Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук»

На правах рукописи

Смирнов Дмитрий Владимирович

МЕТОДЫ ПОИСКА ПРИЗНАКОВ ИНСАЙДЕРА В BIG DATA

05.13.19 – Методы и системы защиты информации, информационная безопасность

Диссертация на соискание учёной степени

кандидата технических наук

Научный руководитель:

доктор физико-математических наук, профессор

Грушо Александр Александрович

Москва – 2021

Содержание

Введение

Компании, которые используют ИТ-платформу собственной разработки для ведения своей деятельности, можно условно назвать цифровыми. В дополнение к собственной ИТ-платформе цифровые компании также могут использовать мобильные рабочие места, облачные сервисы по управлению персоналом, облачную бухгалтерию и т. д., продвигают свои товары и услуги в Интернете, осуществляют транзакции или продажи своих продуктов онлайн. Цифровые компании могут быть представлены в любой отрасли экономики: финансовой, нефтяной, энергетической, транспортной, связи и т. д.

Проникновение информационных технологий, с одной стороны, дает цифровым компаниям преимущество в скорости предоставления услуг, их качестве и цене и т. д., но с другой стороны, растут риски кибербезопасности. Для минимизации рисков кибербезопасности применяют различные технические решения. Рынок технических решений по кибербезопасности состоит из 12 сегментов: от сетевой безопасности до антифрод-решений[1]. Однако на рынке отсутствует такой сегмент как защита от внутренних нарушителей (инсайдерские атаки). Сами решения по защите от инсайдера существуют, но из-за их сырости и неэффективности их не объединяют в сегмент. Наибольший риск представляют инсайдерские атаки, которые реализуются как утечка данных из хранилищ (Big Data).

Актуальность темы. Кража персональных данных совместно другими конфиденциальными данными такими как остаток на счете, портфель ценных бумаг, ИНН с движением средств по счету и т. д. позволяет злоумышленникам производить таргетированные атаки на физические лица, собирать информацию о конкурирующих фирмах и т. д.

При этом крупные организации встроили в свою инфраструктуру достаточно много технических средств защиты и по этой причине относительно

[1] CYBERscape - Momentum Cyber - https://momentumcyber.com/docs/CYBERscape.pdf

хорошо защищены от атак с помощью технических средств. Как было описано во Введении, технические средства защищают от атак, исходящих из-за периметра организации (внешние атаки): DDOS-атаки, побор паролей, обход сетевых экранов, заражение вирусами, эксплуатацию уязвимостей операционной системы и достаточно хорошо изучены. Именно по причине наличия адекватных средств защиты, крупные организации относительно хорошо защищены от внешних атак. Также внешние атаки весьма дорогостоящие и гораздо менее экономически выгодны по сравнению с использованием внутреннего злоумышленника (инсайдера), который может не только, например, своими руками или руками коллег несанкционированно выгрузить конфиденциальные данные на флэш-носитель, переслать через электронную почту, изъять зашифрованный жесткий диск из рабочей станции или сервера и т.д., но и в состоянии (используя свои должностные полномочия и сетевые привилегии) провести ряд легальных процедур по выводу информации из защищенных контуров и хранилищ. Наибольшую ценность для инсайдера представляют объекты концентрации данных – хранилища (Big Data), что является объектом защиты в контексте данной диссертационной работы.

Для борьбы с инсайдерскими активностями используется ряд технических средств такие как DLP, SIEM и т. д., но данные средства производят мониторинг нелегальных каналов, а инсайдеры используют легальные каналы движения информации. Инсайдерские атаки становится не только более экономически выгодным способом кражи информации, а всё чаше и единственно возможным.

В данной работе представлены результаты исследований внутренних инсайдерских угроз, где стандартные технические средства либо не работают, либо не дают результата.

Таким образом, инсайдерские угрозы — это вредоносные для организации активности, которые исходят от сотрудников внутри организации (периметра защиты), в частности – от действующих работников, бывших работников, подрядчиков, деловых партнеров и даже завербованных работников или работников, специально внедренных в организацию, которые обладают доступом

к конфиденциальной информации по своим должностным обязанностям и которые имеют представление о системе управления информационной безопасностью организации.

Как было сказано выше, объект защиты является комплекс Big Data. Журналы аудита действий сотрудников в комплексе Big Data тоже является Big Data, но меньшего объема (~10%). Поэтому для поиска признаков инсайдерской активности (вкраплений в данных) в Big Data необходимо научиться оперировать большими объемами текущей информации, анализируя ее и формируя рекомендации для принятия соответствующих управленческих решений в условиях жестких ограничений по времени. Именно этим и обусловлено использование в подобном анализе компьютерных средств и систем искусственного интеллекта, с помощью которых и обеспечивается интеллектуальный анализ больших данных в режиме требуемых ограничений по времени. При этом необходимо, чтобы сформированные в процессе такого анализа рекомендации были объясняемы и понятны экспертам по противодействию инсайдерским активностям – работникам оперативных служб безопасности (ведь именно на них в итоге ложится ответственность за принятые решения и их последствия).

Таким образом важной задачей обеспечения информационной безопасности является научно-техническая проблема поиска признаков действий инсайдеров в Big Data в условиях больших обновляемых данных при ограничениях на время поиска.

Еще в первой половине 90-х годов прошлого века Усамой Файядом (Usama Fayyad), Григорием Пятецким-Шапиро (Gregory I. Piatetsky-Shapiro) и их коллегами в научный лексикон были введены понятия *data mining & knowledge discovery* (DM&KD) - *поиск зависимостей в данных, результатом которого становится порождение нового знания* – см., например, ряд научных конференций и научный журнал с тем же названием, издаваемый уже более 20

лет[1]. В самом общем виде это направление в компьютерном анализе данных ориентировано на *поиск* всевозможных *зависимостей*[2], извлекаемых из накапливаемых эмпирических данных. При этом, в общем случае, каких-либо исходных ограничений на вид таких зависимостей не накладывается: тот или иной вид порождаемых зависимостей фактически определяется (уточняется) в процессе выполняемого поиска и представляет собою порождение (выявление) нового *знания*[3] из накапливаемых эмпирических данных. Исторически так сложилось, что уже с середины тех же 90-х годов проблематика DM&KD практически повсеместно стала соседствовать вместе с проблематикой накопления, хранения и обработки больших объемов данных, сперва - так называемыми Хранилищами данных (DWH[4]) а далее и собственно Big Data. В таких ситуациях обычно речь ведется о поиске зависимостей («добыче нового знания») из накапливаемых в DWH эмпирических данных. В отличие от ранее уже известных технологий, например, от проблематики разработки систем управления различными типами баз данных, где вопрос о постоянном пополнении уже имеющихся данных новой информацией, как правило, выносится за скобки (за пределы обсуждения в рамках развиваемой формальной конструкции[5]), или же от поиска в данных зависимостей заранее заданного вида (например, широко распространенных технологий OLTP[6], где жестко фиксируется число параметров, связываемых между собою в соответствующем запросе), технологии DM&KD ориентированы на поиск в DWH новых знаний (эмпирических зависимостей) любого вида, релевантного целям осуществляемого поиска. Ввиду аналогий с деятельностью эксперта, анализирующего эмпирические данные с целью выявления нового знания, в целом

---

[1] Data Mining and Knowledge Discovery. - Springer, 1997- н\в. - https://dblp.org/db/journals/datamine/index.html
[2] Data mining (дословно – добыча данных) – анализ данных с целью добычи нового знания.
[3] Knowledge discovery (дословно – открытие знаний) – порождение нового знания (в процессе анализа данных – data mining).
[4] Data WareHouse (хранилище данных).
[5] Например, в реляционных базах данных основной интерес сфокусирован на формировании зависимостей и так называемых ключей, позволяющих сделать эффективным процесс поиска ответа на запрос (избегая при этом полного перебора вариантов). В свою очередь, пополнение уже имеющегося набора данных новой информацией адресуется к процедурам обеспечивающей эффективность поиска перестройки уже сформированных зависимостей и ключей.
[6] On Line Transactional Processing (оперативная обработка транзакций).

ряде случаев такие технологии стали называть интеллектуальным анализом данных.

В данной диссертационной работе фактически рассматривается аналогичная проблематика: предметная область исследования – поиск признаков вредоносных инсайдерских активностей в характеризующих бизнес крупного коммерческого банка больших объемах данных (Big Data,) естественным образом, распадающихся на составляющие трех типов:

- собственно большие (и постоянно пополняемые новой информацией) операционные данные[1], характеризующие как основной бизнес объекта защиты, так и результаты постоянного мониторинга режимов его функционирования службами защиты (Службой Безопасности);
- значительное количество сотрудников, в моменте ведущих активную профессиональную деятельность с этими данными[2];
- огромное количество текущих операций сотрудников с соответствующими данными. (Причем неявная информация о таких взаимодействиях постоянно фиксируется[3] и накапливается в соответствующих ИТ-ресурсах).

Таким образом, необходимо:

- анализировать постоянно накапливаемые (в процессе функционирования объекта защиты) первичные данные на предмет мониторинга неявно содержащихся в них сведений о *взаимодействиях* (сотрудников с ИТ-ресурсами – см. выше),
- выявляя в этих данных такие неявно содержащиеся описания *взаимодействий*, которые несут (явные или потенциальные) риски *вредоносных* последствий.

Фактически речь идет о проблеме фильтрации постоянно накапливаемых данных об имеющихся взаимодействиях, характеризуемой необходимостью:

- постоянно иметь дело с очень большими объемами первичных данных (Big Data),

---

[1] В реальной рассматриваемой в КД ситуации это – терабайты данных ежедневно на тысячах серверов и в сотнях информационных ресурсов.
[2] В рассматриваемой ситуации это – тысячи сотрудников одновременно.
[3] Например, в соответствующих log’ах и др.

- учитывать (например, в части организации целого ряда сервисов - информационного поиска, поддержания в актуальном состоянии текущего профиля угроз и модели нарушителя и т. п.) постоянные пополнения таких данных новой информацией,
- «укладываться» в жесткие рамки ограничений на время анализа данных и принятия соответствующих управленческих решений (например, по организации противодействия выявляемым угрозам и т. п.)
- обеспечивать «прозрачность» формируемых компьютерной системой защиты выводов рекомендаций для экспертов Службы Безопасности, несущих ответственность за последствия принимаемых решений.

Таким образом, **необходимо** разработать методы и программные «инструменты» работы с Big Data, которые позволяют:

- выделять в Big Data описания *взаимодействий* (см. выше), несущих потенциальные или же явные риски вредоносных последствий – **выделять *вредоносные взаимодействия***,
- организовать подобный процесс фильтрации Big Data **эффективным образом**, в частности, i) разработать соответствующие процедуры сокращения объемов детально анализируемых данных, сохраняющие тем не менее в этих данных соответствующие признаки вредоносности, ii) обеспечить реализуемость всего процесса фильтрации в рамках соответствующих ресурсных ограничений (бюджетов, выделяемых на эти цели основным бизнесом банка; сроков выполнения каждого цикла анализа данных и приятия решений; численности персонала соответствующей квалификации в Службе безопасности и др.).

Таким образом:

Объектом исследования диссертационной работы – потенциально или же явно *вредоносные взаимодействия* персонала с ИТ-ресурсами объекта зашиты (Big Data). Описания таких *вредоносных взаимодействий* в неявном виде содержатся в постоянно пополняемых новыми сведениями системе журналирования, которая тоже является Big Data.

Результат – предложен способ (комплекс методов и реализующих их программных «инструментов» анализа данных и поддержки принятия решений) эффективного выявления таких *вредоносных взаимодействий*, позволивший:

- управляемым образом сокращать объемы детально анализируемых данных, сохраняя при этом искомые признаки вредоносности,

- эффективно («вписываясь» в предъявленные ресурсные ограничения [1]) на практике - в обеспечении защиты бизнеса крупного российского коммерческого банка [2] от угроз вредоносных воздействий – решать поставленные перед ним задачи по идентификации инсайдеров,

- обеспечивая при этом своими архитектурными решениями возможности для дальнейшего[3] масштабирования по производительности – по постоянно растущим объемам требующих анализа данных.

Предмет исследования модели, методы и технические возможности реализации поиска признаков инсайдеров в условиях Big Data взаимодействий большого числа пользователей с хранилищем данных, гетерогенности и пополняемости данных и ограничениях времени обработки данных.

Цель работы и задачи исследования – разработка методов выявления признаков инсайдерской активности в условиях Big Data взаимодействий большого числа пользователей с хранилищем данных, гетерогенности и пополняемости данных и ограничениях времени обработки данных, создание системно-технических [4] и архитектурных решений для поддержки профильной деятельности оперативных работников служб безопасности

В соответствии с целью определены следующие **задачи исследования:**

---

[1] И это может быть оформлено в виде необходимых Соглашений об уровне сервиса (Service Level Agreement)

[2] Терабайты данных ежедневно в сотнях информационных ресурсах на тысячах серверов.

[3] Обусловленного планами и потребностями дальнейшего развития основного бизнеса объекта защиты.

[4]Термин *программно-техническое решение* понимается как комплекс ИТ-средств (при необходимости включающий в себя и обеспечивающие системные и аппаратные составляющие), в основу которого положены разработанные автором данной диссертационной работы и защищенные соответствующими авторскими свидетельствами программные продукты. При этом все необходимые аппаратно-программные дополнения (системное ПО, ранее приобретенные на рынке «типовые» инструментальные ИТ-решения и т.п.) задействованы в таких комплексах в соответствии с корпоративными политиками объекта защиты (крупного российского коммерческого банка) в области защиты данных и обеспечения кибербезопасности.

1. Проанализировать представленные в профильной литературе данные по моделям, методам и алгоритмам выявления признаков вредоносных инсайдерских активностей.
2. Исследовать возможности и условия выявляемости признаков деятельности инсайдеров в условиях Big Data, гетерогенности и пополняемости данных и ограничениях времени обработки данных,
3. Разработать методы интеллектуального анализа данных, позволяющие управлять балансом между детальностью представления знаний и объемами при поиске признаков вредоносных инсайдерских активностей.
4. Разработать системно-технические и архитектурные решения, позволяющие эффективно анализировать большие гетерогенные данные с целью выявления вредоносной инсайдерской активности.
5. Разработать комплекс программных средств, реализующих предложенные методы выявления признаков вредоносных инсайдерских действий и экспериментально продемонстрировать работоспособность и результативность разработанных методов.

Методология исследования. Для достижения поставленной цели и решения сформулированных в диссертационной работе задач использовались методы дискретной математики, статистического анализа данных и машинного обучения, интеллектуального анализа данных, теории вычислительных систем и теории алгоритмов. Экспериментальные исследования осуществлялись с помощью моделирования процессов идентификации вредоносных инсайдерских активностей на тестовых стендах (в т.ч. – имитирующих характеристики производительности вычислительной инфраструктуры крупной индустриальной организации).

