

На правах рукописи

Мартьянов Евгений Александрович

**Исследование и разработка методик оценки
защищенности информационных объектов от
потенциальных нарушителей**

Специальность 05.13.19 – Методы и системы защиты информации,
информационная безопасность

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

Москва – 2019

Работа выполнена на кафедре информационной безопасности Московского государственного университета имени М.В. Ломоносова.

Научный руководитель: **Макаров-Землянский Николай Викулович**

доктор технических наук, ведущий научный сотрудник лаборатории Компьютерной безопасности Научно-исследовательского вычислительного центра ФГБОУ ВО «Московский государственный университет имени М.В. Ломоносова»

Официальные оппоненты: **Зегжда Петр Дмитриевич**

доктор технических наук, профессор кафедры «Информационная безопасность компьютерных систем» ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого»

Лукин Владимир Николаевич

кандидат физико-математических наук, доцент кафедры № 806 «Вычислительная математика и программирование» ФГБОУ ВО «Московский авиационный институт (национальный исследовательский университет)»

Ведущая организация:

Федеральное государственное бюджетное учреждение науки «Институт проблем управления имени В.А. Трапезникова Российской академии наук» (ИПУ РАН)

Защита состоится «20» марта 2019 г. в 15 часов 00 мин. на заседании диссертационного совета Д 002.073.02 при Федеральном исследовательском центре «Информатика и управление» Российской академии наук по адресу 119333, г. Москва, ул. Вавилова, д. 44, к. 2.

С диссертацией можно ознакомиться в библиотеке Федерального исследовательского центра «Информатика и управление» Российской академии наук по адресу 119333, г. Москва, ул. Вавилова, д. 44, к. 2 и на сайте www.frccsc.ru.

Автореферат разослан «_____» _____ 2019 г.

Ученый секретарь
диссертационного совета



Р.В. Разумчик

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Необходимость и актуальность данной работы обусловлена следующим. Любая современная организация в той или иной степени использует информационные технологии в своей работе и с каждым годом их роль только возрастает. Разумеется, всегда находятся отдельные люди или коллективы, которые хотят заработать или занять более выгодное положение на рынке за счет чужих усилий и знаний. Отсюда вытекают различные атаки на корпоративную собственность, которая всё чаще представлена в цифровом виде. Подобные атаки могут быть направлены на организацию извне, они еще называются хакерскими атаками, и именно этот вид атак, согласно исследованиям, наиболее распространен на текущий день. Однако, от них есть определенные способы защиты, вплоть до отключения критических областей информационной системы (ИС) от внешней сети. Но есть еще один путь для злоумышленника, который так «просто» не закрыть: это атаки изнутри, совершаемые сотрудниками с легальным доступом к этой информации. Именно такие атаки наиболее болезненны для организации, ведь совершающие их сотрудники, или инсайдеры, могут украсть или модифицировать информацию таким образом, что никто своевременно даже не узнает об этом. Они наносят серьезный ущерб репутации компании, могут приводить к перебоям в работе и серьезным финансовым убыткам, или даже причинять вред жизни и здоровью людей.

Инсайдеры представляют серьезную угрозу для организаций по ряду причин. Во-первых, они могут иметь авторизованный доступ к ИС и являться ее пользователями. Они могут быть знакомы с ее структурой и средствами защиты, могут их обходить или создавать новые скрытые пути доступа для других. Во-вторых, количество нелегальных действий крайне мало по сравнению с общим количеством действий всех пользователей в ИС, из-за чего обнаружить их и правильно отреагировать очень сложно.

Угроза инсайдера всегда была комплексной проблемой и никогда не решалась каким-то одним средством, например, настройками доступов к данным в информационной системе. Она требовала и требует комплексного подхода и ежедневной работы всех ответственных лиц. Если не удалось предотвратить атаку со стороны инсайдера, то важно не только ее своевременно обнаружить, но и правильно отреагировать на нее, чтобы минимизировать потенциальный ущерб для организации.

По данным исследований CERT^{1 2} и ISACA³ количество инцидентов, виновной которым являются бывшие или действующие сотрудники, т.е. инсайдеры,

¹ 2014 State of Cybercrime Survey Presentation. – <https://resources.sei.cmu.edu/library/asset-view.cfm?assetID=298318>. – Accessed: 2017-05-01

² 2016 State of Cybercrime Survey Presentation. – <https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=499782>. – Accessed: 2017-05-01

³ State of Cybersecurity Implications for 2016. – http://www.isaca.org/cyber/Documents/state-of-cybersecurity_res_eng_0316.pdf. – Accessed: 2017-05-01

составляет около трети от всех атак на организации, уступая только хакерским атакам. При этом суммы причиненного ущерба от внутренних и внешних атак примерно равны. Добавив к этому факт, что общее количество подобных атак с годами только увеличивается, можно сделать очевидный вывод о необходимости качественных средств защиты от зловердных действий инсайдера.

