ;
11) отсутствие учёта съемных носителей информации;
12) отсутствие специального хранилища сменных носителей информации.

Множество ДФ:

1) внедрение программ-закладок;
2) запуск программы в качестве системной;
3) подмена динамически загружаемой библиотеки;
4) модификация кода или данных подсистемы защиты ОС;
5) нарушение функционирования;
6) навязывание сообщений;
7) изменение таблиц маршрутизаторов
8) ввод сотрудниками неверных данных;
9) намеренное искажение информации;
10) внедрение вредоносного ПО;
11) нарушение целостности в базах данных;
12) подмена информации на съемных носителях.

Средства СОДИ:

1) Средства шифрования пакетов (трафика): «Тропа», «VIPNet», «Континент», «Шип», «Игла»;
2) Средства шифрования отдельных сообщений (почты, передаваемых файлов, блоков): «Защищённая почтовая служба VIPNet», «Курьер»;
3) Средства криптографической защиты информации на магнитных носителях: «Cryptomania», «StrongDisk»;
4) Межсетевые экраны: «Застава-Джет», «FortE+ (Застава-Элвис)», «FW-1», «Cisco PIX», «Black Hole (SecurlT Fire-Wall)», «Cyber Guard», «AltaVista»;
5) Средства аудита/протоколирования: «RealSecure Network», «Sensor RealSecure OS», «Sensor RealSecure Manager fr Open View».

Экспертиза ИТКС дала следующие первичные оценки отдельных параметров (табл. 3.8-3.17).

Таблица 3.8

Сравнение значимостей СФН

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	1	9	9	9	1/5	1/7	1/7	1	1	1	9	1/9
2	1/9	1	1/7	1/5	1/9	1/9	1/9	1	1/9	1/9	1/9	1/9
3	1/9	7	1	3	1/7	1/7	1/7	1	1/7	1/7	5	1/9
4	1/9	5	1/3	1	1/9	1/9	1/9	1	1/9	1/9	1/9	1/9
5	5	9	7	9	1	1/7	1/7	1	1/9	1/9	7	1/7
6	7	9	7	9	7	1	1	1	1/9	1/9	9	1/9
7	7	9	7	9	7	1	1	1	1/9	1/9	7	7
8	1	1	1	1	1	1	1	1	1	1	1	1
9	1	9	7	9	9	9	9	1	1	1	1/5	3
10	1	9	7	9	9	9	9	1	1	1	1/7	1/9
11	1/9	9	1/5	9	1/7	1/9	1/7	1	5	7	1	9
12	9	9	9	9	7	9	1/7	1	1/3	9	1/9	1

Таблица 3.9

Матрица причинно-следственных
связей между СФН и ДФ

		СФН											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ДФ	**1**	0	0	0	0	0	0	0	0	0	0	1	0
	2	0	0	0	0	0	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0	0	0	1	0	0
	4	0	0	0	0	0	0	0	0	1	0	1	0
	5	0	0	0	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0	0	0	0	0
	8	0	0	0	0	0	0	0	0	0	0	0	0
	9	0	0	0	0	0	0	0	0	0	0	0	0
	10	0	0	0	0	0	0	0	0	0	0	1	0
	11	0	0	0	0	0	0	0	0	0	0	0	0
	12	0	0	0	0	0	0	0	0	0	1	0	0

Таблица 3.10

Вектор доступности СФН

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
3	1	7	5	5	3	3	5	5	9	9	9

Таблица 3.11

Матрица степени воздействия ДФ на ИР

		ИР														
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
ДФ	**1**	0,05	0,05	0,00	0,07	0,06	0,06	0,00	0,07	0,07	0,00	0,00	0,00	0,00	0,07	0,00
	2	0,07	0,07	0,04	0,02	0,02	0,02	0,08	0,02	0,02	0,07	0,17	0,05	0,07	0,02	0,07
	3	0,05	0,05	0,00	0,07	0,06	0,06	0,00	0,07	0,07	0,00	0,00	0,00	0,00	0,07	0,00
	4	0,12	0,12	0,27	0,07	0,04	0,16	0,15	0,05	0,05	0,13	0,25	0,29	0,27	0,09	0,27
	5	0,02	0,02	0,00	0,02	0,02	0,02	0,00	0,02	0,02	0,00	0,00	0,00	0,00	0,02	0,00
	6	0,02	0,02	0,04	0,02	0,02	0,02	0,08	0,02	0,02	0,07	0,04	0,05	0,07	0,02	0,07
	7	0,03	0,03	0,08	0,05	0,04	0,04	0,15	0,05	0,05	0,13	0,08	0,10	0,13	0,05	0,13
	8	0,05	0,05	0,00	0,07	0,06	0,06	0,00	0,07	0,07	0,00	0,00	0,00	0,00	0,07	0,00
	9	0,07	0,07	0,00	0,09	0,09	0,08	0,00	0,10	0,10	0,00	0,00	0,00	0,00	0,09	0,00
	10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,05	0,00	0,00	0,00	0,00	0,05	0,00
	11	0,03	0,03	0,04	0,05	0,04	0,04	0,08	0,05	0,05	0,07	0,04	0,05	0,07	0,05	0,07
	12	0,12	0,12	0,00	0,11	0,06	0,16	0,00	0,07	0,07	0,00	0,00	0,00	0,00	0,07	0,00

Таблица 3.12

Оценка степени соответствия характеристик СОДИ
предъявляемым требованиям

		Характеристики														
Требования		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
	1	0,1						0,1	0,8				0,2			
	2	0,1	0,2		0,1	0,2				0,3	0,1				0,5	0,2
	3	0,1					0,8	0,6	0,8				0,3	0,1		
	4	0,1		0,3					0,1		0,5					0,3
	5					0,1	0,2	0,6	0,8	0,2		0,3			0,1	
	6	0,1			0,2			0,3	0,3				0,4	0,2		0,1
	7			0,3			0,2	0,8	1			0,5				
	8	0,1			0,1			0,8	0,8		0,2				0,8	
	9	0,1	0,3		0,1	0,2		0,5	0,8	0,3		0,7	0,8	0,1		
	10			0,1				1	1						0,3	0,2
	11	0,1	0,3	0,3	0,2		0,4			0,3		0,3				
	12			0,3		0,2		0,1			0,1		0,6	0,3		0,6
	13	0,1	0,3	0,3				0,2	0,2						0,7	
	14				0,3	0,3						0,2		1		

Таблица 3.13

Стоимость всех ИР в расчёте на 1 год

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
100000	100000	25000	30000	25000	28000	75000	65000
9	**10**	**11**	**12**	**13**	**14**	**15**	
75000	100000	75000	100000	75000	765000	80000	

Суммарная стоимость: 1 748 000 руб.

Результаты расчёта производных показателей

Таблица 3.14

Распределение ущерба, наносимого ДФ

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
7322	7322	58572	87858	248931	43929	73215	161073	14643	87858	131787	131787

Таблица 3.15

Распределение частот возникновения ДФ

0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0,01	0,09	0,04	0,08	0,13	0,05	0,04	0,13	0,05	0,05	0,11	0,08	0,14

Таблица 3.16

Вероятности сохранения достоверности информации под действием ДФ

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0,33	0,38	0,30	0,25	0,39	0,38	0,57	0,51	0,23	0,23	0,34	0,65

Таблица 3.17

Достоверность отдельных ИР, ранжированных по значимости

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
0,23	0,21	0,15	0,3	0,35	0,28	0,23	0,45	0,51	0,45	0,58	0,62	0,77	0,83	0,8

Анализ результатов исследования

Исследование ИТКС с учётом исчерпывающего списка ИР, ДФ, СФН и структуры ИП выявило следующие проблемы в обеспечении достоверности ИР:

– показатели достоверности большинства ресурсов ниже уровня 0,5, что в контексте нечётких оценок соответствует описанию «недостоверно»;

– в то время как ИП, ассоциированные с наиболее ценными ИР

защищены от действия ДФ относительно хорошо, уровень достоверности ИР средней ценности достаточно низок;

– возможности нарушения достоверности различных ИР под действием родственных ДФ имеют близкие по величине значения.

Графики, отражающие первые две указанные особенности, приведены на рис. 3.10 для пятнадцати ИР, выстроенных в порядке следования от менее значимых к более значимым:

1) резервные копии ИР,
2) финансовая отчётность,
3) личные дела сотрудников,
4) электронная переписка,
5) информация о системе охраны и пропускном режиме,
6) информация о противопожарной системе,
7) техническая документация,
8) проектная документация,
9) технологическая производственная документация,
10) информация систем шифрования,
11) конфигурации ПО,
12) электронные БД,
13) режимы управления технологическим оборудованием,
14) настройки технологического оборудования,
15) ИР систем управления технологическим оборудованием.

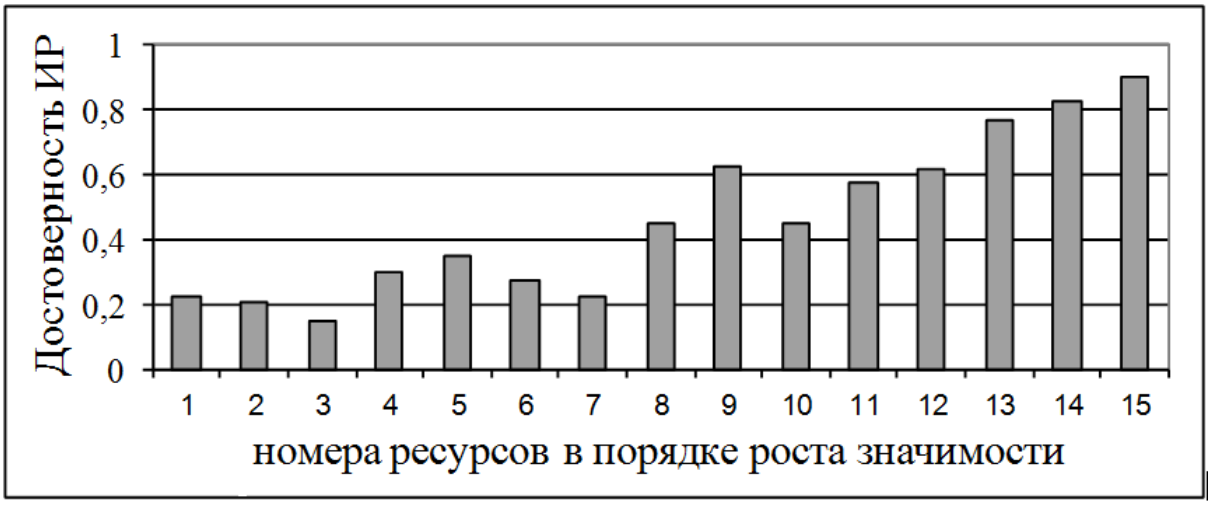


Рис. 3.10. Показатели достоверности ИР

Решения по обеспечению требуемого уровня достоверности

Устранение отдельных типов СФН путём совершенствования СОДИ отражено на рис. 3.11-3.14.

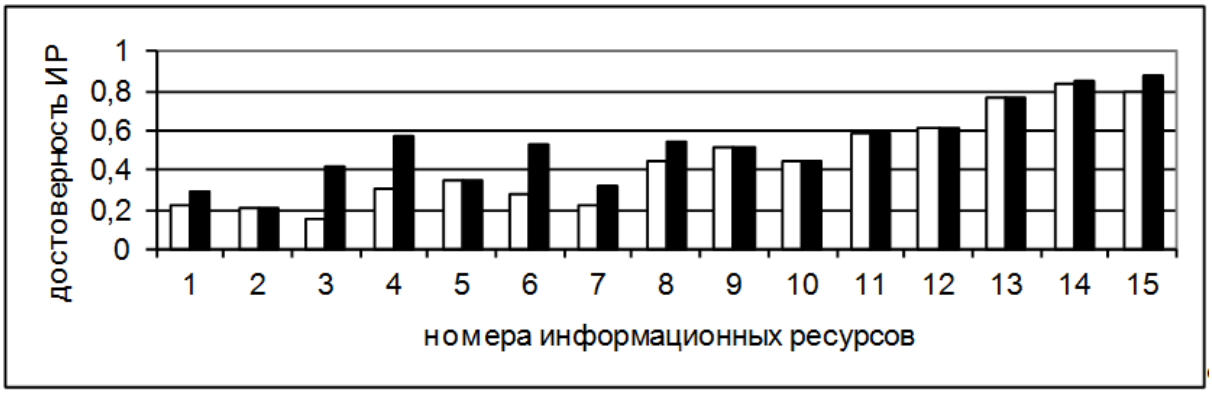


Рис. 3.11. Изменение показателей достоверности при повышении адекватности моделей представления данных

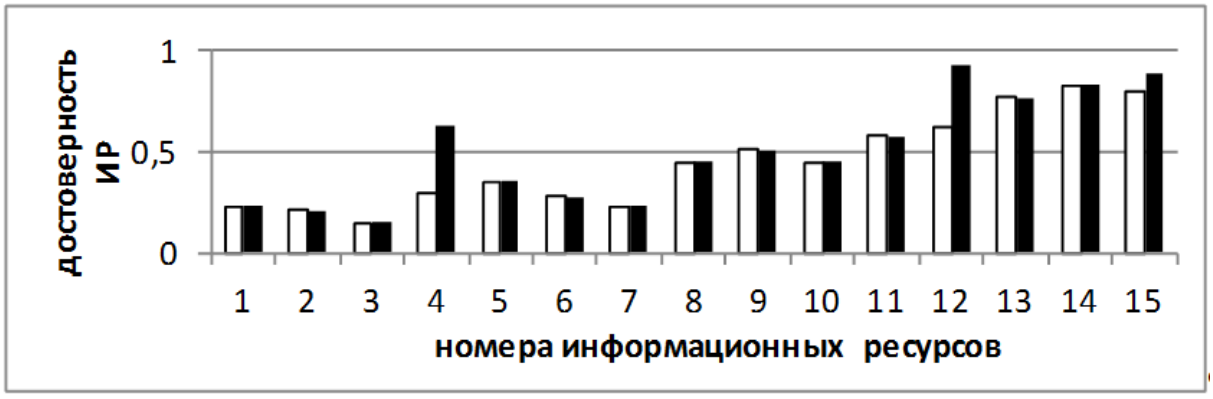


Рис. 3.12. Изменение показателей достоверности при повышении качества организации информационного обмена

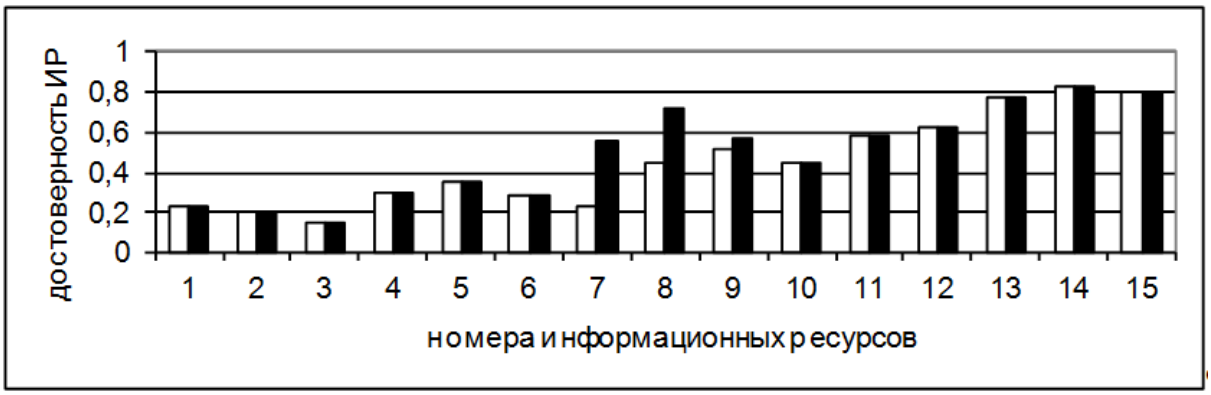


Рис. 3.13. Изменение достоверности ИР при повышении качества контроля ИР

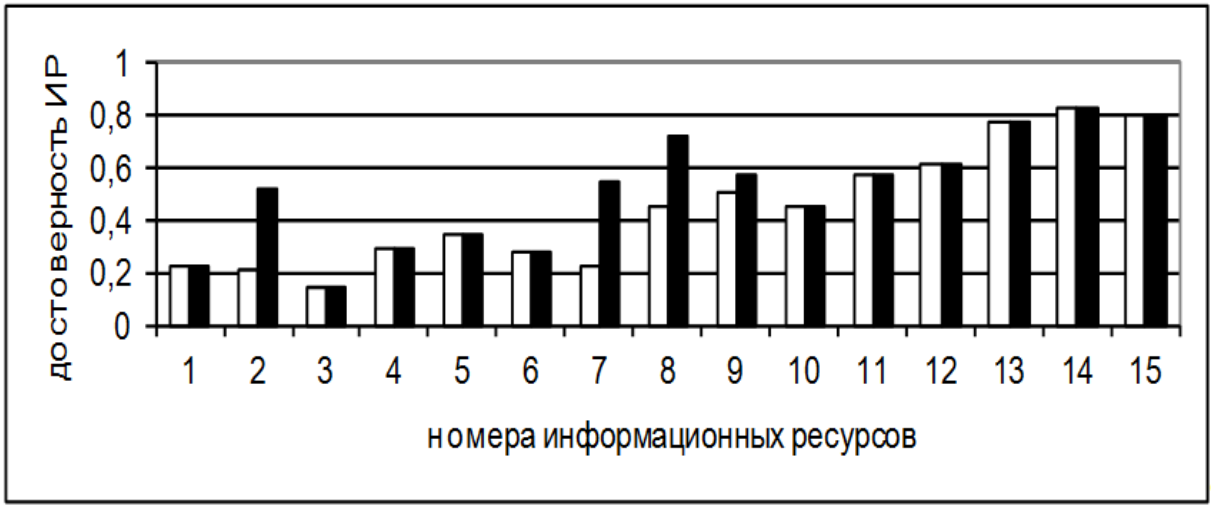


Рис. 3.14. Изменение достоверности ИР при повышении квалификации персонала

Предложено три варианта комплексного изменения структуры системы ОДИ:

– применение дополнительных средств обеспечения целостности электронных массивов информации: использовать средства организации электронной цифровой подписи – «КриптоПро CSP» и «VCERT PKI»;

– дополнительно к первому варианту применить технические средства повышения помехоустойчивости каналов передачи информации и дублирования;

– дополнительно к предыдущим вариантам добавить средства обеспечения целостности информации в неэлектронной форме и на электронных носителях: Storage Central фирмы NETGEAR, Acronis Recovery for MS SQL Server, Enterprise Vault фирмы Symantec (VERITAS), F1 DrLab фирмы DrLab.

Перерасчёт частных показателей для улучшенных вариантов СОДИ

Ниже представлены результаты перерасчета частных показателей, которые позволят получить улучшение вариантов СОДИ (табл. 3.18-3.21).

Таблица 3.18

Переоценка степени соответствия характеристик СОДИ предъявляемым требованиям

		Характеристики														
Требования		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
	1	0,1						0,1	0,8				0,2			
	2	0,1	0,2		0,4	0,2				0,3	0,1				0,5	0,2
	3	0,1					0,8	0,6	0,8				0,3	0,1		
	4	0,1		0,3					0,1		0,5					0,3
	5					0,1	0,5	0,6	0,8	0,2		0,3			0,1	
	6	0,1			0,2			0,3	0,3				0,4	0,2		0,1
	7			0,5			0,2	0,8	1			0,5				
	8	0,1			0,4			0,8	0,8		0,2				0,8	
	9	0,4	0,3		0,1	0,5		0,5	0,8	0,3		0,7	0,8	0,5		
	10			0,1				1	1						0,3	0,4
	11	0,4	0,3	0,3	0,2		0,4			0,4		0,3				
	12			0,3		0,4		0,1			0,1		0,6	0,3		0,6
	13	0,1	0,3	0,3				0,2	0,2						0,7	
	14				0,3	0,7						0,7		1		

Таблица 3.19

Новое распределение ущерба, наносимого ДФ

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
7322	7322	58572	81858	148231	13929	7321	16107	14343	17858	131787	13187

Таблица 3.20

Новые вероятности сохранения достоверности информации под действием ДФ

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0,39	0,38	0,31	0,42	0,39	0,58	0,57	0,71	0,23	0,73	0,54	0,65

Таблица 3.21

Достоверность отдельных ИР, ранжированных по значимости

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
0,48	0,54	0,65	0,65	0,55	0,59	0,56	0,75
9	**10**	**11**	**12**	**13**	**14**	**15**	
0,57	0,65	0,70	0,93	0,77	0,90	0,90	

На рис. 3.15 представлены диаграммы изменения показателей достоверности для выше предложенных вариантов изменения структуры СОДИ.

Все три варианта не выходят за рамки экономической целесообразности. Но только третий вариант обеспечивает и наибольший общий показатель достоверности (0,84), и значения частных показателей всех ИР больше минимально установленной границы в 0,5.

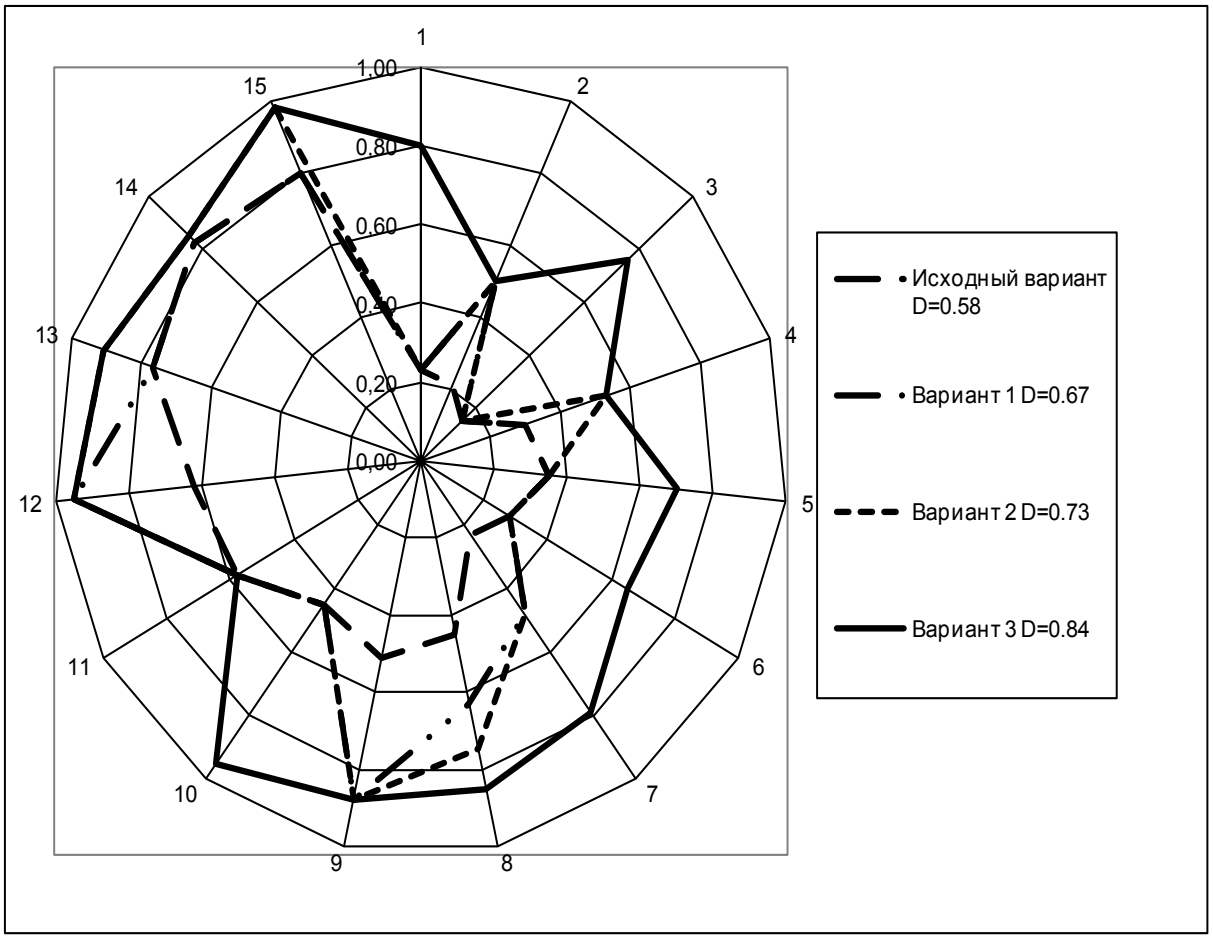


Рис. 3.15. Изменение показателей достоверности ИР

Выводы по главе

Разработана новая общая модель оценки достоверности информации в ИТКС в условиях информационных воздействий с учетом решения сопутствующих задач оценки рисков и экономической эффективности мероприятий по повышению достоверности.

Для решения поставленных задач разработаны:

– алгоритм проведения экспертизы параметров ИТКС для определения текущего уровня достоверности информации;

– методика проведения парных сравнений, существенно

сокращающая количество действий, выполняемых экспертом;

– процедуры получения числовых значений количественных и качественных параметров ИТКС, являющихся исходными данными для оценки достоверности и использующие новые лингвистические таблицы

– методика расчета информационных рисков в ИТКС, защищенности и экономической эффективности, использующая новый алгоритм определения стоимости информационных ресурсов.

– методика расчета вероятностей возникновения и вероятностей устранения угроз безопасности, основанная на оценке ряда частных показателей.

– математическая модель оценки достоверности информации в ИТКС и методика проведения оценки, основанная на анализе характера и последствий воздействия угроз информационным ресурсам.

Проведено экспериментальное исследование ИТКС промышленного предприятия. Выявлено, что для обеспечения достоверности информации используется недостаточное количество технических средств и не проводится регулярных мероприятий. В связи с чем общий уровень достоверности защищаемой информации весьма низкий. Вместе с тем, совершенствование отдельных механизмов обеспечения достоверности информации при незначительных (на общем фоне) затратах позволяет повысить достоверность защищаемой информации.

Список библиографических ссылок

1. Алексеев О.Г. Комплексное применение методов дискретной оптимизации. М.: Наука, 1987. 248 с.

2. Анохин А.М., Глотов В.А., Павельев В.В. Методы определения коэффициентов важности критериев // Автоматика и телемеханика. 1997. №8. С. 3-35.

3. Асаи К., Ватада Д., Иваи С. Прикладные нечёткие системы. М.: Мир, 1995. 418 с.

4. Бешелев С.Д., Гурвич Ф.Г. Математико-статистические методы экспертных оценок. М.: Статистика, 1980. 263 с.

5. Вентцель Е.С. Теория вероятностей. М.: Высшая школа, 1999. 576 с.

6. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. М.: Высшая школа, 2000. 480 с.

7. Виноградов И.М. Основы теории чисел. М.: Наука, 1981. 176 с.

8. ГОСТ Р АСУО/МЕК 15408-1-2001. Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 1. Введение и общая модель.

9. ГОСТ Р АСУО/МЕК 15408-2-2001. Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 2. Функциональные требования безопасности.

10. Грушо А.А., Тимонина Е.Е. Теоретические основы защиты информации. М.: Изд-во Агентства «Яхтсмен», 1998. 192 с.

11. Евланов Л.Г., Кутузов В.А. Экспертные оценки в управлении. М.: Экономика, 1978. 133 с.

12. Защита информационных процессов: Стандарт CobiT. Управление и аудит информационных технологий. Особенности проведения внешнего аудита ИТ. URL: http://citforum.ru/consulting/standart_cobit/article1.1.2003623.html (дата обращения: 18.09.2015)

13. Конеев, И.Р., Беляев А.В. Информационная безопасность предприятия. СПб.: БХВ-Петербург, 2003. 752 с.

14. Кофман А. Введение в теорию нечётких множеств. М.: Радио и связь, 1982. 432 с.

15. Кунакова Н. Современные методы и средства анализа и управления рисками информационных систем компаний. 2005. URL: http://citforum.ru/products/dsec/cramm/ (дата обращения: 18.09.2015)

16. Литвак Б.Г. Экспертные оценки и принятие решений. М.: Патент, 1996. 272 с.

17. Маркевич М. Принципы проведения активного аудита информационной безопасности компании. 2007. URL: http://dsec.ru/ipm-research-center/article/principles_of_carrying_out_active_audit_information_security_company/ (дата обращения: 18.09.2015)

18. Медведовский И., Петренко С. Анализ рисков информационных систем компаний. 1999. URL: http://ais.khstu.ru/Electr_Books/books/Analiz_riskov_inform_kompanii.chm, http://www.classs.ru/library/node/3113 (дата обращения: 18.09.2015)

19. Методология оценки безопасности информационных технологий по общим критериям // Jet Info информационный бюллетень. 2004. №6 (133). 16 с.

20. Монахов М.Ю., Кулаков М.А., Полянский Д.А. Анализ и пути повышения защищенности корпоративной сети предприятия // Вестник Костромского государственного университета им. Н.А. Некрасова. 2009. №1. С. 66–68.

21. Орлов А.И. Экспертные оценки // Заводская лаборатория. 1996. Т.62. № 1. С. 54-60.

22. Панкова, Л.А., Петровский А.М. Шнейдерман М.В. Организация экспертизы и анализ экспертной информации. М.: Наука, 1984. 120 с.

23. Петренко С.А., Симонов С.В. Управление информационными рисками: Экономически оправданная безопасность. М.: ДМК-Пресс, 2004. 384 с.

24. Полянский Д.А. Комплексная защита объектов информатизации. Книга 10. Оценка защищённости: учебное пособие. Изд-во Владим. гос. ун-та, 2005. 80 с.

25. Полянский Д.А. Методика расчёта вероятностей возникновения угроз безопасности информационной системе предприятия // Математические методы в технике и технологиях – ММТТ-20. сб. трудов XX междунар. науч. конф. В 10 т. Т. 6. Секция 12 / под общ. Ред. В.С. Балакирева. Ярославль: Изд-во Ярос. гос. техн. Ун-та, 2007. С. 20-21.

26. Полянский Д.А. Методика расчёта экономической эффективности системы защиты информации предприятия на основе экспертных оценок // Формирование социально-ориентированной экономики: вопросы теории и практики. Межвуз. сб. науч. трудов / филиал ВЗФЭИ в г. Владимире. Владимир, 2007. С. 134–137.

27. Полянский Д.А. Методы оптимизации системы защиты информации предприятия // Труды Владимирского государственного университета. Выпуск 1. Информационно-телекоммуникационные технологии и электроника. Владимир, 2006. С. 53–57.

28. Полянский Д.А. Применение методики экспертных оценок для расчёта вероятностей возникновения угроз безопасности информационной системе предприятия // Проблемы эффективности безопасности функционирования сложных технических и информационных систем. Сборник №1. Труды XXVI Международной научно-технической конференции. Серпу-

хов: Серпуховской ВИ РВ, 2007. С. 68-72.

29. Полянский Д.А. Расчёт вероятностей возникновения угроз безопасности информационной системе // Методы и технологии автоматизации обучения, компьютерной графики и информационной безопасности / под ред. И.Е. Жигалова и М.Ю. Монахова. Владимир: Владим. гос. ун-т. 2007. С. 57–60.