Актуальность исследования. В отличие от традиционных методов борьбы с мошенничеством и кражей информации предлагаемые в данной работе подходы и решения ориентированы на то, что инсайдеры «растворяют» свою незаконную деятельность в потоке выполняемых ими ежедневных легитимных рабочих

процедур, а обычные технические средства защиты не «видят» угроз в их деятельности. При этом те, кто неоднократно и целенаправленно похищают информацию, демонстрируют поразительную изворотливость, чаще всего в основе своих действий используя легальные процедуры, процессы, доступы, предоставленные им по службе (не редко используя в своих целях и беспечность коллег). Именно поэтому подобные действия могут подолгу оставаться незамеченными. Сотрудники из такой категории могут длительное время небольшими и незаметными микродействиями создавать бреши в системах защиты, микропорциями проникать через системы контроля и накапливать данные на своих персональных служебных учетных записях, применять простые методы обратимой трансформации данных, передаваемых с систем хранения на свои неслужебные ресурсы и таким образом эффективно обходить фильтры DLP или других подобных штатных систем защиты.

При этом на рынке (причем – не только отечественном) сегодня отсутствуют готовые коммерческие продукты, обеспечивающие выявление инсайдеров с доказанной эффективностью. Наиболее близкие доступные программные инструменты, моделирующие поведение пользователей - это User-Entity Behaviour Analytics (UEBA), но их позиционируют преимущественно как инструменты анализа данных о текущем состоянии ИТ-инфраструктуры (windows, linux логов и т.д.), а не проблемно-ориентированного анализа данных прикладных систем (CRM, BI, DWH и т.д.).

Научная новизна диссертационной работы определяется в первую очередь разработкой ***оригинальной методики*** аналитической оценки поведенческой активности пользователей и персонала, с помощью которой в режиме ограниченного времени можно обрабатывать большие объемы релевантных цели поиска операционных данных (информацию из журналов действий сотрудников и др.), а также анализировать на предмет выявления признаков вредоносных инсайдерских активностей (в рамках соблюдения прав человека) в поведении – нарушение трудовой дисциплины, чрезмерная закредитованность, конфликтность, немотивированные отклонения от традиционных в компании процедур и т.д.

Методика опирается на разработанные автором диссертационной работы ***теоретические модели*** (использующие методы статистического анализа данных, интеллектуального анализа данных и машинного обучения), а также реализующие их ***алгоритмические конструкции***, позволяющие выявлять аномалии в больших гетерогенных данных, вести в целях идентификации признаков инсайдеров их оперативный анализ, по результатам которого и формируются рекомендации для работников оперативных служб безопасности по целенаправленному противодействию выявленным инсайдерским активностям.

Теоретическая значимость исследования заключается в том, что данная работа не только систематизирует известные методы выявления инсайдеров, но и интегрирует их с новыми, разработанными автором в рамках данного диссертационного исследования, методами анализа больших гетерогенных данных (статистического выявления аномалий в поведении объектов мониторинга, интеллектуального анализа данных и машинного обучения при идентификации угроз и формировании рекомендаций по противодействию их влиянию), что позволило предложить принципиально новые эффективные способы выявления признаков инсайдерских активностей в информационной среде Big Data.

Апробация работы. В процессе экспериментального исследования получено подтверждение работоспособности и результативности разработанных методов и оказана помощь при промышленной реализации программного комплекса поиска признаков инсайдеров в крупной коммерческой организации. Получены в установленном порядке: свидетельство о государственной регистрации программы для ЭВМ № 2021614494 «Аналитическая панель доступов к данным», дата государственной регистрации 25.03.2021 (заявка ЕА-40490) и свидетельство о государственной регистрации программы для ЭВМ № 2021613506 «Поисковая система доступа к данным», дата государственной регистрации 19.04.2021 (заявка ЕА-40486).

Основные результаты работы докладывались и обсуждались на различных научных семинарах и конференциях, в том числе:

1. семинар в подразделении, реализовавшего (проприетарную) промышленную систему поиска признаков инсайдера в крупном коммерческом банке (10 заседаний семинара)
2. совместный семинар 53 и 16 отделов ФИЦ ИУ РАН.

Реализация и внедрение результатов работы: разработанные в диссертации программные инструменты внедрены в промышленный контур крупного отечественного коммерческого банка, что подтверждено соответствующим актом о внедрении.

Практическая значимость. Работоспособность и практическая значимость разработанной методики выявления признаков инсайдерский активности, а также обеспечивающих ее применение программных инструментов анализа данных и поддержки принятия решений подтверждены промышленным внедрением и использованием разработанного инструментария в текущую деятельности крупного отечественного коммерческого банка.

Соответствие паспорту специальности 05.13.19. Диссертационная работа соответствует следующим пунктам паспорта специальности 05.13.19:

(1) *Теория и методология* обеспечения информационной безопасности и защиты информации. Соответствует в части:

- Методики анализа Big Data в условиях ограниченного времени принятия решений в целях выявления признаков инсайдера
- Целенаправленного применения методов искусственного интеллекта (ИАД и машинное обучение для решения задачи восстановления по прецедентам частично-определенного отношения релевантности *ТЕКУЩИЕ ДАННЫЕ МОНИТОРИНГА ~ПРОФИЛЬ УГРОЗ*)
- Методов сокращения перебора при поиске решений (диаграмма сходств) в Big Data
- Методов выявления сговора сотрудников

(2) *Методы, аппаратно-программные и организационные* средства защиты систем (объектов) формирования и предоставления пользователям информационных ресурсов различного вида. Соответствует в части:

- Реализации *защиты* (подсказки\рекомендации офицеру службы безопасности: на какие факты обратить внимание и какой профиль сотрудника содержит какие отклонения в поведении), включая программную реализацию.
- *Разработанного метода* оценки вероятности возникновения аномалии, позволяющего ранжировать выявленные аномалии по вероятности, в том числе, позволяющие определить является ли случайная аномалия рисковым событием.

(3) *Методы, модели и средства выявления, идентификации и классификации угроз* нарушения информационной безопасности объектов различного вида и класса. Соответствует в части:

- *Выявления угроз* (реконструкция – применением машинного обучения на прецедентах – лишь частично заданного конкретными примерами отношения релевантности ТЕКУЩИЕ ДАННЫЕ МОНИТОРИНГА ~ПРОФИЛЬ УГРОЗ), включая программную реализацию этих процедур.
- *Разработанного метода* работы с противоречиями при выявлении аномалии в поведении сотрудников, в том числе, методы, обладающие свойствами изменять параметризацию среды
- *Разработанного метода*, позволяющего работать с гетерогенными данными и выявлять компрометирующие данные в различных информационных пространствах и объединять выявленные данные в единый результат.

(6) *Модели и методы* формирования *комплексов средств противодействия угрозам хищения* (разрушения, модификации) информации и нарушения информационной безопасности для различного вида объектов защиты вне зависимости от области их функционирования. Соответствует в части:

- *Идентификации угроз* и *организации защиты* (используя общую для разных предметных областей технику работы с профилем угроз => статистический анализ аномалий + диаграммы сходств») объектов защиты, характеризуемых Big Data, включая программную реализацию этих процедур.
- *Разработанного метода* выявления сговора сотрудников

(13) *Принципы и решения* (технические, математические, организационные и др.) *по созданию новых* и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности. Соответствует в части:

- *Новизны* (Big Data, ограниченное время, реконструкция отношения релевантности ТЕКУЩИЕ ДАННЫЕ МОНИТОРИНГА~ПРОФИЛЬ УГРОЗ машинным обучение на прецедентах), включая программную реализацию этих процедур, описанную в Главе 4.

Основные научные результаты, выносимые на защиту

1. Разработаны методы работы с противоречиями при выявлении аномалий в поведении сотрудников. В том числе, методы, обладающие свойствами изменять параметризацию среды и позволяющие управлять балансом между детальностью представления знаний и объемами вычислений при поиске признаков вредоносных инсайдерских активностей.

2. Методы, позволяющие работать с гетерогенными данными, выявлять компрометирующие данные в различных информационных пространствах и объединять выявленные данные в единый результат.

3. Методы, позволяющие с помощью кластеризации Big Data и оптимизации применения статистических методов выявлять сговор инсайдеров.

4. Методы оценки вероятности случайного возникновения аномалии, позволяющие ранжировать выявленные аномалии по вероятности и, позволяющие определить, является ли случайная аномалия рисковым событием.

5. Методика анализа в ограниченное время большого потока данных, релевантных цели поиска признаков деятельности инсайдера.

6. Разработан комплекс программных средств, обеспечивающих реализацию предложенной методики и проведены эксперименты, подтверждающие решение поставленной научно-технической задачи.

Публикации. Всего автор опубликовал 7 научных статей. По теме диссертации опубликовано 6 научных статей 1-6], в изданиях рекомендованных ВАК [1-6], 4 статьи– в изданиях, индексированных в базах SCOPUS [2-5]. Одна статья в рецензированном журнале. Получены свидетельства на регистрацию программ для ЭВМ и базы данных № ЕА-40490 «Аналитическая панель доступа данных», № ЕА-40486 «Поисковая система доступа к данным».

Личный вклад. Выносимые на защиту результаты получены соискателем лично. Результаты работ, написанных в соавторстве, получены автором лично, соавторы привлечены как консультанты. Одна работа выполнена без соавторов.

Структура и объём диссертации. Диссертация состоит из введения, 4 разбитых на параграфы Глав, списка литераторы из 98 наименований. Объем текста диссертации составляет 144 листа. Рисунков – 26. Таблиц – 3.

Глава 1 Некоторые ключевые проблемы поиска признаков инсайдера

1.1. Проблема идентификации инсайдеров и источники данных об инсайдерах

В настоящее время имеется целый ряд работ по выявлению и противодействию враждебным инсайдерским атакам. Рассмотрим наиболее интересные из них.

В работе [7] термин инсайдер относят к «законным пользователям, которые злоупотребляют своими привилегиями, и учитывая их знакомство и близость к вычислительной среде, могут легко нанести значительный ущерб или убытки». Согласно [8] можно рассматривать инсайдера «в качестве субъекта базы данных, обладающего персональным знанием информации в конфиденциальной области». В [9] определяют инсайдера как «любое лицо, которое имеет законный привилегированный доступ к внутренним цифровым ресурсам, то есть любому, кому разрешено видеть или изменять настройки компьютера организации, данные или программы таким образом, каким не могут быть произвольные представители общественности. В работе [10] был сделан вывод, что инсайдер «является лицом, которое было законно наделено правом доступа, представления или принятия решения об одном или нескольких активах структуры организации». Авторы [11] определяют инсайдера как «лицо, в настоящее время или одновременно уполномоченное на доступ к IS, данным или сети организации». В [12] дается определение инсайда в отношении политики безопасности, содержащей указанный набор правил. [12] определяет инсайдера как «доверенную сущность, которая имеет право нарушать одно или несколько правил в данной политике безопасности». Небинарный подход, указывающий на «степень инсайдерства» с правилами управления доступом, которые используются для разработки этих степеней, был предложен в [13], в котором определяется кто-то как «инсайдер в отношении доступа к некоторым четко определенным данным или ресурсам».

Согласно [14], инсайдерская угроза – это «действие инсайдера, которое подвергает риску организацию или ее ресурсы».

Инсайдерская угроза с различными типами знаний. В работе [15] злонамеренные инсайдерские угрозы делятся на две группы: предателей и маскировщиков. Эти два класса можно различить, основываясь на объеме имеющихся у них знаний. Предатели имеют полное представление о системах, с которыми они работают ежедневно, а также о фактической политике безопасности. Предатели обычно действуют от своего имени, а потому используют собственные полномочия для злонамеренных действий. С другой стороны, маскировщики могут обладать гораздо меньшими знаниями, чем предатели. Это злоумышленники, которые крадут учетные данные другого законного пользователя, а затем используют украденные учетные данные для совершения злонамеренного действия от имени другого пользователя. Другим примером является получение доступа к учетной записи жертвы путем использования некоторой уязвимости системы. Эти два класса не обязательно являются непересекающимися; предатели могут сначала использовать свой законный доступ для получения доступа другого пользователя, а затем нанести ущерб, используя этот доступ.

Авторы[1] предлагают профилировать вредоносные инсайдерские угрозы по трем типам:

1) *IT-саботаж*, при котором «инсайдер использует IT для причинения конкретного вреда в организации или физическому лицу». Такими инсайдерами обычно являются недовольные сотрудники с техническим образованием, имеющие административные привилегии. Примером этой категории является установка логической бомбы, которая активируется после увольнения сотрудника.

2) *Хищение интеллектуальной собственности*. Обычно речь идет о шпионаже, который обычно совершается техническим персоналом, например, инженерами и разработчиками, а также нетехническим персоналом, например,

[1] D. M. Cappelli, A. P. Moore, and R. F. Trzeciak. 2012. The CERT guide to insider threats: how to prevent, detect, and respond to information technology crimes (Theft, Sabotage, Fraud). Addison-Wesley.

клерками и продавцами. Преступники могут похищать информацию, к которой они имеют ежедневный доступ, и брать ее с собой, когда они покидают организацию (например, используя IP для своего дела, передавая ее новому работодателю или передавая другой организации).

3) *Мошенничество*, при котором «инсайдер использует ИТ для несанкционированного изменения, добавления или удаления данных организации для личной выгоды или кражи». Инсайдерское мошенничество, как правило, совершается сотрудниками низкого уровня с нетехническим опытом, такими как сотрудники отдела кадров или службы поддержки. Причиной этого часто являются жадность или финансовые трудности, и этот вид преступлений, как правило, носит долгосрочный характер. Вербовка таких мошенников внешними структурами также весьма распространена.

Тематические исследования, которые не подпадают под эти три профиля, обозначаются как разные[1].

Для того, чтобы дать единое представление о существующих таксономиях угроз инсайдеров, авторы[2] использовали методологию who, what, where, when, why и how (5W1H). Эти 5W1H представляют собой элементарные вопросы, касающиеся проблемы сбора информации, которые первоначально использовались для сообщения новостей, но имели и другие применения (например, построение онтологий доменов[3]).

В [16] предлагаются ориентированная на человека классификация внутренних нарушителей, рассматривая три измерения: роль инсайдера в системе, причину неправильного использования и последствия для системы.

1) *Роль в системе* – классифицирует людей по «типу и уровню системных знаний, которыми они обладают». В этот класс включены наиболее квалифицированные специалисты, знающие систему, которые имеют полный

[1] M. L. Collins, M. C. Theis, R. F. Trzeciak, J. R. Strozer, J. W. Clark, D. L. Costa, T. Cassidy, M. J. Albrethsen, and A. P. Moore. 2016. Common sense guide to prevention and detection of insider threats 5th edition. Pub-lished by CERT, Software Engineering Institute, Carnegie Mellon University (2016).