Теоретическую базу исследования составили работы известных российских ученых: Ю.В. Прохоров, Ю.А. Розанов, А.А. Грушо, А.Н. Ширяев, В.А. Ильин, В.П. Строгалев, а также зарубежных ученых: D. Cappelli, A. Moore, T. Senator, H. Goldberg, G. Zipf, J. Kleijnen, R. Shannon, которые внесли большой вклад в получение основных результатов в области исследования угрозы инсайдера и защищенности информационных объектов от него и смежных областях. В работах рассматривались различные модели угроз инсайдеров, разрабатывались методы их обнаружения. В ходе проведения научных работ по исследуемой области был создан уникальный научный задел, используемый и в настоящее время. Тем не менее, проведенный в работе обзор разработанных подходов в рамках крупнейших в данной области проектов CERT и ADAMS показал, что на сегодняшний день наиболее исследованный способ выявления инсайдера среди множества пользователей информационной системы основывается на поиске аномалий в его поведении либо по сравнению с другими пользователями, либо с его собственным поведением за другие периоды времени, когда атака не совершалась. Для такого анализа необходим большой объем наблюдений за действиями всех пользователей в рамках исследуемой информационной системы в течение всей их работы, потому что качество сделанных выводов напрямую зависит от возможностей средств мониторинга и детализированности регистрируемых событий. Сегодня существует несколько программных средств для сбора такого рода информации, например Raytheon SureView ⁴, IBM System G ⁵, а также отечественное ПО РАМС ИБ ⁶. Однако использование программных комплексов такого уровня обходится организациям достаточно дорого и требует значительных дополнительных вычислительных ресурсов, поэтому далеко не все могут позволить себе такие расходы, учитывая тот факт, что собранная информация носит конфиденциальный характер и требует защиты.

В результате, рассматриваемая проблема актуальна для организаций всех уровней, как коммерческих, так и государственных, деятельность которых связана с информационными ресурсами.

В диссертационной работе проведены новые исследования в области обнаружения инсайдера в информационной системе организации. Использование разработанного в диссертации нового методического аппарата позволит сократить расходы организации на обнаружение инсайдера, при этом повысить точ-

⁴ SureView. – http://www.raytheoncyber.com/news/feature/sureview_suite.html. – Accessed: 2017-05-01.

⁵ System G: Developed Graph Computing Industry Solutions. – <http://systemg.research.ibm.com/solutions.html>. – Accessed: 2017-05-01.

⁶ Система регистрации, анализа и мониторинга событий информационной безопасности. – <http://www.ssec.ru/2014-rams-ib.htm>. – Accessed: 2017-05-01.

ность такого обнаружения. Полученные в диссертационной работе результаты будут способствовать более полному решению проблемы обнаружения инсайдера и минимизации возможного ущерба организации от его атаки. Диссертационный материал и содержащиеся в нем выводы и предложения могут быть использованы в качестве основы для проведения дальнейших исследований и совершенствования имеющегося программного обеспечения, целью которого является выявление инсайдера.

Объект исследования – методики и методологии обнаружения признаков инсайдера в информационной системе.

Предмет исследования – методы выявления инсайдера в информационной системе, основанные на теории вероятности и математической статистике.

Целью диссертационной работы является разработка статистических методов с наименьшей вероятностью ложного срабатывания для выявления инсайдеров среди пользователей информационной системы, представляющей собой корпоративное хранилище данных.

Для достижения поставленной цели в работе поставлены и решены следующие задачи:

1. Анализ предметной области, включающий в себя исследование угрозы инсайдера в организации, изучение существующих подходов и методик для выявления его действий в информационной системе.
2. Построение модели работы пользователей с корпоративным хранилищем данных, в рамках которой решается задача выявления зловредных действий инсайдера.
3. Разработка метода, позволяющего выявить факт сбора ценной для организации информации инсайдером в описанной модели.
4. Получение оценки возможности выявления инсайдера среди пользователей корпоративного хранилища данных для разработанного метода в зависимости от объема наблюдений и различных сценариев поведения инсайдера.
5. Разработка алгоритма и инструментального программного средства выявления сбора пользователем избыточной для его функциональных обязанностей информации.

Основные научные результаты, выносимые на защиту:

1. Новая формальная модель работы пользователя и инсайдера в информационной системе организации, представляющей собой корпоративное хранилище данных, в рамках которой доказана независимость полученных результатов от количества пользователей и свойств тех случайных

процессов, которыми описывается их работа. Построенная модель позволила использовать теорию запретов вероятностных мер в конечных пространствах для исследования свойств предложенных методов выявления инсайдера. Основным результатом является метод нахождения запретов для данного семейства процессов.

2. Новый метод выявления потенциального инсайдера из множества пользователей, которые работают с хранилищем данных, основанный на выявлении факта сбора им избыточной информации, для которого доказано, что во введенной модели вероятность успешного обнаружения на конечном шаге равна единице, а ложные тревоги исключены.
3. Разработанное программное средство с возможностью имитационного моделирования работы пользователей с хранилищем данных, позволяющее получать экспериментальные результаты применения статистических методов для выявления инсайдера.

Научная новизна диссертационной работы заключается в следующем:

1. Предложен новый подход к обнаружению действий инсайдера, который, в отличие от существующих, в рамках построенной модели не зависит от характеристик случайных процессов, описывающих работу пользователя. Более того, он гарантирует отсутствие ложных срабатываний и обеспечивает вероятность успешного обнаружения равную единице за конечное время за счет использования теории запретов.
2. В разработанной модели построены запреты вероятностной меры, характеризующей работу честного пользователя. С их использованием теоретически обоснована оценка точности алгоритма выявления инсайдера в множестве пользователей.
3. Разработан новый программный комплекс с возможностью имитационного моделирования работы пользователя и инсайдера в хранилище данных, который, в отличие от известных, позволяет получить практические ограничения применимости традиционных методов математической статистики в рамках разработанной модели.

Методология и методы исследования. В качестве методов исследования при решении сформулированных задач использовался аппарат теории вероятностей и математической статистики, а также элементы теории множеств. Для исследования предметной области проводился анализ научных работ в области теории запретов и современных подходов к обнаружению инсайдера. Для практического обоснования работоспособности предлагаемого в работе подхода проводилось имитационное моделирование. Достоверность предлагаемого в диссертации подхода обоснована теоретическими и экспериментальными исследованиями.

Теоретическая и практическая значимость. Предложенный метод позволяет анализировать нарушения информационной безопасности, связанные как с несанкционированными действиями злоумышленников, так и с другими процессами, которые носят вероятностный характер.