30. Полянский Д.А. Экономика защиты информации: учебное пособие. Владимир: Изд-во Владим. гос. ун-та, 2009. – 96 с.

31. Полянский Д.А. Экономическая эффективность систем информационной безопасности // Информационные технологии в образовательном процессе и управлении: Межвузовский сб. статей / Под ред. В.Н. Федосеева. Шуя: Издательство «Весть», 2007. С. 11–12.

32. Полянский Д.А., Монахов М.Ю. Методика определения значимости условий возникновения ошибок при обработке информации в АСУП // Автоматизация в промышленности. 2008. №11. С.10-12.

33. Полянский Д.А., Монахов Ю.М. Модель поиска оптимальной структуры системы защиты информации // Информационные системы и технологии в образовании и экономике / Сб. трудов научно-практической конф. М., Покров: МГПУ им. С.А. Шолохова, 2007. С. 85–86.

34. Полянский Д.А., Устинов В.Н. Обеспечение информационной безопасности процесса принятия управленческих решений // Труды российского научно-технического общества радиотехники, электроники и связи им. А.С. Попова, выпуск LX-1. М.: ООО «Инсвязьиздат», 2005. С. 164-165.

35. Симонов С. Технологии и инструментарий для управления рисками // Jet Info информационный бюллетень. 2003. №2 (117). 32 с.

ГЛАВА 4. ИССЛЕДОВАНИЕ И РАЗРАБОТКА МОДЕЛЕЙ И АЛГОРИТМОВ ОБНАРУЖЕНИЯ И ПРЕДОТВРАЩЕНИЯ РАСПРЕДЕЛЕННЫХ АТАК НА ДОСТУПНОСТЬ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Формальные методы обнаружения и предсказания атак (в первую очередь DoS-атак) практически отсутствуют для широкого использования в реальных системах. В настоящее время среди специалистов сложилось четкое убеждение в том, что анализ информационного сетевого потока является наиболее эффективным методом обнаружения аномального поведения распределенной ИТКС по причине его большой информативности и потенциальной возможности реагирования в реальном масштабе времени. Поэтому наиболее перспективные исследования в настоящий момент [2, 3, 6-9, 23, 26, 27, 29-32, 33-39] направлены на разработку способов и процедур обнаружения атак, основой которых является изучение влияния вредоносного воздействия на характеристики сетевого трафика.

В качестве теоретической базы в основном применяется методология и прикладные результаты теории систем массового обслуживания (теории очередей). В течение десятилетий анализ очередей основывался на предположении о соответствии типа трафика распределению Пуассона. Однако результаты ряда исследований [33-38] показали, что в отдельных случаях трафик является по своей природе самоподобным (self-similar), или фрактальным. При таком трафике системные характеристики не подчиняется формулам анализа очередей, а имеют место большие задержки и снижение пропускной способности. Эти результаты были подтверждены на трафиках самых разных типов. В [36] Верес показал, что возникающие скопления пакетов TCP-трафика ведут себя как хаотические. В данном случае под хаосом понимается сложное поведение, меняющееся во времени, вызванное взаимным влиянием потоков TCP-трафика. Такое поведение трафика чувствительно к малейшим возмущениям, несмотря на то, что описывается простыми и детерминистическими уравнениями. Рассмотрение TCP-трафика с позиций теории хаоса дает возможность учитывать корреляции потоков и предоставляет ту же сложную картину сети, которая наблюдается и в реальности.

В настоящее время достоверно неизвестно, что в точности вызывает

хаотическое поведение TCP-трафика, соответственно неизвестно как данная модель может быть применима в общем случае. В частности, исследования в [37] указывают на то, что хаотичность наблюдалась в сетях, где вероятность потери пакетов составляла значительно более 1% и наблюдался нарастающий процесс отсрочки передачи (вполне вероятно, что данная сеть была под DoS-атакой).

В главе на основе гипотезы, что самоподобность и персистентность сетевого трафика вызвана DoS-атакой, разрабатываются модели и процедуры анализа TCP-трафика на основе теории хаоса, предлагается методика раннего обнаружения информационной атаки, исследуется адекватность предложенных моделей.

4.1. Математическая модель самоподобного процесса

Непрерывный самоподобный процесс

Процесс $X(t)$ считается статистически самоподобным с параметром H ($0{,}5 \le H \le 1$), если для любого положительного числа a, процессы $X(t)$ и $a^{-H}X(at)$ будут иметь идентичные распределения, т.е. иметь одинаковые статистические свойства для всех положительных целых n:

$$\{X(t_1), X(t_2), \ldots, X(t_n)\} \sim^{D} \{a^{-H}X(at_1), a^{-H}X(at_2), \ldots, a^{-H}X(at_n)\}. \quad (4.1)$$

Отношение $\sim^{D}$ обозначает асимптотическое равенство в смысле распределения. Практически статистическая самоподобность подразумевает, что выполняются следующие условия [37]:

– среднее $E[X(t)] = \dfrac{E[X(at)]}{a^{H}}$; (4.2)

– дисперсия $Var[X(t)] = \dfrac{Var[X(at)]}{a^{2H}}$; (4.3)

– автокорреляция $R(t,\tau) = \dfrac{R(at, a\tau)}{a^{2H}}$. (4.4)

H – параметр Херста (Hurst), показывает «степень» самоподобности. Значение H = 0,5 показывает отсутствие самоподобности, а большие значения H (близкие к 1) показывают большую степень самоподобности или длительной зависимости (long-range dependence, LRD) в процессе. Это обозначает, что если LRD-процесс имеет тенденцию к увеличению (или уменьшению) в прошлом, то с большой вероятностью он будет иметь тенденцию к увеличению (или уменьшению) в будущем.

Дискретный самоподобный процесс

Рассмотрим временной процесс $X = \{X_n, n \in Z^+\}$ и определим другой временной процесс путем усреднения оригинального временного процесса на непересекающиеся соседствующие блоки длиной m как

$$X_n^{(m)} = \frac{1}{m} \sum_{i=mn-(m-1)}^{mn} X_i \quad . \qquad (4.5)$$

$X^{(1)}$ представляет в этом случае наибольшее возможное для процесса разрешение. Последующие эволюции процесса $X^{(m)}$ могут быть получены путем m - усреднения процесса X_n, например:

$$X_n^{(4)} = \frac{X_{4n3} + X_{4n2} + X_{4n1} + X_{4n}}{4}.$$

Процесс $X^{(m)}$ представляет собой менее детализированную копию процесса $X^{(1)}$. В случае, если статистические свойства (среднее, дисперсия) сохраняются при усреднении, тогда процесс является самоподобным.

Существует два класса самоподобных процессов, так называемые точно самоподобные и асимптотически самоподобные процессы. Процесс X называется точно самоподобным с параметром β ($0<\beta<1$) если для $m \in Z^+$ выполняются следующие условия:

– дисперсия определяется следующим образом:

$$Var[X^{(m)}] = \frac{Var[X]}{m^{\beta}}; \quad (4.6)$$

– функция автокорреляции

$$R(k, X^{(m)}) = R(k, X). \quad (4.7)$$

Параметр β связан с параметром Херста H соотношением

$$\beta = 2(1 - H). \quad (4.8)$$

Процесс X называется асимптотически самоподобным, если для больших k дисперсия определяется как $Var[X^{(m)}] = \frac{Var[X]}{m^{\beta}}$, а функция автокорреляции $R(k, X^{(m)}) \to R(k, X)$ при $m \to \infty$.

Имеются наблюдения, что для обоих классов самоподобных процессов дисперсия $Var[X^{(m)}]$ уменьшается намного медленнее, чем $1/m$ при $m \to \infty$ по сравнению со стохастическими процессами, где дисперсия уменьшается пропорционально $1/m$ и приближается к 0 при $m \to \infty$.

Наиболее точным свойством самоподобных процессов является то, что функция автокорреляции не вырождается при $m \to \infty$, в отличие от стохастических процессов, где $R(k, X) \to 0$ при $m \to \infty$.

Математическая модель самоподобного сетевого потока

Существует несколько подходов в формировании самоподобного потока. Наиболее известным является метод, первоначально предложенный Мандельбротом [28]. Данный метод основан на суперпозиции нескольких (строго чередующихся) независимых и имеющих одинаковое распределение *ON/OFF* - источников, интервалы между *ON*- и *OFF*-периодами которого обладают эффектом Ноя (Noah effect) [30].

Под строго чередующимися *ON/OFF*-источниками подразумевается модель, где *ON*- и *OFF*-периоды строго чередуются, длительности *ON*-периодов независимы и имеют одинаковое распределение, длительности *OFF*-периодов тоже независимы и имеют одинаковое распределение, и последовательности длительностей *ON*- и *OFF*-периодов не зависят друг от друга. При этом длительности *ON*- и *OFF*-периодов могут иметь разные распределения.

Причем, именно эффект Ноя в распределении длительностей *ON/OFF*-периодов является основной точкой при моделировании самоподобного трафика в отличие от моделей, когда используются стандартные экспоненциальное или геометрическое распределение. Эффект Ноя является синонимом синдрома бесконечной дисперсии, появившейся благодаря эмпирическим наблюдениям того, что многие природные явления могут быть описаны распределением с бесконечной дисперсией.

Математически для достижения эффекта Ноя можно использовать распределение Парето или логарифмически-нормальное распределение. Наиболее популярным является распределение Парето [31]. Распределение Парето имеет функцию распределения

$$F(x) = 1 - \left(\frac{в}{x}\right)^{\alpha}, \tag{4.9}$$

где α – параметр формы, характеризующий, будет ли распределение иметь конечное или бесконечное среднее и дисперсию;
$в$ – параметр нижней границы (минимальное значение случайной величины x).

Плотность распределения Парето задается функцией

$$f(x) = \begin{cases} \dfrac{\alpha}{в}\left(\dfrac{в}{x}\right)^{\alpha+1}, & x > \alpha, \alpha > 0 \\ 0, & x \le \alpha \end{cases}. \tag{4.10}$$

Параметр α определяет среднее и дисперсию x следующим образом:

– для $\alpha \le 1$ распределение имеет бесконечное среднее;

– для $1 \le \alpha \le 2$ распределение имеет конечное среднее и бесконечную дисперсию;

– для $\alpha \le 2$ распределение имеет бесконечную дисперсию.

Существует отношение между параметром α и параметром Херста:

$$H = (3-\alpha) / 2. \tag{4.11}$$

Методы вычисления параметра Херста

Метод агрегирования дисперсии. Множество состояний системы $X = \{X_i, i \geq 1\}$ делится на подмножества мощностью m, и подсчитывается среднее значение для каждого из подмножеств:

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i \quad , \tag{4.12}$$

где k – номер подмножества.

После этого для каждого из подмножеств вычисляется дисперсия, с помощью которой можно оценить значение $Var[X^{(m)}]$.

Поскольку $Var[X^{(m)}] \sim \sigma_0^2 m^{\beta}$ при $m \to \infty$, где $\beta = 2H - 2 < 0$, то выполняя следующие действия можно вычислить параметры β и H.

Пусть множество X содержит N элементов. Тогда множество делится на m подмножеств. Для каждого из подмножеств вычисляется среднее $X^{(m)}(k)$ и дисперсия выборки:

$$\overline{Var[X^{(m)}]} = \frac{1}{N/m} \sum_{k=1}^{N/m} (X^{(m)}(k))^2 \quad - \left(\frac{1}{N/m} \sum_{k=1}^{N/m} X^{(m)}(k) \right)^2 . \tag{4.13}$$

Данные вычисления проводятся для нескольких значений m, после чего строится график соответствия значений логарифмов дисперсии выборки и числа m. На графике выбираются такие значения m, чтобы они были равноудалены друг от друга, т.е. $\frac{m_{i+1}}{m_i} = C$, где C - некоторая константа, которая зависит от размера выборки и желаемого количества точек.

$\overline{Var[X^{(m)}]}$ является оценкой дисперсии $Var[X^{(m)}]$. Выбранные точки должны лежать на прямой с углом наклона $\beta = 2H - 2; -1 \leq \beta < 0$. Если значения X независимы, то угол наклона прямой будет равен $\beta = -1$.

R/S-статистика. Для $X = \{X_i, i \geq 1\}$ с частной суммой $Y(n) = \sum_{i=1}^{n} X_i$ и дисперсией выборки

$$S^2(n) = \frac{1}{n} \sum_{i=1}^{n} X_i \quad - \left(\frac{1}{n}\right)^2 Y(n)^2, \tag{4.14}$$

R/S-статистика вычисляется как

$$\frac{R}{S}(n) = \frac{1}{S(n)} \left[\max_{0 \leq t \leq n} \left(Y(t) \quad \frac{t}{n} Y(n) \right) - \min_{0 \leq t \leq n} \left(Y(t) \quad \frac{t}{n} Y(n) \right) \right]. \tag{4.15}$$

При $n \to \infty$ справедливо соотношение

$$E\left(\frac{R}{S(n)} \right) \sim C_H n^H, \tag{4.16}$$

где C_n – некоторая константа, независящая от n.

4.2. Моделирование распределенной атаки

Математическая модель DDoS-атаки

Моделированию посвящены работы [4, 5, 14-16, 18, 21]. Положим, что атакуемый узел имеет входящий канал с пропускной способностью C бит/с, а пограничный маршрутизатор – входной буфер размером B бит. Ситуация атаки может быть симулирована с помощью модели статистического мультиплексирования трафика от N атакующих узлов, которые могут находиться в двух состояниях: отсылки пакетов (*ON*-состояние) и бездействия (*OFF*-состояние).

Обозначим периоды времени (в секундах) функционирования и бездействия как $t_1 = \beta^{-1}$ и $t_2 = \alpha^{-1}$ соответственно. Если источник является активным (*ON*-состояние), то он генерирует r пакетов в секунду. Размер посылаемого пакета в битах обозначим как L, а объем полученных пакетов в момент времени t как $Q(t)$.

Тогда вероятность перегрузки буфера может быть аппроксимирована формулой [32]:

$$P(Q(t) > B) = e^{-\gamma B}, \tag{4.17}$$

где

$$\gamma = \left(\frac{\beta}{NrL} - \frac{\alpha}{C} \right) N. \tag{4.18}$$

Чтобы вызвать перегрузку канала передачи, атакующему необходимо выбрать такие параметры атаки, чтобы значение γ было близко к нулю или отрицательное. Поэтому число узлов для осуществления атаки должно удовлетворять неравенству

$$N \geq \frac{C(\beta + \alpha)}{\alpha r L}. \tag{4.19}$$

Назовём коэффициентом занятости источника трафика τ отношение продолжительности отсылки пакетов источником ко всему периоду функционирования и бездействия

$$\tau = \frac{\beta^{-1}}{\alpha^{-1} + \beta^{-1}}. \tag{4.20}$$

Тогда неравенство, ограничивающее количество атакующих узлов, можно выразить как $N \geq \frac{C}{r\tau L}$. В типичных *ICMP-flood* атаках атакующие узлы постоянно находятся в *ON*-состоянии, направляя паразитный трафик жертве. В этом случае коэффициент τ = 1 и неравенство для количества атакующих узлов упрощается до $N \geq \frac{C}{rL}$.

Чтобы атака стала трудно распознаваема, злоумышленник маскирует ее под обычную перегрузку в сети. Для этого ему нужно подобрать довольно малые значения для параметров τ и r. Так, если злоумышленник выберет значения τ = 0,05 и r = 20 пак/с, а целью атаки является сервер с

каналом пропускной способностью C = 10 Мбит/с, то понадобится $1{,}7 \cdot 10^4$ атакующих узлов для проведения атаки.

Кроме того, злоумышленник может выполнять несколько процессов на каждом из узлов атаки, а те, используя фиктивные адреса отправителей, смогут выступать для атакуемого в качестве различных источников трафика. Таким образом, осуществить распределенную *DoS*-атаку, маскируя ее под естественные перегрузки канала, достаточно реально.

Сделанные выводы можно обобщить и на различные типы источников атаки. Пусть у атакующего имеется M различных типов атакующих узлов. Тогда атакующий должен так подобрать параметры атаки, чтобы выполнялось соотношение

$$\sum_{m=1}^{M} N_m r_m L_m \tau_m \geq C . \qquad (4.21)$$

Имея в своём распоряжении достаточное число N подконтрольных узлов, злоумышленник может генерировать трафик, не вызывающий подозрений и аналогичный трафику от обычного пользователя. На основании некоторой собранной статистики о динамике состояния сети (например, уровне потерь пакетов) – процессе $\{X_t\}$ – можно дать оценку будущим состояниям изучаемой системы.

Описание моделирующей системы

Одним из распространенных методов проведения DDoS-атаки является перегрузка входящего канала жертвы [24]. Атакующие узлы пытаются загрузить всю полосу пропускания канала пакетами с данными, не несущими никакого смысла, с подставным адресом отправителя. Примером такой атаки служит *ICMP-flood* атака [25].

Наша модель будет описывать ИТКС, состоящую из N вычислительных машин (ВМ) - источников *TCP*-трафика, соединенных с маршрутизатором (B), который в свою очередь соединен с сервером – целью атаки.

На рис. 4.1 представлена топология моделируемой системы. Каждая ВМ работают по протоколу *TCP NewReno* (данная модификация использу-

ется в большинстве систем). Посылаемый трафик от источников будут генерироваться *FTP*-приложениями. В модели $C_0 = 1$ Мбит/с; отсылка пакетов (*ON*-состояние) - $t_1 = \beta^1 = 5$ мс, бездействие (*OFF*-состояние) – $t_2 = \alpha^{-1} = 50$ мс; каждый из атакующих узлов соединен с маршрутизатором каналом $C_i = 10$ Мбит/с ($i = 1..N$).

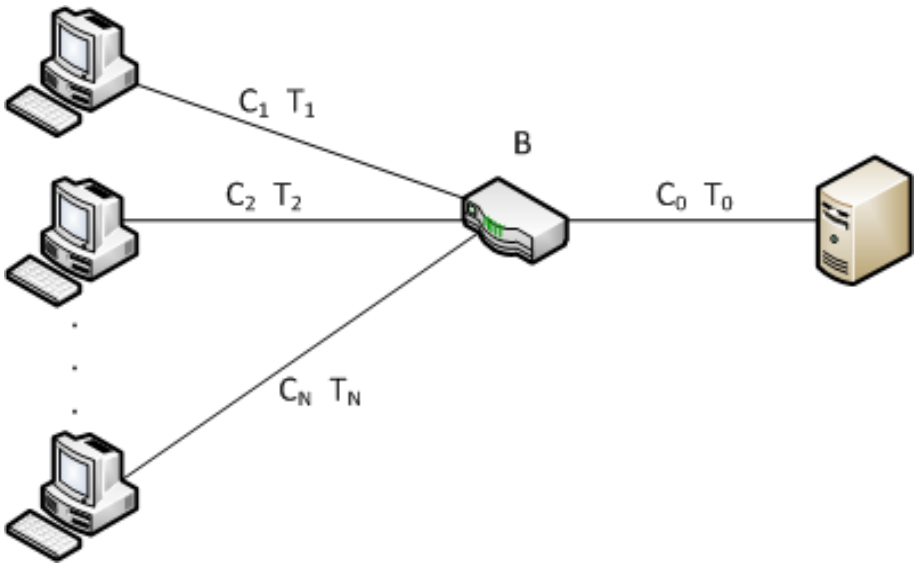


Рис. 4.1. Топология моделируемой ИТКС

Размер передаваемого файла 10 Кбайт (типичный для трафика *FTP*-приложений [10, 11]), значение параметра распределения равно 1,2.

Для моделирования рассматриваемой сети удобно воспользоваться существующим ПО. Наиболее подходящей программой здесь является Network Simulator [35], функционирующая под управлением ОС Linux, поскольку предоставляет возможность провести имитационное моделирование поведения транспортных и прикладных протоколов.

Программа предоставляет описание узлов сети, задание их характеристик, характеристик каналов передачи, связывающих узлы между собой. ПО Network Simulator (NS) реализует все модификации протокола TCP (Vegas, Tahoe, Reno, NewReno), а также протокол UDP. С помощью скриптового языка TCL можно разработать свой протокол сетевого взаимодействия.

Для исследования поведения ИТКС в условиях DDoS-атаки был разработан пакет прикладных программ, включающий:

– скрипт для приложения NS, описывающую моделируемую сеть;

– программу для составления статистики по потерянным пакетам из файла трассировки, созданного программой NS;

– программу вычисления размерности Минковского фрактального объекта.

Результаты моделирования

Моделирование производилось в течение 2000 секунд. Процессы начинает передачу в моменты времени t_1 = 0,072 с и t_2 = 0,0669 с соответственно. Синтезированный трафик, показывающий загрузку канала передачи между сервером и маршрутизатором двумя потоками *TCP*-соединений, изображен на рис. 4.2.

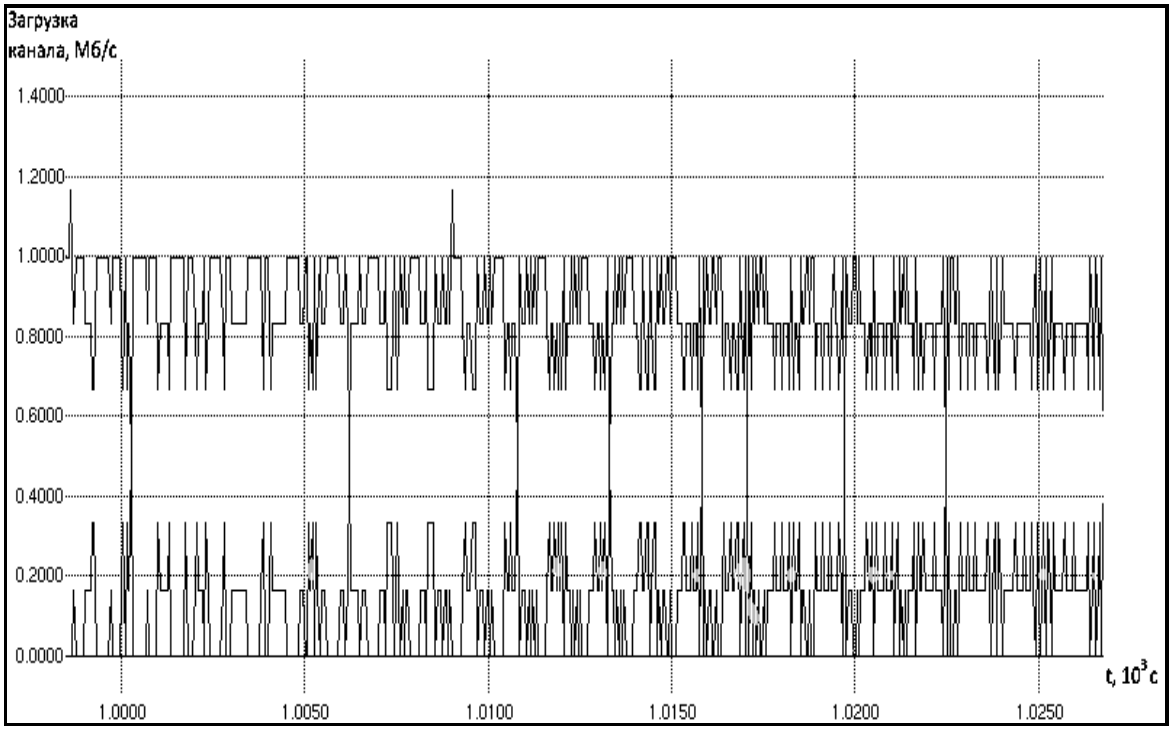


Рис. 4.2. Загрузка канала передачи между сервером и маршрутизатором двумя потоками TCP-соединений

На рис. 4.3 изображена динамика изменения значения окон контроля перегрузки каждого из TCP-соединений. Как видно из графика, в течение 300 с окна перегрузки TCP-соединений адаптировались к предоставленному каналу передачи, после чего перешли в более-менее стабильную фазу передачи.

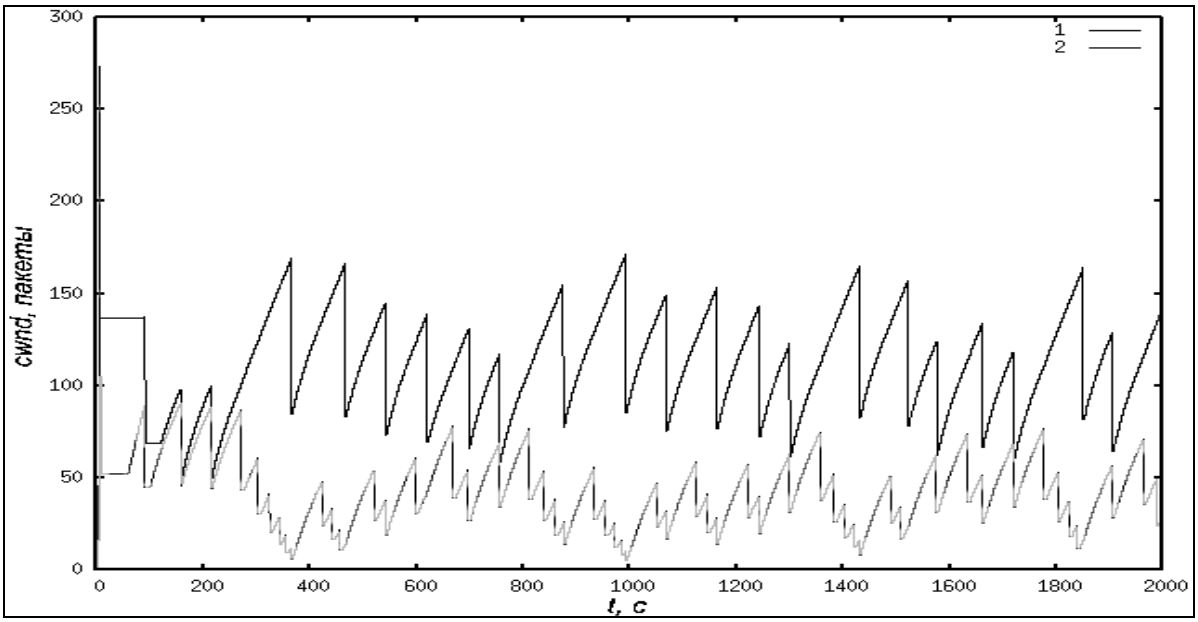


Рис. 4.3. Изменение величины окон перегрузки TCP-соединений

На графике видно, что колебания 1-ого и 2-ого соединений следуют одному и тому же сценарию: подъем – алгоритм избегания затора, после следует потеря пакета и возврат – снижение окна контроля перегрузки.

Рассмотрим подробнее один из интервалов (рис. 4.4).

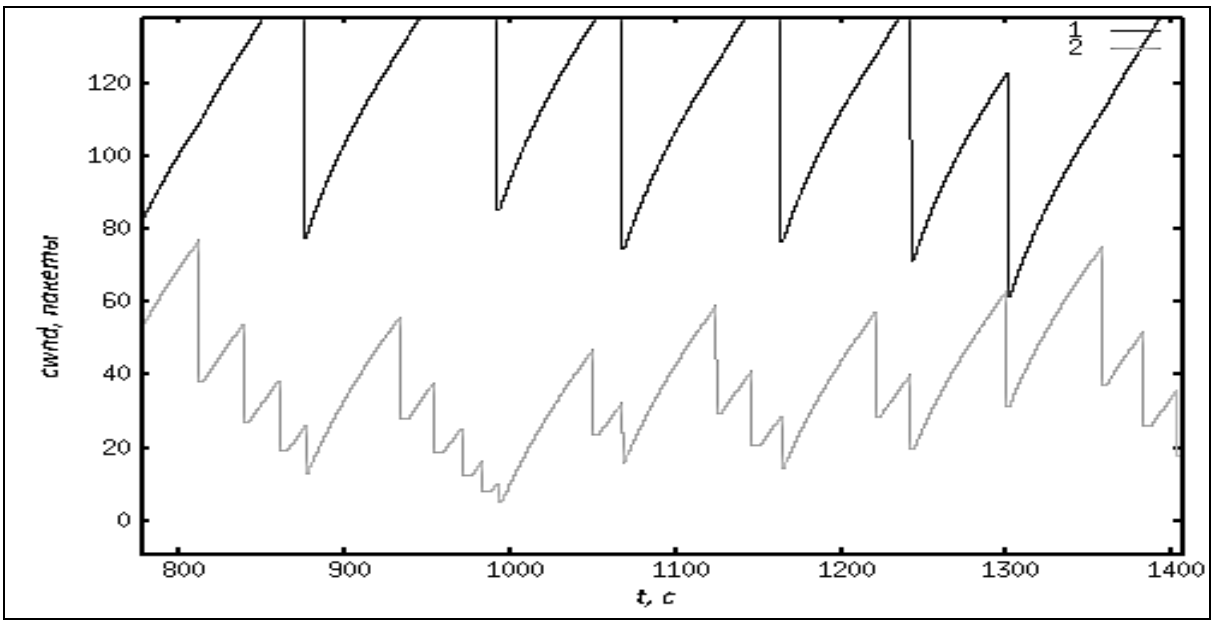


Рис. 4.4. Интервал графика окон перегрузки с 800 по 1400 с

В момент времени t_1 = 932,17 с вторая ВМ обнаруживает потерю пакета при передаче, в это время окно контроля перегрузки имело значение *cwnd* = 55,9293. В соответствии со спецификацией протокола *TCP* версии *New Reno* после детектирования потери пакета передающая сторона должна уменьшить порог разгона *ssthresh* и окно *cwnd*в два раза, что и наблюдается на графике (значение *cwnd* падает с 55,9293 до 27,964).

После этого передача возобновляется сразу в режиме избегания затора: клиент работает в режиме быстрого наверстывания. При этом значение *cwnd* возрастет на 1 за время полного оборота пакета *RTT*, но к 952-ой секунде происходит очередная потеря пакета, тем самым начиная процесс быстрого наверстывания заново.

Параметры моделируемой системы можно рассматривать не только в зависимости от времени. Построим траекторию системы в фазовом пространстве. Фазовое пространство – многомерное пространство, где каждое из измерений представляет собой значение одной из системных переменных. Таким образом, каждая точка в этом пространстве представляет уникальное состояние рассматриваемой системы.

Изобразив эволюцию системы в данном пространстве, получаем (так как система детерминирована), что если система возвратится в одну из уже построенных точек, то в дальнейшем она повторит уже отображенное множество состояний, и в итоге получится замкнутый цикл. Получается, что если система периодична, то ее траекторией является замкнутый цикл, и наоборот.

Изучаемая система может иметь достаточно большое количество переменных: количество переданных пакетов, значения порогов разгона каждого из соединений, значения окон приема и контроля перегрузки и др.

Однако если траектория системы в фазовом пространстве замкнута, то также замкнута будет ее проекция в пространстве с другим количеством системных переменных.

Построим проекцию фазового пространства на плоскость, где каждому из измерений будет соответствовать размер окна перегрузки одного из *TCP*-соединений (рис. 4.5).

При построении изображения не учитывался период до 300-ой секунды (т.н. переходный процесс [37]). Размерность изображенного на рис. 4.5 объекта дробная. Чтобы определить ее значение воспользуемся клеточным методом.