[2] I. Homoliak, F. Toffalini, J. Guarnizo, Y. Elovici, and M. Ochoa. In-sight into Insiders and IT: A Survey of Insider Threat Taxonomies, Analysis, Modeling, and Countermeasures. arXiv:1805-01612v2 [cs.CR] 10 Nov 2018.

[3] L. Yang, Z. Hu, J. Long, and T. Guo. 2011. 5W1H-based conceptual modeling framework for domain ontology and its application on STPO. In Int. Conference on Semantics Knowledge and Grid. IEEE, 203-206.

контроль над большинством ресурсов ИС (например, системные администраторы); а также наиболее продвинутые пользователи, которые, хотя и не имеют привилегированного доступа, приобрели большой объем знаний о системах и сетях организации, способные выявлять системные уязвимости (например, программисты, администраторы баз данных); и пользователи приложений, которые используют определенные стандартные приложения, такие как браузеры Интернета, офисные пакеты и клиенты электронной почты, но обычно не имеют дополнительных привилегий для доступа к ресурсам, кроме тех, которые требуются их приложениям.

2) *Причина неправильного использования* – это измерение описывает атрибуты инцидентов с инсайдерской угрозой. Рассматривая это измерение, авторы классифицируют инсайдеров на две группы: умышленные проступки, которые действуют по разным причинам (например, умышленное невежество, месть), и случайные проступки, которые также могут быть классифицированы по фактической причине, негативно влияющей на поведение законного пользователя (например, отсутствие подготовки, чрезмерная нагрузка, личные проблемы).

3) *Последствия для системы* – это измерение различает различные способы неправильного использования, что проявляется в определенных трассировках в ИТ-инфраструктуре на уровне системы. Авторы описывают три уровня, которые приписываются этим последствиям:

- ОС – модификации структуры файловой системы, установка неавторизованного программного обеспечения и т. д.;
- сеть – сетевые пакеты могут содержать несанкционированное содержимое, может осуществляться утечка конфиденциальных данных через электронную почту или службы обмена файлами и т. д.;
- и оборудование – вандализм или удаление аппаратных компонентов, установка ключевых регистраторов, модификации конфигураций по умолчанию для критически важных аппаратных компонентов (например, с целью саботажа или кражи IP).

Интересные тематические исследования также описаны в работах [13, 14, 16-18]. Все эти авторы имели дело либо с утечкой данных, либо с кражей данных, либо с саботажем, особенно в финансовом и военном секторах. Кроме того, некоторые из них также были сосредоточены на непреднамеренной инсайдерской угрозе, например, в [18] рассматривали фишинговые атаки, и в [14] описан эпизод непреднамеренного отказа в обслуживании.

Рассмотрим использование исходных данных в задачах выявления инсайдеров маскировщиков. Несмотря на то, что было проведено значительное количество исследований, касающихся проблемы обнаружения маскировщика, только в нескольких исследованиях использовались наборы данных, специально разработанные для этой цели. В следующих примерах используются наборы данных, а соответствующие вредоносные сценарии направлены на нарушения политики путем получения несанкционированного доступа.

Наиболее широко известные наборы данных, которые исследователи предлагают использовать для выявления инсайдеров:

Набор данных RUU. RUU – набор данных маскировщиков, который был представлен в работах [1] [2]. Набор данных был собран с ПК 34 обычных пользователей и состоит из событий, полученных из доступа к файловой системе, процессов, реестра Windows, динамических загрузок библиотек и графического интерфейса системы. Набор данных содержит маскарадные сеансы, выполняемые 14 людьми на основе заданной задачи нахождения любого фрагмента данных, имеющего прямую или косвенную финансовую ценность; пользователи не ограничиваются какими-либо конкретными средствами или ресурсами.

Набор данных WUIL. Набор данных Windows-Users and Intruder Simulations Logs (WUIL) был разработан и реализован в работе [3] и содержит общие взаимодействия файловой системы независимо от их типа (например, открытие, запись, чтение). Набор данных WUIL содержит записи от 20

[1] M. B. Salem and S. J. Stolfo. 2009. Masquerade attack detection using a search-behavior modeling approach. Columbia University, Computer Science Department, Technical Report CUCS-027-09 (2009).
[2] M. B. Salem and S. J. Stolfo. 2011. Modeling user search behavior for masquerade detection. In Int. Workshop on Recent Advances in Intrusion Detection. Springer, 181-200.
[3] B. Camiña, R. Monroy, L. A. Trejo, and E. Sánchez. 2011. Towards building a masquerade detection method based on user file system navigation. In Mexican Int. Conference on Artificial Intelligence. Springer, 174-186.

пользователей-добровольцев (увеличенный до 76 в дальнейших исследованиях[1]), которые контролировались в разные периоды времени во время их повседневной деятельности. Несмотря на то, что некоторые пользователи выпускали журналы на один час, другие выпускали журналы на несколько недель. Данные были собраны с помощью внутреннего инструмента для аудита файловой системы машин Windows различных версий (то есть XP, 7, 8 и 8.1). В то время как данные законных пользователей были собраны от реальных пользователей, сессии маскарада были смоделированы с использованием пакетных сценариев, учитывающих три уровня квалификации пользователей: базовый, промежуточный и расширенный.

Набор данных DARPA 1998. Набор данных DARPA 1998 по оценке обнаружения вторжений был синтезирован лабораторией Линкольна MIT на основе статистических параметров «правительственного сайта, содержащего 100» пользователей на 1000 «различных хостах»[2], и его основной целью было оценить и улучшить системы обнаружения вторжений. Однако он также использовался в исследованиях проблемы обнаружения инсайдерской угрозы [19]. Набор данных DARPA 1998 состоит из сетевых трассировок и журналов системных вызовов, созданных на атакуемых машинах. Выполненные атаки делятся на четыре группы:

1) «отказ в обслуживании»,

2) «удаленный для пользователя»,

3) «пользователь»

4) «наблюдение».

С точки зрения инсайдерской угрозы, единственной интересной группой атак является группа «пользователь», которую можно рассматривать как маскарадные атаки. Тем не менее, релевантны только системные журналы вызовов, поскольку прямые последствия этих атак могут отслеживаться только на хостах-

[1] B. Camiña, R. Monroy, L. A. Trejo, and M. A. Medina-Pérez. 2016. Temporal and spatial locality: an abstraction for masquerade detection. IEEE transactions on information Forensics and Security 11, 9 (2016), 2036-2051.
[2] R. P. Lippman, D. J. Fried, I. Graf, J. W. Haines, K. R. Kendall, D. McClung, D. Weber, S. E. Webster, D. Wyschogrod, R. K. Cunningham, and others. 2000. Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation. In DARPA Information Survivability Con-ference and Exposition, DISCEX'00, Vol. 2. IEEE, 12-26.

жертвах. Набор данных DARPA 1998 получил критику[1] и в настоящее время считается устаревшим.

Наборы данных по выявлению инсайдеров-предателей. Исследования, посвященные обнаружению предателей, были более ограниченными. Это различие можно объяснить предположением, что обнаружение маскировщиков проще, чем обнаружение предателя, как утверждается в [15], где утверждается , что «маскировщик, вероятно, будет совершать действия, несовместимые с типичным поведением жертвы». С другой стороны, учитывая, что злоумышленник является движущейся мишенью, он может в некоторой степени имитировать поведение жертвы наряду с совершаемыми им злонамеренными действиями. В следующем исследовании используются наборы данных, которые включают вредоносные намерения в данных, и направлены на нарушения политики с использованием законного доступа.

Enron Dataset. Набор данных Enron[2] состоит из 500 000 реальных электронных писем (с 1998 по 2002 год), связанных со 150 пользователями, в основном с высшим руководством корпорации Enron. Хотя некоторые из электронных писем были удалены, поскольку они содержали вложения или конфиденциальную информацию, набор данных содержит интересную информацию, которая может быть использована для анализа текста в электронных письмах и анализа социальных сетей, направленных на обнаружение инсайдерской угрозы с участием сотрудничающих предателей.

APEX 2007. Набор данных APEX «07 был собран, согласно работе[3], Национальным институтом стандартов и технологий (NIST) с намерением смоделировать задачи аналитиков в разведывательном сообществе. Набор данных APEX «07 состоит из действий и исследовательских отчетов восьми доброкачественных аналитиков, в то время как вредоносная инсайдерская угроза

[1] J. McHugh. 2000. Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by Lincoln laboratory. ACM Transactions on Information and System Security (TISSEC) 3, 4 (2000), 262-294.
[2] CALO Project. 2015. Enron Email Dataset. (2015). http://www.cs.cmu.edu/~enron/. Accessed on December 2020.
[3] E. Santos, H. Nguyen, F. Yu, K. Kim, D. Li, J. T. Wilkinson, A. Ol-son, and R. Jacob. 2008. Intent-driven insider threat detection in intelligence analyses. In Int. Conference on Web Intelligence and Intelligent Agent Technology. IEEE, 345-349.

была смоделирована пятью аналитиками, задачи которых основывались на задачах доброкачественных аналитиков, чтобы сделать обнаружение более сложным.

Успешные методы обнаружения инсайдерских угроз требуют сочетания различных подходов. В [20] смежные работы разделены на шесть категорий, вероятно, основанных на наиболее значимых тенденциях в области:

1) «психологические и социальные теории»,

2) «подходы, основанные на аномалии»,

3) «подходы, основанные на honeypot»,

4) «подходы, основанные на графах»,

5) «подходы к теории игр»,

6) «мотивирующие исследования».

В целом, существующие исследования касаются либо категоризации гетерогенных исследований [20, 21, 22] или однородные исследования, которые ограничены подходами к обнаружению [15, 23, 24, 25]. Характерной особенностью однородных исследований является то, что они ограничиваются конкретными областями применения: подходами обнаружения [23], подходами обнаружения инсайдерских угроз [24, 25] и профилирование инсайдеров на основе хоста/сети [15].

Однако последствия реализации инсайдерских атак могут быть непредсказуемые. Например, из-за инсайдерской атаки деятельность цифровой компании может быть приостановлена на какой-нибудь период, органы исполнительной власти могут наложить штраф, ввести ограничения деятельности против компании или инициировать расследование в отношении должностных лиц.

1.2. Поиск признаков инсайдера в больших данных и хранилищах данных

Наибольший риск имеют инсайдерские атаки, которые реализуются как утечка данных из хранилищ (Big Data).

Существует два риска утечки в среде Big Data:

1. Риск несанкционированного копирования конфиденциальных данных (“утечка”)
2. Риск несанкционированного просмотра конфиденциальных данных (“просмотр”)

Риск несанкционированного копирования конфиденциальных данных (“утечка”) пытаются устранить за счет создания безопасной инфраструктуры Big Data.

Наилучшие практики по созданию безопасной ИТ-инфраструктуры Big Data:

1. Создание защищенного контура.
2. Контроль выгрузок данных из контура.
3. Блокировка портов на рабочих местах.

Цель создания безопасной ИТ-инфраструктуры – вести всю обработку с конфиденциальными данными не на рабочем месте сотрудника, а на серверах и исключить возможность попадания любых данных на компьютер сотрудника. Таким образом, у сотрудника пропадет физическая возможность, например, извлечь жесткий диск с данными из компьютера в конце рабочего дня. Защищенный контур создается с помощью терминальных технологий, таких как виртуальные АРМ, бастион-хосты и т. д.

Риск фотографирования экрана сотрудником пытаются устранять с помощью “белых комнат” т. е. помещений с входным режимом, который запрещает пронос любых электронных устройств.

Риск несанкционированного просмотра конфиденциальных данных (“просмотр”) невозможно устранить за счет создания безопасной инфраструктуры т. к. мы имеем дело с намерениями сотрудника просматривать конфиденциальные данные. Такой риск снижают за счет систем маскирования данных, ролевых моделей доступа к маскированным/демаскированным данным и мониторинга, который выявляет расхождение между ролевыми моделями и фактическими конфигурациями доступа (которые часто динамически меняются). Но всегда остаются сотрудники, например разработчики витрин, ИТ-администраторы, сотрудники, отвечающие за качество данных и т. д., которые имеют по своим

должностным полномочиям легальный доступ к конфиденциальным данным. Сотрудник, имеющий доступ к конфиденциальным данным, может быть завербован, может совершать преступления из личной выгоды. В этой связи требуется технология, которая превентивно “подсветит” рисковое событие.

Большинство практик и ИТ-технологий, которые позволяют выявить инсайдера, функционируют после совершения им преступления. Например, система аудита, в которой регистрируется последовательность действий злоумышленника и это упрощает проведение расследования инцидента. База совершенных преступлений может использоваться для обучения системы мониторинга, чтобы формировать систему признаков рискованных событий.

Таким образом риск несанкционированного просмотра конфиденциальных данных (“просмотр”) минимизируется за счет проактивного мониторинга, который трудно реализовать в условиях Big Data.

Проблема утечек из хранилищ Big Data еще недостаточно изучена. Принципиально в среде Big Data существуют следующие проблемы.

1. Аудит Big Data является Big Data, поэтому просмотреть журналы аудита “глазами” в разумное время уже невозможно, требуются акселераторы интеллектуального анализа данных для аудиторов.
2. В Big Data могут работать тысячи пользователей, что порождает свои проблемы класса Big Data.
3. Данные постоянно расширяются и меняются (эффект «Open»).

Чтобы выявить инсайдера на ранних стадиях по слабозаметным следам следователи производили корреляция множества источников данных – кадровая база, база физических доступов в помещения, база доступов к информационным системам, роли в информационных системах. Такая работа является очень сложной в системах класса Big Data.

Методы обнаружения вторжений, основанные на аномалиях, для обнаружения вредоносного доступа к сканированию базы данных следует различать в зависимости от того, какие функции они извлекают из журнала аудита СУБД запросов SQL для моделирования поведения. Модели поведения

представляются в виде поведенческого профиля [26, 27]. Эти функции могут быть синтаксис-ориентированными, контекстно-ориентированными и data-ориентированными, что иногда в литературе называется результат-ориентированными [27, 28, 29].

Методы, использующие синтаксис-ориентированные функции, создают поведенческие профили, используя синтаксические функции SQL-запроса, включенные, но не ограничивающиеся атрибутами в предложении проекции, запрошенными отношениями, атрибутами в предложении выбора и/или типом команды SQL [27]. Методы, использующие функции, ориентированные на данные или результаты, строят поведенческие профили, используя данные, возвращенные в ответ на SQL-запрос, или любое другое статистическое измерение возвращенных данных, например, минимальное и максимальное значение в случае числовых данных. Например, можно использовать объем информации (процент возвращенных данных), возвращаемой в ответ на запрос, или возвращенные значения атрибутов для поведения модели или пользователя [29].

Контекстно-ориентированные методы создают поведенческие профили с использованием контекстных элементов. Контекстные особенности связаны с контекстом запроса, например, время, в которое был сделан запрос, идентификатор пользователя лица, выполняющего запрос, или количество запросов, выполненных за заданный период времени, и т. д. [30].