Предлагаемый подход можно использовать как для повышения качества анализа имеющихся систем обнаружения инсайдера, так и как самостоятельный способ выявления факта сбора и накопления ценной информации в компании теми сотрудниками, кому она не предназначена.

В качестве практической реализации автоматизированного поиска инсайдеров разработано программное обеспечение, выполняющее на каждом шаге проверку принадлежности очередной выборки пользователя множеству, описывающему его функциональные обязанности. Приложение способно в реальном времени выполнять такую проверку для порядка 50 пользователей информационной системы, каждый из которых регулярно использует в своей работе около 20-25 атрибутов, а описывающее обязанности каждого пользователя множество строится на основе 10-20 функциональных задач, что подтверждается 5 актами о внедрении:

1. «Kraftway Service Desk» для сопровождения собственных информационных ресурсов и в ходе проведения аттестационных испытаний объектов информатизации.
2. АО «Научно-исследовательский институт технической физики и автоматизации» для оценки защищенности внутренних информационных ресурсов.
3. ОАО НПО «Физика» в ходе аттестационных испытаний собственных информационных ресурсов.
4. ОАО «Швабе-Фотосистемы» для сопровождения и развития информационных ресурсов.
5. ООО «РУСБЕЛГАЗ» для поиска инсайдера среди внутренних пользователей информационной системы и нарушений информационной безопасности, связанных с доступом к корпоративной информации.

Степень достоверности и апробация результатов. Основные теоретические и практические результаты диссертации докладывались и обсуждались на XX Международной научной конференции студентов, аспирантов и молодых ученых «Ломоносовские чтения», секция «Вычислительная математика и кибернетика» (г. Москва, 2013г.), на заседании семинара кафедры информационной безопасности ВМК МГУ им. М.В. Ломоносова под руководством заведующего кафедрой информационной безопасности академика И.А. Соколова (г. Москва, 2017г.), на Всероссийском семинаре «Математические проблемы

информационной безопасности» Московского университета МВД России имени В.Я. Кикотя (г. Москва, 2018г.), на регулярных семинарах лаборатории компьютерной безопасности и анализа информационных ресурсов НИВЦ МГУ им. М.В. Ломоносова.

Публикации. Материалы диссертации опубликованы в 6 печатных работах [1–6], из них 4 статьи в рецензируемых журналах [1–4], 1 статья в сборниках трудов конференций [6].

Личный вклад автора. Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы. Подготовка к публикации полученных результатов проводилась совместно с соавторами, причем вклад диссертанта был определяющим. Все представленные в диссертации результаты получены лично автором.

Работы [2–4, 6] написаны единолично. В работе [1] Мартьянову Е.А. принадлежат результаты исследования по оценке защищенности информационных ресурсов, Макарову-Землянскому Н.В. принадлежит постановка задачи и проверка результатов. В работе [5] Мартьянову Е.А. принадлежит постановка задачи исследования по оценке защищенности информационных ресурсов и методика оценки, Быстрицкому Н.Д. принадлежит обзор проблем корректного взаимодействия пользователя с информационным ресурсом в сети Интернет.

Диссертация состоит из введения, трех глав с выводами по каждой из них, заключения, списка сокращений, списка цитируемой литературы и приложения. Общий объем работы составляет 136 страниц, включая 16 рисунков и 9 таблиц. Список литературы включает 109 наименований на 11 страницах.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулированы цель и основные задачи, научная новизна исследований, показана теоретическая и практическая значимость полученных результатов, представлены выносимые на защиту научные положения.

В первой главе приводится обзор исследований на тему защиты от инсайдеров в организациях и анализу существующих способов их выявления. В ней выделяются ключевые проблемы разработанных решений и осуществляется постановка задачи диссертационного исследования.

Рассмотренные разработки в области защиты информационной системы от инсайдера можно разделить на два типа:

- методические документы, в которых консолидированы требования и рекомендации, позволяющие снизить риски успешной атаки инсайдера, а также представлены психологические и социальные аспекты, приводящие к появлению такого феномена, как инсайдер;
- математические модели информационных систем с действующим в них

инсайдером, которые основаны на разных подходах к построению с использованием различных методов и техник.

Относящиеся к первому типу документы очень важны, так как именно они формируют общий взгляд на исследуемую проблему и могут быть использованы в качестве основы при формировании внутренней политики безопасности организации. Они во многом способствуют лучшему пониманию нормативных документов в области защиты информации и содержат практические указания по выполнению требований для защиты корпоративной информации. В них представлены причины появления угрозы инсайдеров и успешности их атак с подробным разбором возможных сценариев действий, что способствует лучшему пониманию проблемы и в дальнейшем ложится в основу всех технических средств обнаружения действий инсайдеров. Однако, каким бы существенным не был вклад методических документов в решение общей задачи, в одиночку они не позволяют защитить собственность организации от целенаправленных подготовленных атак инсайдеров.

В диссертационной работе подробно рассмотрена программа CERT и центр изучения угроз инсайдера (Insider Threat Study) при CERT, основная цель которого заключается в построении общей картины, описывающей угрозу инсайдера. За несколько лет исследователи провели детальный анализ более 700 инцидентов с участием инсайдера, выделили и описали их общие закономерности и предложили классификацию возможных угроз⁷:

- саботаж и шпионаж – использование ИТ для нанесения ущерба организации или ее конкретным сотрудникам;
- воровство интеллектуальной собственности;
- мошенничество – изменение или удаление части информации организации для личных целей.

Кроме того авторы привели отдельные рекомендации по противодействию каждой из них. На основании проведенного анализа, они пришли к выводу, что не существует универсального профиля инсайдера, по которому его можно однозначно выявить, а предотвратить или минимизировать возможный ущерб от успешной атаки могут лишь продуманные и согласованные действия всех сотрудников организации.