В клеточном методе [22] область, содержащая фрактальный объект, разбивается на клетки (коробки), в двумерном случае это – квадраты, со стороной *L*.

Затем подсчитывается число клеток, необходимых для покрытия всего фрактала. Такой подсчет производится для нескольких размеров стороны клетки.

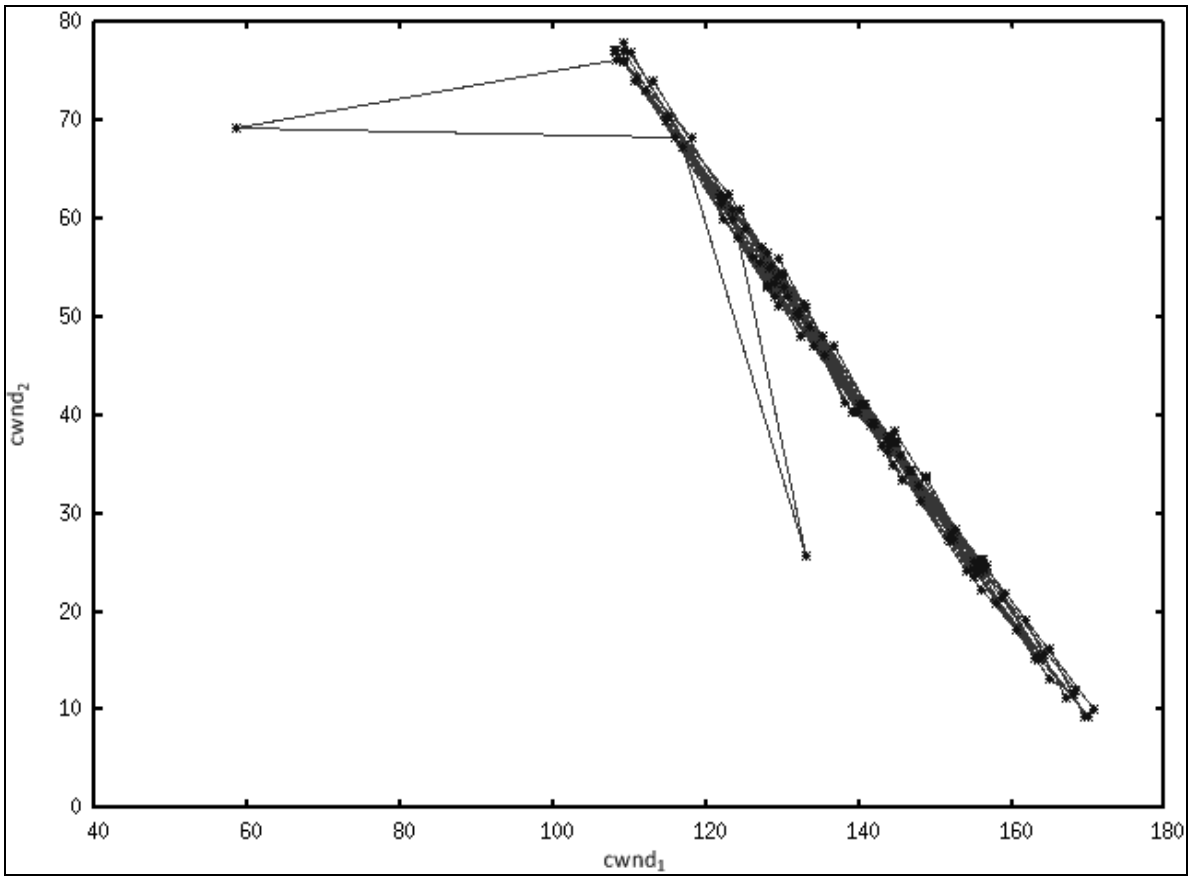


Рис. 4.5. Проекция фазового пространства

Программа вычисления размерности Минковского для нашей системы выдает следующие результаты (рис. 4.6):

0,01 N:135	0,58 N:104	7 N:24	26 N:8
0,04 N:135	0,61 N:107	8 N:19	27 N:7
0,07 N:133	0,64 N:104	9 N:17	28 N:7
0,1 N:128	0,67 N:104	10 N:16	29 N:6
0,13 N:121	0,7 N:102	11 N:15	30 N:6
0,16 N:124	0,73 N:107	12 N:14	31 N:6
0,19 N:119	0,76 N:99	13 N:12	32 N:7
0,22 N:122	0,79 N:99	14 N:13	33 N:6
0,25 N:120	0,82 N:92	15 N:12	34 N:6
0,28 N:115	0,85 N:102	16 N:10	35 N:7
0,31 N:112	0,88 N:97	17 N:10	36 N:6
0,34 N:116	0,91 N:101	18 N:11	37 N:6
0,37 N:116	0,94 N:95	19 N:10	38 N:6
0,4 N:117	0,97 N:94	20 N:8	39 N:5

Рис. 4.6. Фрагмент результатов программы вычисления размерности Минковского

Здесь первое число – длина стороны клетки L, второе – количество клеток, покрывающих фрактал – N. В выводе программы видно, что не всегда клетки с меньшим размером оптимальнее охватывают объект. Это обуславливается тем, что в применяемом алгоритме клетки следуют друг за другом, не выравниваясь относительно границ объекта. Этот момент был бы существенен при единичном измерении размерности.

Полученную ломаную аппроксимируем прямой (на рис. 4.7 она обозначена точками), которая задается уравнением $\log N(L)=\log C - d \log L$, где d – фрактальная размерность, C – некоторая константа.

Отсюда размерность можно вычислить как

$$d = \frac{\log C}{\log L} - \frac{\log N(L)}{\log L}. \qquad (4.22)$$

Так как $\log L \to -\infty$ при $L \to 0_+$, то формулу для вычисления размерно-

сти перепишем как $d = (\log N(L))/\log (1/N)$, и d здесь угол наклона аппроксимирующей прямой. Таким образом, получаем d = 0,8625.

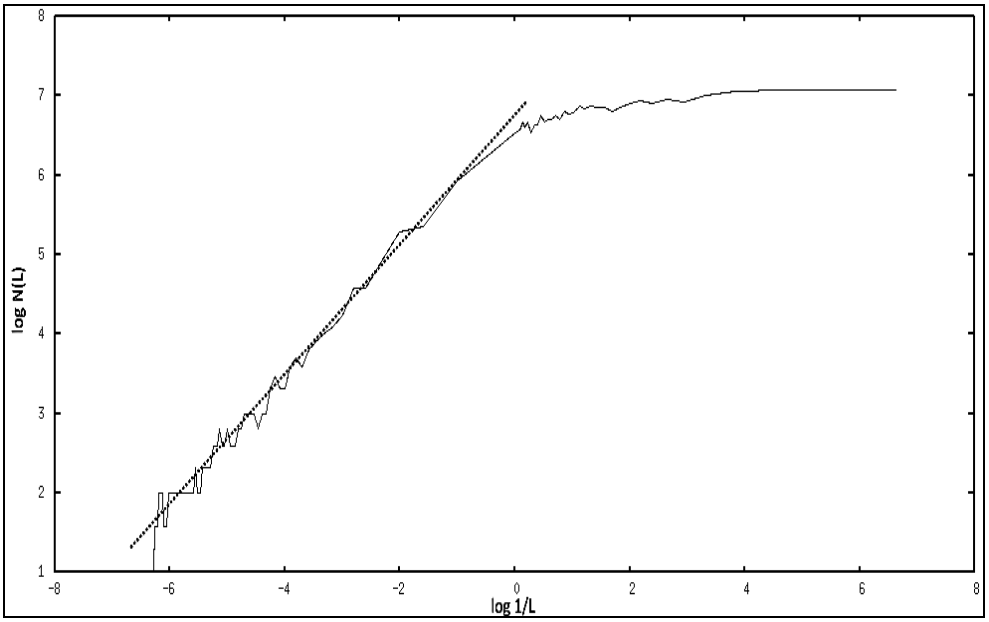


Рис. 4.7. Вычисление размерности Минковского

Чувствительность к атакующему воздействию

Покажем, что рассматриваемая система имеет значительную зависимость от атакующего воздействия, которое в данном случае представляет собой изменение начальных условий. Т.е. даже при незначительном изменении входных параметров уже через некоторое время характеристики измененной системы будут иметь значительное отличие от исходных, что и является одним из главных признаков хаотической системы.

Эксперимент заключался в том, что программным путем было внесено изменение в модель: на 300-ой с окно одного из *TCP-с*оединений было увеличено на 1 пакет.

Динамика изменения окон перегрузки для данного случая приведена на рис. 4.8. До 300-ой с TCP-сессии ведут себя одинаково, значения окон перегрузки совпадают, но после внесения изменений картина меняется. На рис. 4.9 представлена разность значений окон контроля перегрузки для каждой сессии.

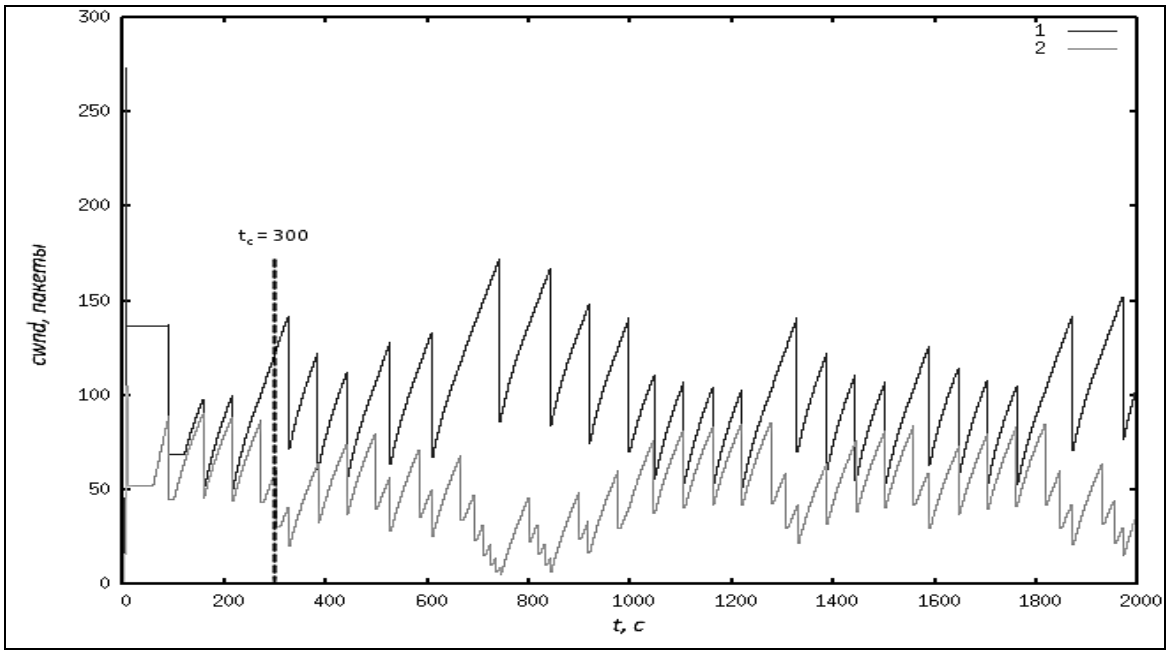


Рис. 4.8. Динамика изменения окон контроля перегрузки с измененными входными данными

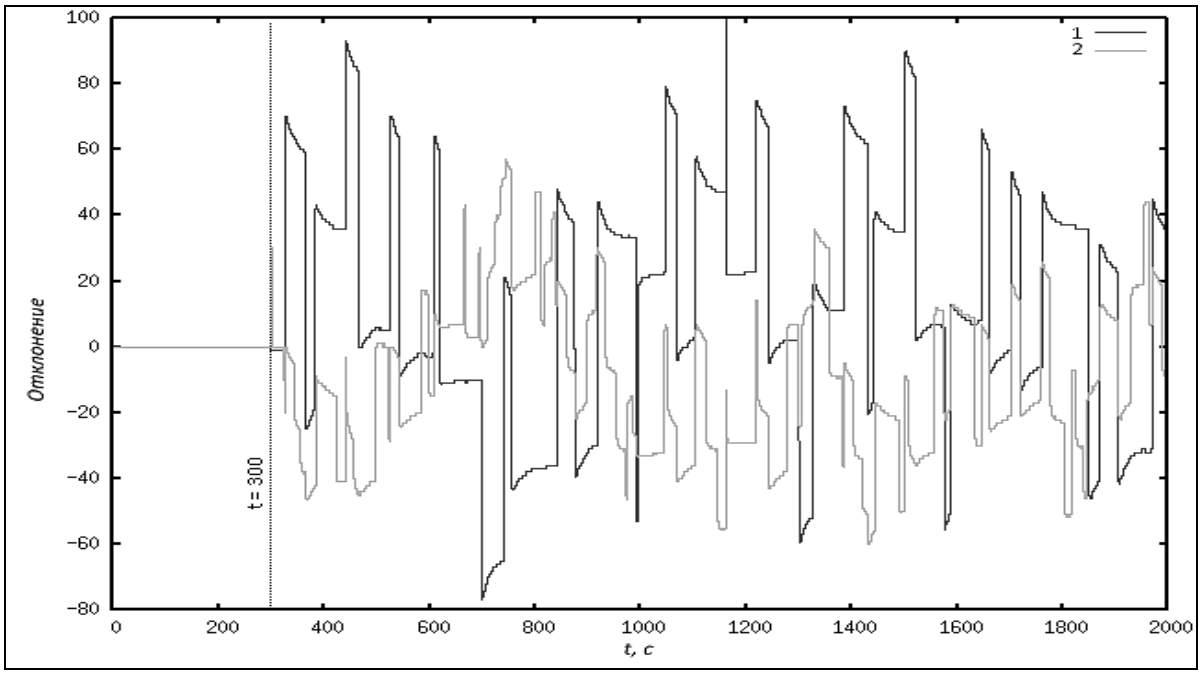


Рис. 4.9. Разность значений окон контроля перегрузки

Покажем, что данная система функционирует в режиме детерминированного хаоса. Для этого необходимо вычислить показатель Ляпунова – степень экспоненциального разбегания траекторий системы. Показатель Ляпунова является коэффициентом растяжения системы

$$\lambda(t_0) = \frac{1}{\Delta t} ln \left| \frac{f^{\Delta t}(t_0 + \xi) - f^{\Delta t}(t_0)}{\xi} \right|, \qquad (4.23)$$

где ξ – расстояние между двумя точками траектории в начальный момент времени t_0;

Δt – время, за которое системы разбегаются на расстояние $E(t_0 + \Delta t) = f^{\Delta t}(t_0 + \xi) - f^{\Delta t}(t_0)$.

Евклидово расстояние вычисляется по следующей формуле

$$E(t) = \sqrt{\sum_{i=1}^{N} (cwnd_i^{(1)}(t) - cwnd_i^{(2)}(t))^2}, \qquad (4.24)$$

где $cwnd_i^{(1)}(t)$ – окно перегрузки 1-го *TCP*-соединения в момент времени t в первом случае;

$cwnd_i^{(2)}(t)$ – окно перегрузки 2-го *TCP*-соединения в момент времени t во второй системе с измененными входными данными.

В нашем эксперименте ξ = 1 и положим Δt = 50. Расстояние между двумя состояниями системы за время Δt = 10

$$E = \sqrt{61,9996^2 + (\ \ 24,4766)^2} = 66,6562.$$

Показатель Ляпунова

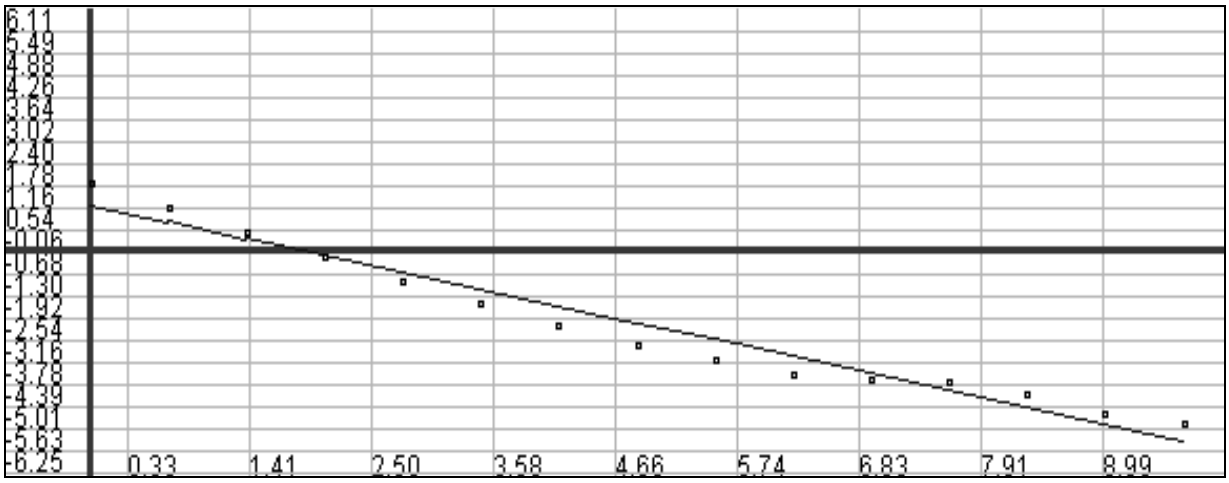
$$\lambda(t_0) = \frac{1}{50} ln\, 66,6562 = 0,08399.$$

Таким образом, различие в двух системах нарастает каждую секунду со скоростью $e^{\lambda} = 1,087$.

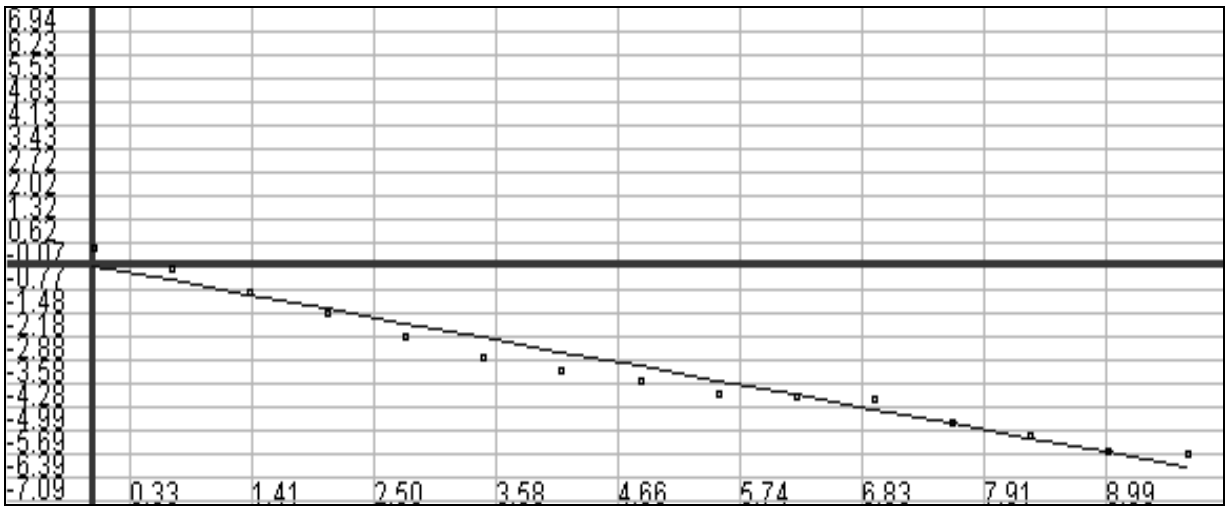
Тот факт, что размерность системы (d = 0,8625) является дробным числом и показатель Ляпунова положителен, позволяет говорить о том, что аттрактор рассматриваемой системы является странным.

Вычислим параметр Херста для нашей системы. Важно получить значение параметра для того, чтобы определить, является ли наблюдаемый процесс персистентным, т.е. имеет память.

Для вычисления параметра Херста были получены значения пропускной способности отдельно для каждого *TCP*-соединения и агрегированного трафика. Вычисление производилось в программе *SELFIS* по двум методам: метод агрегированной дисперсии (*AVM*) и метод *R/S*-статистики (нормированного размаха). Результаты вычислений приведены на рис. 4.10 - 4.12.

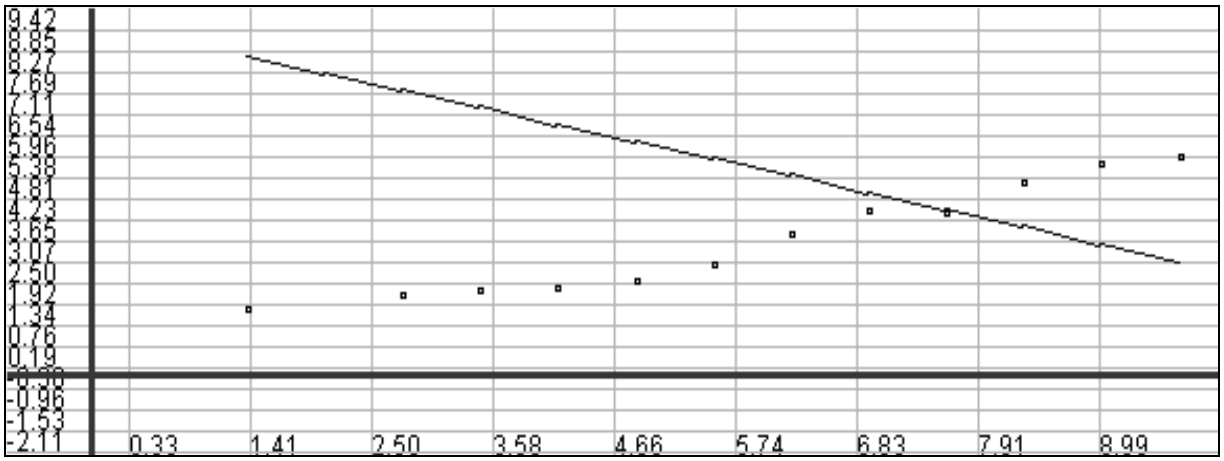


а (*AVM; H* = 0,66)

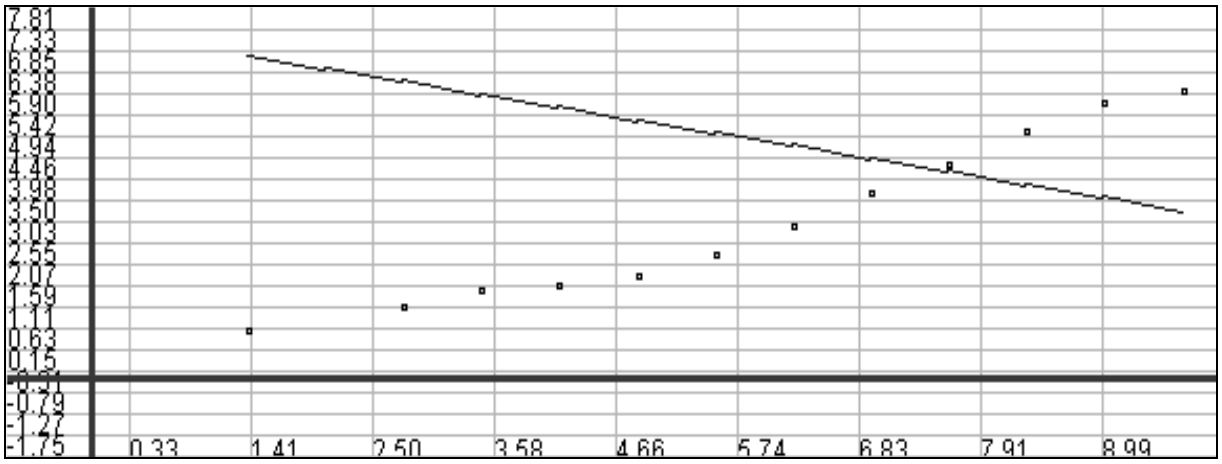


б (*R/S; H* = 0,673)

Рис. 4.10. Определение параметра Херста для 1-го соединения



а (*AVM*; *H* = 0,691)

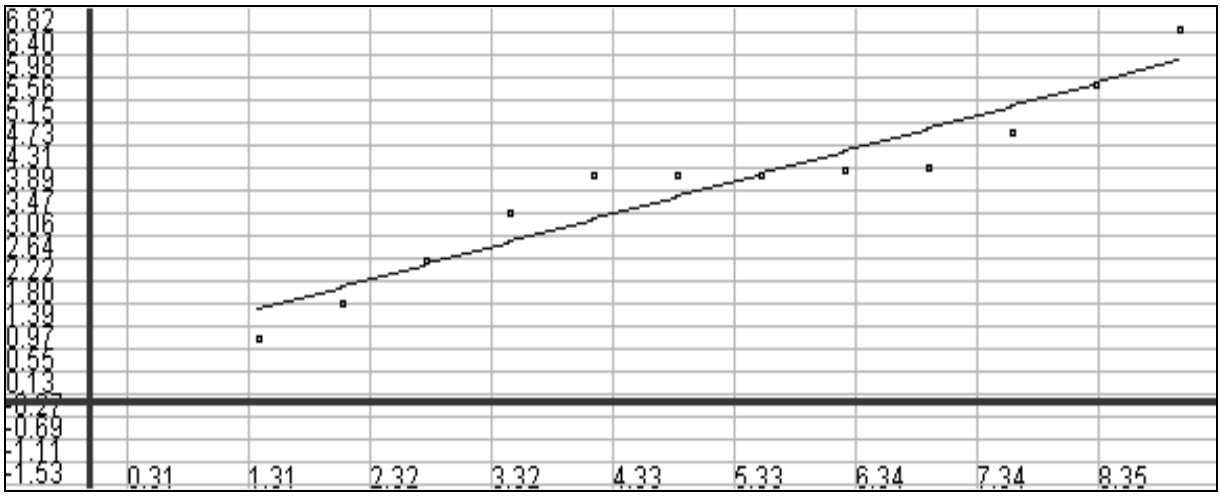


б (*R/S; H* = 0,515)

Рис. 4.11. Определение параметра Херста для 2-го соединения

а (*AVM; H* = 0,888)

Рис. 4.12. Определение параметра Херста для 3-го соединения (начало)



б (*R/S; H* = 0,605)

Рис. 4.12. Определение параметра Херста для 3-го соединения (окончание)

Как видно из рисунков выше для всех рассматриваемых трафиков параметр Херста превышает 0,5, что говорит о том, что система является персистентной и самоподобной. Близость параметра Херста к 1 для агрегированного трафика указывает на то, что система имеет длительную память (*LRD*).

4.3. Алгоритм предсказания *DDoS*-атаки на основе *FARIMA*-модели агрегированного трафика

Моделирование показало, что трафик в условиях *DoS*-атаки обладает свойствами персистентности и самоподобия, и, как следствие, имеет медленно убывающую зависимость. Важнейшим параметром, характеризующим степень самоподобия, является параметр Херста.

Дополнительно следует отметить, что самоподобный процесс часто носит взрывной (*burst*) характер, что выражается в возможности наличия выбросов во время относительно низкой скорости поступления событий. Применительно к трафику самоподобие выражается в неизменности поведения при изменении временных масштабов наблюдения и сохранения склонности к всплескам при усреднении по шкале времени.

Близость параметра Херста к 1 для агрегированного трафика, указывает на то, что трафик в ИТКС, находящейся в условиях перегрузок, имеет длительную зависимость от начальных условий (*LRD*). Такие системы могут быть с достаточной точностью описаны и предсказаны *FARIMA*-

моделью (смешанная модель авторегрессии и скользящего среднего) [12, 17, 20, 39].

FARIMA (*p, d, q*)-процесс – это процесс $\{X_t,\ t = \ldots -1, 0, 1, \ldots\}$, где параметры p и q – неотрицательные целые числа, d лежит в пределах $-0{,}5 < d < 0{,}5$. Процесс описывается разностным уравнением

$$\Phi(B)\Delta^d X_t = \Theta(B)a_t, \qquad (4.24)$$

где

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p,$$

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q;$$

B – оператор взятия предыдущего значения в множестве $\{X_t\}$, $\Delta^d = (1-B)^d$ – оператор дробной разности и $\{a_t,\ t = \ldots -1, 0, 1, \ldots\}$ – белый шум с дисперсией σ^2.

В нашем случае процесс $\{X_t\}$ представляет собой динамику изменения потерь пакетов в ИТКС. Если полиномы $\Phi(z)$ и $\Theta(z)$ не имеют общих корней, и все корни лежат вне единичной окружности, то процесс $\{X_t\}$ является стационарным и обратимым.

Если параметр d удовлетворяет неравенству $0 < d < 0{,}5$, то описываемая системы имеет длительную зависимость с параметром Херста $H = d + 0{,}5$.

Применение модели *FARIMA* для процесса является более трудоёмким, чем его моделирование с помощью *ARMA*-процесса [29]. В то же время обсчёт *FARIMA* - процесса можно разбить на две составляющие – генерацию дробно-разностного шума *FARIMA(0,d,0)* и моделирование *ARMA(p,q)*-составляющей.

Для *FARIMA*-процесса мы можем записать

$$W_t = \Delta^d X_t, \qquad (4.25)$$

где $W_t = \dfrac{\Theta(B)}{\Phi(B)} a_t$.

Тогда, если можно определить величину степени различия d и значения для оператора дробной размерности, то $\{W_t\}$ – дробно-разностная составляющая $\{X_t\}$ – может быть представлена *ARMA*-моделью.

Как показано в [29], показатель степени различия и параметр Херста для *LRD*-процессов связаны соотношением $H = d + 0{,}5$. Таким образом, вычислив параметр Херста для рассматриваемой системы, применяя *R/S*-статистику или какой-либо другой метод, можно найти и показатель степени различия d.

Вычисление $\{W_t\}$ из $\{X_t\}$ требует конечного числа шагов аппроксимации

$$\Delta^d (X_t - u) = \sum_{k=0}^{\infty} (-B)^k (X_t - u) = \sum_{j=0}^{\infty} \overline{\omega_j} (X_{t-j} - u), \qquad (4.26)$$

где $\overline{\omega_j} = \dfrac{(-1)^j \Gamma(1+d)}{\Gamma(1+j)\Gamma(1+d-j)}$.

Как видно, формула вычисления значения оператора дробной размерности включает в себя значения X_0, X_{-1}, X_{-2}, которые не могут быть замерены системой обнаружения (мониторинга) в ИТКС. Вследствие этого их необходимо заменить средним значением u.

Теперь становится возможным вычислить значения параметров p и q для *ARMA*-модели, проанализировав составляющую $\{W_t\}$.

Это можно сделать, применив информационный критерий Акаике (*Akaike Information Criterion*) [29] или информационный критерий Байеса (*Bayesian Information Criterion*). Появляется возможность вычислить все параметры $\phi_1, \phi_2, \ldots, \phi_p$ и $\theta_1, \theta_2, \ldots \theta_q$, а также σ^2. Так как $\{a_t,\ t = \ldots -1,\ 0,\ 1,\ \ldots\}$ – белый шум с постоянной дисперсией σ^2, то теперь, вычислив все параметры *FARIMA*-модели, мы можем предсказать значение X_t в момент времени t согласно следующим формулам:

$$X_t = \Phi^{-1}(B) \cdot \Theta(B) Y_t, \qquad (4.27)$$

где $Y_t = \Delta^d a_t$.