Сочетание контекста, синтаксиса и элементов, ориентированных на данные, может использоваться при моделировании нормативного поведения. Один из таких методов обнаружения аномалий, в котором используются синтаксис и информационные особенности, предложен в [28].

Система обнаружения аномалий, формирующая поведенческие профили, которые не понятны людям значимым образом, известна как система обнаружения аномалий Black-Box. С другой стороны, система обнаружения аномалий, формирующая поведенческие профили, которые понятны людям, известна как система обнаружения аномалий White-Box. Понятность подразумевает, что фактическая причина аномалии может быть идентифицирована администраторами

(офицером безопасности, администратором базы данных и т.д.) при проверке аномалии. Интуитивно белые подходы могут помочь объяснить аномалии.

Синтаксис, ориентированный на абстракцию SQL-запросов, подразумевает использование только факторов, связанных с синтаксисом инструкции SQL. Возникает вопрос, какую часть инструкции SQL можно рассматривать при построении поведенческой модели, то есть следует ли рассматривать всю инструкцию или некоторые части инструкции - задачу выбора соответствующей абстракции запроса SQL. Абстракция представляет собой кортеж инструкции SQL и состоит из таких функций запроса, как имя связи, имена атрибутов, объем возвращенных данных или любая статистика по возвращенным данным.

Был предложен ряд методов [27, 28, 31-36], которые преобразуют синтаксис инструкции SQL в более абстрактный fingerprint, который может использоваться для сравнения запросов. В [31] абстракции запросов SQL также называются fingerprints запросов SQL [32, 33], сигнатурами запросов SQL [27] или скелетами запросов SQL. Абстракции запросов, используемые существующими подходами обнаружения на основе аномалий, также описаны в этой статье. Помимо исследования [27], использование абстракции на практике также изучалось с помощью журналов аудита, поскольку обычно журналы аудита охватывают большое количество запросов и методов для значимого суммированиязнаний в журналах [1]. Например, в недавнем исследовании сообщалось, что за период времени в 19 часов в крупном банке США было сделано около 17 миллионов SQL-запросов [31].

Несколько подходов обнаружения аномалий для обнаружения вредоносных обращений к СУБД используют синтаксис-ориентированные функции SQL-запросов для построения поведенческих профилей [27, 37, 38]. Например, подход DetAnom [27] обнаруживает вредоносные обращения СУБД с помощью прикладных программ. SQL-запросы, выполняемые прикладной программой, представлены в виде SQL-абстракций запросов для формирования нормативного

[1] G. Kul, D. T. A. Luong, T. Xie, V. Chandola, O. Kennedy, and S. Upadh-yaya. Similarity metrics for SQL query clustering. IEEE Transactions on Knowledge and Data Engineering, 30(12):2408–2420, Dec 2018.

профиля приложения. В [27] абстракция запроса SQL состоит из следующих элементов (*c*, *t*, *r*, *q*, *n*), где *c* – тип команды SQL, например, SELECT. Атрибут *t* – это список идентификаторов атрибутов, спроецированных в запросе и относящихся к отношению, к которому они принадлежат. Атрибут *r* – это список идентификаторов связей. Атрибут *q* – это список идентификаторов атрибутов в предложении WHERE, а *n* – количество предикатов в предложении WHERE.

Предложенный в [26] подход демонстрирует, что можно построить системы обнаружения аномалий на основе поведения путем рассмотрения последовательности запросов (корреляций запросов) для моделирования поведения внутренних запросов для обнаружения вредоносных обращений, проявляющихся в последовательностях запросов, а не запроса в изоляции, к СУБД. Подход [26] моделирует поведение запроса инсайдера с использованием n-граммов, которые фиксируют краткие корреляции запросов SQL. Модель использовала абстракции журналов аудита запросов SQL для построения инсайдерских профилей, нормативный профиль с использованием безопасных журналов и профилей выполнения с использованием журналов аудита выполнения. Динамический профиль сравнивается с нормативными профилями и отклонения являются показателем аномалий. Эта модель, предложенная в [39], ввела представление о том, что поведение, которое редко (нечасто) представляет собой потенциально вредоносный доступ инсайдера, а частое поведение, возможно, является безопасным поведением. Была изучена область разработки набора элементов. Алгоритмы интеллектуального анализа набора элементов, включая PrePost +, Apriori-Inverse и Apriori-Rare, были приняты для разработки частых и редких наборов запросов для моделирования поведения запросов (с точки зрения абстракций запросов SQL). Результаты указывают на потенциальную эффективность моделирования поведения злонамеренных запросов инсайдерами как редкого поведения, которое также позволяет обнаруживать доступ к базам данных инсайдерами как аномалии. Ориентированный на синтаксис подход, в общем, полезен при обнаружении атаки маскировщика, а также атак SQL инъекции, обе эти атаки приводят к структурным изменениям в SQL-операторе.

Данные (результат) -центрические функции. Эти методы мало исследованы, было сообщено об использовании данных-ориентированных функций в качестве основы для обнаружения аномалий в контексте баз данных. Элементы, ориентированные на данные, включают в себя объем данных, возвращаемых в ответ на запрос, или возвращенные значения атрибутов или любой другой статистики, выполняемой с возвращенным набором значений атрибутов. В [29] утверждается, что только синтаксис-ориентированные особенности запроса являются плохим дискриминатором намерений. Эти запросы потенциально могут давать один и тот же результат, в то время как семантически подобные запросы могут давать различные результаты. Таким образом, инсайдер может создать законный SQL-запрос для извлечения результатов из базы данных, которую он уполномочен запрашивать, однако система обнаружения, основанная исключительно на синтаксисе, может не заметить аномальное поведение. В подходе, предложенном в [29], пользовательские профили представляют собой кластеры, составленные на основе *S*-вектора, который обеспечивает статистическую сводку результатов (кортежи/строки). На этапе обнаружения были приняты алгоритмы кластеризации, то есть, в качестве контролируемых методов обучения, кластеризация евклидовых *k*-средних, поддержка векторных машин (SVM), классификатор на основе дерева решений и наивного метода Байеса, а также методы кластеризации для обнаружения отклонений (на основе кластеризации евклидовых расстояний) и др.

Если запрос принадлежит кластеру, то он считается нормальным, иначе он рассматривается как аномальный. Представленный подход подходит для обнаружения изолированных запросов и не учитывает последовательность запросов.

Подход в [40] рассматривал моделирование поведения с точки зрения СУБД. Ориентированный на запись/СУБД подход (также называемый семантическим подходом) представлен в [40], который рассматривает частотные корреляции для обнаружения вредоносных доступов инсайдера как аномалии. При построении

профилей используются контрольные диаграммы из статистической модели процессов в качестве способа обнаружения аномалий.

Были рассмотрены два сценария, в первом сценарии обучающие данные для моделирования нормативного поведения содержат отклонения. Во втором сценарии обучающие данные для моделирования нормативного поведения не содержат отклонений. Эксперименты продемонстрировали эффективность подхода в обнаружении наблюдаемых атак инсайдеров как аномалий. Было обнаружено, что семантический подход не только идентифицирует невидимое поведение, но и идентифицирует поведение, которое должно было присутствовать в текущем поведении, как аномалии, которые мы называем надзорными аномалиями. Аномалии надзора – это аномалии, возникшие из-за небрежности человека или человеческих ошибок, например, случай, когда врач или медсестра пропускали ежедневный осмотр пациента. Было также продемонстрировано, что предлагаемая модель построения профилей, ориентированных на запись, может быть преобразована в модель построения профилей, ориентированных на роль. Подходы, ориентированные на данные, также способны обнаруживать сложные атаки. Например, атаки сбора данных, связанные с извлечением большого количества данных, поэтому превышают то, что извлекается законным пользователем.

В литературе сообщается о нескольких чисто контекстно-ориентированных подходах. Один такой контекстно-ориентированный подход представлен в [30], в котором контекстуальные особенности рассматриваются при моделировании поведения пользователя. В рамках этого подхода в качестве примера использования в медицинском секторе было использовано применение IDS на основе аномалии, и была изучена процедура Break-The-Glass (BTG), которая представляет собой процедуру, нарушающую традиционный механизм контроля доступа и обеспечивающую доступ к данным пациентов в случае чрезвычайной ситуации для сотрудников различных отделений. При таком подходе в [30] пользователи, которые должны вести себя аналогично, делятся на группы, а профили строятся для групп. Пространство элементов состоит из контекстных

элементов, таких как тип доступа, время, деление, дата. Профили были представлены в виде последовательности гистограмм и были построены с использованием концепции ячеек. Ячейки представляют частоту элементов. В фазе обнаружения измеряют расстояние между гистограммой пользователя и существующим профилем, и большее расстояние является указанием в качестве аномалии. Подход представляет профиль пользователя и группы в виде последовательности гистограмм признаков, которые могут быть легко интерпретированы заинтересованным человеком, как нарушение безопасности. Поэтому подход, приведенный в [30], можно классифицировать как подход обнаружения аномалий в условиях белого ящика. Контекстно-ориентированные подходы обычно увеличивают эффективность обнаружения ID-подхода в сочетании с синтаксисом или информационно-ориентированным подходом.

Пример подхода с использованием данных и синтаксис-ориентированных функций для построения поведенческих профелей представлен в [28]. Методы машинного обучения, в частности Native Bayes classifiers и multi-labeling classifiers, также были развернуты в процессе генерации профилей. Пользовательские профили строятся на фазе обучения из журналов, содержащих действия пользователей.

Подход преобразует SQL-запрос в абстракцию SQL-запроса, называемую 4-вектором. Этот вектор представляется (c, PR, PA, SR) состоит из элементов, ориентированных на данные и синтаксис, включая тип команды c; список связей, к которым обращается PR запроса; список атрибутов, к которым обращается запрос относительно отношения PA; и объем выбранной информации из отношения SR.

Этот гибридный подход демонстрируется в двух установках, которые связаны с обнаружением аномалии на основе ролей и обнаружением аномалии без контроля. На этапе обнаружения роль запросов прогнозировалась с использованием наивного классификатора Байеса с несколькими метками в случае дублирования ролей (что приводит к появлению нескольких ролей). Если прогнозируемая роль отличается от фактической роли, то запрос помечается как

аномальный. Запрос рассматривается как аномальный, если он попадает в кластер, который не содержит запроса, сделанного этим пользователем. Этот подход перспективен для обнаружения одного изолированного вредоносного запроса; однако он игнорирует последовательности запросов при моделировании поведения. Подход, представленный в работе [1], использует контекстно-ориентированные и информационно-ориентированные функции. Нормативные профили создаются путем обнаружения правил ассоциации между контекстно-ориентированными функциями и функциями, ориентированными на данные, с помощью частого интеллектуального анализа набора элементов[2]. Основная идея подхода заключается в том, чтобы связать результаты, полученные SQL-запросом, с контекстом, в котором они были получены. Например, транзакция, совершенная в Лондоне утром, обычно извлекает записи для сотрудников отдела людских ресурсов. Поэтому записи сотрудников отдела кадров связаны с контекстом, который они обычно извлекают утром из Лондона. На этапе обнаружения для любого входящего запроса были извлечены контекстно-ориентированные элементы и сопоставлены правила, соответствующие этим элементам, а затем результат запроса сопоставляется с результатами, связанными с извлеченными правилами. Недостатком такого подхода является то, что большие базы данных приводят к большим профилям. Кроме того, этот подход является слишком ограничительным и менее вероятным для масштабирования. Другим недостатком этого подхода является общий недостаток контекстно-ориентированных подходов, который представляет собой контекст, который можно легко имитировать.

Подход в работе [3], также использующий контекстно-синтаксические особенности для моделирования поведения и формирует некоторые предположения, например, что каждый отдел в организации имеет уникальное IP-пространство, сотрудники работают посменно (есть три смены в день). Функции,

---

[1] M. Gafny, A. Shabtai, L. Rokach, and Yu. Elovici. Poster: Applying unsupervised context-based analysis for detecting unauthorized data disclosure. In Proceedings of the 18th ACM Conference on Computer and Communica-tions Security, CCS'11, pages 765–768, New York, NY, USA, 2011. ACM.

[2] R. Agrawal, T. Imieli´nski, and A. Swami. Mining association rules be-tween sets of items in large databases. SIGMOD Rec., 22(2):207–216, June 1993.

[3] G. Zhiping Wu, S. L. Osborn, and X. Jin. Database Intrusion Detection Using Role Profiling with Role Hierarchy, pages 33–48. Springer Berlin Hei-delberg, Berlin, Heidelberg, 2009.

собранные для моделирования, включают идентификатор сотрудника, идентификатор роли, время, IP-адрес, тип доступа (прямой или через приложение). Подход также записывает SQL-запрос, связанный с контекстными функциями. Этот профиль состоит из вероятностей появления каждой функции, наблюдаемой для каждого пользователя. На этапе обнаружения новая транзакция сравнивается с построенной в профиле для проверки ближайшего эмитента (пользователя) этой транзакции. Транзакция помечается как аномалия в случае, если эмитент транзакции отличается от вычисленного. Этот подход также учитывал иерархию ролей, что означает, что если роль 1 выше роли 2 в иерархии, то привилегии доступа роли 2 являются подмножеством роли 1. Например, если запрос, сделанный 1, маркирован как злонамеренный, но тот же вопрос законен для 2 тогда, это не рассматривают как злонамеренный вопрос. Этот подход направлен на расширение управления доступом на основе ролей (RBAC) для обеспечения выполнения запроса только авторизованными пользователями. Подобно подходам, описанным выше, этот подход также обнаруживает только одну вредоносную транзакцию СУБД, где позже в документе утверждается, что один запрос может быть правомерным, однако группа из них, составленная вместе, может привести к злонамеренным или нелегитимным действиям.

В [41] рассматривается моделирование контекстуальных данных и синтаксис-ориентированных признаков. Подход обнаружения вторжений в базу данных в [41] адаптирован для хранилищ, в которых приложения получают доступ к хранилищам данных через Интернет. Эти профили строятся с учетом различных признаков и представлены с точки зрения вероятностного распределения каждого признака для каждого пользователя и для всей совокупности. На этапе обнаружения для этого подхода проводится тестирование для согласования распределения признаков с распределениями, полученными на этапе обучения, с использованием статистических тестов, таких как тесты Хи-квадрата и Колмогорова Смирнова [41]. В случае несоответствия деятельность, связанная с этим признаком, маркируется как аномалия. Подход ориентирован на

веб-вредоносный доступ к хранилищам данных, хотя инсайдеры остаются без внимания.

Обнаружение аномалий – это единый подход к обнаружению инсайдерской угрозы. Методы обнаружения аномалий можно разделить на три категории в отношении того, какой объем маркированных данных необходим: неконтролируемые, полуконтрольные и контролируемые.