Второй тип документов используется в качестве основы для алгоритмов в программно-аппаратных средствах обнаружения и предотвращения враждебных действий инсайдера. Исследования различаются глубиной проработки вопроса, используемой моделью работы пользователей, количеством и набором изучаемых характеристик и подходом к обнаружению зловредных действий.

⁷ Cappelli Dawn, Moore Andrew, Trzeciak Randall. The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud). – Addison-Wesley Professional, 2012.

В диссертационной работе подробно рассмотрен проект ADAMS (Anomaly Detection at Multiple Scales) от DARPA, основная цель которого разработать и внедрить технологию классификации и обнаружения аномалий в больших объемах данных, что нашло своё применение в рассматриваемой задаче. В проекте приняло участие большое количество коллективов, наиболее интересные работы которых рассмотрены в первой главе.

Главная трудность для решения задачи выявления инсайдера заключается в отсутствии информации о том, что необходимо искать. Более того, количество враждебных действий, как правило, крайне мало по сравнению с общей активностью всех пользователей. Всё это приводит к серьезным сложностям для построения методов обнаружения: для основанных на машинном обучении детекторах существенными оказываются требования к обучающей выборке, которую очень сложно получить, и к вычислительным ресурсам; для методов на основе математической статистики – необходимость делать предположение относительно распределения, которым характеризуются анализируемые данные. Результаты экспериментов показывают, что второй класс детекторов обладает более высокой точностью при условии выполнения сделанных допущений, однако, именно они довольно сильно ограничивают возможность применения на практике таких детекторов.

Изучив проблему угрозы инсайдера и проанализировав существующие подходы, сделан вывод, что нужен довольно простой и эффективный алгоритм поиска инсайдера, использующий только ту информацию о пользователях, которую можно легко получить и обработать. Метод должен обладать теоретически обоснованной высокой точностью обнаружения, чтобы его применение было оправдано.

Во второй главе предложена новая формальная модель работы пользователя в информационной системе организации для которой построено описание нового метода выявления потенциального инсайдера и доказаны его отличительные свойства.

В диссертационной работе под информационной системой организации понимается хранилище данных большого объема со всей имеющей ценность корпоративной информацией. Под информационным объектом понимается та информация, с которой работает пользователь и которую он может использовать против работодателя. Пользователями информационной системы являются аналитики, работающие с информацией. Инсайдером будем называть легального пользователя информационной системы, у которого есть доступ к недоступной широкой публике информации. В качестве основной цели зловредных действий инсайдера рассматривается воровство интеллектуальной собственности в соответствии с выделенной в рамках исследований CERT угрозой. Защищенностью информационного объекта называем методы выявления инсайдера. Для анализа действий пользователей используется только информация об обращениях к хранилищу данных.

Работа каждого пользователя – это случайный процесс, о котором ничего

неизвестно, его не удастся детально описать простыми способами, не накладывая существенных ограничений. Вместо этого в диссертационной работе она представляется как последовательность множеств атрибутов, выбираемых из хранилища данных. При этом не рассматриваются дальнейшие действия с полученной информацией. Другими словами, если $A = \{a_1, \dots, a_m\}$ – множество всех атрибутов в хранилище данных, то работа пользователя – это последовательность множеств $X_i, i = 1, 2, \dots$, при этом $X_i \in \mathfrak{X}$, где \mathfrak{X} – множество всех подмножеств A .

У каждого пользователя есть конкретные функциональные обязанности, и для работы ему необходим определенный набор атрибутов, что представляется как множество всевозможных разрешенных обращений к хранилищу данных $X = \{X_1, \dots, X_k\}$. В отличие от честного пользователя инсайдер накапливает избыточную для своих обязанностей информацию. При таком представлении информационной системы и работы пользователя в ней для обнаружения зловредных действий достаточно проверять каждую выборку на соответствие его текущим обязанностям. Даже однократный выход за разрешенное множество атрибутов будет означать, что пользователь – потенциальный инсайдер и требуется более детальное изучение его активности с применением других технических и организационных средств.

В диссертационной работе вводится понятие ошибки первого и второго рода как ложное срабатывание и пропуск инсайдера соответственно. Обозначим вероятностную меру для честного пользователя через P , а для инсайдера через Q . Доказано ⁸, что

$$\Delta(P) = \lim_{n \rightarrow \infty} \Delta_n(P) = \bigcap_{n=1}^{\infty} \Delta_n(P),$$

является носителем меры P , где $\Delta_n(P)$ цилиндрическое множество. Если инсайдер выбирает избыточную для его обязанностей информацию, то в его выборке присутствует атрибут $y \notin \bigcup_{i=1}^k X_i$. В диссертационной работе показано, что добавление такого атрибута в последовательности выборок приводит к тому, что

$$\Delta(Q) \cap \Delta(P) = \emptyset.$$

В таком случае доказано, что подобное средство обнаружения принесет результат за конечное время наблюдения с вероятностью 1. Для этого необходимо воспользоваться понятием запрета меры ⁹.

Определение. Запретом меры P_n называется вектор $\bar{\omega}_s \in \mathfrak{X}^s, s \leq n$, такой, что $P_n(\bar{\omega}_s \times X^{n-s}) = 0$.

⁸ Grusho A., Grusho N., Timonina E. Consistent sequences of tests defined by bans // Springer Proceedings in Mathematics & Statistics, Optimization Theory, Decision Making, and Operation Research Applications.—New York, Heidelberg, Dordrecht, London: Springer, 2013. – P. 281–291.

⁹ Грушо А., Тимонина Е. Запреты в дискретных вероятностно-статистических задачах // Дискретная математика. — 2011. — С. 281–291.