Таким образом, моделирование процессов, происходящих в ИТКС, с помощью *FARIMA*-модели, учитывая их самоподобную природу и долго-

временную зависимость от начальных условий, предоставляет возможность оценить будущее состояние системы на основе её текущих параметров.

Система обнаружения атак на основе анализа текущей динамики изменения уровня потери пакетов в ИТКС, построенная с применением *FARIMA*-модели (в качестве изучаемого процесса которой выступает изменение уровня потерь), может оценить угрозу реализации *DDoS*-атаки в сети. В данном случае система обнаружения атак по росту предсказанного значения X_t может сделать вывод о проведении в сети распределённой *DoS*-атаки.

4.4. Пример внедрения системы раннего обнаружения аномалий в АСУ ОАО ВЗ «Электроприбор»

ОАО Владимирский завод «Электроприбор» – ведущее предприятие по производству средств радиосвязи. На заводе задачи АСУ поддерживаются корпоративной информационно-телекоммуникационной системой, которая носит территориально распределенный характер и объединяет все структурные подразделения завода. Развитие АСУ, переход на электронный документооборот, производство новых изделий приводят к постоянному росту конфиденциальной информации, хранящейся и обрабатываемой в ИТКС завода. В последнее время подразделением информационной безопасности (ИБ) предприятия фиксировался устойчивый рост инцидентов ИБ. Только за 2013 год было зарегистрировано более тысячи атак из внешней сети и более 3000 попыток нарушения ИБ внутри ИТКС.

Негативное влияние на уровень защищенности ИТКС предприятия оказывает ее распределенность на площади в несколько квадратных километров и постоянный процесс модернизации, наличие подразделений за пределами территории, хранение информации различной степени конфиденциальности, большое количество специализированного программного обеспечения. Сюда также отнесем недостаточность финансирования IT-подразделения. Все эти факторы в итоге приводят к значительному возрастанию рисков безопасности.

Предложенное размещение механизмов раннего обнаружения атак [1, 13, 19] в ИТКС приведено на рис. 4.13.

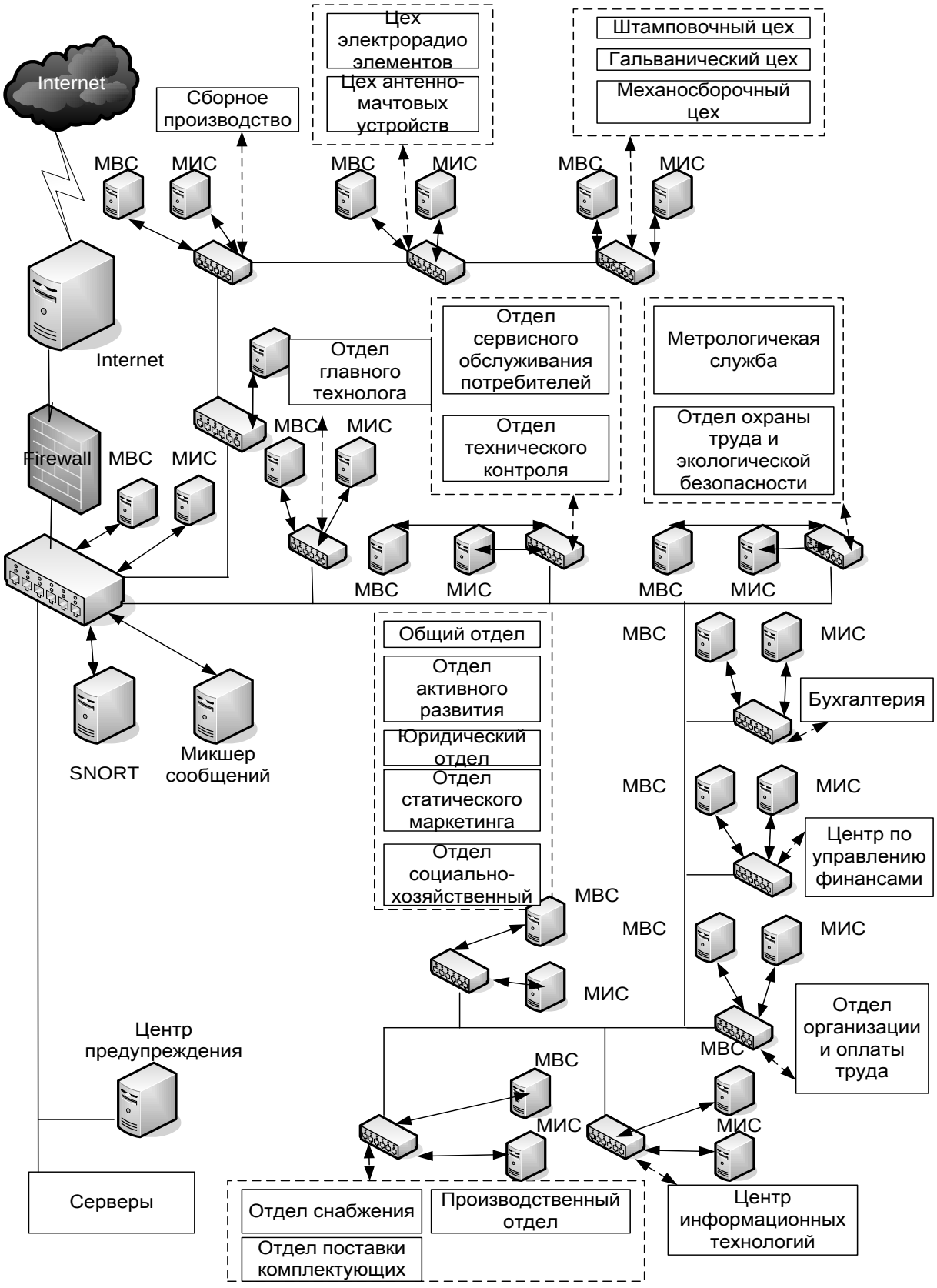
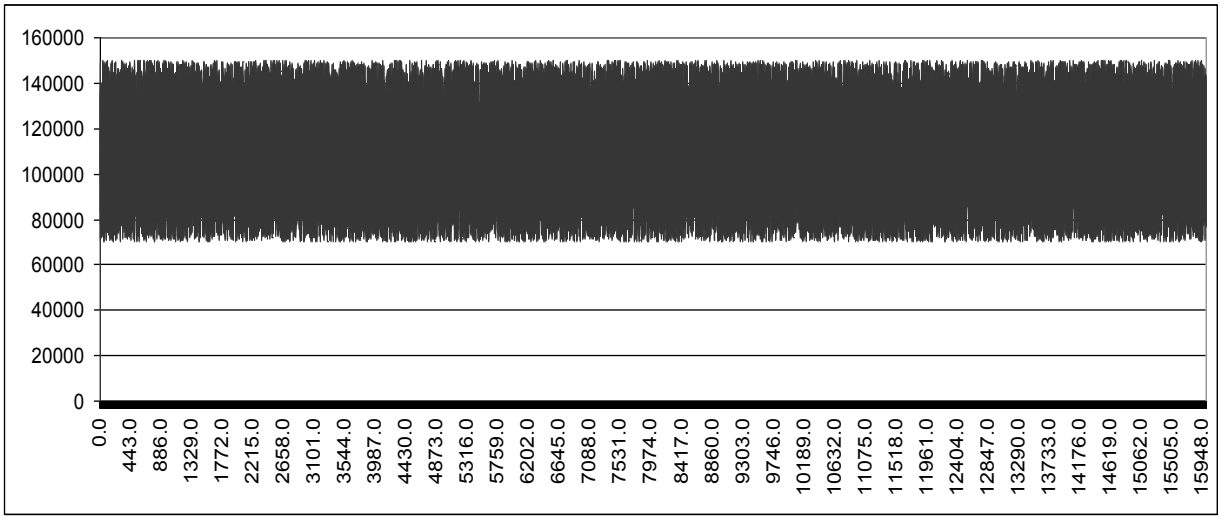


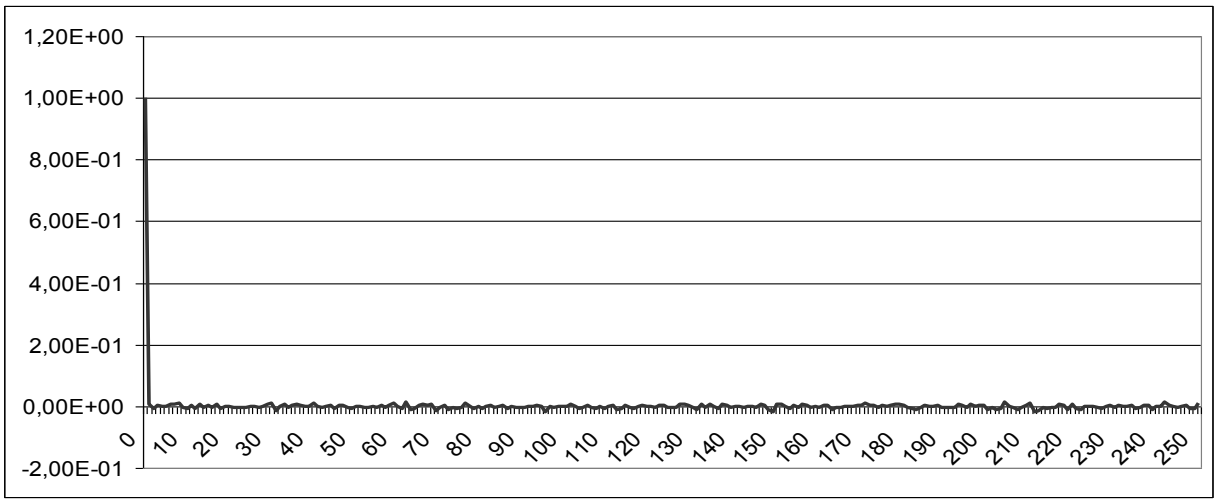
Рис. 4.13. Размещение механизмов раннего обнаружения атак

В сети была синтезирована распределенная атака. Далее показан факт ее обнаружения штатными и предложенными средствами.

1) Система без атаки (7 часов, 8 921 422 пакетов). Данная реализация показывает поведение сети в отсутствие вредоносных воздействий. На графике, представленном на рис. 4.14, а, показывающем агрегацию, мы наблюдаем отсутствие взрывного характера протекания процесса и (относительно) нормальное распределение размеров пакетов. Автокорреляционная функция (АКФ) эквидистантного процесса, показанного выше, резко убывает, и на протяжении всей выборки остаётся на очень низком уровне, что демонстрирует отсутствие зависимости между замеренными длинами пакетов (рис. 4.14, б).



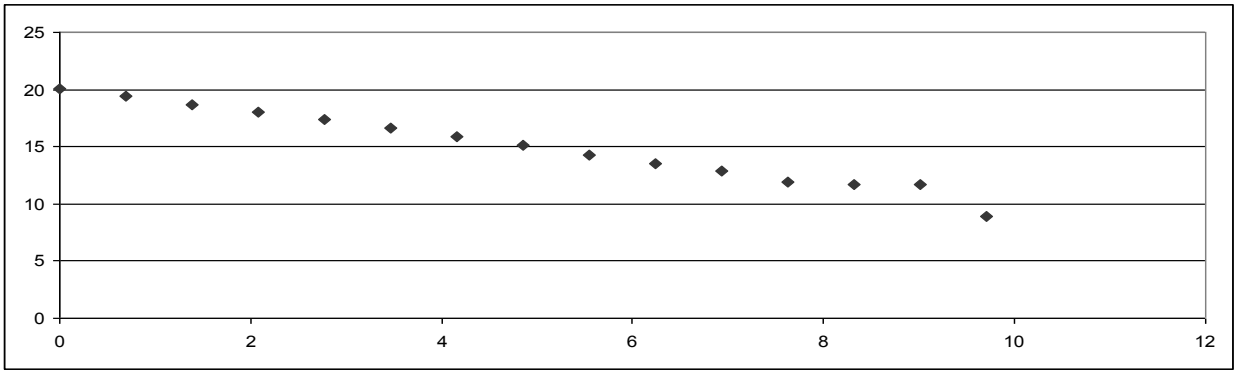
а) агрегация длин пакетов по времени 0,5 с



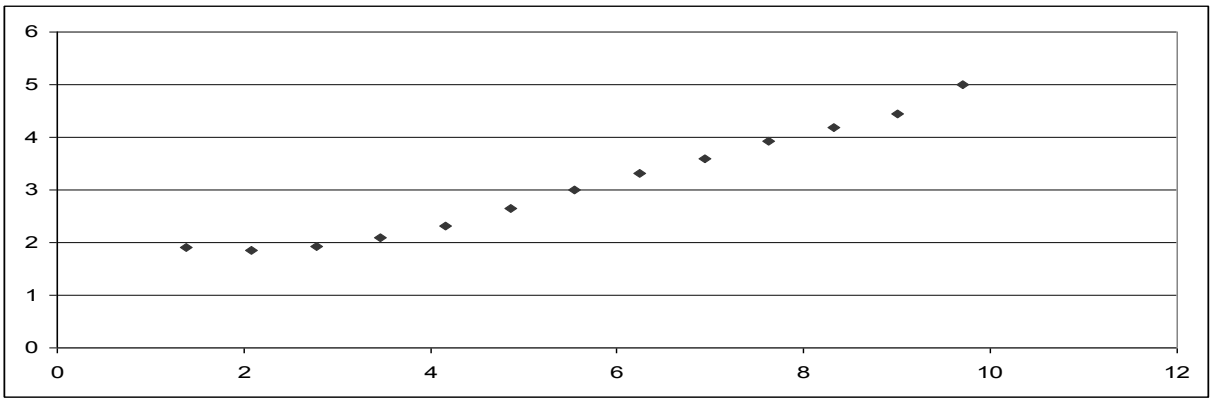
б) АКФ трафика, k_{max} = 250

Рис. 4.14. Система без атаки

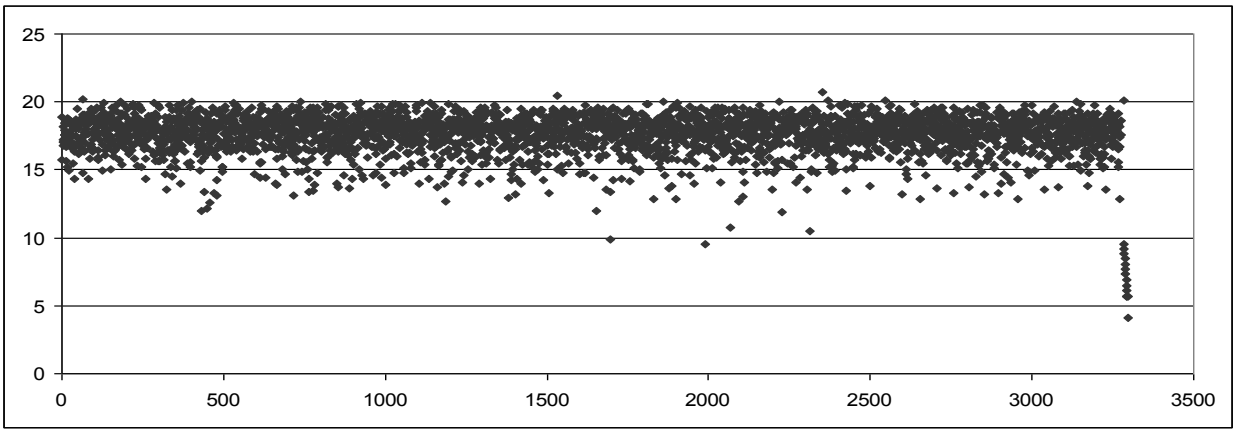
Оценка параметра Хёрста тремя методами (рис. 3.15) дала низкий (< 0,5) результат, что говорит об отсутствии персистентности и самоподобия в полученной в ходе эксперимента реализации сетевого трафика. *IDS* в течение всего эксперимента не сообщала о подозрительной активности.



а) AVM, H = 0,47 (корреляционная точность 99,23 %)



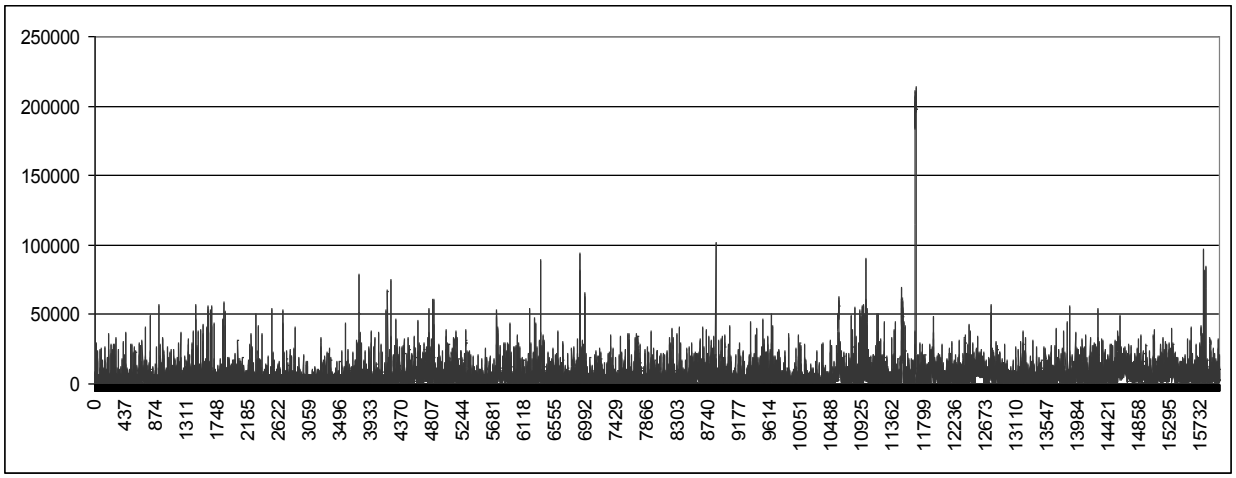
б) R/S – статистика, H = 0,39 (корреляционная точность 98,41 %)



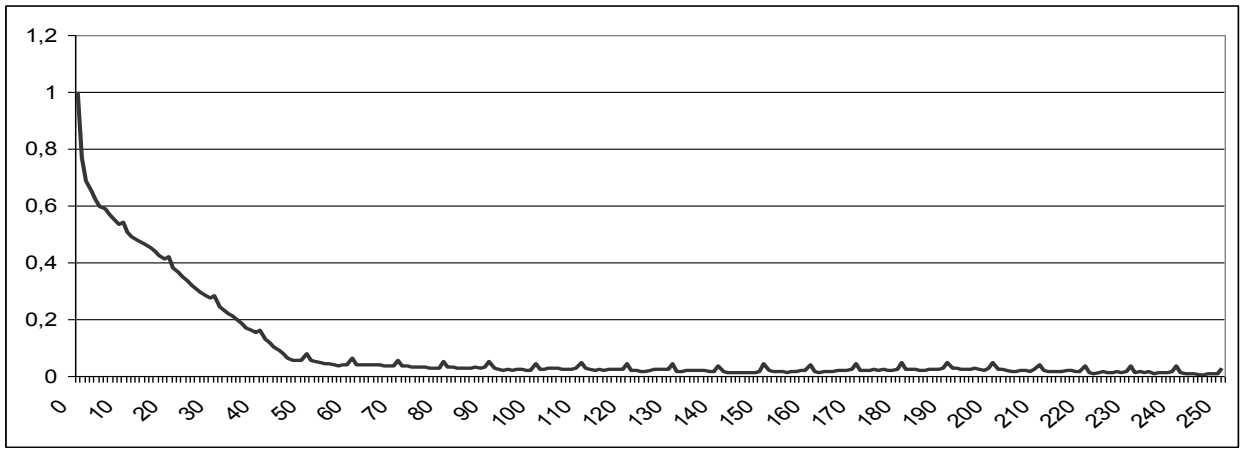
в) периодограмма. H = 0,48 (корреляционная точность 70,54 %)

Рис. 4.15. Оценка параметра Хёрста

2) Начало распределённой атаки, 5 часов, 8 840 891 пакет. Данная реализация трафика показывает поведение сети в начале обнаруженной распределённой атаки. На графике агрегированного по времени процесса мы уже можем наблюдать неравномерности; имеются выбросы достаточно сильной амплитуды, в то время как среднее значение трафика относительно мало. Данное поведение характерно для процессов с *LRD* (рис. 4.16, а). Автокорреляционная функция трафика (рис. 4.16, б) достаточно медленно убывает, долговременную зависимость можно проследить в течение 50-60 событий. В то же время на графике АКФ отчетливо прослеживаются периодические пульсации. Эти факты говорят нам о наличии четко выраженной детерминистской составляющей в данном процессе, что является прямым следствием вредоносного воздействия.



а) агрегация длин пакетов по времени 0,5 с



б) АКФ трафика, k_{max} = 250

Рис. 4.16. Начало атаки

Расчёт параметра Хёрста (рис. 4.17) показал значения, превышающие 0,7. Это значит, что можно сделать вывод о наличии вредоносного воздействия на систему, поскольку трафик приобретает свойства персистентности и самоподобия. На этом этапе мы уже делаем вывод о наличии атаки, а *IDS* не делает даже предупреждений.

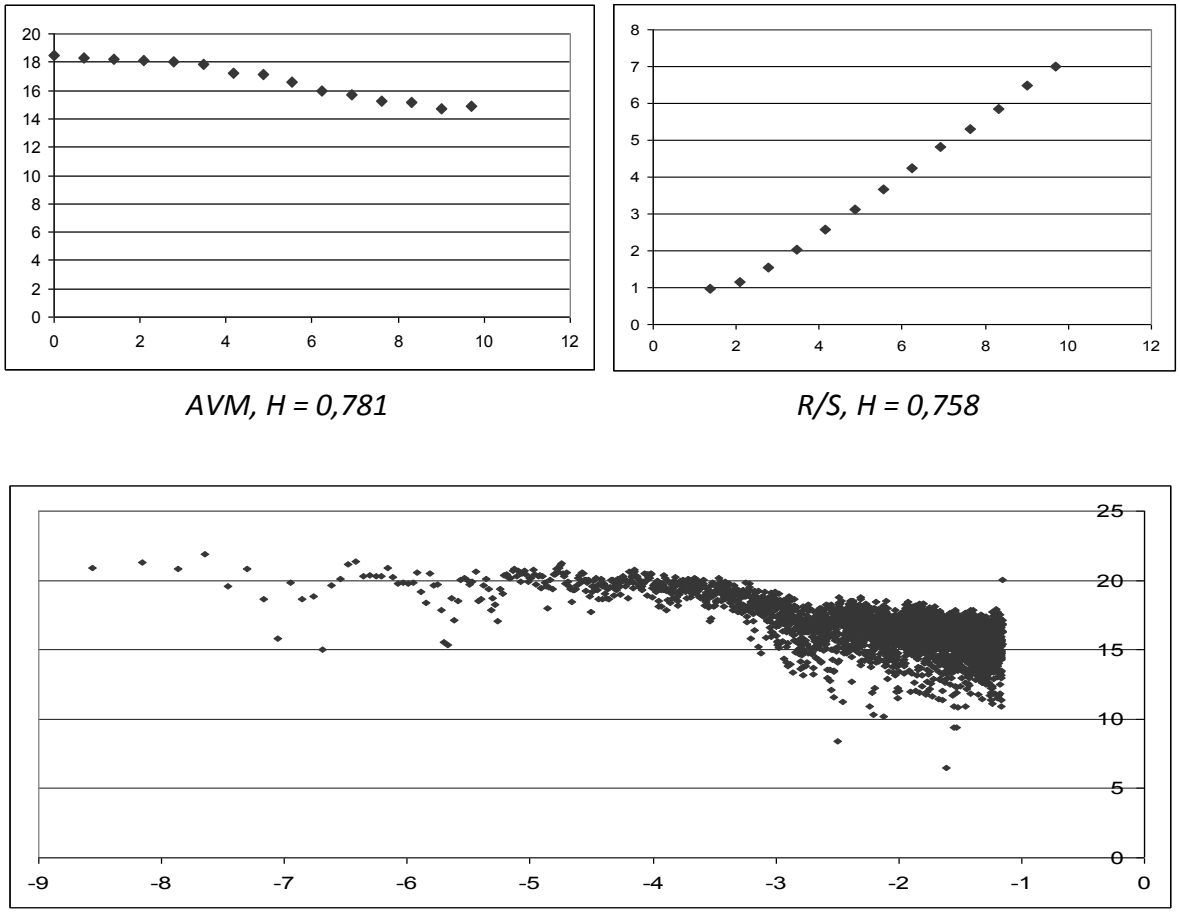
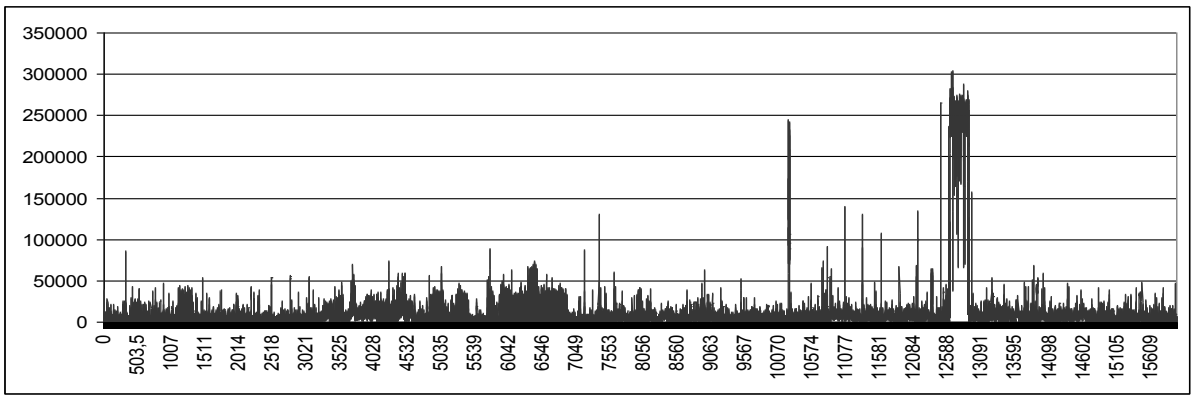
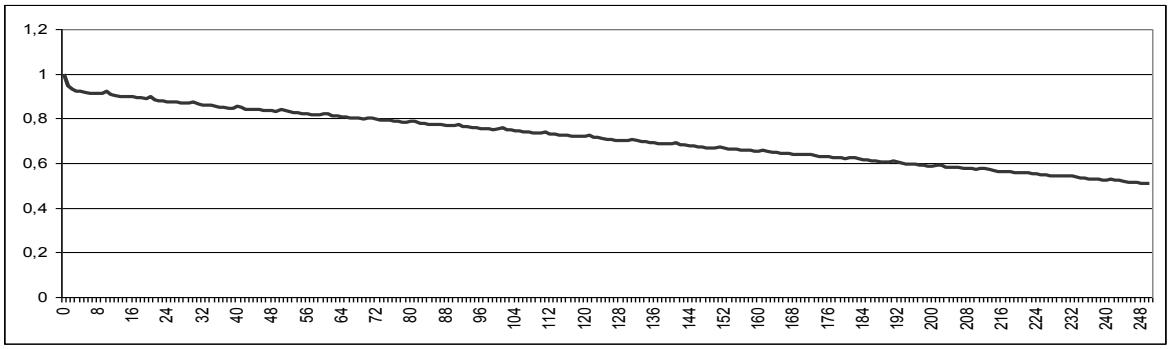


Рис. 4.17. Оценка параметра Хёрста

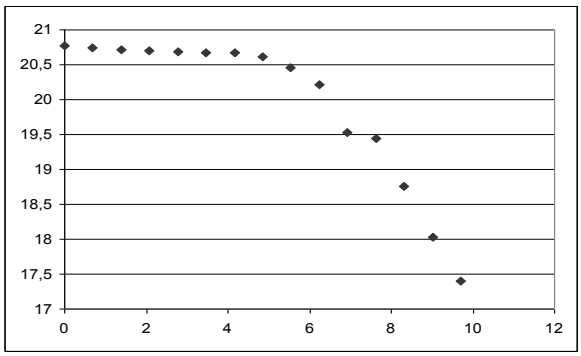
3) Пик атаки, 5 часов, 9 000 213 пакетов. Данная реализация показывает время, когда атака была обнаружена штатной *IDS*. Произошло это в период максимальной сетевой активности. На графике это серия высоких значений с 12590 по 13080 с эксперимента. По сравнению с предыдущим графиком свойства, характеризующие процесс с долговременной зависимостью от начальных условий, выражены ещё более явно (рис. 4.18).