Получение точных меток всех типов инцидентов для контролируемого обучения часто дорого и нецелесообразно. Неконтролируемые методы не требуют маркированных данных, но они имеют высокий ложноположительный уровень, потому что они оперируют предположением, что аномалии реже, чем номинальные. Этот уровень можно уменьшить, введя обратную связь, известную как экспертная обратная связь или активное обучение. Это позволяет аналитику маркировать подмножество данных. Другая проблема заключается в том, что модели часто не интерпретируются, поэтому неясно, почему модель решила, что экземпляр данных является аномалией.

1.3 Обнаружение аномалий

Аномалии считаются шаблонами в данных, которые отклоняются от ожидаемых.

Инсайдеры склонны адаптировать свое поведение, чтобы оно выглядело номинальным. Понятие номинального поведения может развиваться в окружающей среде, и нынешнее понятие номинального поведения может не отражать номинальное поведение в будущем. Это называется концептуальным дрейфом [42].

Кроме того, серьезной проблемой является отсутствие маркированных данных для подготовки и проверки моделей. Кроме того, нет такого понятия, как лучший детектор аномалий во всех доменах. Производительность детектора

зависит от того, насколько хорошо определение аномалии детектором соответствует концепции аналитика об "интересной аномалии"[1].

Из-за этих проблем обнаружение аномалий затруднено в общем виде. Большинство существующих методов сосредоточено на решении конкретной формулировки проблемы. Обычно факторы определяются доменом, в котором обнаруживаются аномалии[2].

Кроме того, при определении того, какие методы могут использоваться в реальном мире, следует учитывать вычислительную мощность, память и ограничения реального времени. Обнаружение аномалий может выполняться в автономном режиме (также известном как пакетный режим), если доступны все экземпляры данных. Большинство методов обнаружения аномалий работают в автономном режиме. В оперативном режиме экземпляры данных часто поступают в реальном времени последовательно в потоках данных [42, 43].

Типы методов обнаружения аномалий

Методы обнаружения аномалий используют контролируемый, полунадзорный или неконтролируемый подход. Контролируемые подходы имеют более высокие показатели обнаружения и более низкие объемы ложных тревог (FPR), чем неконтролируемые подходы; однако неконтролируемые подходы могут обнаруживать неизвестное поведение, но контролируемые подходы не могут. Получение точных и репрезентативных данных обо всех видах поведения часто обходится дорого. Маркировка обычно выполняется экспертом. Получение данных аномального поведения сложнее, чем получение данных номинального поведения. Зачастую данные носят динамичный характер; например, могут возникать новые аномалии и старые аномалии могут больше не считаться аномалиями[3].

Далее представлены несколько наиболее популярных методов обнаружения неконтролируемой точечной аномалии.

---

[1] Das S., Wong W.K., Fern A., Dietterich T.G. & Siddiqui M.A. (2017) Incorporating feedback into tree-based anomaly detection. arXiv preprint arXiv:1708.09441.

[2] Chandola V., Banerjee A. and Kumar V. (2009) Anomaly detection: A survey. ACM Computing Surveys 41, pp. 1–58. URL: https://doi.org/10.1145/ 1541880.1541882.

[3] Gogoi P., Bhattacharyya D.K., Borah B. & Kalita J. (2011) A survey of outlier detection methods in network anomaly identification. Comput. J. 54, pp. 570–588.

1. Локальный коэффициент отклонения

Локальный коэффициент отклонения (LOF), введенный в работе[1], является наиболее известным алгоритмом обнаружения локальных отклонений на основе плотности и первым ввел понятие локальных отклонений [44]. Этот алгоритм присваивает степень отклонения каждому экземпляру данных. Алгоритм является локальным в том смысле, что он рассматривает только ограниченное соседство экземпляра данных. Локальные отклонения являются отклонениями относительно плотностей их района. Локальные плотности оцениваются с использованием k-ближайших соседей для каждого экземпляра данных.

LOF требует построения окрестности вокруг каждой точки данных, включая вычисление попарного расстояния с каждой точкой данных, что представляет собой процесс со сложностью времени O(n2). Подвыборка может использоваться для уменьшения сложности LOF. Миннесотская система обнаружения вторжений [45] осуществляет выборку набора данных и сравнивает все точки данных с этим меньшим набором, что уменьшает сложность времени до $O(n - m)$, где n – размер данных, а m – размер подвыборки.

Преимущество LOF заключается в том, что он может использоваться для обнаружения всех видов отклонений, включая те, которые не могут быть обнаружены с помощью алгоритмов на основе расстояния [45].

Согласно оценке, проведенной в [44], методы на основе LOF плохо работают с наборами данных, содержащими глобальные аномалии, генерируя множество ложных срабатываний; следовательно, этих методов следует избегать, если цель состоит в обнаружении глобальных аномалий. Эти методы не приспособлены для выявления признаков сговора инсайдеров.

2. Объединение в кластеры

Кластеризация - популярный метод, используемый для обнаружения аномалий. Обнаружение аномалий на основе кластеризации обычно проводится неконтролируемым или полунадзорным образом.

---

[1] Breunig M.M., Kriegel H.P., Ng R.T. & Sander J. (2000) Lof: identi-fying density-based local outliers. In: ACM sigmod record, vol. 29, ACM, vol. 29, pp. 93–104.

В первом подходе модель обучается с использованием как номинальных, так и аномальных данных; в последнем случае его обучают, используя только номинальные данные, а затем обучаемую модель используют в качестве профиля, описывающего номинальное поведение. Если предположение гласит, что аномальные экземпляры данных являются меньшинством и сами по себе не образуют больших кластеров, то можно использовать неконтролируемый подход[1].

Существует много способов кластеризации и их вариантов. Наиболее распространенной метрикой расстояния, используемой в кластеризации, является евклидово расстояние. В этой метрике каждый элемент вносит одинаковый вклад в расчет расстояния. Это может быть нежелательно во многих применениях; например, когда признаки имеют различные вариации или коррелированы. Это приведет к тому, что элементы с более высокой изменчивостью будут доминировать над элементами с более низкой изменчивостью. Альтернативой является метрика расстояния, известная как расстояние Махаланобиса[2].

K-means является одним из наиболее известных алгоритмов кластеризации. Обычно используется потому, что имеет линейную сложность по времени [44]. Одним из таких основанных на аномалиях IDS является ADMIN[3], который контролирует использование терминала пользователя, создает базовый профиль, соответствующий номинальному использованию для пользователя, и проверяет будущие экземпляры данных относительно профиля. K-means разбивает данные на k кластеров. Вместо k-means, ADMIN использовал подход динамической кластеризации, потому что они не хотели устанавливать k.

Кластеризация к-means также использовалась для обнаружения аномалий в сетевом трафике. Munz et al.[4] использовали k-means кластеризацию для разделения трафика на номинальные и аномальные кластеры. Затем они вычисляли центроиды кластера для использования в масштабируемом

[1] Bhuyan M.H., Bhattacharyya D.K. & Kalita J.K. (2014) Network anomaly detection: Methods, systems and tools. IEEE Communications Sur-veys & Tutorials 16, pp. 303–336. URL: https://doi.org/10.1109/surv.2013. 052213.00046.
[2] Mimmack G. M., Mason S. J., and Galpin J. S. (2001) Choice of dis-tance matrices in cluster analysis: Defining regions. Journal of climate 14, pp. 2790–2797.
[3] Sequeira K. & Zaki M. (2002) Admit: anomaly-based data mining for intrusions. In: Proceedings of the eighth ACM SIGKDD international confer-ence on Knowledge discovery and data mining, ACM, pp. 386–395.
[4] Münz G., Li S. & Carle G. (2007) Traffic anomaly detection using k-means clustering. In: GI/ITG Workshop MMBnet, pp. 13–14.

обнаружении аномалий в реальном времени. Они выбрали k = 2, потому что предположили, что номинальный и аномальный трафик образуют два кластера.

Стандартный SVM был представлен Boser et al.[1] в 1992 году. Исходный SVM представляет собой двухклассовую методику контролируемой классификации. Одноклассный SVM (OC-SVM) является разновидностью широко используемого подхода SVM при обнаружении аномалий.

Он был введен Schölkopf et al. [46], и идея использования OC-SVM для обнаружения аномалий была предложена Schölkopf et al.[2] в 2001 году. OC-SVM обычно используется полууправляемым образом. Предполагается, что данные обучения относятся только к одному классу[3]. OC-SVM Шёлькопфа создает гиперплоскость[4] с максимальным запасом от начала координат в пространстве признаков, отделяющем все точки данных от начала координат [46]. OC-SVM может использоваться для обнаружения аномалий, поскольку он может преодолеть проблему чрезмерной подгонки, создав мягкий запас, используя переменные ослабления[5].

Когда набор данных не может быть разделен линейно, функции ядра могут использоваться для отображения его в более высокое размерное пространство, где он может быть разделен линейно. Это называется "хитрость ядра"[6].

Подход OC-SVM Tax and Duin [7] использует гиперсферу вместо гиперплоскости. Он направлен на минимизацию объема гиперсферы. Он использует переменные слабости для создания мягкой границы аналогично подходу Шёлькопфа.

1.3.1 Методы на основе ансамбля

[1] Boser B.E., Guyon I.M. & Vapnik V.N. (1992) A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on Computational learning theory, ACM, pp. 144–152.
[2] Schölkopf B., Williamson R.C., Smola A.J., Shawe-Taylor J. & Platt J.C. (2000) Support vector method for novelty detection. In: Advances in neural information processing systems, pp. 582–588.
[3] Wagner C., François J., State R. & Engel T. (2011) Machine learning approach for IP-flow record anomaly detection. In: NETWORKING 2011, Springer Berlin Heidelberg, pp. 28–39. URL: https://doi.org/10.1007/ 978-3-642-20757-0_3.
[4] Erfani S.M., Rajasegarar S., Karunasekera S. & Leckie C. (2016) High dimensional and large-scale anomaly detection using a linear one-class svm with deep learning. Pattern Recognition 58, pp. 121–134.
[5] Vapnik V.N. (1999) An overview of statistical learning theory. IEEE transactions on neural networks 10, pp. 988–999.
[6] Jakkula V. (2006) Tutorial on support vector machine (svm). School of EECS, Washington State University 37.
[7] Tax D.M. & Duin R.P. (2004) Support vector data description. Ma-chine Learning 54, pp. 45–66. URL:https://doi.org/10.1023/b:mach. 0000008084.60811.49.

Основная идея ансамблевых методов заключается в объединении нескольких детекторов для создания детектора, который превосходит каждый отдельный детектор в ансамбле[1].

В 2008 году Лю и др.[2] предложен iForest, который представляет собой алгоритм дерева принятия решений на основе ансамбля для обнаружения неконтролируемой аномалии. Позже он был улучшен в 2012 году [47]. iForest — это современный способ обнаружения неконтролируемой аномалии.

Большинство основанных на модели подходов к обнаружению аномалий сначала создают профиль номинальных экземпляров, а затем идентифицируют как аномалии любые экземпляры, которые не соответствуют этому профилю. По этой причине эти подходы оптимизированы для профилирования номинальных экземпляров вместо обнаружения аномалий. Это может привести к ложным срабатываниям (т.е. обнаружению номинальных экземпляров как аномальных).

iForest изолирует аномалии вместо профилирования номинальных экземпляров. Этот метод предполагает, что аномалии находятся в меньшинстве и что их значения сильно отличаются от номинальных экземпляров. Метод строит совокупность изолированных деревьев для данного набора данных, а затем рассматривает как аномальные те экземпляры данных, которые превышают средние длины пути. Идея заключается в том, что отклонения легче отделить от остальных данных, чем потери. Оценка аномалии выводится из средней длины пути [47].

iForest строит модель с несколькими подвыборками, чтобы уменьшить влияние маскировки и демпфирования. Маскирование происходит, когда слишком много аномалий, когда скопления аномалий становятся большими и плотными. С другой стороны, ошибка происходит, когда номинальные экземпляры данных идентифицируются как аномальные. Обнаружение аномалии с помощью iForest протекает в два этапа. Первоначальная подготовка осуществляется путем построения изолированных деревьев с использованием подвыборок обучающего

[1] Rokach L. (2010) Ensemble-based classifiers. Artificial Intelligence Re-view 33, pp. 1–39. URL: https://doi.org/10.1007/s10462-009-9124-7.
[2] Liu F. T., Ting K. M. & Zhou Z. H. (2008) Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining, IEEE, pp. 413–422.

комплекта. В изолированных деревьях в каждом узле выбирается случайный элемент и случайная точка разделения для этого элемента. На стадии тестирования баллы аномалий для проверочных экземпляров вычисляются путем их пропускания через изолированные деревья [47].

Согласно эмпирической оценке [47], iForest работает лучше, чем LOF, с точки зрения ROC-AUC и имеет линейную сложность по времени. iForest хорошо справляется с проблемами высоких размеров и в ситуациях, когда данные обучения не имеют каких-либо аномалий. iForest также имеет низкую потребность в памяти. Из-за этих факторов он хорошо масштабируется до больших наборов данных с высокомерными проблемами и большим количеством неактуальных атрибутов.

Существует много вариаций iForest. Zhiguo и Minrui[1] расширили iForest для работы над потоковой передачей данных с использованием метода скользящего окна. Marteau et al.[2] предложил гибридный iForest, который решает ограничения в iForest, известном как "слепые зоны", и распространяет его на полуконтрольное и контролируемое обучение. Контролируемое расширение позволяет включить знания об известных аномалиях. "Слепые зоны" существуют в iForest, поскольку алгоритм предполагает, что аномалии имеют более короткие длины пути, чем номинальные данные. Это справедливо для нормально распределенных данных; однако в целом это не так. Например, это не верно для данных, распределенных в вогнутом наборе, как тор, как показано Marteau et al. [44].

В случае тора "слепое пятно" находится внутри него и iForest выступает не лучше в "слепом пятне", чем случайный классификатор. Гибридный iForest использует расстояния до соседних номинальных и аномальных данных в качестве дополнительных источников информации для преодоления "слепой зоны".

Помимо тестирования гибридного iForest на синтетических данных, он также был протестирован на наборах данных обнаружения вторжений. Было

[1] Ding Z. & Fei M. (2013) An anomaly detection approach based on iso-lation forest algorithm for streaming data using sliding window. IFAC Pro-ceedings Volumes 46, pp. 12–17. URL: https://doi.org/10.3182/20130902-3-cn3020.00044.
[2] Marteau P.F., Soheily-Khah S. & Béchet N. (2017) Hybrid isolation forest application to intrusion detection. arXiv preprint arXiv:1705.03800.

обнаружено, что он выгодно сравнивается со стандартными iForest и одно- и двухклассными SVM с точки зрения ROC-AUC.

Das et al.[1] также расширил iForest до полуконтролируемого обнаружения аномалий путем включения обратной связи от аналитика. По сравнению со стандартным iForest их методика выполнялась аналогично или лучше в зависимости от набора данных.

Сходным с iForest способом является надежный случайный разрез iForest, предложенный Guha et al.[2]. В стандартном iForest размеры, подлежащие разделению, выбираются равномерно случайным образом. Преимущество этого заключается в том, что размеры обрабатываются независимо и на них не влияет различное масштабирование размеров.