Построение множества запретов – очень сложная задача. Тем не менее, разработанная модель позволяет их строить. Если определить запрет как произвольное множество из X , к которому добавлен атрибут y , то справедливо следствие¹⁰ о том, что существует конечное число N такое, что существует последовательность критериев проверки гипотез $H_{0,n}$ (пользователь честный) против $H_{1,n}$ (пользователь – инсайдер), $n = 1, 2, \dots$ с критическими множествами S_n , $n = 1, 2, \dots$, определяемыми запретами, что для всех $n \geq N$ функция мощности критерия проверки $H_{0,n}$ против $H_{1,n}$ принимает значение 1.

Данный результат говорит о следующем. Если для честного пользователя, работающего в рамках своих функциональных обязанностей, задать множество запретов, то с вероятностью 1 за конечное число проверок его можно отличить от инсайдера с теми же функциональными обязанностями. А из определения деятельности инсайдера следует, что ложные тревоги исключены.

В отличие от имеющихся подходов к выявлению инсайдера разработанный метод обладает рядом уникальных особенностей:

- не зависит от количества пользователей информационной системы и специфики их работы;
- не накладывает ограничений на то распределение, которому подчинена их работа;
- в качестве данных, на основе которых проводится анализ, использует только информацию об обращениях пользователей к хранилищу данных.

В заключительном параграфе второй главы построенная модель обобщена для случая, когда существуют ограничения на набор разрешенных для пользователя атрибутов и строк данных. Это сильно усложняет описание возможностей инсайдера, но, тем не менее, получены аналогичные результаты.

В третьей главе описан разработанный программный продукт, с использованием которого проведено имитационное моделирование трех различных сценариев работы пользователя с корпоративным хранилищем данных. Выполненный в первой главе анализ показал, что для исследования и проверки всех методов выявления инсайдера в определенной степени используются искусственные данные о действиях пользователей в информационной системе. Исходя из этого, были сформулированы требования к программному обеспечению, основными характеристиками которого являются:

- возможность выполнять анализ всех запросов пользователей к ХД по подготовленной входной информации с целью обнаружения выхода за разрешенное множество атрибутов;

¹⁰ Grusho A., Grusho N., Timonina E. Power functions of statistical criteria defined by bans // ECMS 2015: Proceedings of 29th European Conference on Modelling and Simulation. — Germany: Digitaldruck Pirrot GmbH, 2015. — P. 617–621.

- возможность моделировать хранилище данных, близкое по своим характеристикам к используемым в крупных организациях;
- возможность моделировать работу пользователей хранилища данных, представляющую собой последовательные выборки определенных строк и столбцов хранилища данных в соответствии с их функциональными обязанностями.

Таким образом, на основе заданных требований выбрана среда разработки, сформулированы требуемые характеристики программного обеспечения, выделены необходимые структурные компоненты и разработан прикладной алгоритм для проведения анализа работы пользователей.

Работу приложения на верхнем уровне можно описать следующим образом. После запуска происходит заполнение внутренних структур данных и инициализация используемых модулей на основе введенных пользователем параметров. Далее, в зависимости от выбранного режима работы, может запускаться модуль, который воссоздает работу пользователей с ХД и сохраняет результаты в файл. Затем анализатор считывает файл с данными о действиях пользователя (или полученный в результате ИМ, или на основе реальных данных о действиях пользователей) и производит их проверку согласно разработанному алгоритму. По окончании работы приложение сохраняет результаты. Для уменьшения необходимого для моделирования времени предусмотрен многопоточный режим работы, в котором отдельные повторения независимых экспериментов выполняются параллельно.

В диссертационной работе проведено имитационное моделирование трех основных сценариев работы пользователя для верификации разработанного алгоритма и получения граничных условий его применения на практике:

1. Все пользователи хранилища данных делают выборки большого объема, при этом вероятность попадания всех атрибутов в выборку одинаковая;
2. У разных пользователей разная вероятность попадания атрибутов в их выборку, что соответствует различным функциональным обязанностям в описанной модели;
3. В процессе работы пользователя вероятность попадания одних и тех же атрибутов в его выборку изменяется от запроса к запросу, что соответствует сценарию, при котором инсайдер маскирует свою деятельность.

По итогам изучения каждого из них сформулированы результаты. Для первого сценария получается следующая картина (рис. 1), из которой можно сделать вывод о том, что для больших выборок пользователь в любом случае получит ценную информацию и отличить его от инсайдера таким способом нельзя.