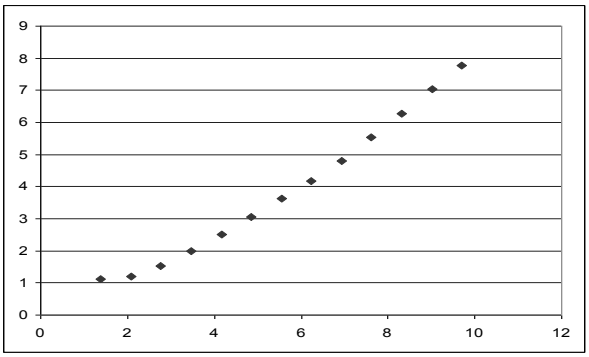
Агрегация длин пакетов по времени 0,5 с



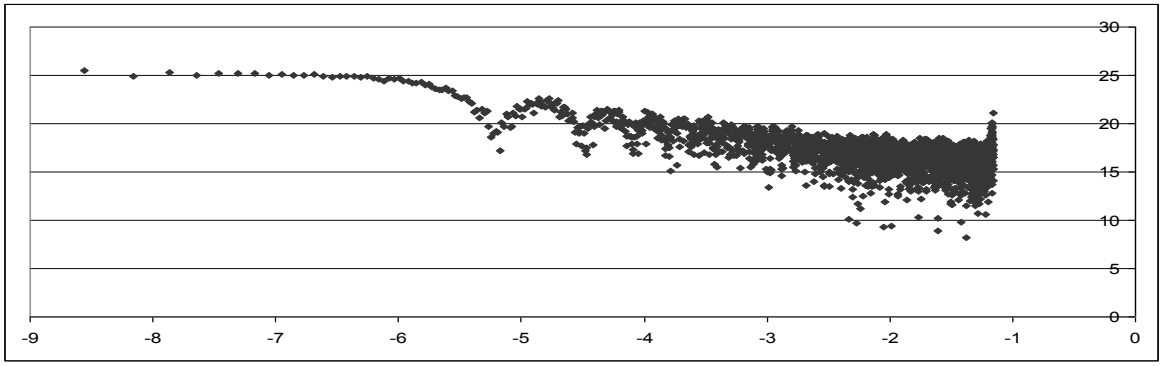
АКФ трафика, k_{max} = 250



AVM, H = 0,847



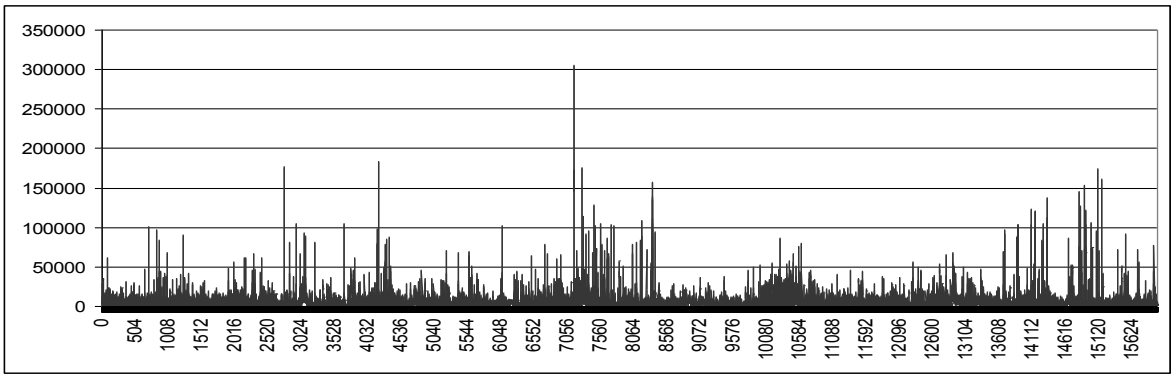
R/S, H = 0,826



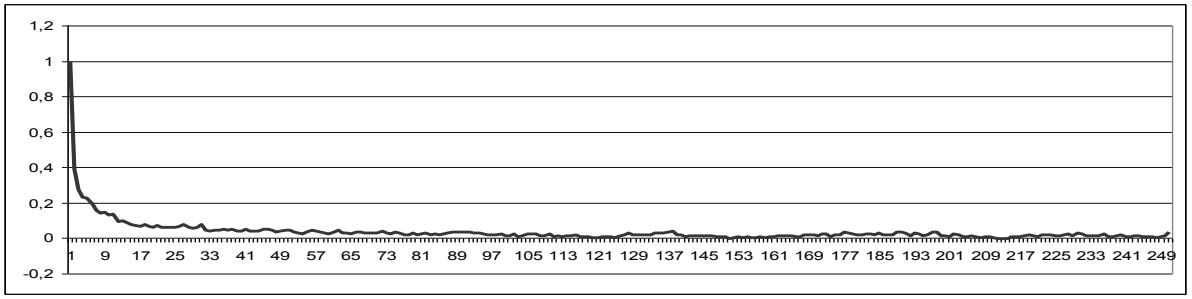
Периодограмма, H = 1,186

Рис. 4.18. Пик атаки

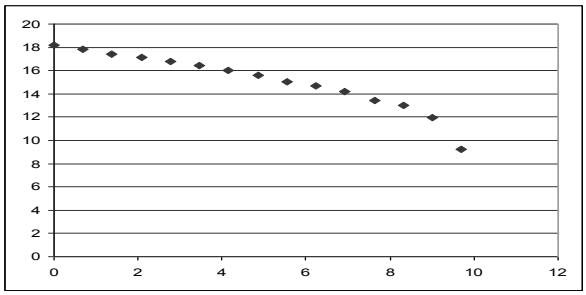
4) Спад вредоносной активности, 5 часов, 7 826 356 пакетов (рис. 4.19).



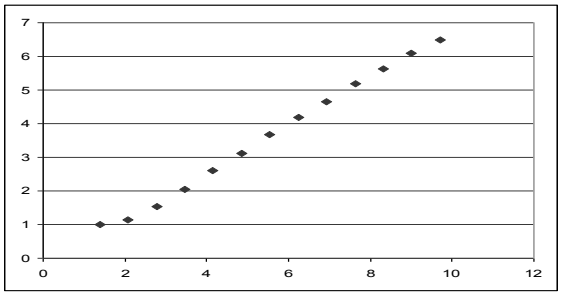
Агрегация длин пакетов по времени 0,5 с



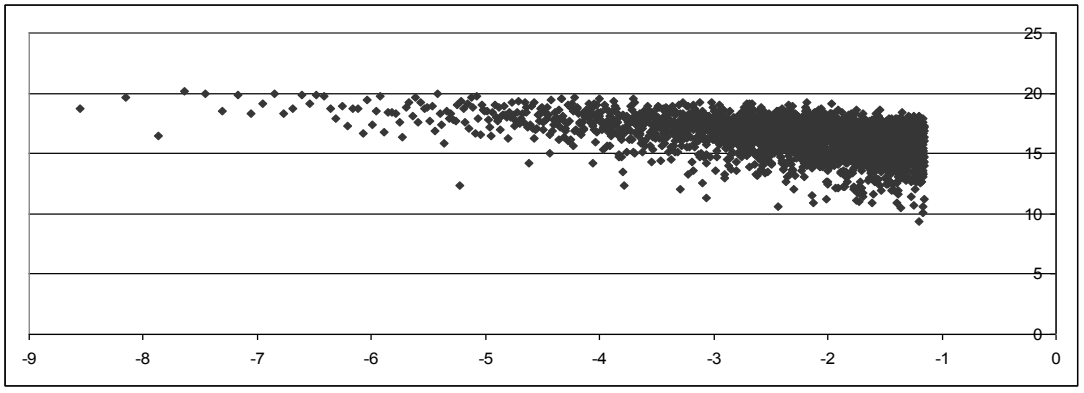
АКФ трафика, k_{max} = 250



AVM, H = 0,621



R/S, H = 0,702



Периодограмма, H = 0,807

Рис. 4.19. Спад атаки

Выводы по главе

Теоретическое и экспериментальное исследование подтвердило гипотезу о том, что в условиях *DoS*-атак агрегированный сетевой трафик ИТКС становится персистентным и самоподобным, позволило разработать механизм раннего обнаружения аномального поведения ИТКС, вызванного начавшейся *DoS*-атакой. Для достижения этого:

– на основе анализа системных характеристик и процессов ИТКС, а также возможностей (активности) злоумышленника разработана математическая модель распределенной *DoS*-атаки на ИТКС, позволяющая выполнять аналитическое и экспериментальное исследование процессов распределенного информационного воздействия на ИТКС, рассчитывать возможную перегрузку телекоммуникационной подсистемы ИТКС;

– разработано программное обеспечение, позволяющее моделировать в ИТКС сетевой трафик в условиях *DoS*-атак, вычислять параметры хаотического самоподобного процесса вредоносного сетевого потока;

– показано, что описание самоподобных и имеющих долговременную зависимость от начальных условий процессов, происходящих в ИТКС, с помощью *FARIMA*-модели, предоставляет возможность оценивать ее будущее состояние на основе текущих параметров. Разработан алгоритм предсказания *DDoS*-атаки на основе *FARIMA*-модели агрегированного трафика ИТКС.

Примеры эффективного апробирования механизмов и средств раннего обнаружения аномального поведения реальных ИТКС дают возможность констатировать адекватность и функциональность основных теоретических построений и разработанных на их основе программных средств.

Список библиографических ссылок

1. Анализатор Sniffer Pro LAN фирмы Sniffer Technologies. 2015. URL: http://www.securitylab.ru/software/233623.php (дата обращения: 18.09.2015)

2. Андреев Ю.В., Балабин А.М., Дмитриев А.А и др. Использование динамического хаоса в коммуникационных системах и компьютерных сетях // Препринт ИРЭ РАН. М., 2000. № 2(626). 76 с.

3. Андреев Ю.В., Балабин А.М., Дмитриев А.А и др. Стратегии использования динамического хаоса в коммуникационных системах и компьютерных сетях. Разделение хаотического кодера и кодера канала // Зарубежная радиоэлектроника. 2000. № 11. С. 4-26.

4. Вредоносные программы в компьютерных сетях / Монахов Ю.М., Груздева Л.М., Монахов М.Ю. Владимир: Изд-во Владим. гос. ун-та, 2010. 96 с.

5. Груздева Л.М., Монахов Ю.М. Об одной математической модели динамики распространения вредоносных программ // Математические методы в технике и технологиях. Сб. трудов XX междунар. науч. конф. В 10 т. Т. 6. Секция 12 / под общ. ред. В.С. Балакирева. Ярославль: Изд-во Ярос. гос. техн. ун-та, 2007. С. 65-66.

6. Дмитриев А. С., Панас А. И., Старков С. О. Динамический хаос как парадигма современных систем связи // Зарубежная радиоэлектроника. 1997. № 10. С. 4-26.

7. Дмитриев А. С., Старков С. О. Передача сообщений с использованием хаоса и классическая теория информации // Зарубежная радиоэлектроника. 1998. №11. С. 4-32.

8. Дружинин Е.Л., Родин А.В., Самохин А.М., Чернышев Ю.А. Выявление статистических закономерностей поведения сетевых устройств. URL: http://www.eyeadmin.com/ru/articles/27 (дата обращения: 18.09.2015)

9. Ильницкий С.В. Работа сетевого сервера при самоподобной (self-similar) нагрузке. 2004. URL: http://314159.ru/ilnickis/ilnickis1.pdf (дата обращения: 18.09.2015).

10. Котенко И.В., Степашкин М.В., Богданов В.С. Архитектуры и модели компонентов активного анализа защищенности на основе имитации действий злоумышленников // Проблемы информационной безопасности. Компьютерные системы. 2005. № 4.

11. Котенко И.В., Степашкин М.В. Модели действий хакеров-злоумышленников при реализации распределенных многошаговых атак // КИИ-2006. X Национальная конференция по искусственному интеллекту с международным участием. Труды конференции. Т 2. М.: Физматлит, 2006.

12. Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории. М.: Постмаркет, 2000. 352 с.

13. Мишин Д.В., Монахов Ю.М. Экспериментальное исследование эффективности Distributed Network IDS // Алгоритмы, методы и системы

обработки данных: сб. нач. статей / под ред. С.С. Садыкова, Д.Е. Андрианова. М.: ООО «Центр информационных технологий в природопользовании», 2009. С.95-100.

14. Монахов Ю.М Уязвимости протокола транспортного уровня TCP // Алгоритмы, методы и системы обработки данных: Сборник научных статей / под ред. С.С. Садыкова, Д.Е. Андрианова. М.: Горячая линия – Телеком, 2006. С.203-210.

15. Монахов Ю.М. Атака на информационную систему предприятия // Формирование социально-ориентированной экономики: вопросы теории и практики. Межвуз. сб. науч. трудов / филиал ВЗФЭИ в г. Владимире. Владимир, 2007. С.105-109.

16. Монахов Ю.М. Динамика протокола TCP в условиях сетевых атак и перегрузок // Математические методы в технике и технологиях: сб. трудов XXI междунар. науч. конф. в 10 т. Т. 7. Секция 6 / под общ. Ред. В.С. Балакирева. Саратов: Сарат. гос. техн. ун-т, 2008. С. 264.

17. Монахов Ю.М. Использование FARIMA-модели для описания и предсказания поведения сети передачи данных в условия атак типа «отказ в обслуживании» // Горный информационно-аналитический бюллетень. 2008. №10. С.133-137.

18. Монахов Ю.М., Груздева Л.М. Теоретическое и экспериментальное исследование распределенных телекоммуникационных систем в условиях воздействия вредоносных программ: Монография. Владимир: Владим.гос.ун-т., 2012. 168 с.

19. Монахов Ю.М., Макаров Р.И. Автоматизированная система обнаружения аномального функционирования распределенной вычислительной среды АСУ // Системный анализ: теория и практика. 2009. №3. С. 86 – 89.

20. Монахов Ю.М. Модели обнаружения аномального функционирования информационно-вычислительной среды интегрированных АСУ : диссертация ... кандидата технических наук : 05.13.06 [Место защиты: Владимир. гос. ун-т]. Владимир, 2009. 129 с.

21. Оценка сетевых характеристик компьютерных сетей в условиях информационного вредоносного воздействия / Груздева Л.М., Монахов Ю.М., Монахов М.Ю. Владимир: Изд-во Владим. гос. ун-та, 2010. 112 с.

22. Полянский Д.А. Монахов Ю.М. Оценка безопасности информационно-вычислительной сети на основе формальных моделей // XIX

Международная научная конференция «Математические методы в технике и технологиях». Воронеж, 2006. Т. 10. С. 196-198.

23. Треногин Н.Г., Соколов Д.Е. Фрактальные свойства сетевого трафика в клиент-серверной информационной системе // Вестник НИИ СУВПТ. 2003. С. 163-172.

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

25. Carl G., Kesidis G., Brooks R. R., Rai S. Denial-of-Service Attack-Detection Techniques // IEEE Internet Computing. 2006. V. 10, № 1. P. 82-89.

26. Chandrashekhar G., Wakde D. G. On the prediction of packet process in network traffic using FARIMA time-series model // J. Indian Inst. Sci. 2004. V.84. P. 31-39.

27. Crovella M., Bestavros A. Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes // In IEEE/ACM Transactions on Networking. 1997. V.5. № 6. P. 835-846.

28. Devaney R.L. An Introduction to Chaotic Dynamical Systems. Second Edition. Addison-Wesley Publishing Company, 1989.

29. Fei X. Modeling and Predicting Long-range Dependent Traffic with FARIMA Processes. Dept. of Information Engineering The Chinese University of Hong Kong. URL: http://www.ensc.sfu.ca/~ljilja/cnl/papers/mflrd.ps (дата обращения: 18.09.2015).

30. Guo L., Crovella M., Matta I. TCP congestion control and heavy tails // Tech. Rep. BUCS-TR-2000-017, Computer Science Dep., Boston University, 2000. 11 p.

31. Haining W., Zhang D., Shin G. Change-Point Monitoring for Detection of DoS Attacks //IEEE Transactions on Dependable and Secure Computing. 2004. V.1. URL: http://www.cs.wm.edu/~hnw/paper/tdsc.pdf (дата обращения: 18.09.2015).

32. Kocarev L., Vattay G. Complex Dynamics in Communication Networks. Springer, 2005. 361 p.

33. Steger J., Vaderna P., Vattay G. On the Propagation of Congestion Waves in the Internet. 2004. URL: http://arxiv.org/abs/cond-mat/0401283 (дата обращения: 18.09.2015).

34. Vattay G. Self-similarity in bottleneck buffers // Proceedings of

Globecom 2001. December 2001. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.527&rep=rep1&type=pdf (дата обращения: 18.09.2015).

35. Vattay G., Fekete A., Steger J., Marodi M. Modeling Competition, Fairness and Chaos in Computer Networks // ASTED International Conference on Communications, Internet, and Information Technology. St. Thomas, US Virgin Islands, 2002.

36. Veres B.M. The chaotic nature of TCP congestion control // IEEE INFOCOM'2000. 2000. URL: http://www2.ensc.sfu.ca/~ljilja/ENSC835/Fall03/Assignments/papers/74.pdf (дата обращения: 18.09.2015).

37. Veres B.M., Kenesi Z., Molnar S., Vattay G. On the propagation of long-range dependence in the Internet // ACM SIGCOMM 2000, Stockholm, Sweden, 2000.

38. Willinger W., Taqqu M., Sherman R., Wilson D. Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level // IEEE/ACM Transactions on Networking. 1997. V. 5, № 1. P. 71-86.

ГЛАВА 5. АНАЛИТИЧЕСКИЕ И ИМИТАЦИОННЫЕ МОДЕЛИ РАСПРОСТРАНЕНИЯ НЕДОСТОВЕРНОЙ ИНФОРМАЦИИ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Один из основных способов изучения процессов в информационно-телекоммуникационных системах (ИТКС) – моделирование, которое затрагивает проблему изучения процессов, проходящих в ней. В нашем случае это распространение недостоверной информации по узлам информационно-телекоммуникационной сети. В силу того, что данный процесс социально нежелателен, он «угрожает» нормальному (устойчивому) функционированию системы, а значит, с точки зрения информационной безопасности относится к классу угроз. Далее будем употреблять как синонимы понятия «распространение недостоверной информации» и «угроза недостоверной информации» (УгНдИ).

В главе представлены результаты построения моделей УгНдИ в ИТКС, их экспериментальное исследование и практическая апробация.

Были поставлены следующие задачи исследования:

– создать имитационную модель распространения угрозы недостоверной информации в ИТКС: разработать алгоритм УгНдИ в ИТКС; на его основе создать имитационную модель УгНдИ в ИТКС и провести ее экспериментальное исследование;

– на основе экспериментальных данных по имитационной модели создать аналитическую модель УгНдИ в ИТКС, провести экспериментальное исследование аналитической модели и проверить ее адекватность;

– смоделировать процесс распространения угрозы недостоверной информации на крупномасштабной ИТКС.

5.1. 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 [12]. В данной работе рассматриваются модели влияния, так как они наиболее адаптивны к решаемым задачам. На рис. 5.1 представлена обобщенная классификация

моделей влияния. Коротко охарактеризуем представленные классы моделей влияния.

Рис. 5.1. Классификация 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.

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 не-

который 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) [27].

Модели просачивания и заражения являются популярным способом изучения распространения информации и инноваций в социальных системах.

Модель Изинга – математическая модель, описывающая возникновение намагничивания материала. В [38] 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 [10].

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 [42] 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 [12]. Есть агент, входящий в некоторую социальную сеть. Агент информирован о текущей ситуацион-

ной обстановке (действиях и представлениях других агентов, параметрах среды – так называемом состоянии природы (stateofnature) 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 [18] 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) [29] – 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.

Эпидемиологические модели

В [28] 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 [41] 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 [30] 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 [35] 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 [32] 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.

5.2. Имитационная модель распространения недостоверной 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 УгНдИ занимает несколько часов даже при использовании распределенных вычислительных ресурсов. Аналитическая модель может дать прогноз УгНдИ почти мгновенно. С ее помощью можно получить актуальные данные (до того момента, когда количество атакующих абонентов будет максимальным) по динамике УгНдИ.

Процесс УгНдИ характеризуется следующими особенностями [1-7]. В сети существуют узлы трех типов. Первый тип – атакующие узлы, это узлы, распространяющие (недостоверную) информацию. Второй тип – защищенные узлы, характеризующиеся тем, что не принимают участие в распространении недостоверной информации и никогда не будут этим заниматься. Третий тип – потенциально уязвимые. Узлы такого типа не участвуют в процессе распространения, но могут быть подвержены негативному влиянию со стороны атакующих узлов и могут начать распространять недостоверную 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 [13-15]), φ – коэффициент топологической уязвимости сети, отражающий внутреннее свойство ИТКС, основанное на характеристиках ее топологии, которое способствует распространению недостоверной информации, t – время процесса (в условных единицах времени).

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

$$\begin{cases} I(t) = f(N, \beta, \gamma, \varphi, t) \\ R(t) = g(N, \beta, \gamma, \varphi, t) \end{cases} \tag{5.1}$$

Методика разработки аналитической модели включает в себя последовательность следующих действий:

1) формирование имитационной модели для исследования характера и параметров процесса УгНдИ;

2) синтез аналитических зависимостей параметров процесса;

3) проведение экспериментов с целью проверки точности (адекватности) модели.

Приведем алгоритм реализации УгНдИ, основываясь на описании процессов, протекающих в реальных ИТКС. Схема реализации угрозы представлена на рис. 5.2.

Алгоритм УгНдИ в ИТКС

Шаг 1. Распространение недостоверной информации (НдИ) (далее процесс «атаки») инициирует какой-либо абонент-злоумышленник (на рис. 5.2 – узел 1), распространяя сообщения с НдИ (реализует угрозу) по списку контактов. Атаку может начинать один злоумышленник или группа.

Шаг 2. Абоненты-получатели (узлы 2, 3, 4), приняв сообщение с НдИ, 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 НдИ в сообщении, заинтересованности абонента и т.д.

Шаг 3. Абоненты могут перестать воспринимать и, соответственно, распространять НдИ (узел 5) (далее процесс «защиты»), вследствие воздействия механизмов защиты (например, предупреждение о ней), поэтому сообщения с НдИот атакующих абонентов будут постоянно отвергаться.

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

Таким образом, распространение НдИ 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 распространения НдИ в ИТКС, которая состоит из разработанной программы ModelGraph и данных, которые могут быть сгенерированы спомощью ПО Pajek [36].

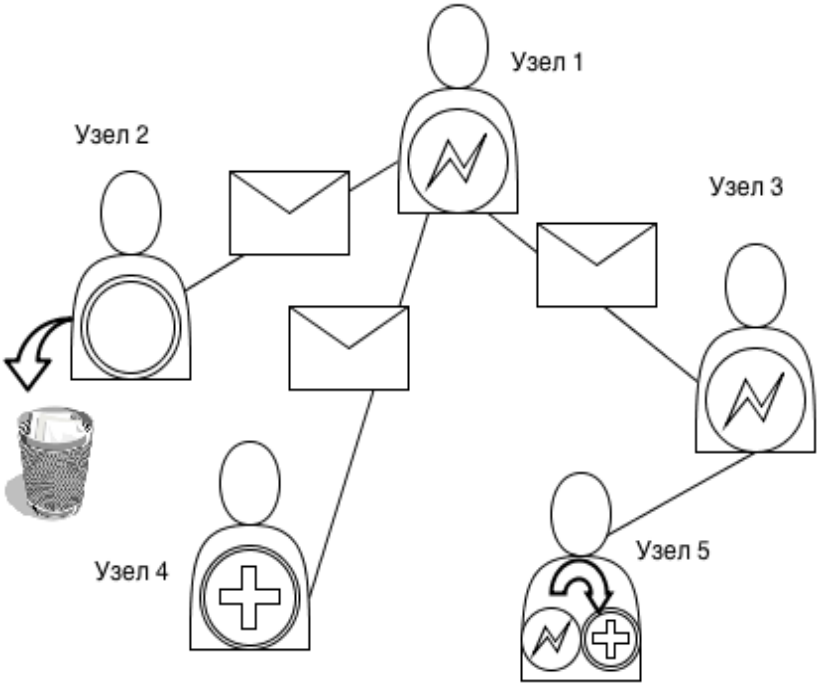


Рис. 5.2. Схема распространения НдИ

Имитационная модель распространения НдИ

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 УгНдИ (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 уз-

лов ($|V^R|=R_0$), $V^S=\{v_i^S\}$ – 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 βвыполнить:$V^S \backslash v_j$ и $V^I \cup v_j$; с вероятностью γ выполнить: $V^I \backslash v_i$, $V^R \cup v_i$.

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

ModelGraph – программа для имитационного моделирования распространения НдИ 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 УгНдИ в ИТКС*. 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.

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 VladimirBatagelj и AndrejMrvar.

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

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 (рис. 5.3).

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 (рис. 5.4).

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 (рис. 5.5).

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

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);

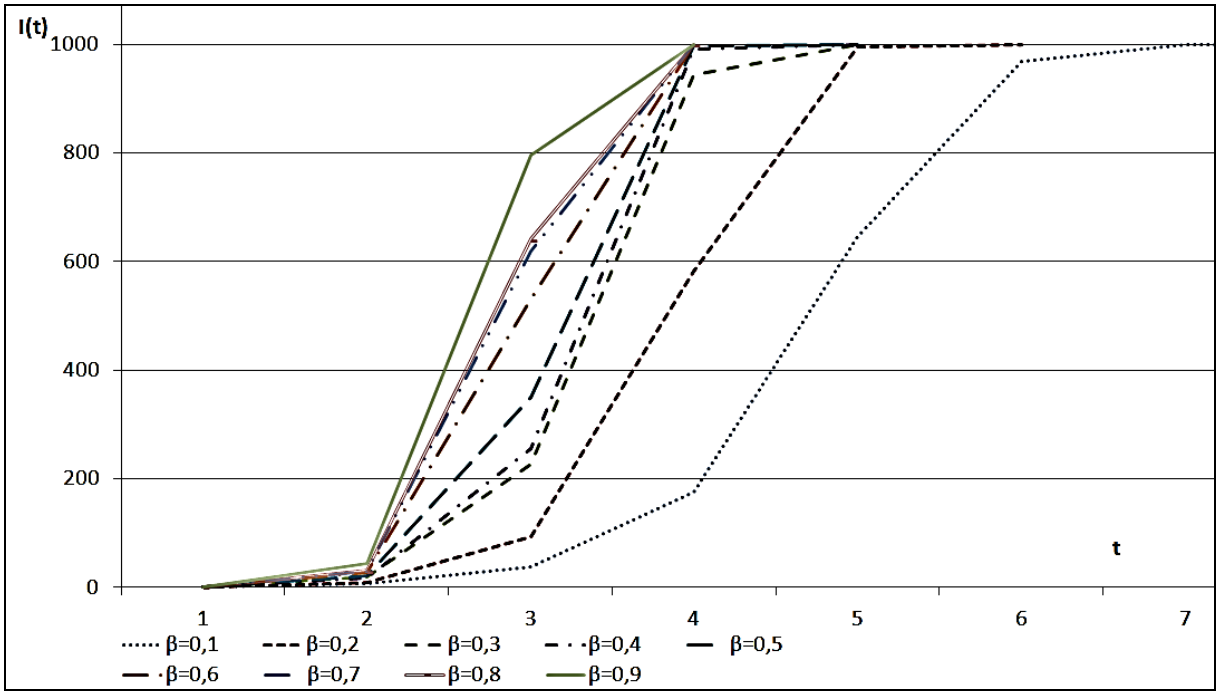


Рис. 5.3. Влияние β на процесс атаки

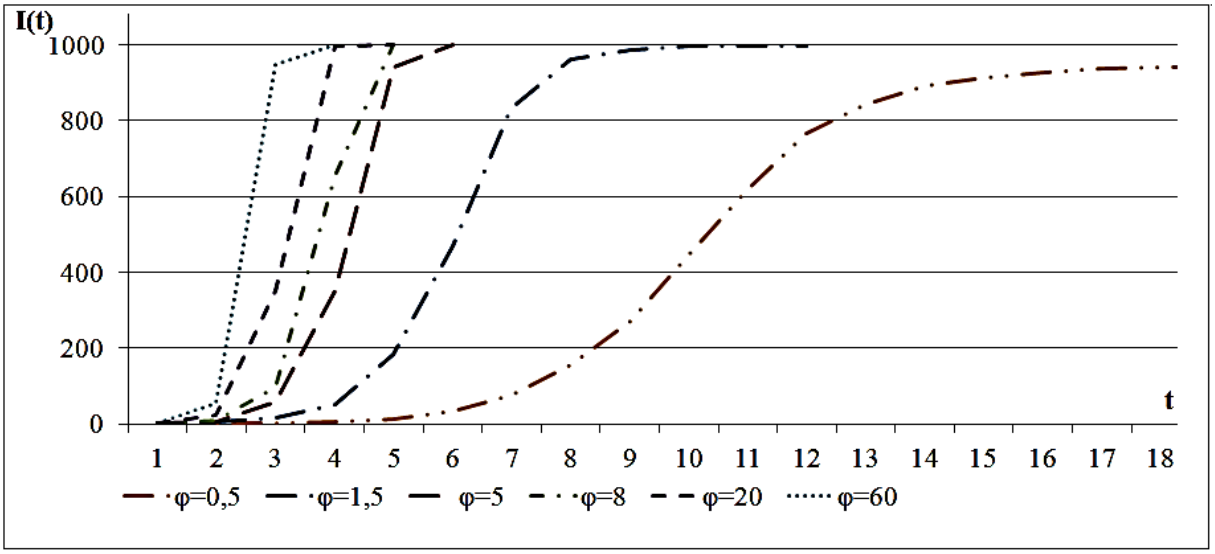


Рис. 5.4. Влияние φ на процесс атаки

– 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).

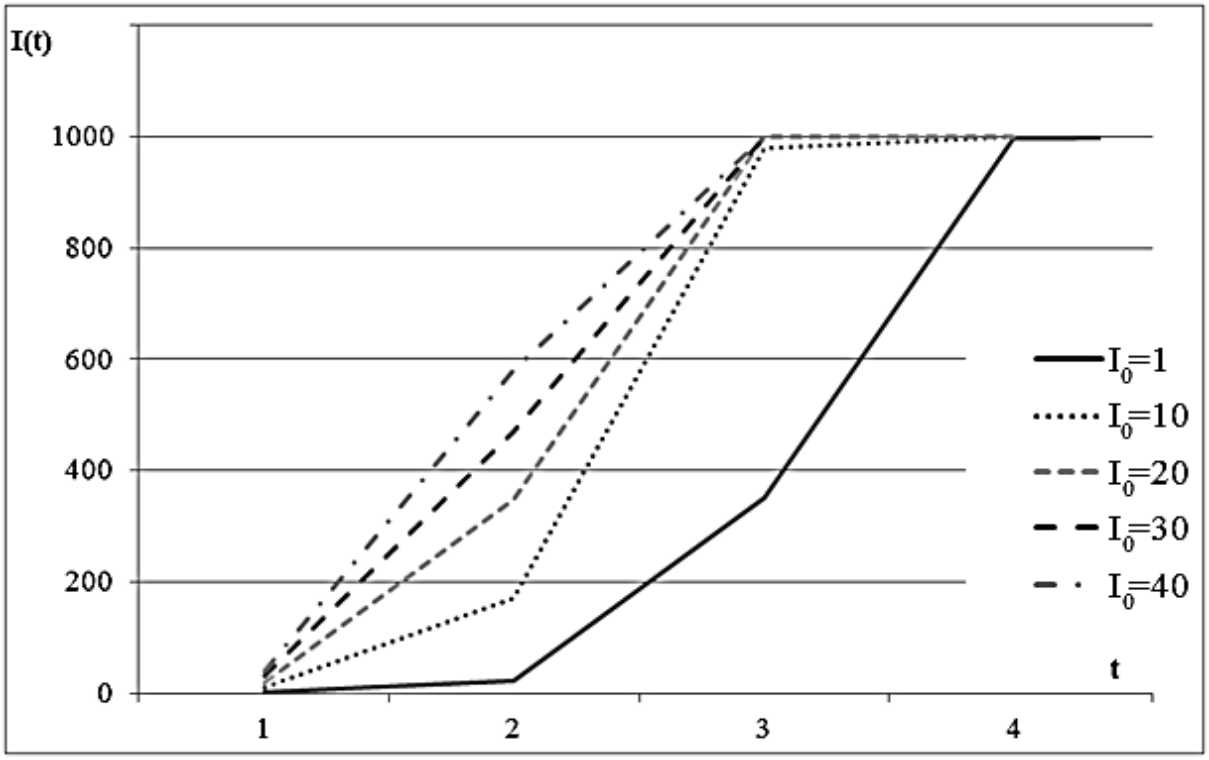


Рис. 5.5. Влияние I_0 на процесс атаки

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

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. (рис. 5.6).

Эксперимент 5. Влияние начального количества защищенных узлов.

Эксперименты проводились при следующих значениях параметров: $N = 1000$, $\varphi = 20$, $I_0 = 1$, $\beta = 0{,}5$, $\gamma = 0{,}5$, $R_0 = 0..200$. (рис. 5.7).