Недостаток заключается в том, что, поскольку разрезы выбираются равномерно по всем измерениям, когда в наборе данных много несущественных измерений, большинство разделов находятся в несущественных измерениях, и это приводит к плохому выполнению алгоритма.

Надежный произвольный вырез iForest также предназначен для работы с потоками и может быть динамически обновлен путем вставки и удаления точек данных.

Интерпретируемые модели необходимы для поддержки аналитика в проведении различия между вредоносной и не вредоносной деятельностью. Исследования подтвердили пользу объяснений для аналитика[3].

Предоставление описаний имеет решающее значение для обратной связи, поскольку будущая производительность детектора аномалий зависит от способности аналитика правильно классифицировать экземпляры.

С точки зрения автора диссертации, объяснимость результатов автоматического анализа основывается на знании причинно-следственных связей,

[1] Das S., Wong W.K., Fern A., Dietterich T.G. & Siddiqui M.A. (2017) Incorporating feedback into tree-based anomaly detection. arXiv preprint arXiv:1708.09441.
[2] Guha S., Mishra N., Roy G. & Schrijvers O. (2016) Robust random cut forest based anomaly detection on streams. In: International conference on machine learning, pp. 2712–2721.
[3] Wagstaff K.L., Lanza N.L., Thompson D.R., Dietterich T.G. & Gilmore M.S. (2013) Guiding scientific discovery with explanations using demud. In: Twenty-Seventh AAAI Conference on Artificial Intelligence.

предшествовавших результату или вероятностным оценкам случайного появления результата в данных.

Доши-Велез и Ким [48] определили интерпретируемость как "способность объяснять или представлять в понятных терминах человеку". Das et al. [49] определили объяснения как "наиболее важную причину (причины), которая (ые) повлияла на прогнозируемый выход алгоритма", и интерпретируемость как "представление предсказаний лаконичным и простым в понимании образом"

Необходимость интерпретируемости возникает из-за неполноты формализации проблемы [48].

Чем более интерпретируема модель, тем легче понять, почему она сделала определенный прогноз. Следовательно, существует потребность в интерпретируемости, когда имеется ценность в предоставлении обоснования для предсказания. Кроме того, модели могут быть отлажены и проверены только в том случае, если их можно интерпретировать. Когда что-то в модели идет не так, объяснение ошибочного прогноза может быть использовано для понимания причины ошибки. Кроме того, пользователям проще доверять системе, которая дает хорошие объяснения [49, 50].

Существуют модельно-агностические и модельно-специфические методы интерпретации. Обычно используемым решением проблемы объяснимости является использование специфичных для модели методов, так называемых "интерпретируемых" моделей, таких как деревья решений и правила принятия решений[1]. В этих подходах используются модели, в которых объяснения могут быть созданы путем проверки компонентов модели, таких как одно правило или путь в дереве решений. Они работают до тех пор, пока модель точна и внутренние компоненты модели достаточно ограничены [51].

Модельно-агностические методы являются гибкими. Модель рассматривается как черный ящик, и объяснения отделяются от модели. Следовательно, объяснения работают с любой моделью. Модели-агностические

[1] Letham B., Rudin C., McCormick T.H., Madigan D. et al. (2015) In-terpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. The Annals of Applied Statistics 9, pp. 1350–1371.

методы используют подходы, такие как изучение интерпретируемой модели из прогнозов модели черного ящика, путем изменения входных данных и наблюдения за тем, как модель реагирует или и то, и другое [51]. Недавно введенный модельно-агностический метод является локальным интерпретируемым модельно-агностическим объяснением[1].

Если модель очень сложна, получить глобальное представление о ней может быть очень трудно. Локальные объяснения могут быть несовместимы друг с другом, поскольку модель может использовать функцию по-разному в зависимости от других функций [51].

Модельно-агностические методы сталкиваются с некоторыми проблемами. Рудин [52] утверждал, что модельно-агностические методы не должны использоваться при принятии решений с высокими ставками. По словам Рудина [52], широко распространено мнение, что более сложные модели черных ящиков более точны; однако это часто не соответствует действительности, особенно когда данные представлены с точки зрения естественно значимых признаков. Объяснения из модельно-агностических методов могут быть более неточными, чем объяснения из модельно-специфических методов. Вместо этого, когда это применимо, следует разработать интерпретируемые модели [52].

1.4 Использование графовой аналитики для задач выявления инсайдера

Подчеркивая важность объединения многих аспектов инсайдерской угрозы (в том числе технических и поведенческих событий/показателей, а также человеческих факторов, таких как ускоряющее событие, личностные характеристики, историческое поведение, мотивация к нападению и способность атаковать [53], [54]), сообщество исследователей инсайдерской угрозы сосредоточило свое внимание на высокоразмерных, разнородных методах анализа

[1] Ribeiro M.T., Singh S. & Guestrin C. (2016) Why should i trust you?: Ex-plaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp. 1135–1144.(

данных. Чтобы преодолеть трудности в анализе высокоразмерных, разнородных данных, некоторые исследователи, связанные с инсайдерской угрозой, сосредоточились на подходах к машинному обучению и подходах на основе графов. Возможности представления многомерных данных, улучшенные возможности визуализации и возможность обнаружения неконтролируемых аномалий являются основными движущими силами, чтобы продолжить исследования по обнаружению инсайдерских угроз на основе графовых подходов.

Основываясь на том, что эта исследовательская работа сосредоточена на обнаружении аномалий в моделях графов в области инсайдерских угроз, только литература, связанная с основанными на графах моделями обнаружения инсайдерских угроз, обнаружением аномалий в моделях графов и ранжированием отклонений в этих графах, обсуждается в следующих подразделах.

1.4.1 Подход к обнаружению инсайдерских угроз.

Структура, предложенная Ченом и Малином[1], является графовым подходом к обнаружению инсайдерских угроз в совместных информационных системах (CIS). Их модель содержит компонент извлечения реляционной структуры и блок обнаружения аномалий. Журналы доступа CIS добываются для сообществ пользователей с использованием двухстороннего отображения графов и моделей уменьшения размерности. Основанная на сообществе система обнаружения аномалий (CADS) предлагаемой инфраструктуры обнаруживает ближайшие соседи каждого пользователя и вычисляет отклонение каждого пользователя от его ближайших соседей. Даже несмотря на то, что они определили требование обнаружения сообщества для идентификации возможных аномальных пользователей, отображение графов не расширяется до модельных графов.

Структура, предложенная Эберле и Холдером [2], является еще одним основанным на графах подходом для обнаружения вредоносных инсайдерских угроз. Аномалии обнаруживаются с помощью подструктур графа, которые не

[1] Y. Chen and B. Malin, “Detection of anomalous insiders in collabora-tive environments via relational analysis of access logs,” in Proceedings of the first ACM conference on Data and application security and privacy. ACM, 2011, pp. 63–74.
[2] W. Eberle and L. Holder, “Applying graph-based anomaly detection ap-proaches to the discovery of insider threats,” in Intelligence and Security In-formatics, 2009. ISI ’09. IEEE International Conference on, 2009, pp. 206–208.

являются изоморфными нормативной подструктуре графа, с использованием принципа минимальной длины описания для обнаружения аномальных активностей. Они сосредоточились на трех широких типах аномалий графов, выявленных в их работе, а именно на вставках, модификациях и удалениях. В этом подходе отсутствует учет многих атрибутов из входных данных, связанных с инсайдерской угрозой, и использование нескольких алгоритмов для различных аномалий может привести к сложной структуре обнаружения угрозы.

Система упреждающего обнаружения внутренних угроз, предложенная Brdiczka et al [53], использует графическое обучение и психологическое моделирование пользователей. Эта модель представляет собой комбинацию модели обнаружения структурных аномалий и модели психологического профилирования, которая изучает возможность включения динамических свойств узловых атрибутов. Альтебян и Панда[1] также предложили использовать теорию графов для формулирования двух компонентов, графа знаний и графов зависимостей объектов. Граф знаний представляет единицы знаний для данного инсайдера, и они обновляются с течением времени. Граф зависимостей — это глобальный иерархический граф, который показывает все зависимости между различными объектами. Несмотря на то, что эта модель пытается включить накопленные знания об инсайдере с течением времени по системам и объектам, ее можно улучшить, включив несколько других параметров, таких как модели поведения пользователя и психологические аспекты.

В другом исследовании, проведенном Нэнсом и Марти [49] было представлено использование двудольных графов для выявления и визуализации инсайдерской угрозы. Они пытались установить приемлемые модели поведения инсайдеров на основе классификаций ролей рабочих групп. Высокие ложноположительные показатели являются одним из недостатков этого метода, даже несмотря на то, что он способен обнаруживать определенные инсайдерские

[1] Q. Althebyan and B. Panda, “A Knowledge-Base Model for Insider Threat Prediction,” in Information Assurance and Security Workshop, 2007. IAW ’07. IEEE SMC, 2007, pp. 239–246.

угрозы. Кент и др.[1] предложили использование подграфов аутентификации для анализа поведения пользователей в корпоративной сети с использованием набора атрибутов подграфа при профилировании пользователей. Анализ временных рядов подграфов и использование двудольных графов также вводятся в их работу, которая сосредоточена на гораздо более всеобъемлющем анализе в их текущей работе.

Вклад вышеописанной исследовательской работы ясно указывает на целесообразность использования основанных на графах подходов для представления данных, связанных с инсайдерской угрозой. Кроме того, ясно, что исследователи были сосредоточены на различных методах обнаружения аномалий для выявления возможных вредоносных пользователей. Но многие из этих предложенных методов не расширяются, чтобы отразить неоднородность и высокую размерность, связанные с проблемой инсайдерской угрозы. Наконец, предложенные выше методы обнаружения аномалий на основе графов не учитывают одновременно и топологию графов, и атрибуты графов.

1.4.2 Кластеризация графов в графовых моделях

Помеченные графы обеспечивают способ получения более богатого представления графа, в котором узлы и рёбра демонстрируют свои свойства как через топологию графа, так и через атрибуты графа. Эти типы графов могут быть применимы во многих приложениях, таких как социальные сети, сети транзакций, технологические сети, биологические сети и так далее, в которых большая часть связанной информации хранится как атрибуты соответствующих вершин и рёбер. Учитывая разнообразие связанной входной информации с проблемой инсайдерской угрозы, мы считаем, что помеченные графы можно рассматривать как подходящее средство для представления данных.

[1] A. D. Kent et al., “Authentication graphs: Analyzing user behavior within an enterprise network,” Computers & Security, vol. 48, pp. 150–166, feb 2015.

Обнаружение аномалий в помеченных графах можно разделить на методы, основанные на структуре, и методы, основанные на сообществе[1]. Основная идея методов на основе структуры состоит в том, чтобы идентифицировать необычные подструктуры на основе связности графа, а также атрибутов. Методы обнаружения аномалий на основе сообщества фокусируются на узлах, в которых значения атрибутов значительно отличаются от других членов сообщества, к которому они принадлежат. Основываясь на рассуждениях, лежащих в основе вышеупомянутых методов обнаружения аномалий, мы считаем, что обнаружение аномалий на основе сообщества в помеченных графах более подходит для рамок обнаружения инсайдерской угрозы, так как это был бы лучший подход для анализа индивидуального поведения по отношению к коллегам. Приписываемая кластеризация графов получила гораздо больше внимания в недавнем прошлом с выявлением требования анализа высокоразмерных, гетерогенных данных. Метод, предложенный Чжоу и др.[2], генерирует новые вершины ("вершины атрибута") для каждого атрибута и новые ребра ("ребра атрибута") между вершиной и вершиной атрибута, если соответствующая вершина имеет выбранное значение атрибута. Затем они используют унифицированную модель случайного обхода окрестностей на дополненном графе, чтобы найти кластеры.

В отличие от вышеописанного способа Мозера и др.[3], Gunnemann и др.[4][5] предложили кластеризацию приписываемых графов в качестве метода двоякой кластеризации, который одновременно представляет поднаборы атрибутов и плотные подграфы. Кластер подпространств — это набор объектов с соответствующим размером, в котором атрибуты объекта очень похожи друг на друга. Плотные подграфы представляют собой совокупность узлов, плотно

[1] L. Akoglu, H. Tong, and D. Koutra, “Graph based anomaly detection and description: a survey,” Data Mining and Knowledge Discovery, vol. 29, no. 3, pp. 626–688, 2015.
[2] Y. Zhou, H. Cheng, and J. X. Yu, “Graph Clustering Based on Structural/Attribute Similarities,” Proc. VLDB Endow., vol. 2, no. 1, pp. 718–729, aug 2009.
[3] F. Moser et al., “Mining Cohesive Patterns from Graphs with Feature Vectors,” in Proceedings of the 2009 SIAM International Conference on Data Mining, pp. 593–604.
[4] S. Gunnemann et al., “Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms,” pp. 845–850, 2010.
[5] S. Gunnemann et al., “Efficient mining of combined subspace and subgraph clusters in graphs with feature vectors,” in Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, 2013, pp. 261–275.

связанных друг с другом на основе свойства "квазиклика" [1] . Среди вышеупомянутых приписанных алгоритмов кластеризации графов мы считаем, что методы кластеризации подпространства и подграфа более подходят в контексте инсайдерской угрозы, по сравнению с дополненной кластеризацией графов атрибутов.

1.4.3 Методы ранжирования превышения

Алгоритмы, упомянутые в вышеупомянутых подразделах, в основном сосредоточены на кластеризации графов по помеченным графам.

Но объем этих алгоритмов не охватывает обнаружение аномалий и ранжирование отклонений одновременно. Методы, предложенные Gao et al. [2], одновременно находят сообщества, а также идентифицируют выбросы сообщества с использованием неконтролируемого алгоритма обучения, называемого "CODA". Но использование глобального пространства атрибутов для обнаружения сообществ создаст ограничения на прямое решение проблемы инсайдерской угрозы. Другим недавним механизмом кластеризации является "FOCUSCO", предложенный Perozzi et al.[3], который соединяет как кластеризацию графов, так и обнаружение отклонений и использует ориентированную на пользователя методику выбора атрибутов, которая значительно отличается от других подходов.

"GOutRank" - первый подход к ранжированию отклонений в подпространствах меточной кластеризации[4]. Они использовали существующие методы выбора подграфов и подпространств. Предлагаемый механизм ранжирования выбросов использует три индикатора, которые включают в себя измерение подпространства, измерение кластера и структуру графа.

[1] G. Liu and L. Wong, “Effective Pruning Techniques for Mining Quasi-Cliques,” in Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases - Part II, ser. ECML PKDD ’08. Ber-lin, Heidelberg: Springer-Verlag, 2008, pp. 33–49.

[2] J. Gao et al., “On Community Outliers and Their Efficient Detection in Information Networks,” in Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD ’10. New York, NY, USA: ACM, 2010, pp. 813–822.