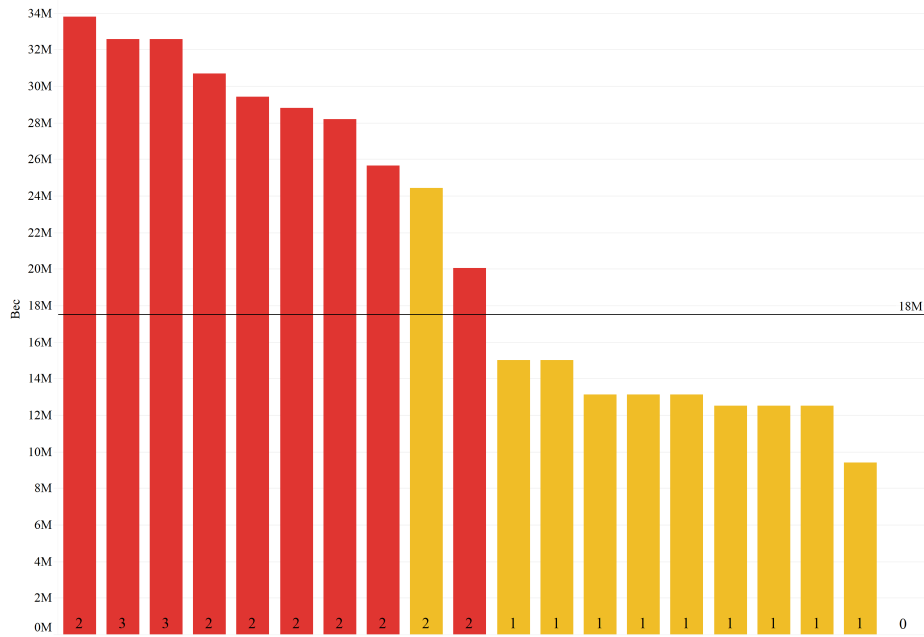


Рис. 1. Веса выборок пользователя в первом сценарии его работы. По оси абсцисс расположены выборки пользователей, по оси ординат – вес соответствующей выборки. Число в нижней части каждого столбца обозначает количество значимых атрибутов в выборке. Красным обозначены выборки, в которых есть пары значимых атрибутов; желтым – где таких пар не оказалось.

Во втором сценарии, в отличие от имеющихся исследований в области обнаружения инсайдера, для вероятности попадания атрибута в выборку пользователя используется закон Ципфа¹¹ (что справедливо для описанной модели работы пользователя), согласно которому вероятность попадания упорядоченных по убыванию частот атрибутов в выборку приблизительно обратно пропорциональна порядковому номеру атрибута. Т.е. вероятность попадания атрибута с номером i в выборку пользователя равна $P_{\alpha}(i) = \frac{c_{\alpha}}{i^{\alpha}}$, где α – характеристика пользователя, c_{α} – нормировочный коэффициент.

В данном сценарии предъявляются следующие требования к точности модели: необходимо, чтобы наблюдаемая и теоретическая частота отличались не более чем на 0.02 с вероятностью 0.95. Несложно показать, что для этого достаточно 2 400 экспериментов. Затем на модель накладываются еще 2 ограничения, чтобы больше соответствовать реальной работе ИС: ограничение на время наблюдения за пользователем; из результатов исключаются те атрибуты, которые попадают практически во все выборки пользователя (порядка 20-30). Таким образом образуется интервал, на котором имеет смысл изучать поведение пользователей.

Зная характеристики честного пользователя (α) и инсайдера (γ), можно теоретически рассчитать тот интервал, на котором инсайдер накапливает цен-

¹¹ Zipf G.K. Human Behavior and the Principle of least Effort. // Addison-Wesley Press, Reading, MA.—1949.—P. 484–490.

ную информацию, выходящую за рамки его функциональных обязанностей. Однако, результаты моделирования показывают, что на практике он меньше. Например, для пары $\alpha = 0.8$, $\gamma = 0.5$ теоретически рассчитанный отрезок равен $[101; 183]$, а полученный в результате моделирования равен $[120; 183]$. Это иллюстрируют следующие графики (Рис. 2):