По результатам экспериментов 4 и 5 можно сделать выводы:

– введение подпроцесса защиты увеличивает время всего процесса УгНдИ (эксперимент 4, 5);

– при небольших значениях вероятности защиты ($\gamma < 0{,}3$) угроза реализуется практически на всех узлах в сети (эксперимент 4);

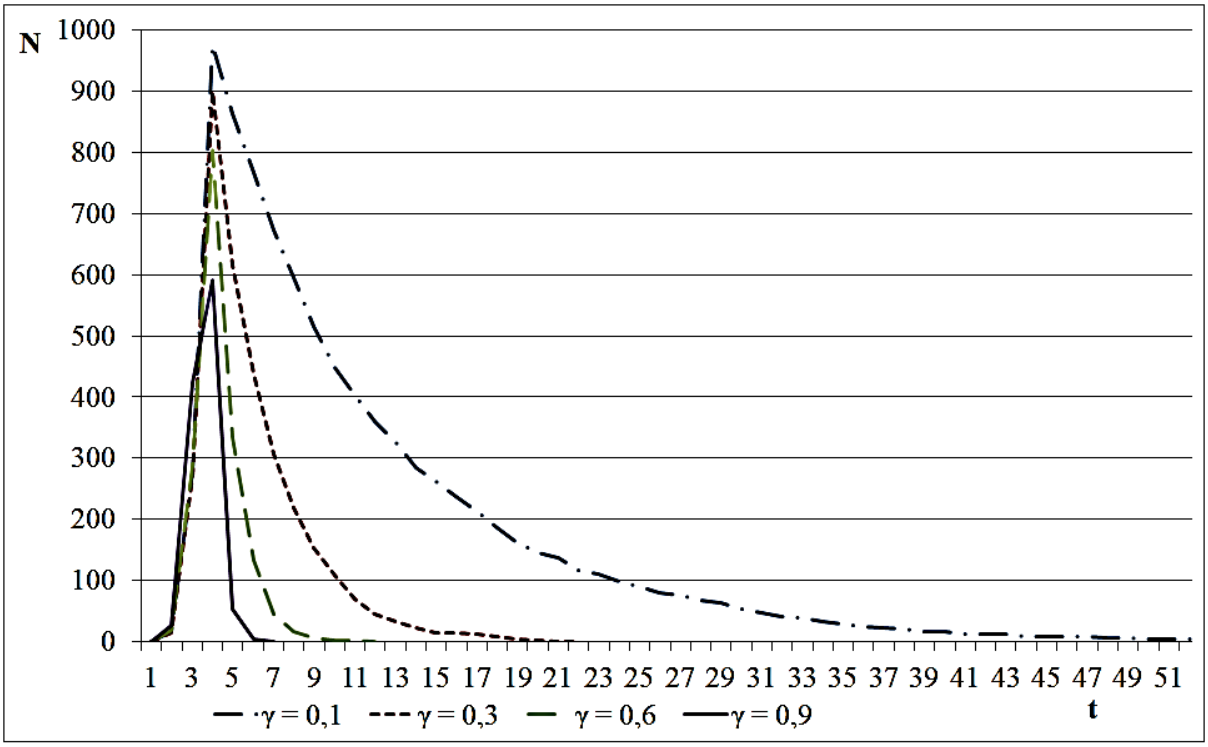


Рис. 5.6. Влияние γ на процесс атаки

– при небольших значениях вероятности защиты ($\gamma < 0{,}3$) 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).

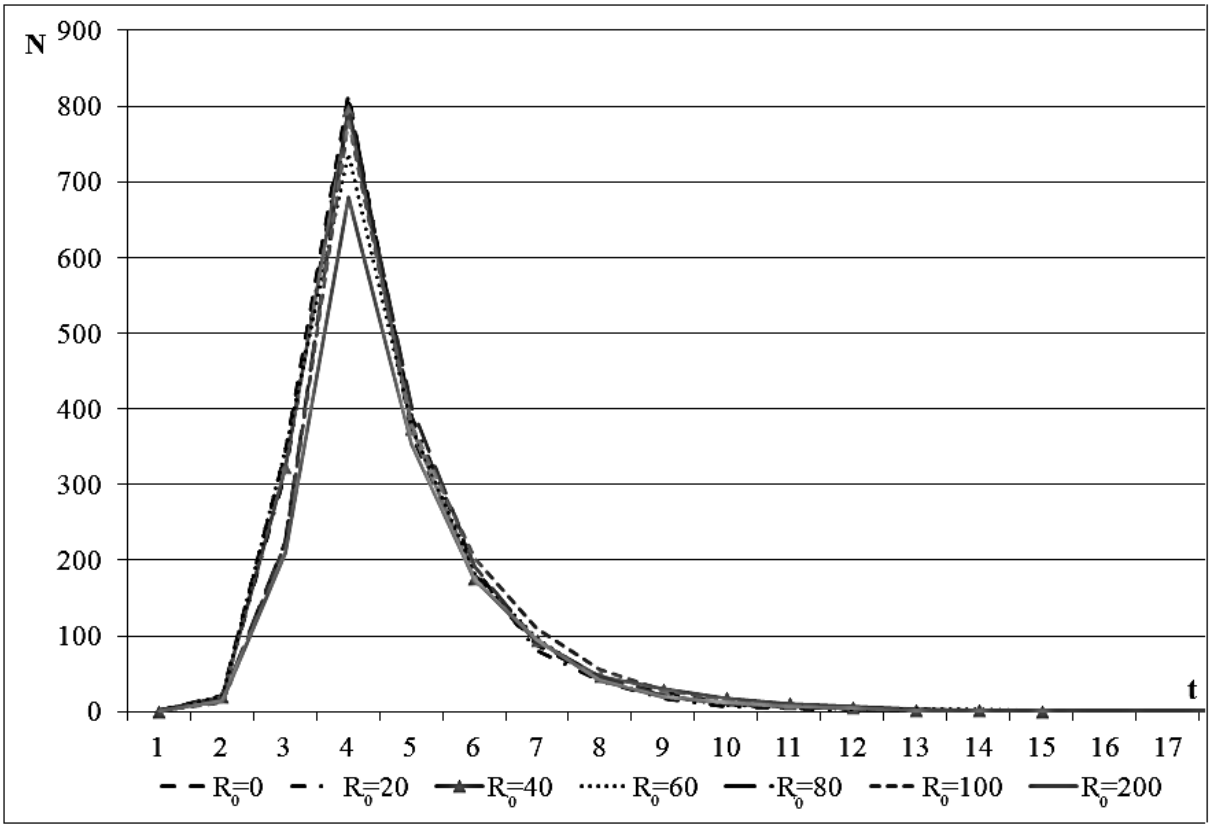


Рис. 5.7. Влияние R_0 на процесс атаки

5.3. Разработка аналитической модели

Анализируя процесс информационного взаимодействия абонентов при распространении недостоверной информации в ИТКС, можно сделать следующие выводы. Имеем дело с тремя типами абонентов: атакующие абоненты, которые распространяют недостоверную информацию, защищенные абоненты, характеризующиеся тем, что не принимают участие в распространении и никогда не будут этим заниматься, и потенциально уязвимые абоненты, которые могут быть подвержены негативному влиянию со стороны атакующих узлов и могут начать распространять недостоверную информацию. При этом мы наблюдаем два противоборствующих подпроцесса атаки и противоборства (защиты) абонентов сети. Для моделирования таких явлений часто применяют эпидемиологические модели [17, 30, 31, 33, 34 и др.], в частности нашему описанию точно соответствует SIR-модель Кермака-Маккендрика [20, 36, 28, 37 и др.]. Характер графиков, полученных в результате имитационного моделирования (рис. 5.8), схож с результатами, которая дает данная модель.

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

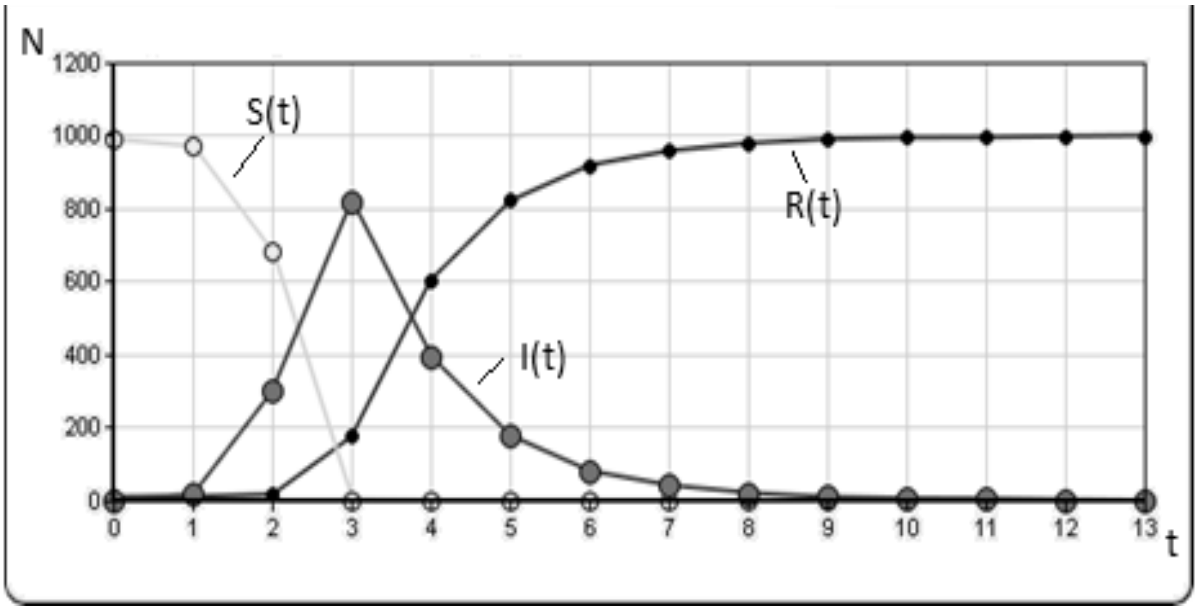


Рис. 5.8. Имитационное моделирование (N = 1000, φ = 20, I_0 = 1, β = 0,5, γ = 0,5, R_0 = 10), $S(t)$ – количество подверженных атаке узлов

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 [20, 28]

$$\begin{cases} \dfrac{dI}{dt} = \beta \dfrac{S(t) \cdot I(t)}{N} - \gamma I(t) \\ \dfrac{dR}{dt} = \gamma I(t) \\ \dfrac{dS}{dt} = -\beta \dfrac{S(t) \cdot I(t)}{N} \end{cases}, \qquad (5.2)$$

где $I(t)$ – количество зараженных (инфицированных) особей, $S(t)$ – количество восприимчивых особей, $R(t)$ – количество «исключенных с иммуниза-

цией» (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 (5.2) для анализа УгНдИ в ИТКС получаем результаты в виде графиков (рисунок 5.9), которые хотя и правильно описывают характер процесса, но не дают нужной точности прогноза.

Была выдвинута гипотеза о том, что система (5.2) 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 (5.2) под прогнозирование УгНдИ в ИТКС путем интегрирования в нее параметра топологической уязвимости сети φ.

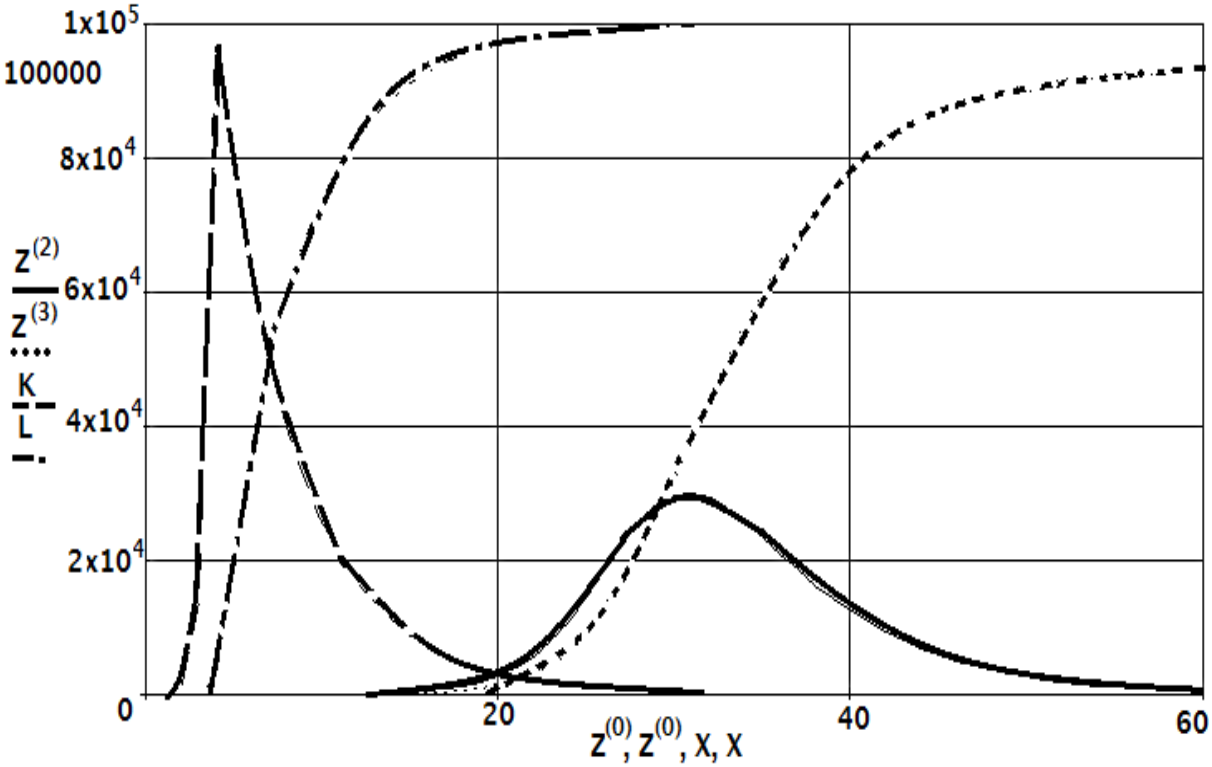


Рис. 5.9. Результаты имитационного моделирования

Проанализировав графики, полученные по результатам имитационного моделирования и аналитического решения системы 5.1, и проследив физический смысл уравнений в данной системе, можно прийти к следую-

щему 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 φможет влиять на $I(t)$ через коэффициент β.

В общем виде адаптированную систему 5.1 можно представить в следующем виде:

$$\begin{cases} \dfrac{dI}{dt} = C\beta \dfrac{S(t)\cdot I(t)}{N} - \gamma I(t) \\ \dfrac{dR}{dt} = \gamma I(t) \\ \dfrac{dS}{dt} = -C\beta \dfrac{S(t)\cdot I(t)}{N} \end{cases}, \qquad (5.3)$$

где C – коэффициент, зависящий от параметра φ.

Отметим, что в [37] 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. Результаты серии экспериментов по имитационному моделированию УгНдИ в ИТКС (рис. 5.9, 5.10) позволили получить зависимость параметра C от φ в виде $(2\cdot\ln\varphi)$. Аппроксимация проводилась методом наименьших квадратов с использованием пакета Math CAD.

Итоговая система имеет вид

$$\begin{cases} \dfrac{dI}{dt} = 2\ln\varphi\cdot\beta \dfrac{S(t)\cdot I(t)}{N} - \gamma I(t) \\ \dfrac{dR}{dt} = \gamma I(t) \\ \dfrac{dS}{dt} = -2\ln\varphi\cdot\beta \dfrac{S(t)\cdot I(t)}{N} \end{cases}. \qquad (5.4)$$

Система дифференциальных уравнений (5.4) позволяет получить прогноз УгНдИ в крупномасштабной ИТКС ($N = 10^5..10^8$) с погрешностью до 20 %.

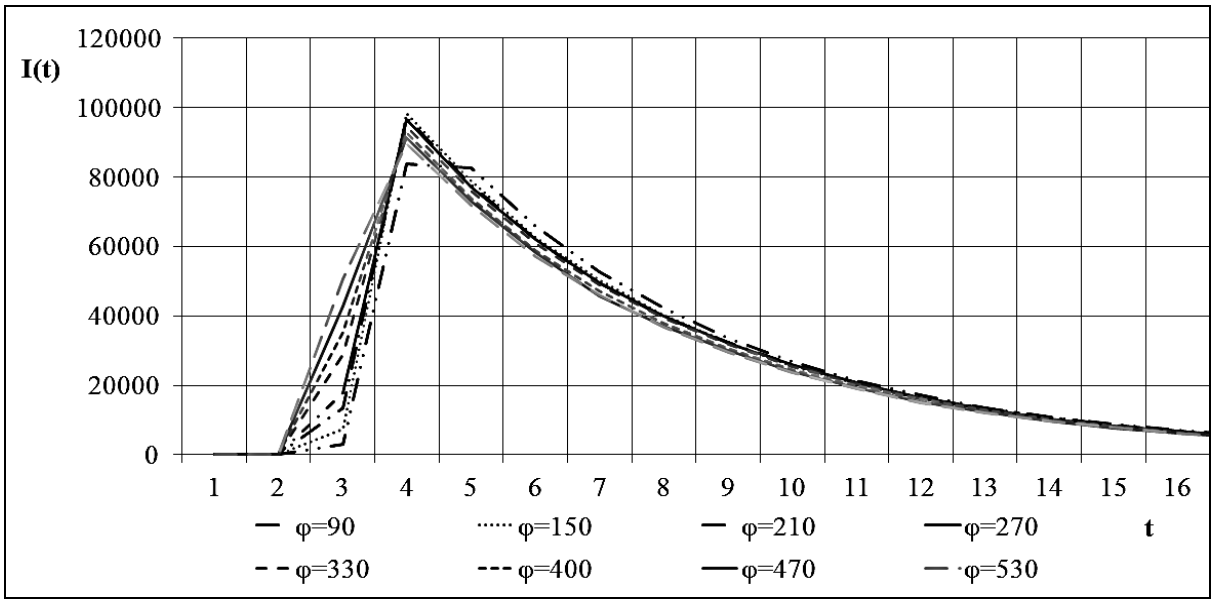


Рис. 5.10. Результаты имитационного моделирования
($N = 10^5$, $I_0 = 1$, $\beta = 0{,}3$, $\gamma = 0{,}2$, $R_0 = 0$)

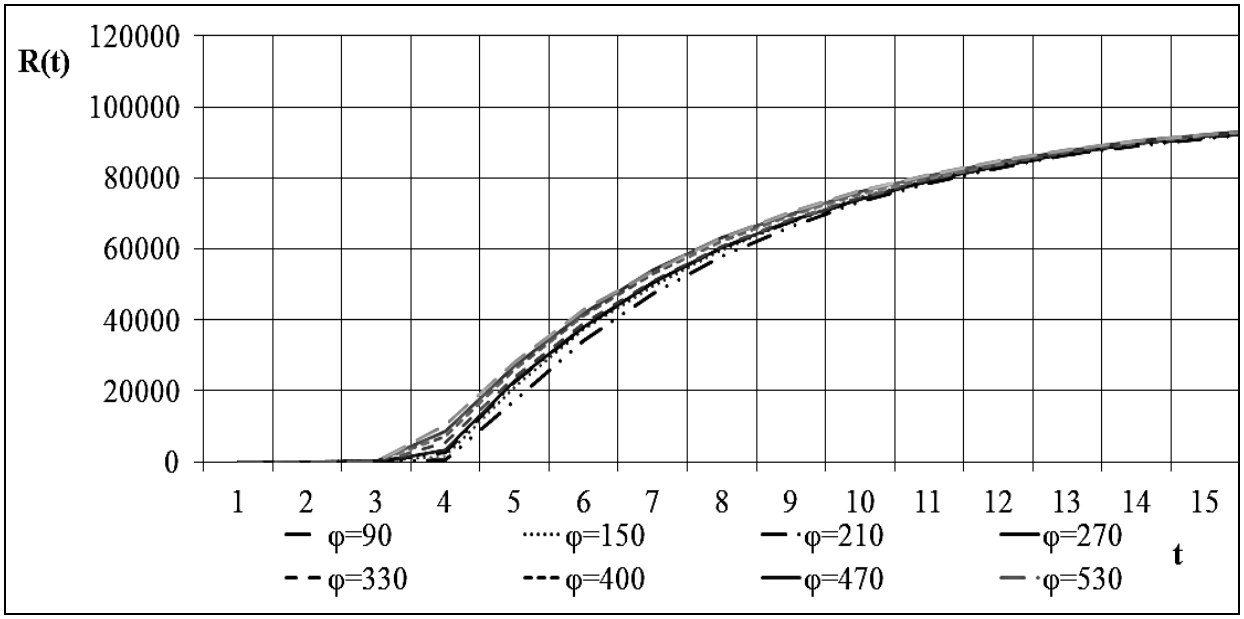


Рис. 5.11. Результаты имитационного моделирования
($N = 10^5$, $I_0 = 1$, $\beta = 0{,}3$, $\gamma = 0{,}2$, $R_0 = 0$)

5.4. Экспериментальное исследование аналитической модели

Результаты аналитической модели сравнивались с результатами имитационного моделирования процесса УгНдИ на топологии реальной

сети. («ВКонтакте»). Эксперимент 1. На рис. 5.12 приведены результаты имитационного моделирования и аналитического решения для β=0,5, γ=0,51, R_0=0, I_0 =1.

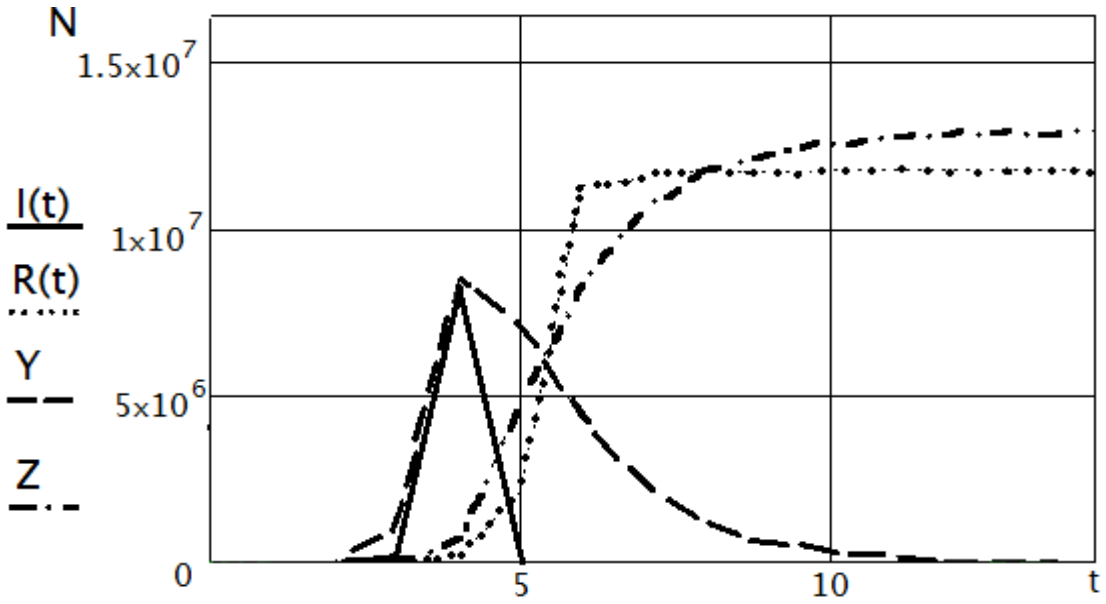


Рис. 5.12. Результаты аналитической и имитационной модели (I и R – аналитическое решение, Y и Z – результаты имитационной модели)

Эксперимент 2. На рис. 5.13 приведены результаты имитационного моделирования и аналитического решения для $\beta = 0{,}5$, $\gamma = 0{,}51$, $R_0 = 0$, $I_0 \approx 24000$.

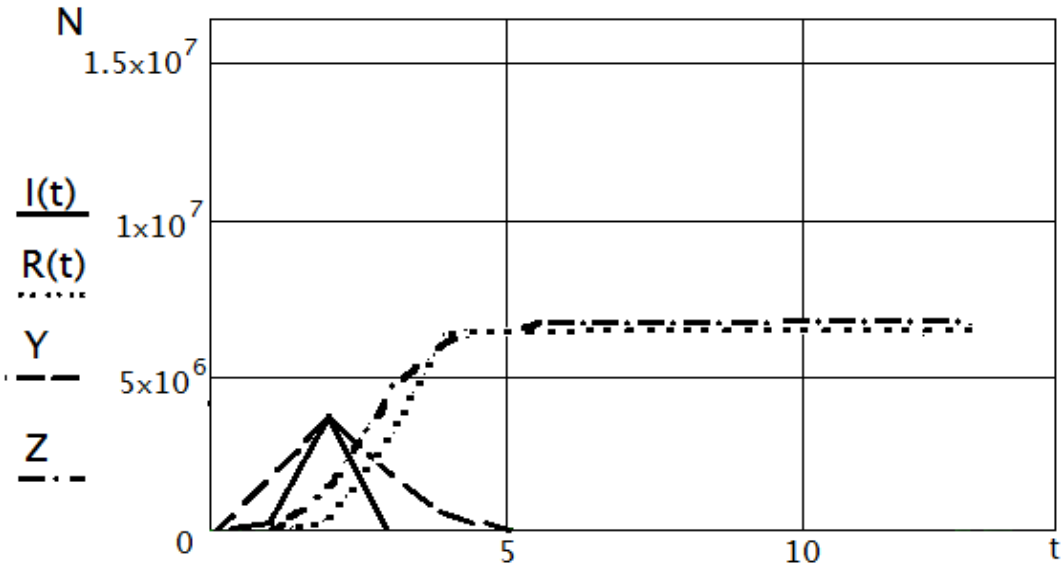


Рис. 5.13. Результаты аналитической и имитационной модели (I и R– аналитическое решение, Y и Z – результаты имитационной модели)

Эксперимент 3. На рис. 5.14 приведены результаты имитационного моделирования и аналитического решения для $\beta = 0{,}5$, $\gamma = 0{,}51$, $R_0 \approx 4 \cdot 10^6$, $I_0 = 1$.

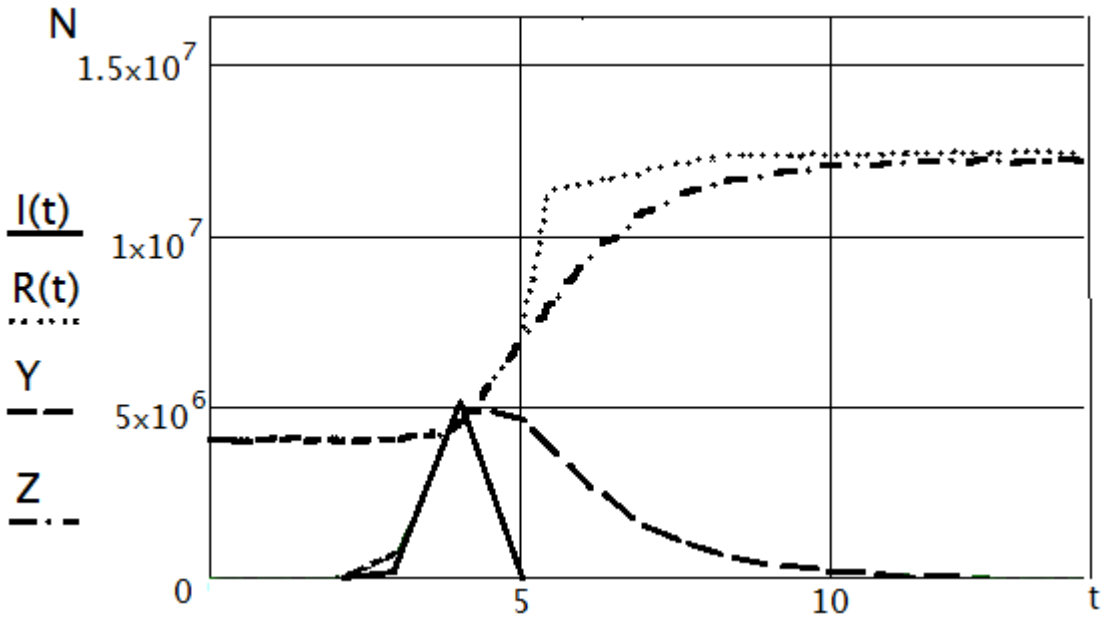


Рис. 5.14. Результаты аналитической и имитационной модели
(I и R – аналитическое решение, Y и Z – результаты имитационной модели)

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 ($\approx 4 \cdot 10^6$) 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 (5.2), можно говорить о значи-

тельном увеличении точности прогнозирования процесса УгЗИ в ИТКС за счет учета влияния на процесс топологической уязвимости сети.

5.5. Разработка методики формирования сетевой топологии

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 т.д., отраженные в различных трудах [8-11] в данной главе не рассматриваются.

Для моделирования УгНдИ 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 [21-26], 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 $16 \cdot 10^6$ 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 [29]) 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 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:

– 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.

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 рис. 5.15.

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.

Алгоритм формирования графа доступной части сети

Шаг 1 (блок 2). Начальная установка. Обнулить множества вершин $V = \varnothing$ и связей $E = \varnothing$. Инициализировать счетчик узлов ($k = 1$). Добавить вершину v_1 во 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 $Z = Z \backslash z_i$ 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 $a_{n+i} = id(z_i)$.

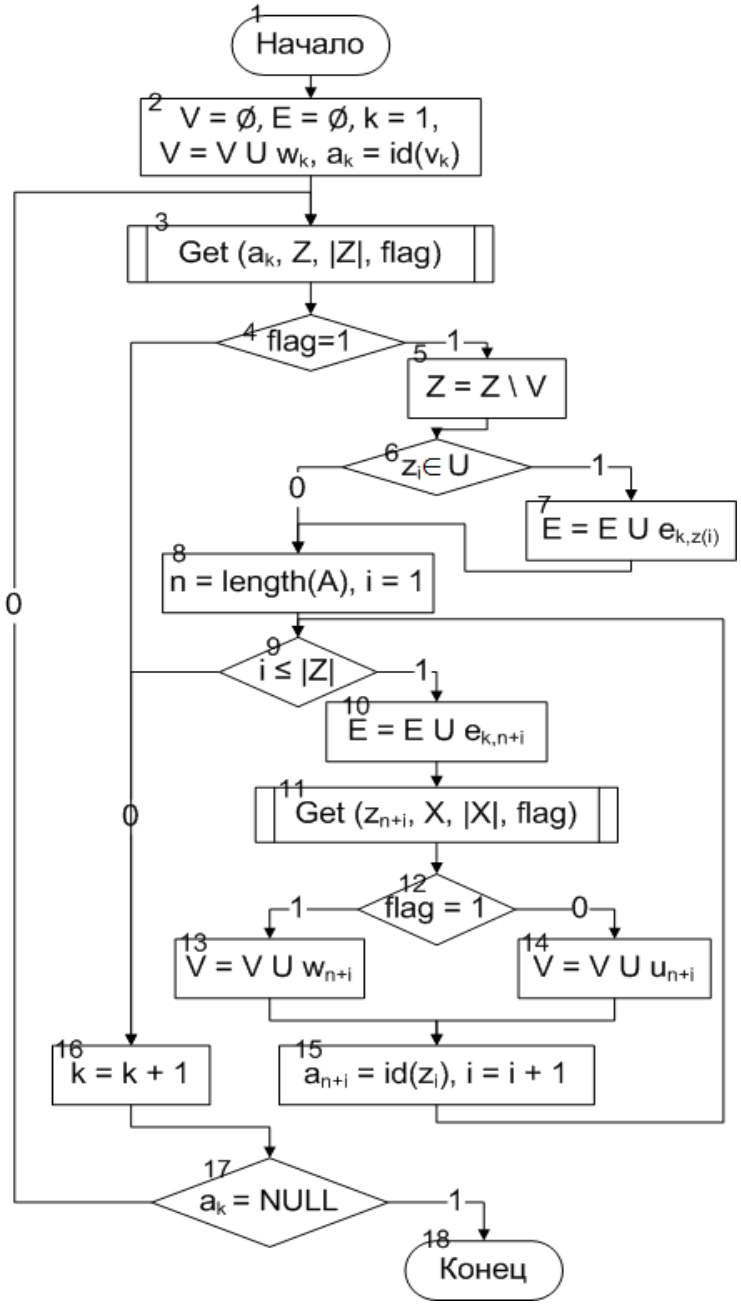


Рис. 5.15. Блок-схема алгоритма формирования
графа доступной части сети

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.