[3] B. Perozzi et al., “Focused Clustering and Outlier Detection in Large Attributed Graphs,” in Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD ’14. New York, NY, USA: ACM, 2014, pp. 1346–1355.

[4] E. M¨uller et al., “Ranking outlier nodes in subspaces of attributed graphs,” pp. 216–222, 2013.

Нормализованная степень узла и нормализованная центральность собственных векторов были использованы в качестве индикаторов структуры графа в ранжировании отклонений.

На наш взгляд, более высокое ранжирование в контексте проблемы инсайдерской угрозы может быть связано с более высоким ранжированием с кластеризацией подпространства/подграфа в помеченных графах.

Отношения между корпоративными пользователями отображаются в неориентированный невзвешенный граф G (V; E; A), где V - набор вершин (пользователей), E - набор рёбер, а A - набор атрибутов. Построение графа, основанного на связях пользователей, осуществляется следующим образом:

1) Пользователи представлены в виде вершин

2) Ребра между вершинами строятся на основе:

_ организационной иерархии (представлять супервизора - подчиненные отношения)

_ сообщений электронной почты (при наличии сообщений электронной почты между двумя пользователями)

При создании структуры графа отношение "супервизор - подчиненный" сотрудников сначала отображается в неориентированном ребре между соответствующими пользователями. Кроме того, использовались журналы обмена сообщениями электронной почты для получения из сети информации о дружбе пользователей в корпоративной сети. Это отношение фиксируется путем анализа всех адресов электронной почты "TO", "CC" или "BCC" в домене предприятия. Затем создается граница между отправителем и получателем.

В рамках этого исследования исключены направленность и веса рёбер, поскольку алгоритмы кластеризации подпространства и подграфа не были разработаны для применения на ориентированных графах. Также были исключены связи с внешними пользователями в созданном графе, но использовали связь с внешними пользователями как отдельный атрибут.

Извлечение атрибутов

В этой работе статистические методы использовались для извлечения связанных атрибутов из различных информационных потоков, таких как записи входа/выхода, записи веб-доступа, сообщения электронной почты, записи использования съемных носителей и операции копирования файлов. В основном это включает извлечение максимального, минимального и среднего значений выбранных параметров.

1.4.4 Проблемы создания системы поиска признаков инсайдеров

В целом аналитик по безопасности несет критическую ответственность за осмысление выхода от многочисленных инструментов, которые порождают информацию, которую можно извлечь из киберданных. Это еще раз подчеркивает проблему объяснимости результатов автоматического анализа данных.

Существующие подходы к контролю доступа в основном направлены на то, чтобы не допускать посторонних лиц и являются неэффективными с точки зрения предотвращения неправомерного использования инсайдером. Кроме того, средства обнаружения и системы контроля доступа обычно не интегрированы и могут нарушать требования соответствия и политики. Существующие системы контроля доступа и обнаружения аномалий не обеспечивают адекватной защиты ресурсов от вредоносных инсайдеров. Эти существующие системы обычно решают проблему либо со строго физической, либо со строго киберпространственной точки зрения, что часто делает невозможным как точное определение желаемых политик контроля доступа, так и точное обнаружение потенциально вредоносного поведения инсайдеров.

Современные подходы к обнаружению инсайдерских угроз основаны на судебной экспертизе и обычно ограничиваются изучением журналов кибербезопасности для применения алгоритмов обнаружения аномалий или сопоставления сигнатур. Эти алгоритмы, хотя и необходимы для обнаружения инсайдерских угроз, являются лишь одной из частей полного решения. Что еще более важно, они недостаточно учитывают ограничения информации (глубина представления информационных объектов и совсем не рассматривают

возможности сговора, нарушающего правила доступа) , которую можно извлечь из киберданных, сокращение ложных срабатываний (FPR) и эффективные политики доступа для предотвращения инсайдерских угроз. Интегрированное решение для обнаружения и предотвращения инсайдерских угроз использует семантически помеченные данные из инфраструктур контроля кибернетического и физического доступа и обеспечивает упреждающее, раннее и надежное обнаружение инсайдерской эксплуатации или уничтожения данных. Фактически речь идет о необходимости формализации работы с многими гетерогенными информационными пространствами

Различные источники данных анализируются для определения возможностей и намерений злоумышленников. Алгоритмы машинного обучения используют наблюдения из систем кибернетического и физического доступа для построения моделей нормализации. Иерархическая организация минимальных моделей позволит распознавать тонкие нештатные ситуации (система уточнения характеристик информационных объектов). Несколько алгоритмов обеспечивают подтверждение, необходимое для минимизации ложных тревог. В одном варианте осуществления механизм рассуждения на основе семантических графов модифицируется для агрегирования выходных сигналов детекторов аномалий для идентификации вероятных вредоносных ситуаций и устранения доброкачественных аномалий. Эталонная модель механизма рассуждения модифицируется для интерпретаций машинного обучения аномальных результатов, отношений между аномалиями и важности гипотетического вредоносного поведения. Детализированный механизм спецификации политики управления доступом должен быть изменен, чтобы выразить нюансы политики, направленные не только на удержание посторонних от защищенных активов, но и на предотвращение неправомерного использования вредоносными инсайдерами без ущерба для доступности активов для законного доступа и использования. Разработка и уточнение политики должны основываться на выходных данных детекторов аномалий, с тем чтобы включить факторы, которые привели к тому, что доступ был отмечен как необычный, т. е. предлагается разработать методы

динамического отслеживания опасностей, способных привести к враждебному воздействию на активы.

Динамическое добавление новых типов данных и источников требует использовать машинное обучение и научиться ожидаемому нормальному поведению новых данных, что позволяет новым источникам данных вносить вклад в представление о рассматриваемом «мире» без изменения механизма рассуждений. Однако это может породить случайные закономерности, которые могут настолько исказить результаты анализа, что сделают его бесполезным. Открытая архитектура позволяет системе развиваться вместе с современными методами обнаружения аномалий, алгоритмами машинного обучения и технологиями управления доступом. Механизм изменения детализации политик доступа позволит значительно сократить разрыв между желаемыми политиками контроля доступа и реализуемыми существующими механизмами политиками.

В некоторых вариантах осуществления система обнаружения угрозы предсказывает и выдвигает гипотезы поведения, указывающие на текущие инсайдерские атаки, путем сбора необработанных кибер- и физических данных, анализа наблюдений, полученных из необработанных данных, и обнаружения подозрительного поведения.

Подозрительное поведение — это интерпретации намерений и действий, основанные на наблюдениях, возможно, представляющих вредоносную деятельность. В некоторых вариантах осуществления вывод индикатора основан на обнаружении аномалии и требует изучения потока наблюдений на предмет необычных закономерностей, которые указывают на изменение привычек/роли человека, преднамеренное или непреднамеренное нарушение политики, неправильную конфигурацию системы управления доступом или активное вредоносное поведение, то есть возможно возникновение эмпирических закономерностей. Уменьшение числа ложных тревог осуществляется с помощью подтверждения и корреляции подозрительного поведения, обнаруженного различными алгоритмами. Система обнаружения угроз использует перекрытия в подозрительном поведении.

Система обнаружения угроз исследует пространство гипотез, чтобы определить, какое подозрительное поведение согласуется с вредоносным поведением, а какое имеет доброкачественные объяснения. Процесс распространения убеждений о подозрительном поведении может использовать систему рассуждений, такую как система рассуждений PNNL CHAM PION (Columnar Hierarchical Auto-associative Memory Processing in Ontological Networks), которая содержит 80 иерархических структур, позволяющих строить рассуждения на основе модифицированных случаев (CBR), расширенную предсказательную функциональность. CBR используют логику описания, чтобы решить, согласуются ли наблюдаемые данные и Подозрительное поведение, распространяемое снизу в иерархии, с предполагаемым вредоносным поведением, и если да, то новые "утверждения динамически сохраняются в рабочей памяти систем (семантическая графическая структура). В отличие от классических подходов, сравнивающих гипотезу со всеми "случаями" или всем семантическим графом (который может быть непомерно большим), более прослеживаемая система CHAMPION анализирует подмножество семантического графа. Для обоснования инфраструктур контроля физического доступа, которые должны быть интегрированы с киберданными, добавляются онтологические представления. Явным преимуществом подхода к рассуждению CHAMPION является его способность интегрировать данные из нескольких источников и распространять анализ на абстракции более высокого уровня. Онтологии также добавляются, чтобы учесть ожидаемые пространственные/временные отношения и даже асинхронное поступление пространственных и киберданных и подозрительного поведения.

Семантическая маркировка при создании политики также может использоваться. Семантическое маркирование определяет часть политики на основе содержимого защищаемых ресурсов. Это позволяет конкретизировать политику, которая более точно соответствует тому, чего пытаются достичь разработчики политики, делая спецификацию политики более быстрой, удобной и более точной. Такой подход также облегчает понимание намерений директивных

органов путем ознакомления с проводимой политикой. Администраторы могут пересматривать и отлаживать политики и автоматически предлагать переформулировать реализованные политики, чтобы эти политики явно включали факторы, которые заставляли систему обнаружения аномалий отмечать необычный доступ. Таким образом, любой аномальный доступ, который является результатом чрезмерно разрешительной политики, будет служить руководством для исправления или пересмотра политики, чтобы запретить другой такой доступ.

Производительность системы обнаружения угроз не ограничивается точностью и охватом одной модели. Вместо этого можно использовать ряд алгоритмов машинного обучения, которые выводят различные аспекты инсайдерского поведения, используя разнообразный набор данных. В некоторых вариантах осуществления используются явные инсайдерские модели, где подозрительное поведение из множества алгоритмов усиливается (используется более детализированное описание) для обеспечения предупреждений с более важной причиной или объясняется для уменьшения ложных срабатываний.

Динамические политики используются для обеспечения контроля доступа в зависимости от контекста, обеспечивая доступность активов для законного, но, возможно, необычного использования, а также предотвращая внутренние атаки.

Реакция и защита с помощью механизма точной спецификации политики доступа могут использоваться для обеспечения высокой гарантии правильности благодаря его формальным основам, легкой расширяемости для включения новых политик или парадигм политики, а также, благодаря доказательствам доступа, которые объясняют, почему был предоставлен доступ, отличной проверяемости и поддержки постепенного пересмотра.

Выводы по Главе 1

В рассмотренных в обзоре методиках предлагается использовать лишь однородные данные (например, только данные по сетевой активности сотрудника, его доступах к базам данных и т. д.), что влечет наследование ошибок и коллизий, зацикливание или размножение ошибок и др. Как следствие, профильные службы

ищут данные из *иных* по природе источников для *подтверждения* или выявления *значимых расхождений*. Следовательно, для выявления признаков инсайдера, требуется только *совокупность* данных, описывающих *различные* области деятельности и жизни человека в разных сферах, его поведения в работе, как пользователя ИС и т. д. Особую значимость при оценке возможной причастности к утечкам информации имеет учет характеристик личности сотрудников, попадающих в круг подозреваемых. Эффективность анализа только однородны данных в таких случаях сомнительна, т. к. инсайдер обычно достаточно умен, профессионален, и именно поэтому ему доверен доступ к обработке ценной информации, следовательно, он знает способы, позволяющие обойти системы контроля. Таким образом, анализ инсайдерской активности нужно проводить по всему спектру сведений - как из эксплуатируемых технических систем и систем контроля, так и с учетом информации из реальной жизни персонала - данных HR-служб, оперативной информации служб собственной безопасности или детективных агентств, баз данных различного назначения (разного рода справочных систем, в том числе баз правоохранительных и других гос. органов, социальных сетей, данных кадровых агентств и т.д.). Чем разнообразнее исходные данные - тем точнее работа моделей, и тем более тонкие, малозаметные отклонения можно проанализировать, чем “дальше” по своей природе данные друг от друга, тем ценнее их взаимосвязанный анализ. Следует подчеркнуть, что обозреваемых методиках не ставится задача анализа Big Data в процессно-реальное время, хотя наиболее ценная для обсуждаемых целей – поиска признаков вредоносной инсайдерской активности – информация аккумулируется именно в хранилищах Big Data.

Таким образом, характеризуя открытые, на текущий момент, проблемы и задачи в рассматриваемой области исследования можно заключить:

1. Приведенный в предыдущих параграфах обзор сосредоточен вокруг ряда выделенных важных для практики проблем, которые связаны с выявлением

признаков наличия инсайдеров, признаков направлений их деятельности, снижением размерности поисковых задач, оценкой надежности выводов.

2. Выделенные проблемы определяют техническую целесообразность разработки средств автоматизации решения задачи поиска признаков инсайдеров. Такая постановка научно-технической проблемы, на решение которой направлена данная диссертационная работа, связана с появлением массивов и хранилищ больших данных (БД). К таким хранилищам имеют доступ много пользователей, относительно которых множество искомых инсайдеров мало. Стандартная статистическая обработка данных в условиях БД порождает большие уклонения и ведет увеличению множества ложных тревог (FPR) и соответственно к высокой вероятности потери искомых признаков, относящихся к инсайдерам. То есть необходимо ограничивать объемы данных, для разумного использования методов статистики, или оценивать полученные закономерности на возможность их случайного происхождения.

3. Возможности снижения объемов достигаются кластерными методами, в рамках которых необходимо сохранять условия использования статистических или иных методов. Однако возможности снижения объемов данных имеют естественные ограничения снизу, также связанные с работоспособностью используемых методов. Поэтому в диссертации поставлена и решена задача построения метода оценки объема данных, при котором последующие методы анализа дают наилучший результат. Предложенный в диссертации метод основан на простейших вероятностных моделях и использует предельные теоремы в схеме серий. В качестве практического примера при иллюстрации метода взята задача выявления сговора инсайдеров. По этой практически важной задаче в литературе найдены только формулировки задачи, но не найдено ни одного содержательного результата.

4. Как отмечалось в литературе, использованной в обзоре все используемые методы анализа должны обеспечивать объяснимость полученных выводов. При поиске инсайдеров используются методы выявления аномалий и обнаружение эмпирических закономерностей. Автору известны только три метода

автоматического получения выводов, позволяющих построить адекватные объяснения. Первый метод связан с выявлением причины аномалии (по крайней мере эмпирической причины) путем выявления общих характеристик объектов, обладающих заданным свойством. Второй метод основан на принципе абдукции, который на данном множестве фактов проверяет непротиворечивость вывода. Третий метод позволяет оценивать возможности случайного порождения закономерности, лежащей в основе полученного вывода. В первом и втором методах ищется противоречие между описаниями фактов, обладающими требуемым свойством и фактами, не обладающими требуемым свойством. Наличие таких противоречий связано с недостаточной глубиной описания характеристик наблюдаемых объектов. Поэтому для более глубокого описания полученных фактов необходима большая детализация или уточнение описания рассматриваемых объектов. В диссертационной работе построены методы такой детализации. (К сожалению, автору не удалось доказать, что предложенные методы исчерпывают все возможности). Однако метод деления характеристик объектов и метод расширения множества характеристик описания объектов позволяют по-новому организовать поиск причин. Первый позволяет вести уточнение описания начиная с грубого набора характеристик, что позволяет использовать более быстрые алгоритмы. Для больших данных это обстоятельство имеет большое значение. Второй подход связан с описанием аномалий в нескольких информационных пространствах. Этот подход рассматривался в научной литературе и отражен в обзоре, но автором в диссертации предложен простой логический метод объединения результатов анализа разнородных информационных пространств.