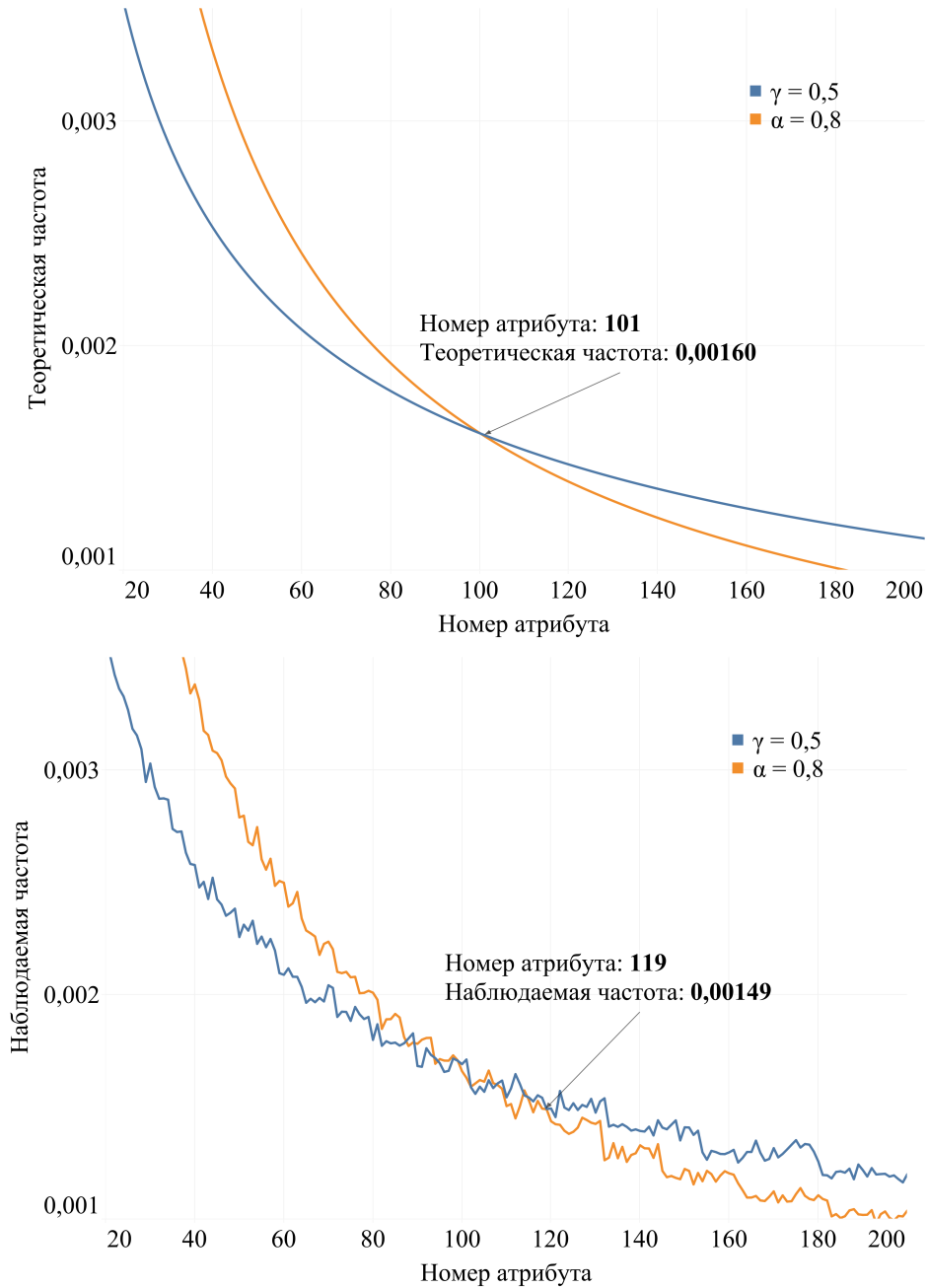


Рис. 2. Сравнение наблюдаемой и теоретической частоты попадания атрибутов в выборку пользователей с параметрами α и γ для 2 400 наблюдений

В третьем сценарии инсайдер маскирует свою враждебную деятельность и в основное время ведет себя как честный пользователь, однако, определенная доля его выборок призвана собрать ценную информацию из хранилища данных. Считается, что с вероятностью $1 - p$ инсайдер ведет себя как честный

пользователь. Проведя моделирование с аналогичными требованиями к точности такого поведения для параметров $\alpha = 0.8$, $\gamma = 0.5$, а также всевозможных значений $p \in (0; 1)$ с шагом 0.1, получено граничное значение, при котором, в рамках используемой модели, невозможно однозначно выявить инсайдера в множестве пользователей. Для данной пары ($\alpha = 0.8$, $\gamma = 0.5$) это значение $p = 0.2$ (Рис. 3). Это означает, что, если инсайдер с вероятностью не более 20% выбирает лишние для него атрибуты, то его невозможно отличить от честного пользователя в описанной модели.

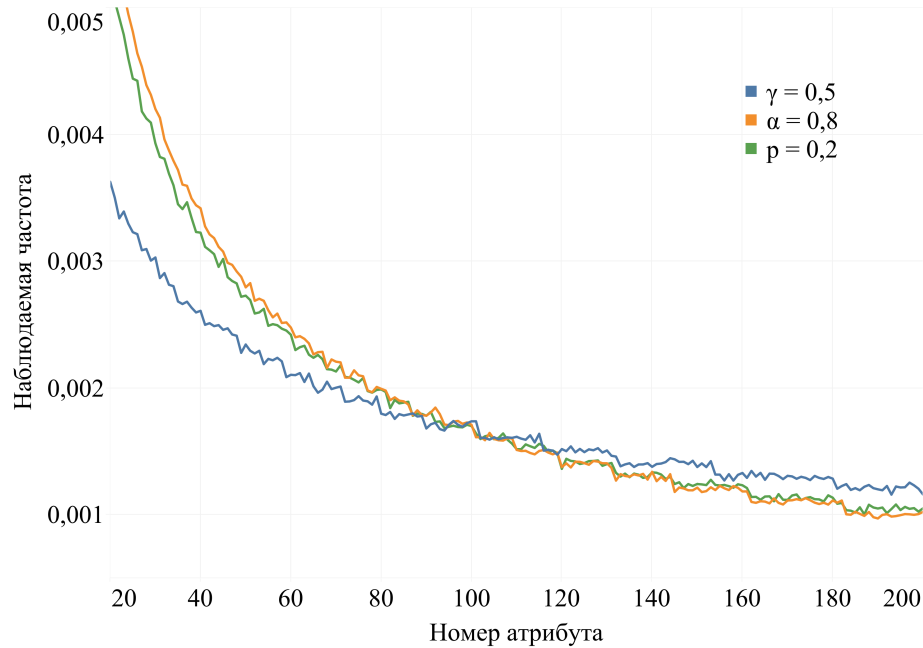


Рис. 3. Пользователь с вероятностью $p = 0.2$ выбирает лишние для него атрибуты

Для последнего, наиболее приближенного к реальной работе пользователей, сценария проведен анализ влияния как количества пользователей, так и мощности разрешенного множества атрибутов на общее время работы детектора. Результаты проиллюстрированы на рисунках 4 и 5.