Этап 1. Выполняем начальные установки согласно первому шагу алгоритма: $k = 1$, $V = \{w_1\}$, A[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 рис. 5.16, на котором закрытые узлы выделены серым цветом, а открытые – белым.

Этап 6. Выполняем функцию Get(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.

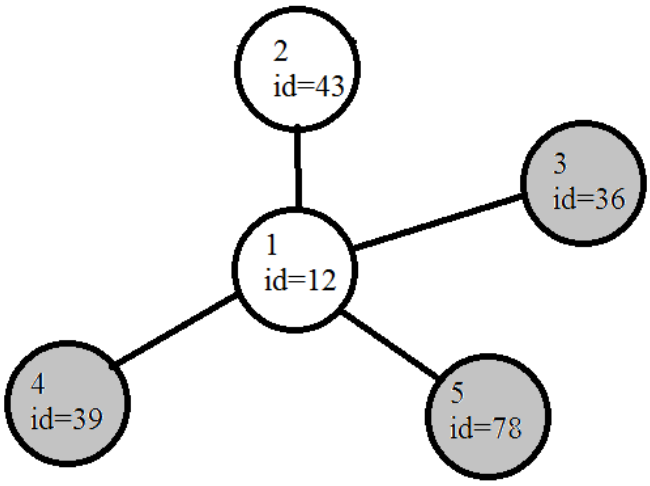


Рис. 5.16. Результат работы алгоритма
(1-5 этапы)

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 рис. 5.17.

Этап 10. Выполняем функцию *Get*(36, Z, $|Z|$, *flag*). Получаем $Z = \emptyset$, $|Z| = 0$, *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 $Z = \emptyset$, $|Z| = 0$, *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 $Z = \emptyset$, $|Z| = 0$, *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.

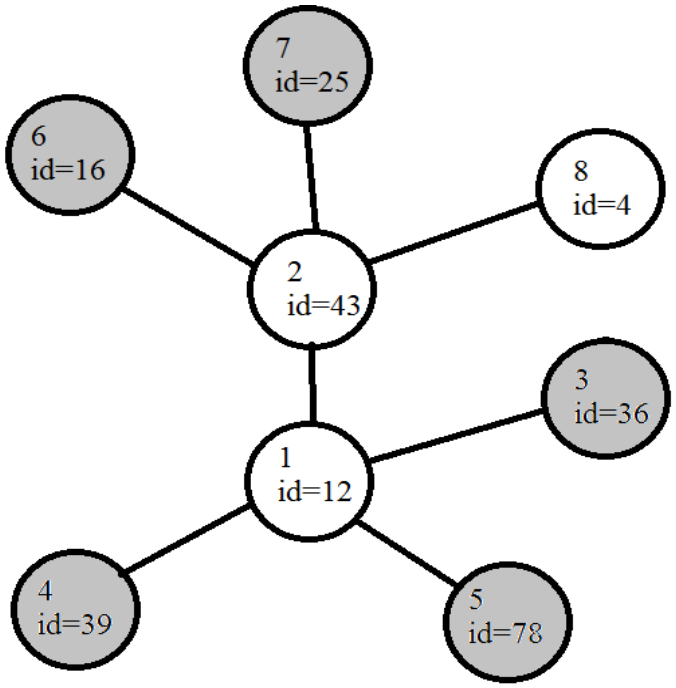


Рис. 5.17. Результат работы алгоритма (1-9 этапы)

Этап 16. Выполняем функцию *Get*(16, *Z*, |*Z*|, *flag*). Получаем $Z = \varnothing$, $|Z| = 0$, $flag = 0$. 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 $Z = \varnothing$, $|Z| = 0$, $flag = 0$. 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 $Z = \varnothing$, $|Z| = 0$, $E = \{e_{1,2}, e_{1,3}, e_{1,4}, e_{1,5}, e_{2,6}, e_{2,7}, e_{2,8}, e_{8,3}\}$.

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

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 рис. 5.18. Результат работы после каждого этапа отражены в табл. 5.1.

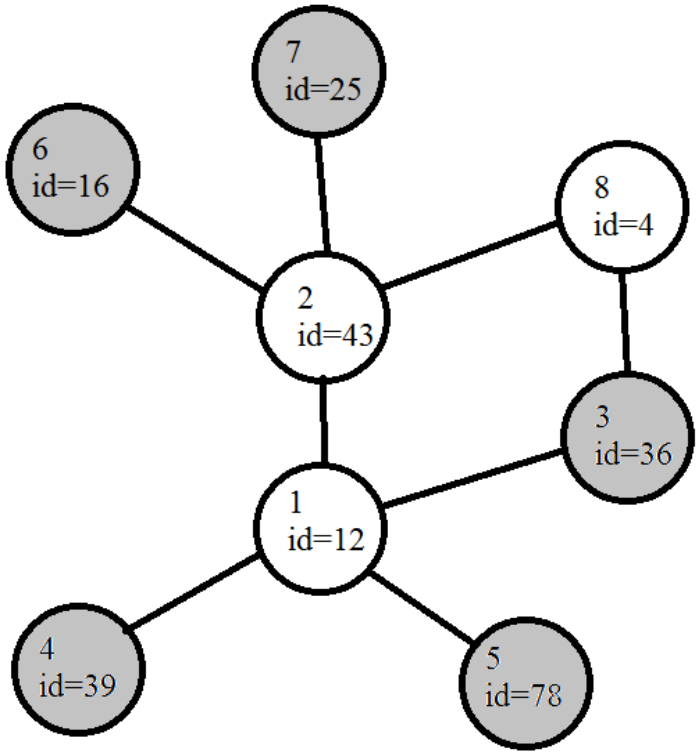


Рис. 5.18. Итоговый результат работы алгоритма

Таблица 5.1

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	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	16,25,4	2	5
9	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	16,25,4	3	5
10	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	3	5
11	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	4	5
12	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	4	5
13	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	5	5
14	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	5	5
15	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	6	5
16	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	6	5
17	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	7	5
18	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	7	5
19	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	Ø	8	5
20	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$	12,43,36,39,78,1625,4	43,36	8	5
21	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$, $e_{8,3}$	12,43,36,39,78,1625,4	Ø	8	5
22	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$, $e_{8,3}$	12,43,36,39,78,1625,4	Ø	8	8
23	w_1, w_2, w_8, u_3, u_4, u_5, u_6, u_7	$e_{1,2}$, $e_{1,3}$, $e_{1,4}$, $e_{1,5}$, $e_{2,6}$, $e_{2,7}$, $e_{2,8}$, $e_{8,3}$	12,43,36,39,78,1625,4	Ø	9	8

5.6. 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. 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. 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. Кластерный коэффициент узла вычисляется как отношение числа существующих связей между смежными узлами к значению общего количества возможных таких связей:

$$C_i = \frac{2n_i}{k_i(k_i - 1)}, \qquad (5.5)$$

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

Рассмотрим пример вычисления кластерного коэффициента для узла 1

(рис. 5.19). Сплошными линиями показаны существующие связи, пунктирными – потенциальные. Степень связности $k = 4$. Число возможных связей между его смежными узлами равно $k(k - 1)/2 = 4(4 - 1)/2 = 6$. Число существующих связей – 2. Кластерный коэффициент $C = 1/3$.

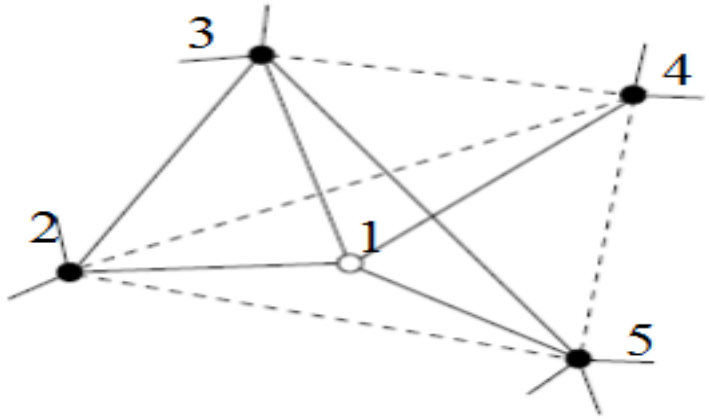


Рис. 5.19. Схематичный рисунок для определения кластерного коэффициента

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 формирования полного графа сети с учетом недоступной части представлена на рис. 5.20.

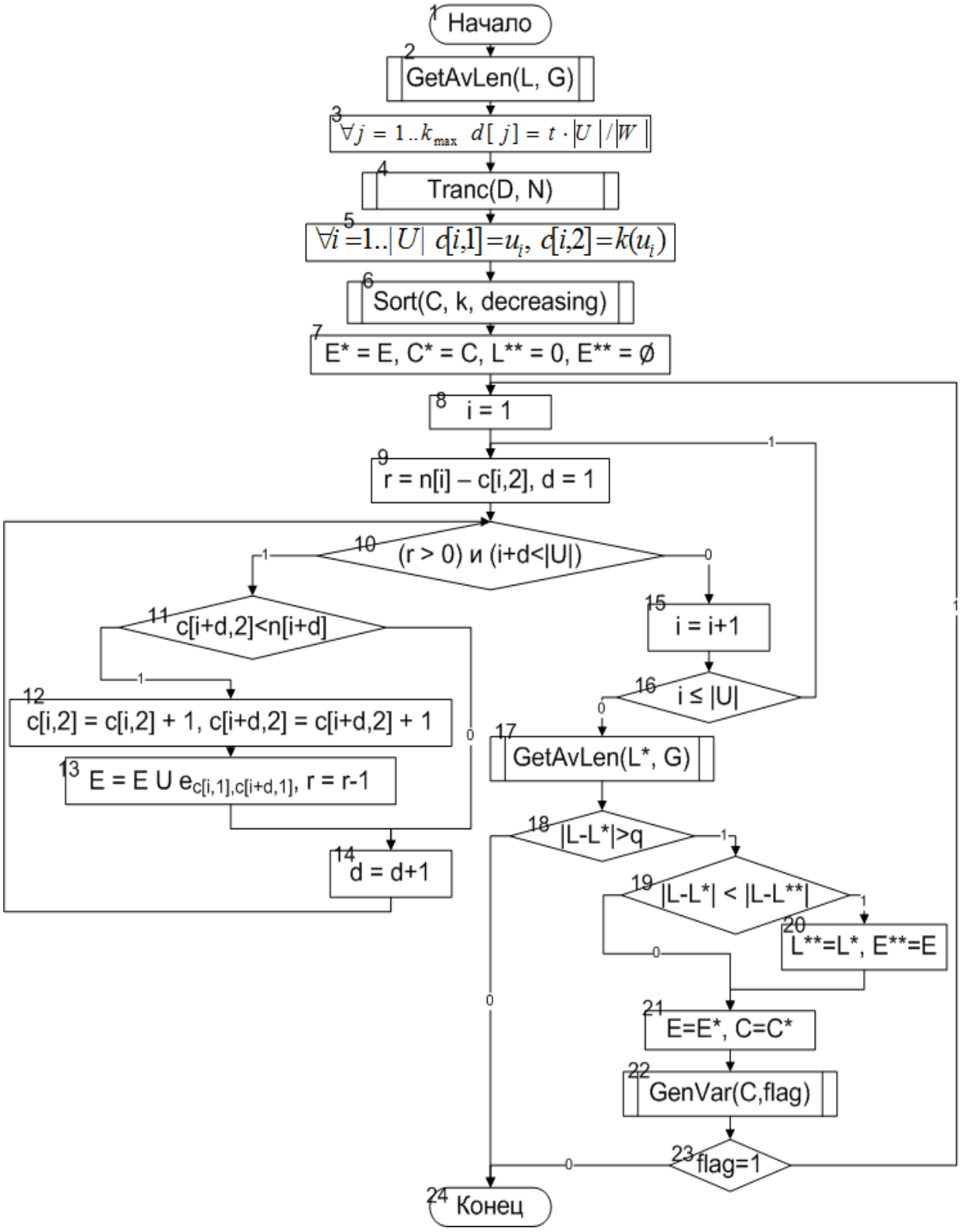


Рис. 5.20. Алгоритм генерации недоступной части сети

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 $E^{**} = \varnothing$.

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. На рис. 4.21 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 табл. 5.2.

Таблица 5.2

Распределение степеней связности по закрытым узлам

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

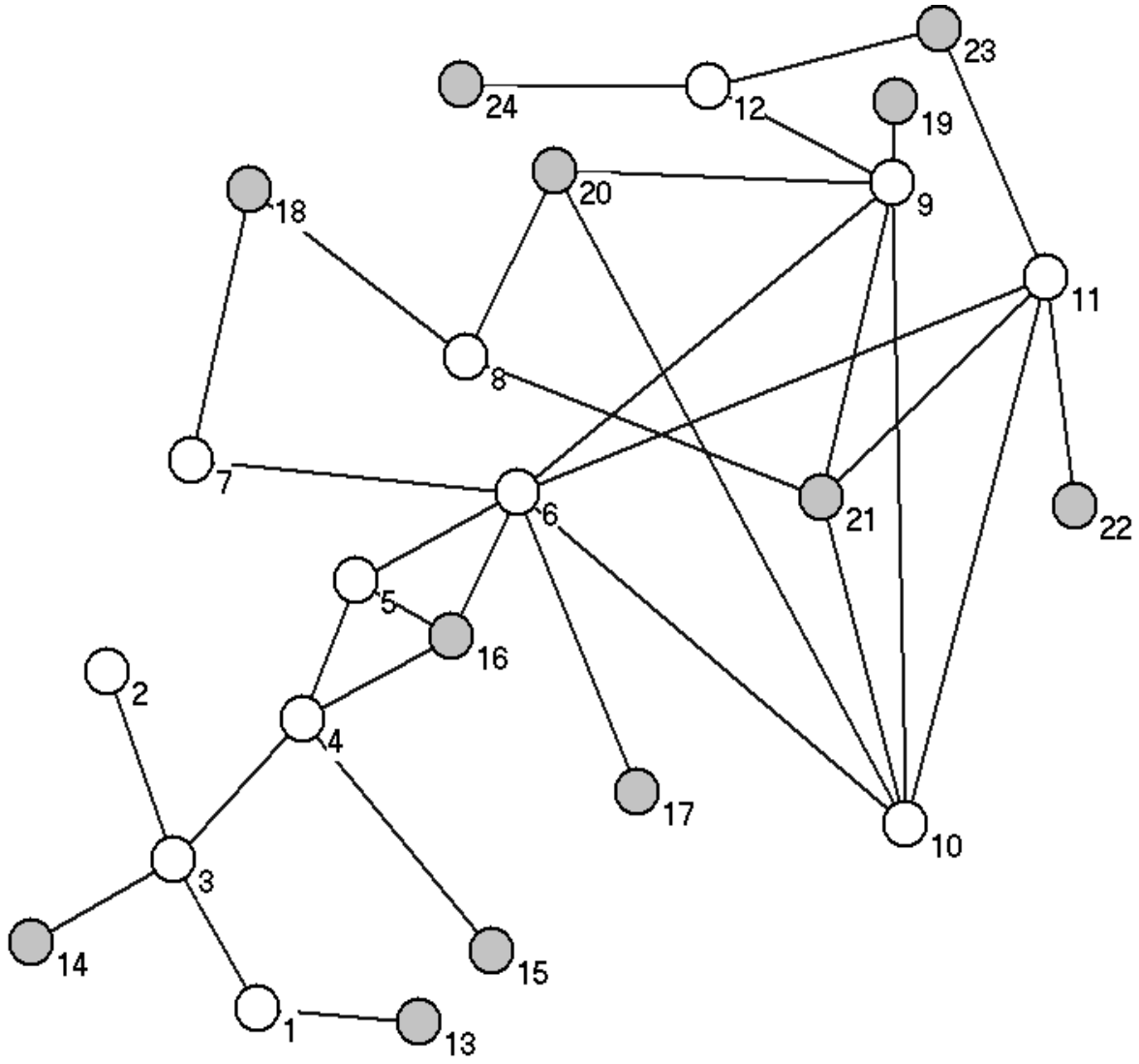


Рис. 5.21. Граф исходной сети

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 табл. 5.3.

Таблица 5.3

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 табл. 5.4.

Таблица 5.4

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*

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. 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 рис. 5.22 представлен полный граф ИТКС, полученный в результате работы алгоритма.

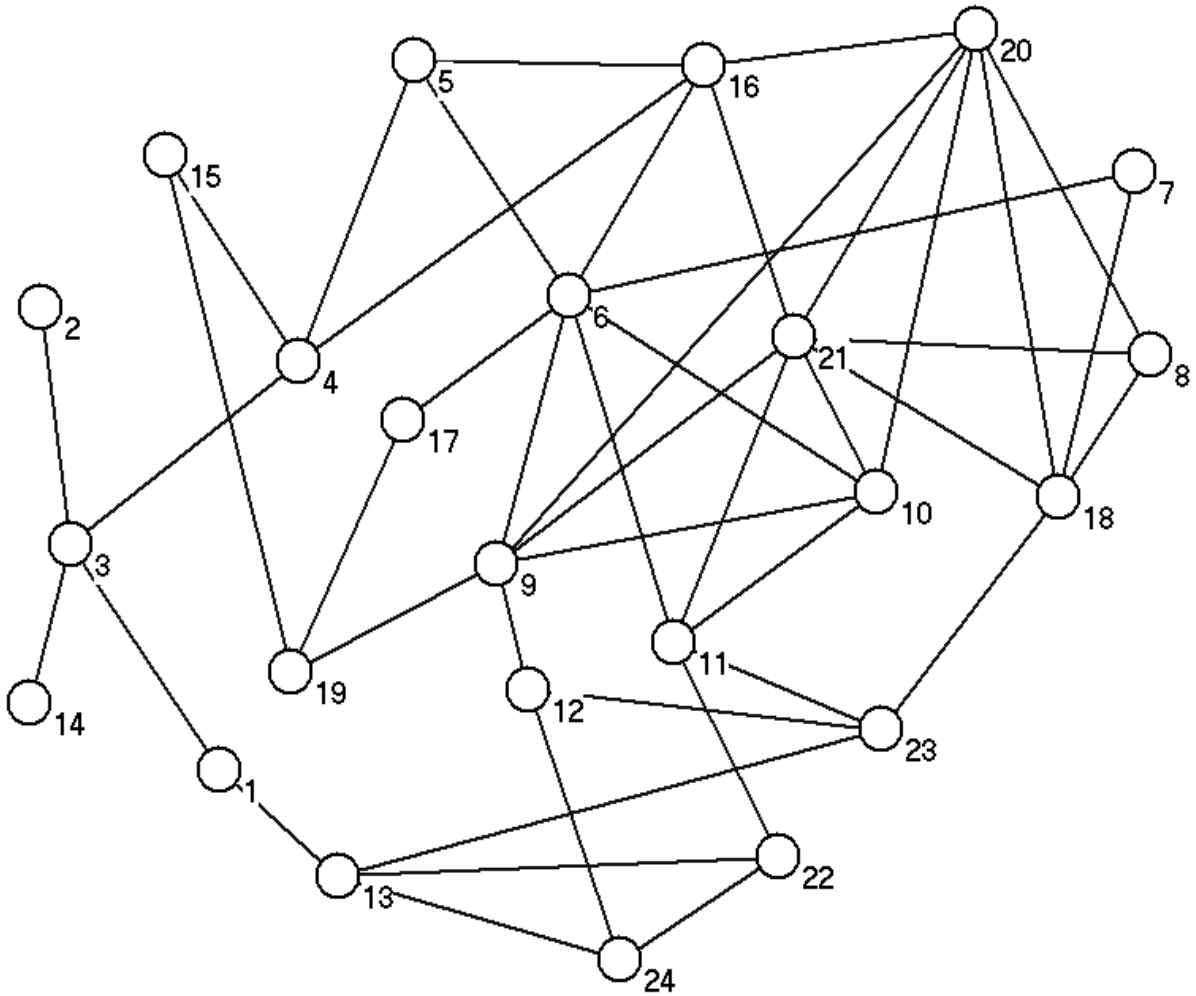


Рис. 5.22. Полный граф ИТКС

5.7. 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:

$$\varphi_i = \frac{k_i(C_i + 1)}{L_i}, \tag{5.6}$$

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 (4.4) – 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 $\varphi_i = 2(N - 1)$. Крайний случай (минимальное значение) – граф из двух узлов. При этом $k_i = 1$, $L_i = 1$ $C_i = 0$. Следовательно, в этом случае $\varphi_i = 1$.

2) С увеличением φ, возрастает уязвимость узла.

Подсчет коэффициента топологической уязвимости для всей сети осуществляется по формуле:

$$\varphi = \frac{k \cdot (C + 1)}{L}. \tag{5.7}$$

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 [А 4, 138];

3) коэффициент кластеризации, как правило, варьируется в значениях от 0,01 до 0,2 [40].

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 (5.5).

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 [21-26], а также тем, что данная сеть крупнейшая и, действительно, в целом менее уязвимая, чем сеть «ВКонтакте».

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 следующего вида:

Номер 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.

5.8. 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 узлов и 16270504 закрытых. Фрагмент выходного файла представлен на рис. 5.23.

На рис. 5.24 показан фрагмент (1000 узлов) полученной топологии, построенный с помощью ПО Pajek.

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 VisualStudio 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 BoostGraphLibrary) и распределенный (ParallelBoostGraphLibrary).

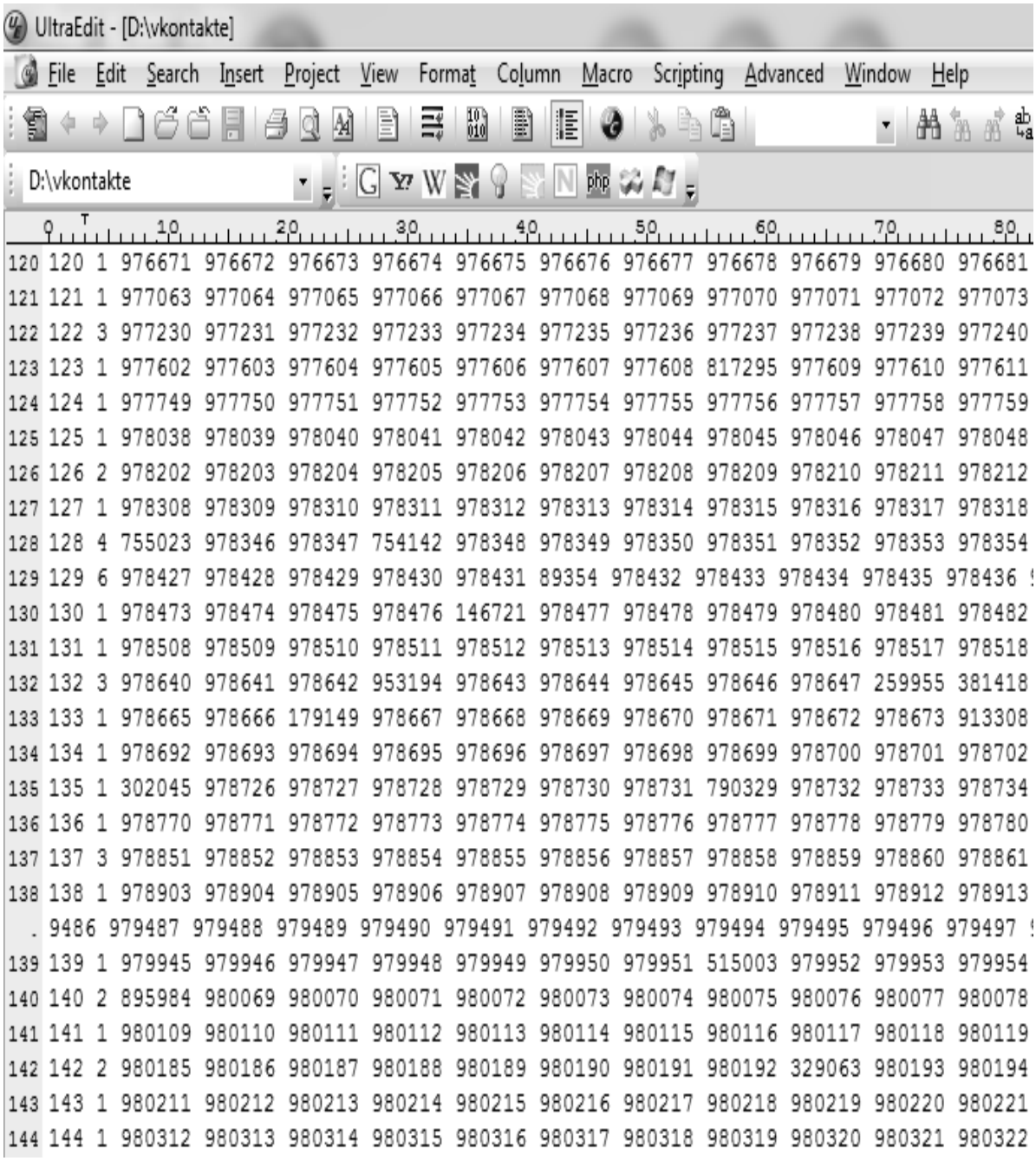


Рис. 5.23. Фрагмент выходного файла программы

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.

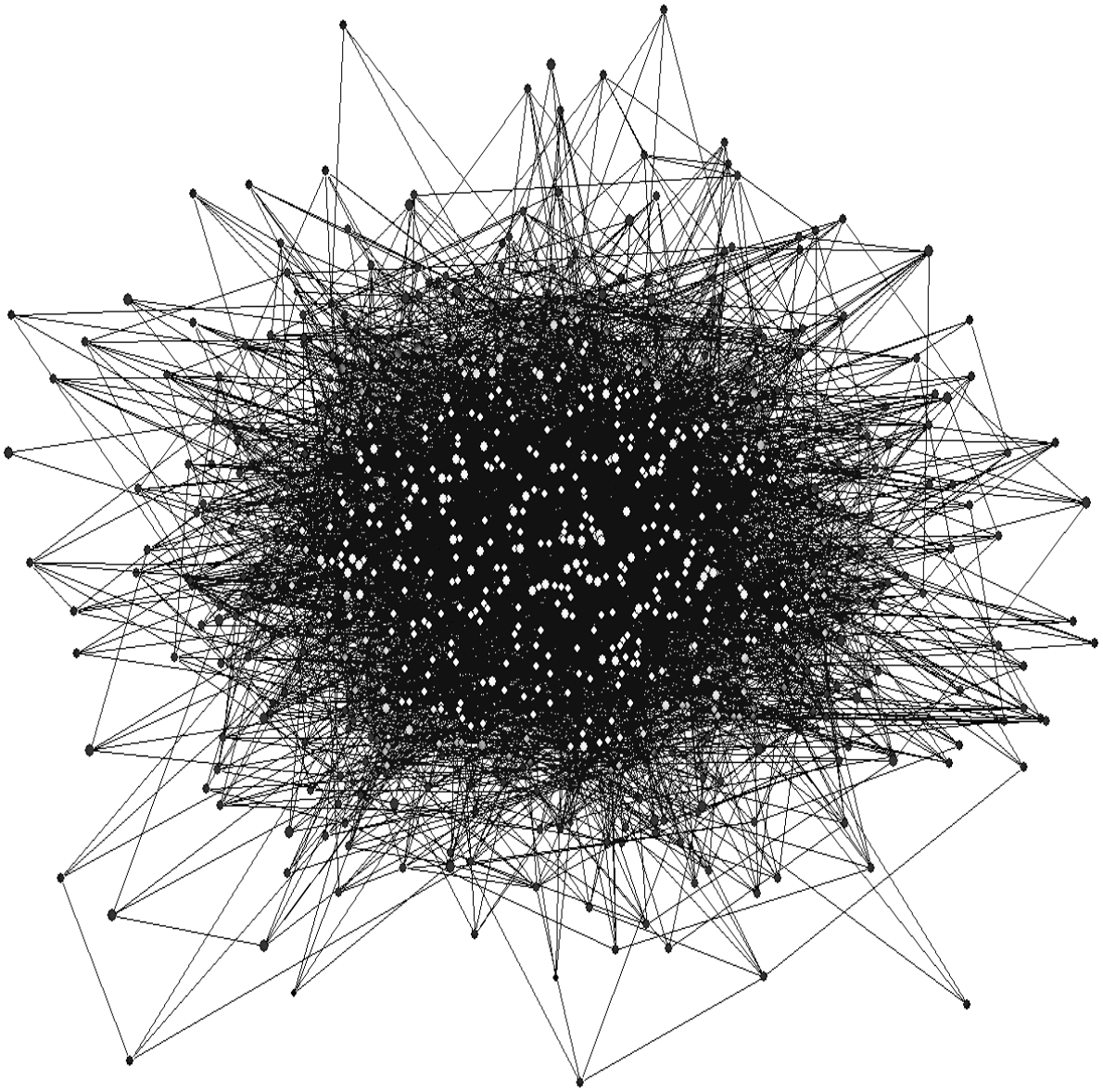


Рис. 5.24. Визуализированный фрагмент топологии

5.9. Экспериментальное исследование

Экспериментальные исследования проводились на двух фрагментах ИТКС. Первый (фрагмент социальной сети «ВКонтакте») получен в рамках данной научной работы, а второй (фрагмент из 16163521 узла социальной сети «Facebook») получен независимо американскими учеными Minas Gjoka, Maciej Kurant и др.

Анализ результатов моделирования

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 рис. 5.25 и 5.26 («ВКонтакте»).