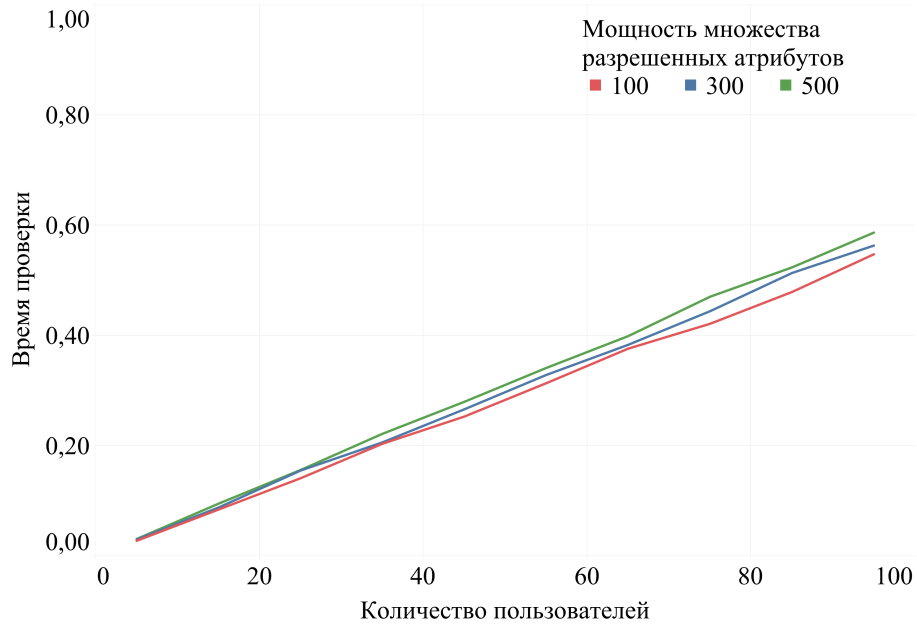


Рис. 4. График зависимости времени проверки от количества пользователей

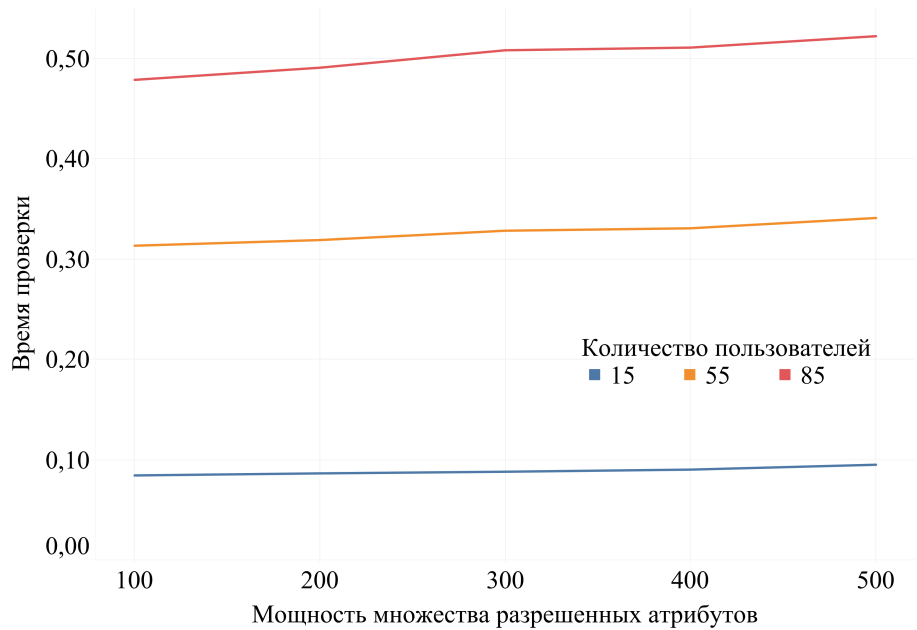


Рис. 5. График зависимости времени проверки от мощности разрешенного множества атрибутов

Разработанное программное обеспечение и метод выявления инсайдера, основанный на определении факта сбора избыточной для пользователя информации, опробованы и внедрены в несколько организаций, в т.ч. в качестве средства постоянного мониторинга для поиска инсайдера среди внутренних пользователей информационной системы.

В Заключение сформулированы основные научные результаты и выводы, полученные в диссертационной работе:

1. Проведен анализ методов и методик выявления инсайдера в информационных системах, а также различных методологических указаний и рекомендаций по снижению риска успешного осуществления инсайдером враждебных действий. На его основе сделан вывод относительно области применения имеющихся подходов и их точности, выделены их достоинства и недостатки, которые существенно влияют на качество выявления инсайдера.
2. Построено описание нового метода выявления потенциального инсайдера в множестве пользователей информационной системы, для которого теоретически обоснована как точность обнаружения на конечном шаге, так и полное отсутствие ложных срабатываний.
3. Построена новая формальная модель работы пользователя в информационной системе организации, представляющей собой хранилище данных. Отличительной особенностью модели является отсутствие ограничений как на количество пользователей, так и на особенность их работы, а в качестве информации для исследования используется только информация об обращении пользователей к хранилищу данных. Несмотря на то, что работа каждого пользователя – это отдельный случайный процесс, о котором ничего неизвестно, обобщить результаты и избежать их детального описания удалось за счет использования критических множеств, определенных запретами. Основным результатом является то, что модель позволила использовать теорию запретов в конечных пространствах и построить метод нахождения запретов для данного семейства процессов.
4. Доказано, что в описанной модели вероятность ошибки первого рода равна нулю, а поиск потенциального инсайдера принесет положительный результат за конечное время наблюдения в описанной модели.
5. Для автоматизированного проведения анализа работы пользователя с хранилищем данных и верификации построенного метода разработано программное обеспечение с возможностью имитационного моделирования различных сценариев работы сотрудников. С его помощью получены экспериментальные результаты применения статистических методов для выявления инсайдера.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Мартьянов Е. А., Макаров-Землянский Н. В. Оценка защищенности информационных объектов от потенциальных нарушителей // Естественные и технические науки. — 2013. — Т. 6. — С. 442–444.

2. Мартьянов Е. А. Возможность выявления инсайдера статистическими методами // Системы и средства информатики. — 2017. — Т. 27, №2. — С. 41–47.
3. Мартьянов Е. А. Имитационная модель поиска инсайдера статистическими методами // Системы и средства информатики. — 2017. — Т. 27, №2. — С. 48–59.
4. Мартьянов Е. А. Запреты вероятностных мер в задаче поиска инсайдера // Системы и средства информатики. — 2017. — Т. 27, №4. — С. 144–149.
5. Мартьянов Е. А., Быстрицкий Н.Д. Получение оценки защищенности web-ресурсов // Аспирант и соискатель. — 2014. — Т. 6. — С. 81–84.
6. Мартьянов Е. А. Методика оценки защищенности информационного объекта // Сборник тезисов XX Международной научной конференции студентов, аспирантов и молодых ученых «Ломоносов-2013». Секция «Вычислительная математика и кибернетика». — 2013. — С. 56–58.