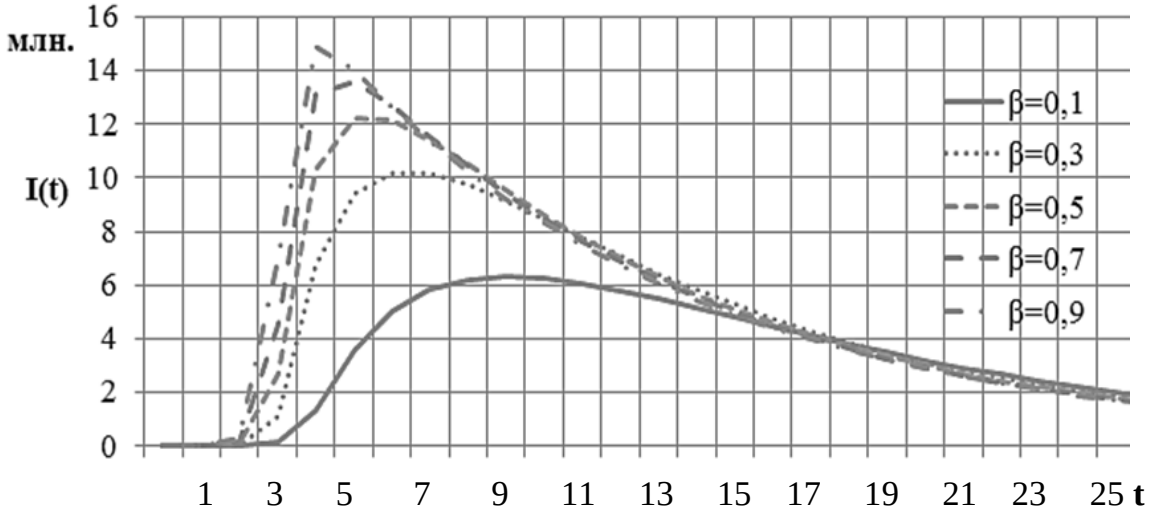


Рис. 5.25. Результаты моделирования с параметрами γ = 0,1, I_0 = 1, R_0 = 0

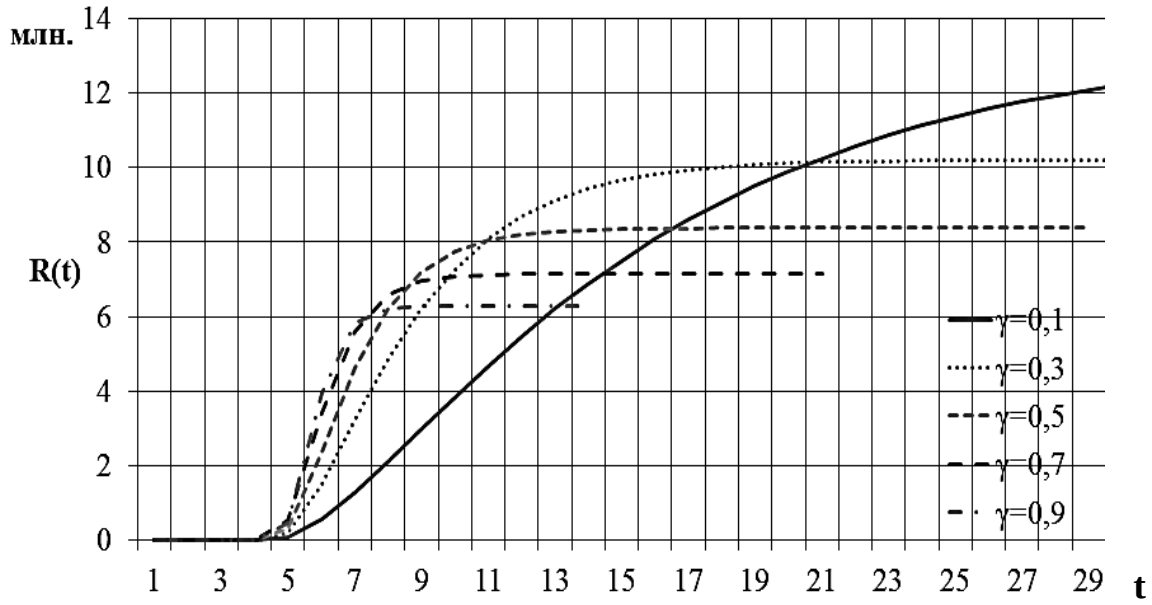
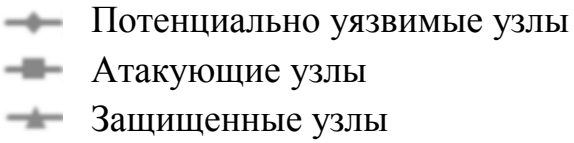


Рис. 5.26. Результаты моделирования с параметрами β = 0,2, I_0 = 1, R_0 = 0

При моделировании использовались следующие частные случаи: $\beta = 0{,}2$ и $\gamma = 0{,}8$, $\beta = 0{,}5$ и $\gamma = 0{,}5$. 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 моделирования распространения недостоверной информации на топологическом фрагменте социальной сети «ВКонтакте» приведены на рис. 5.27-5.38. Общая легенда для рисунков следующая:



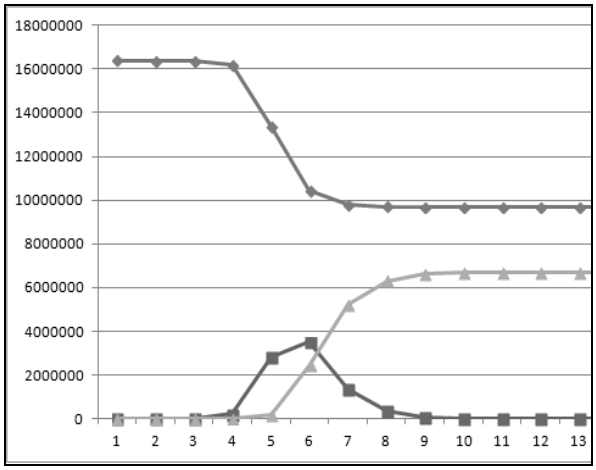


Рис. 5.27. Результаты эксперимента 1 ($\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0$)

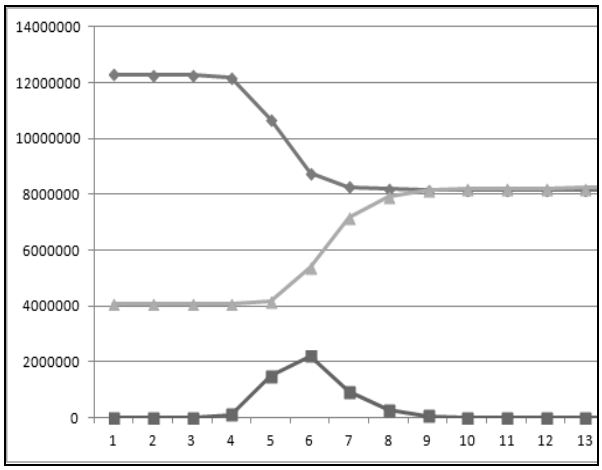


Рис. 5.28. Результаты эксперимента 2 ($\varphi = 200$, $\beta = 0{,}2$, $\gamma = 0{,}8$, $I_0 = 1$, $R_0 = 0{,}25N$)

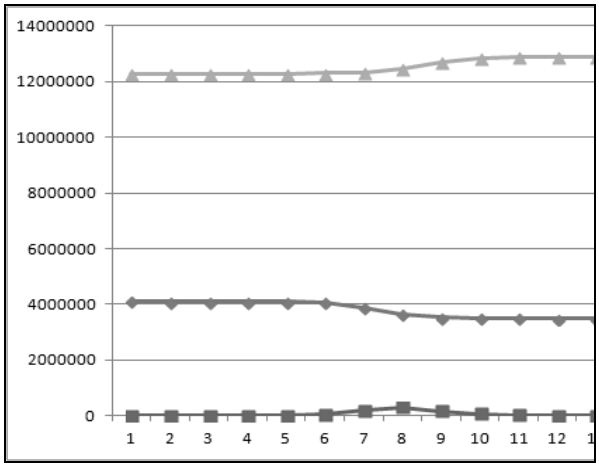


Рис. 5.29. Результаты эксперимента 3 (φ = 200, β =0,2, γ = 0,8, I_0 = 1, R_0 = 0,75N)

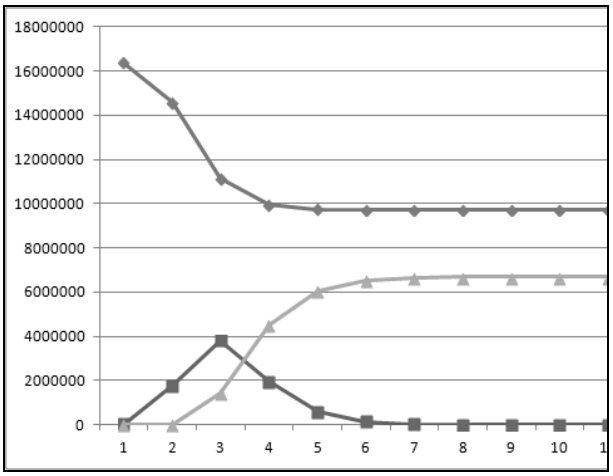


Рис. 5.30. Результаты эксперимента 4 (φ = 200, β = 0,2, γ = 0,8, I_0 = 0,001N, R_0 = 0)

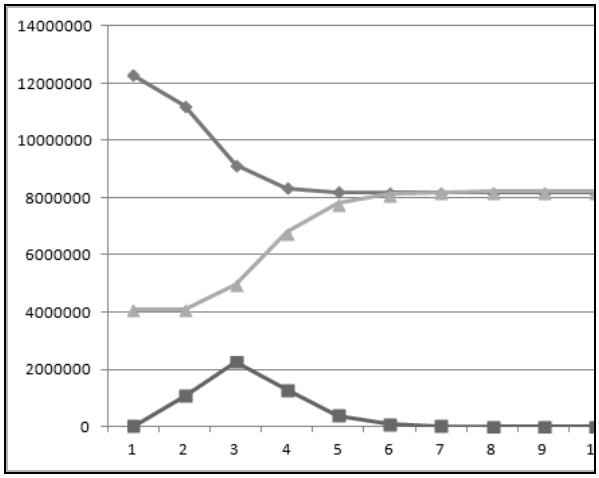


Рис. 5.31. Результаты эксперимента 5 (φ = 200, β = 0,2, γ = 0,8, I_0 = 0,001N, R_0 = 0,25N)

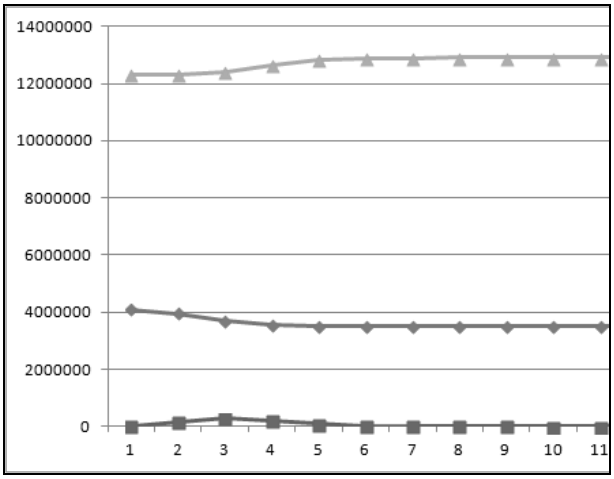


Рис. 5.32. Результаты эксперимента 6 (φ = 200, β = 0,2, γ = 0,8, I_0 = 0,001N, R_0 = 0,75N)

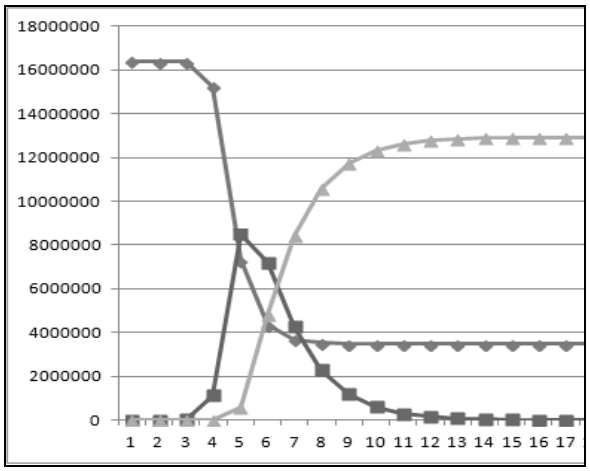


Рис. 5.33. Результаты эксперимента 7 (φ = 200, β = 0,5, γ = 0,5, I_0 = 1, R_0 = 0)

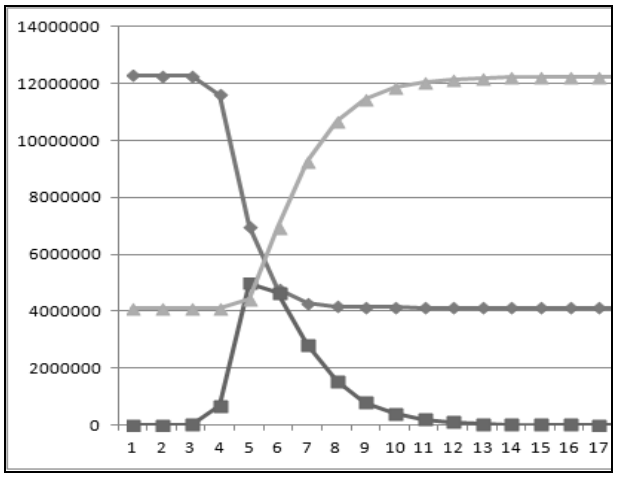


Рис. 5.34. Результаты эксперимента 8 (φ = 200, β = 0,5, γ = 0,5, I_0 = 1, R_0 = 0,25N)

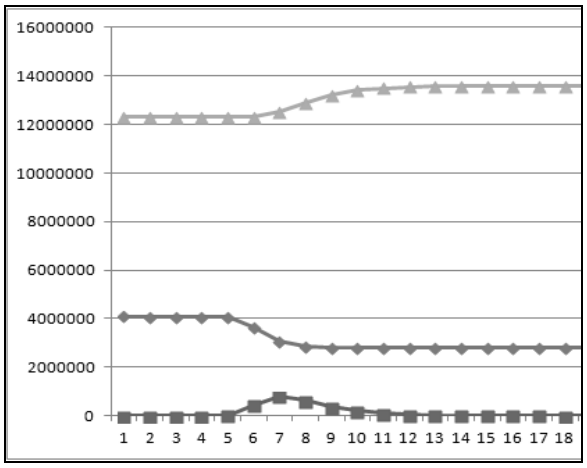


Рис. 5.35. Результаты эксперимента 9 (φ = 200, β =0,5, γ = 0,5, I_0 = 1, R_0 = 0,75N)

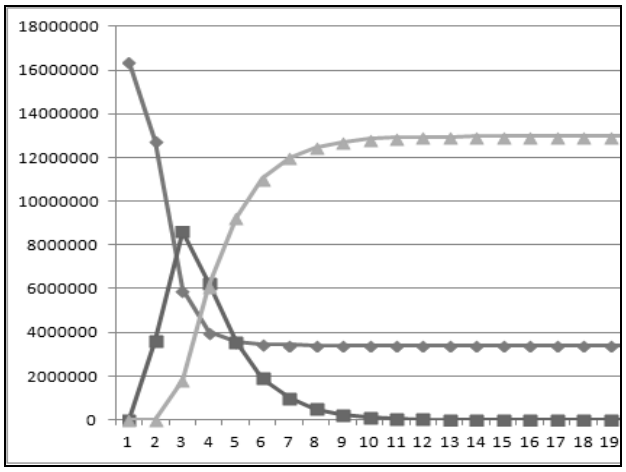


Рис. 5.36. Результаты эксперимента 10 (φ = 200, β =0,5, γ = 0,5, I_0 = 0,001N, R_0 = 0)

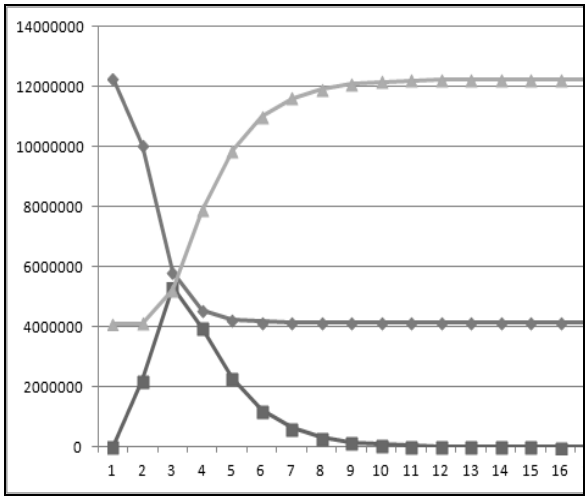


Рис. 5.37. Результаты эксперимента 11 (φ = 200, β = 0,5, γ = 0,5, I_0 = 0,001N, R_0 = 0,25N)

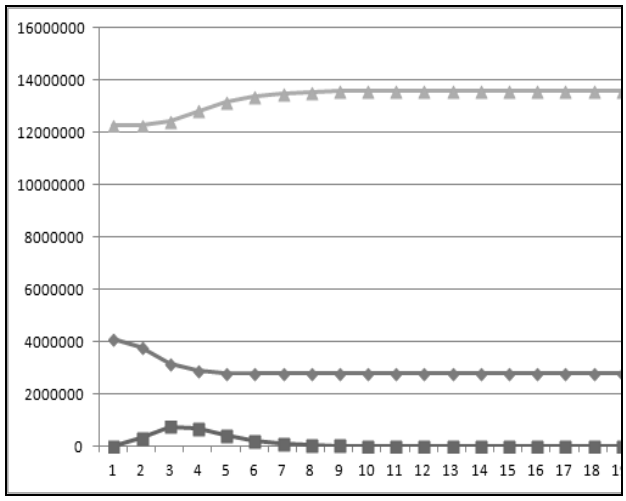


Рис. 5.38. Результаты эксперимента 12 (φ = 200, β = 0,5, γ = 0,5, I_0 = 0,001N, R_0 = 0,75N)

По результатам первых трех экспериментов можно сделать следующие выводы:

– уже один атакующий узел может вызвать «вспышку» в сети, даже при большом значении вероятности защиты; с ростом числа изначально защищенных узлов, максимальное число атакующих узлов падает (эксперименты 1-3, 10-12);

– при росте числа изначально атакующих узлов наблюдается «вспышка» уже на первых этапах (1-6 тики), параметр R_0 влияет на пик также как и в 1-3 экспериментах (эксперименты 4-6);

– при изменении параметров β и γ резко меняется характер угрозы, 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 – эксперименты 7-12).

Характер процесса распространения недостоверной информации сети Facebook такой же, как и на сети «ВКонтакте». Это факт указывает на то, что разные социальные сети имеют схожую топологию.

Выводы по главе

Разработан алгоритм реализации УгНдИ в ИТКС, основанный на характерах процессов, протекающих в реальных условиях. Создана имитационная модель УгНдИ 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 УгНдИ от топологической уязвимости сети.

Разработана аналитическая модель УгНдИ 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 %.

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.

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

Список библиографических ссылок

1. Абрамов К.Г. Влияние перколяционного кластера на распространение нежелательной информации в социальных медиа // Проблемиінформатики і моделювання. Тезиси одинадцятоі міжнародноі науковотехнічноі конференціі. Харків-Ялта, 2011. С. 4-5.
2. Абрамов К.Г., Монахов Ю.М. Моделирование распространения нежелательной информации в социальных медиа // 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.
3. Абрамов К.Г., Монахов Ю.М. Стохастические модели распространения нежелательной информации в социальных сетях // Сборник научных трудов SWorld. Материалы международной научно-практической конференции «Современные проблемы и пути их решения в науке, транспорте, производстве и образовании '2011». Одесса: Черноморье, 2011. №4. С. 42-46.
4. Абрамов К.Г., Монахов Ю.М. Модели распространения вредоносных программ в топологически гетерогенных социальных сетях // Труды НТС. Комитет по информатизации, связи и телекоммуникациям Администрации Владимирской области. 2010.
5. Абрамов К.Г., Монахов Ю.М. Некоторые аспекты безопасности Интернета в условиях инфраструктуры web 2.0 // Труды X Российской научно-технической конференции "Новые информационные технологии в системах связи и управления". Калуга: Изд. "Ноосфера", 2011. С. 593-595.
6. Абрамов К.Г., Монахов Ю.М., Никиташенко А.В. К вопросу об уточнении моделей распространения нежелательной информации в социальных сетях Интернета // Информационные системы и технологии ИСТ-2011: материалы XVII международной научно-технической конференции. Н. Новгород: Электронное издание, 2011. 149 с.

7. Абрамов К.Г., Монахов Ю.М. Распространение нежелательной информации в социальных сетях Интернета // 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.

8. Алешин Л.И. Защита информации и информационная безопасность. М.: МГУК, 1999. 176 с.

9. Биячуев Т.А. Безопасность корпоративных сетей: учеб. пособие / под ред. Осовецкого Л.Г. СПб.: СПбГУ ИТМО, 2004. 161 с.

10. Бреер В.В. Стохастические модели социальных сетей // Управление большими системами. 2009. № 27. С. 169-204.

11. Брэгг Р., Родс-Оусли М., Страссберг К. Безопасность сетей. Полное руководство. М.: Эком, 2006. 912 с.

12. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. М.: Издат. физико-математической литературы, 2010. 228 с.

13. Монахов Ю.М., Медведникова М.А. Аналитическая модель дезинформированности узла социальной сети // Всероссийский конкурс научно-исследовательских работ студентов и аспирантов в области информатики и информационных технологий в рамках всероссийского фестиваля науки. Белгород, 2011. Т. 1. С. 595-597.

14. Монахов Ю.М., Медведникова М.А. Аналитическая модель дезинформированности узла социальной сети // ИММОД-2011. СПб., 2011. Т. II. С. 178-180.

15. Монахов Ю.М., Медведникова М.А., Абрамов К.Г., Бодров И.Ю. Аналитическая модель дезинформированности узла социальной // Комплексная защита объектов информатизации: Труды НТС. Владимир: ВлГУ, 2012. URL: http://sntk.vlsu.ru (дата обращения: 18.09.2015)

16. Программное обеспечение Pajek / VladimirBatagelj, AndrejMrvar. URL: http://pajek.imfm.si/doku.php (дата обращения: 18.09.2015)

17. Amaral L.A.N., Scala A., Barthelemy M., Stanley H.E. Classes of small-world networks // Proceedingsof the National Academy of Sciences of the United States of America. 2000. V.97(21). P. 11149-11152.

18. Chwe M.S. Communication and Coordination in Social Network //

Review of Economic Studies. 2000. V. 67. P.1-16.

19. Dorogovtsev S.N., Mendes J.F.F., Samukhin A.N. Giant strongly connected component of directed networks // Phys. Rev. E. 2001. V. 64. № 2.

20. Frauenthal J.C. Mathematical Models in Epidemiology. New York: Springer-Verlag, 1980. 335 p.

21. Gjoka M., Kurant M., Butts C.T., Markopoulou A.A Walk in Facebook: Uniform Sampling of Users in Online Social Networks. 2011. URL: http://arxiv.org/abs/0906.0060 (дата обращения: 18.09.2015).

22. Gjoka M., Kurant M., Butts C.T., Markopoulou A. Multigraph Sampling of Online Social Networks // IEEE J. Sel. Areas Commun. on Measurement of Internet Topologies. 2011.

23. Gjoka M., Kurant M., Butts C.T., Markopoulou A. Walking on a Graph with a Magnifying Glass: Stratified Sampling via Weighted Random Walks // Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems. ACM, 2011. P. 281-292.

24. Gjoka M., Kurant M., Butts C.T., Markopoulou A. Walking in Facebook: A Case Study of Unbiased Sampling of OSNs // INFOCOM–2010 Proceedings IEEE. 2010. P.1-9.

25. Gjoka M., Kurant M., Wang Y., Almquist Z.W., Butts C.T., Markopoulou A. Coarse-Grained Topology Estimation via Graph Sampling. 2011. URL: http://arxiv.org/abs/1105.5488 (дата обращения: 18.09.2015).

26. Gjoka M., Sirivianos M., Markopoulou A., Yang X. Poking facebook: characterization of osn applications // Proceedings of the first workshop on Online social networks. 2008. P. 31-36.

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

28. Hethcote H.W. The Mathematics of Infectious Diseases // SIAM REVIEW. 2000. V. 42. № 4. P. 599-653.

29. Janky B., Takacs K. Social Control, Participation in Collective Action and Network Stability // HUNNET Working Paper. 2002. URL: http://www.academia.edu/9263540/Social_Control_Participation_in_Collective_Action_and_Network_Stability (дата обращения: 18.09.2015).

30. Kephart J.O., White S.R. Directed-Graph Epidemiological Models of Computer Viruses // Proceedings of the IEEE Computer Society Symposium

on Research in Security and Privacy.1991. P. 343-359.

31. Kuperman M., Abramson G. Small world effect in and epidemiological model // Physical Review Letters. 2001. V. 86. № 13.

32. Leskovec J., Adamic L.A., Huberman B.A. The Dynamics of Viral Marketing. 2008. URL: http://snap.stanford.edu/class/cs224w-readings/leskovec07viral.pdf (дата обращения: 18.09.2015).

33. Leveille J. Epidemic Spreading in Technological Networks // Information Infrastructure Laboratory HP Laboratories Bristol. 2002. P. 65-76.

34. Newman M.E.J. The spread of epidemic disease on networks // Physical Review E. 2002. P. 16-28. URL: http://arxiv.org/pdf/cond-mat/0205009.pdf (дата обращения: 18.09.2015).

35. Pastor-Satorras R., Vespignani A. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks // Journal of Theoretical Biology. 2005. V. 235. P. 275-288.

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

37. Roberts M.G., Heesterbeek J.A.P. Mathematical models in epidemiology // In JA. Filar (Ed.) *Mathematical Models.* Oxford: EOLSS Publishers Ltd, 2004. V. III.

38. Tarnow E. Like Water and Vapor, Conformity and Independence in the Large Group // Behavioral Science. 1996. V. 41. P. 136-151.

39. Tictrac [Electronic resource]. 2015. URL: https://www.tictrac.com (дата обращения: 18.09.2015).

40. Ugander J., Karrer B., Backstrom L., Marlow K. The Anatomy of the Facebook Social Graph. 2011. URL: http://arxiv.org/abs/1111.4503 (дата обращения: 18.09.2015).

41. Williamson M.M., Léveillé J. An epidemiological model ofvirus spread and cleanup // Information Infrastructure Laboratory HP Laboratories Bristol HPL. 2003. URL: http://www.hpl.hp.com/techreports/2003/HPL-2003-39.pdf (дата обращения: 18.09.2015).

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

ЗАКЛЮЧЕНИЕ

Сформулируем основные результаты исследования:

1. Разработана онтология понятийного аппарата, методического обеспечения и признакового пространства при определении достоверности информации, позволяющая уточнить / согласовать существующий понятийный аппарат в данной предметной области. С практической точки зрения онтология в комбинации с многоагентным подходом и сетями «потребности-возможности» (ПВ-сети) может стать решением оптимизационной задачи по подбору такого минимального набора конкретных средств для обеспечения достоверности информации в ИТКС, который обеспечит перекрытие всех потенциальных угроз, сохранить требуемый уровень ее (информации) качества.

2. Предложен концептуальный подход к обеспечению достоверности информации в информационно-телекоммуникационных системах, функционирующих в условиях информационного противодействия. Описана новая модель управления процессом обеспечения достоверности, отличающаяся учетом работы ИТКС в условиях дестабилизирующих факторов, активного противодействия, мониторинга, динамического изменения уровня достоверности источников информации, ограничений ресурсов различных классов в ИТКС. Структура и параметры модели закладываются подмножествами вариантов (многовариантная модель). Настройка данной системы под особенности конкретного предприятия позволит получить инструмент для прогноза развития ситуации и оценки рисков снижения достоверности информации в ИТКС. Кроме того, предложенная модель может стать частью автоматизированной системы мониторинга и управления процессами обеспечения достоверности информации в ИТКС конкретных предприятий.

3. Разработана общая модель оценки показателей достоверности информации в ИТКС в условиях информационных воздействий с учетом решения сопутствующих задач оценки рисков и экономической эффективности мероприятий по повышению достоверности. В основе модели алгоритмы проведения экспертизы параметров ИТКС для определения текущего уровня достоверности информации, процедуры получения числовых

значений количественных и качественных параметров ИТКС, являющихся исходными данными для оценки достоверности, методики расчета информационных рисков, защищенности, вероятностей возникновения и устранения угроз безопасности, методика оценки экономической эффективности мероприятий по обеспечению достоверности. Результаты экспериментального исследования ИТКС промышленного предприятия показали адекватность и эффективность предложенных средств анализа и обеспечения достоверности производственных информационных ресурсов.

4. Показано, что наиболее существенным фактором, снижающим доступность информационных ресурсов и процессов в ИТКС, являются информационные атаки злоумышленников, среди которым наибольший вред наносят распределенные атаки «отказ в обслуживании». Теоретическое и экспериментальное исследование подтвердило гипотезу о том, что в условиях атак данного вида агрегированный сетевой трафик становится персистентным и самоподобным. На основе данного подхода разработаны механизмы раннего обнаружения аномального поведения ИТКС, вызванного начавшейся атакой. На основе анализа системных характеристик и процессов ИТКС, а также возможностей злоумышленников разработаны модели и алгоритмы распределенной атаки «отказ в обслуживании». В качестве теоретической основы прогноза угрозы предложено использовать смешанную модель авторегрессии и скользящего среднего агрегированного трафика ИТКС. Примеры апробирования механизмов и средств раннего обнаружения аномального поведения реальных ИТКС дают возможность констатировать адекватность и функциональность основных теоретических построений и разработанных на их основе программных средств.

5. Разработаны аналитические и имитационные модели процессов реализации угрозы распространения недостоверной информации в ИТКС, а также формирования ее топологии, как крупномасшабного графа, состоящего из открытых и закрытых узлов. Имитационная модель угрозы построена с учетом топологических характеристики сети, а также особенностей информационного взаимодействия абонентов как человеко-машинных систем. С ее помощью проведены эксперименты, результаты которых позволили уточнить модели Кермака-Маккендрика применительно для данной предметной области. Релевантность результатов аналитиче-

ского прогноза подтверждена серией экспериментов на известных социальных сетях интернета. При этом погрешность для процесса защиты составила не более 10 %, для процесса атаки – не более 15 %. Примеры практического применения механизмов прогнозирования угрозы распространения недостоверной информации в ИТКС дают основание констатировать адекватность и функциональность основных теоретических построений и разработанных на их основе алгоритмических и инструментальных средств.

Научное издание

МОНАХОВ Михаил Юрьевич
МОНАХОВ Юрий Михайлович
ПОЛЯНСКИЙ Дмитрий Александрович
и др.

МОДЕЛИ ОБЕСПЕЧЕНИЯ
ДОСТОВЕРНОСТИ И ДОСТУПНОСТИ ИНФОРМАЦИИ
В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ
СИСТЕМАХ

Монография

Печатается в авторской редакции

Подписано в печать 07.10.15.
Формат 60×84/16. Усл. печ. л. 12,09. Тираж 500 экз.
Заказ
Издательство
Владимирского государственного университета
имени Александра Григорьевича и Николая Григорьевича Столетовых
600000, г. Владимир, ул. Горького, 87.