Третий подход основан на том, что появление случайных эмпирических закономерностей можно оценивать в простых вероятностных моделях. Этот подход имеет эмпирическое обоснование и часто используется на практике. Например, метод FMEA оценки надежности технических устройств с помощью простейших вероятностных распределений включен в международные стандарты. Поэтому использование этого подхода в оценках случайности выявленных

признаков инсайдеров представляется вполне обоснованным. Вместе с тем простейшие вероятностные оценки, включенные в алгоритмы фильтрации, значительно снижают уровень ложных положительных тревог.

4. Архитектура автоматизированной системы поиска признаков инсайдеров состоит из подсистемы сбора данных из различных источников и подсистемы аналитики, позволяющей облегчить ручной анализ собранных данных. Перечисленные выше методы войдут во вторую подсистему, которая предполагает работу со многими информационными пространствами.

Таким образом, в данной диссертационной работе можно выделить две группы задач, направленных на решение технической проблемы автоматизации поиска признаков инсайдеров в BIG DATA.

Задачи ***практического плана***:

- подходы к выявлению сговора инсайдеров и описание сговора методами теории графов;
- методы детализации описания информационных объектов с целью эффективного поиска причин выделенных свойств;
- методы объединения результатов анализа в гетерогенных информационных пространствах;
- методы автоматической фильтрации эмпирических закономерностей, появляющихся при анализе больших данных;
- методы представления знаний по поиску признаков инсайдеров.

Задачи ***технического плана***:

- определение объемов выборки для эффективного и результативного[1] применения статистических методов;

[1] При этом термин *результативность* понимается как демонстрируемая на практике – способность разработанных программно-технических решений при решении конкретных прикладных задач обсуждаемого характера на реальном объекте защиты – в крупном отечественном коммерческом банке – обеспечить достижимость поставленных перед ними задач по организации противодействия вредоносным инсайдерским активностям. В свою очередь, термин *практическая эффективность* понимается как способность разработанных программно-технических решений «вписываться» в предъявляемые требованиями основного бизнеса объекта защиты ресурсные ограничения – на размеры бюджетов, выделяемых на эти цели основным бизнесом банка; на сроки выполнения каждого цикла анализа данных и приятия решений; на численности персонала соответствующей квалификации в Службе безопасности и др. Все эти ограничения могут быть в установленном порядке оформлены в виде необходимых корпоративных Соглашений об уровне сервиса (Service Level Agreement), соответствующие их нарушению риски захеджированы заранее согласованными объемами резервов на такие риски.

- кластерные подходы для управления объемом выборок;
- методы поиска причин выделенного свойства в множестве информационных объектов;
- методы снижения множества ложных тревог (FPR);
- методы распараллеливания анализа по нескольким информационным пространствам;
- интеграция методов доступа к гетерогенным данным и различным методам анализа в автоматизированной системе поиска признаков инсайдеров.

Глава 2 Методы анализа данных при поиске признаков инсайдера

2.1 Параметризация в прикладных задачах поиска эмпирических причин

Система выявления признаков инсайдера время от времени будет выявлять аномалии, причины которых необходимо объяснять оперативному работнику. Однако могут возникнуть случаи, когда при объяснении аномалии возникнет противоречие. Например, аномалия с вероятностью 50% является нарушением политики и с вероятностью 50% является легальным действием сотрудника. Для устранения таких противоречий необходимо подключить и проанализировать дополнительные данные или “глубже” проанализировать существующие данные. Таким образом, важная задача – предложить метод, который определит, как система будет работать с противоречиями при анализе и объяснении аномалий.

Рассматривается следующая проблема. Если в классе объектов часть их обладает свойством P, а часть не обладает этим свойством, то возникает задача выявления причины появления свойства P у части объектов. Объекты описываются множеством характеристик этих объектов или параметров по Эшби [56]. Поэтому можно отождествлять объекты и соответствующие множества их характеристик.

Множество характеристик полно, если каждый объект однозначно выделяется по своему подмножеству характеристик. Однако при исследовании причины появления свойства P в некоторых объектах можно прийти к выводу, что полнота множества характеристик не гарантирует описания причин появления свойства P. Тогда необходимо изменять параметризацию класса объектов (множество параметров описания объектов), уточняя ее в такой степени, чтобы в этом описания содержались причины появления свойства P.

В работе рассматривается несколько способов расширения множества характеристик рассматриваемого класса объектов при сохранении свойства полноты.

2.1.1 Эмпирические причины

Рассмотрим простейшую модель ДСМ-метода [57] интеллектуального анализа данных (ИАД), построенную на языке теории множеств. Пусть $U = \{u_1,\ldots u_m\}$ является множеством характеристик наблюдаемых объектов $O_1,\ldots, O_n$. То есть объекты полностью описываются наборами характеристик из U. Можно считать, что любой объект O – это подмножество U, и множество всех возможных объектов – это множество всех подмножеств множества U. Само множество U будем называть параметризацией наблюдаемых объектов. Кроме характеристик определим понятие свойства объекта. Свойство P объекта O отражает некоторую интегральную характеристику объекта O, которое может присутствовать в O или отсутствовать в O. Если объекты появляются последовательно, то вновь появившийся объект должен проверяться на наличие свойства P.

Предположим, что рассматриваемое свойство P удовлетворяет следующим условиям.

(A) Если свойство P выявлено в объектах O_1 и O_2, то оно есть в объекте $O = O_1 \bigcap O_2$.

(B) Если объект O обладает свойством P, то для любого объекта O_1, содержащего O, объект O_1 также обладает P.

Определим понятие эмпирической причины некоторого свойства P в наблюдаемых данных. Основываясь на понимании причины Д.С. Миллем [58], в нашей модели эмпирической причиной свойства P для множества объектов $O_1, O_2, ..., O_l$ будем называть производный объект $O = \bigcap_i O_i$ по всем индексам i, для которых O_i обладает свойством P, но ни один из остальных O_j свойством P не обладает. Причина должна быть единственной и от нее ничего нельзя убавить или добавить [59].

Пример 1. Пусть U = {а, б,…, я} – русский алфавит. Свойство P означает множество букв, из которых можно собрать слово, выражающее понятие «дом».

Пусть есть два объекта O_1 = {д, о, м} и O_2 = {с, у, к}. Ясно, что O_1 является эмпирической причиной свойства *P*. При этом O_2 не обладает свойством *P*.

2.1.2 Различные параметризации

Однозначное представление объектов в виде подмножеств множества характеристик не означает, что не существует другого множества характеристик для описания того же множества наблюдаемых объектов и свойств.

Пример 2. Понятие «дом» можно выразить на английском языке словом “home”. Тогда множество характеристик есть английский алфавит, а аналогичное свойство примера 1 выражается другой причиной, выраженной другим множеством характеристик.

Представление объектов из множества характеристик можно строить с помощью упорядоченных наборов характеристик. Тогда множество описываемых объектов совпадает с множеством слов конечной длины *U** из исходного множества характеристик. Это множество бесконечное и *U*** = *U**. Эмпирическая причина для *U** определяется словом, описывающим интегральное свойство *P* в языке *U**. При этом свойство *P* примера 1 выражается проще. Объект «дом» обладает свойством *P*, а объект «мода» свойством *P* уже не обладает. Однако в данном представлении объектов может возникнуть ситуация, когда эмпирическая причина отсутствует.

Пример 3. Объект «дом» соответствует понятию, описанному словом «дом», и является эмпирической причиной свойства *P*, а объект «строение» не является словом «дом», но соответствует смыслу свойства *P*. Таким образом возникает объект, не обладающий эмпирической причиной, и одновременно обладает свойством *P*.

Чаще всего добавление новых объектов фальсифицирует эмпирическую причину, найденную по исходному набору объектов. Фальсификация происходит в следующих формах.

1. Найденная эмпирическая причина появляется в новом объекте, который не обладает свойством *P*.

2. Найденная эмпирическая причина не появляется в новом объекте, который, как предполагается, обладает свойством *P*.

Рассмотрим некоторые другие примеры изменения параметризации, позволяющие тоньше идентифицировать причины и их фальсификации.

2.1.3 Схема аутентификации как поиск причины в модифицированной параметризации

В данном разделе рассматривается задача изменения параметризации за счет добавления множеств характеристик из различных информационных пространств.

Пусть *A* – субъект, который должен подтвердить субъекту *W* свое имя. Эта процедура называется аутентификацией [60] и может быть проведена следующим образом. Пользователь предъявляет пароль получает допуск к данным. Если инсайдеры-маскировщики (см. классификации инсайдеров в главе 1) получают чужой доступ к системе, то при появлении простейших признаков таких действий необходима более глубокая аутентификация пользователя [61]. То есть исходная аутентификация должна быть подтверждена или усилена дополнительной информацией, возможно, из других источников. Если дополнительная аутентификация не подтверждает подлинности пользователя, то служба безопасности получает серьезное свидетельство вредоносных действий работающего пользователя. Используемый подход означает изменение параметризации информации о работающем пользователе. При этом без наличия простейших признаков действий инсайдера не целесообразно проводить изменение параметризации потому, что это приведет к усложнению системы работы с данными.

Пусть Σ_1 — это информационное пространство, содержащие объекты $O_1^{(1)}, O_2^{(1)},..., O_{s_1}^{(1)}$. Каждый из этих объектов классифицируется двумя значениями *t* (truth) и *f* (false). Свойство *P* состоит в знании для этих объектов правильного вектора $x^{(1)} = (x_1^{(1)}, x_2^{(1)},..., x_{s_1}^{(1)})$, где для $i = 1,..., s_1$, $x_i^{(1)} = \begin{cases} t, \\ f. \end{cases}$

Пусть вектор $x^{(1)}$ известен *A* и *W*. Предположим, что существует субъект *B*, который назвал себя именем *A* и также пытается подтвердить это имя. Возможны два случая.

1. *B* случайно выбирает значение вектора $x^{(1)} = (x_1^{(1)}, x_2^{(1)},..., x_{s_1}^{(1)})$.

2. Пространство Σ_1 известно *B* и он знает вектор $x^{(1)}$.

Субъект *W* проверяет знание вектора $x^{(1)}$, считая, что он общается с некоторым субъектом *Ф*, который может принимать значения *Ф* = *A* или *Ф* = *B*. Субъект *W* предъявляет объекты $O_1^{(1)}, O_2^{(1)},..., O_{s_1}^{(1)}$ субъекту *Ф*. Тогда, получив вектор $x^{(1)}$, Субъект *W* может считать, что субъект *Ф* знает эмпирическую причину свойства *P*. При этом эту причину может знать как *A* так и *B*. Причем *B* может случайно угадать вектор $x^{(1)}$ (ложная аутентификация).

Проверка имени *A* может быть продолжена, если к информационному пространству Σ_1 добавить информационные пространства $\Sigma_2,...,\Sigma_m$. Из характеристик каждого из добавленных пространств можно сформировать объекты $O_1^{(i)}, O_2^{(i)},..., O_{s_i}^{(i)}$, $i = 2,..., m$, для которых *W* и *A* знают вектора $x^{(i)}$, $i = 2,..., m$. Таким образом, пара субъектов *W* и *A* расширяют параметризацию для определения причин свойства *P* до знания векторов $x^{(1)}$, …, $x^{(m)}$.

Если субъект *B* выбирает значение вектора $x_j^{(i)}$ случайно, то с вероятностью, как угодно, близкой к 1, он на каком-то шаге ошибается. Это позволит субъекту *W* понять, что субъект *Ф* не обладает свойством *P*.

Если субъект *B* скомпрометировал пространства $\Sigma_{i_1},...,\Sigma_{i_r}$, $r < m$, то субъект *W* знает о возможной компрометации каких-то информационных пространств, но не знает каких. Тогда субъект *W* предъявляет *Ф* объекты, созданные в информационных пространствах $\Sigma_1, \Sigma_2,..., \Sigma_m$, попадает на объекты нескомпрометированных пространств, и определяет фальсификацию

эмпирической причины *P*. Если субъект *Ф* определяет правильно все вектора $x^{(1)}, \ldots, x^{(m)}$, то причина *P* подтверждена и аутентификация прошла успешно.

Рассмотренный пример показывает, что расширение множества характеристик за счет привлечения дополнительной информации позволяет уточнять причину исследуемого свойства *P*. Идея подтверждения эмпирической причины за счет расширения исходной параметризации с помощью добавления характеристик рассматривалось в работах [61, 2].

2.1.4 Изменение параметризации за счет разбиения параметров

Пусть $U = \{u_1,\ldots u_m\}$ – это хосты сети. Пусть свойство P соответствует обработке информационной технологии за время $\tau > T_0$. Пусть объект *O* – это множество хостов, выделяемое провайдером для реализации информационной технологии. При повторах технологии выделяется множество объектов $\{O^+\}$, на которых наблюдаются задержки $\tau > T_0$ выполнения информационной технологии, и множество объектов $\{O^-\}$, для которых $\tau \le T_0$. Т.е. в сети возникает так называемая «мерцающая» ошибка [62, 63].

Предположим, что причина *o* свойства *P* ищется с помощью пересечения объектов из множества $\{O^+\}$. Эмпирическая причина является подобъектом *o*, если *o* не встречается в множестве $\{O^-\}$. Пусть к множеству объектов добавляется новый объект *O*, и пусть $o \subseteq O$, но информационная технология реализуется так, что $\tau \le T_0$. Таким образом, происходит фальсификация эмпирической причины.

При исследовании характеристик *U* оказалось, что u_1 – это локальная сеть из двух машин $u_1^{(1)}$ и $u_1^{(2)}$ с прокси-сервером для выхода в общую сеть. Для реализации информационной технологии любая из этих машин выбирается случайно. Оказывается, что $u_1^{(2)}$ работает всегда быстро, а $u_1^{(1)}$ - всегда медленно. Рассмотрим новое пространство характеристик $U' = \{u_1^{(1)}, u_1^{(2)}, u_2, \ldots, u_m\}$. Тогда в новом множестве характеристик причина свойства $P = (\tau > T_0)$ определяется тем же методом, что и ранее, но эмпирическая причина не фальсифицируется.

Пример 5 «Просмотр персональных данных физических лиц по заказу».

Криминальные структуры, нечестные коллекторские агентства или просто знакомые, если имеют выход на сотрудника предприятия с доступом к адресам, геолокации и пр. данным физических лиц, могут заниматься незаконным розыском людей или их имущества. Криминальным структурам может потребоваться установить, где проживает физическое лицо, т. е. адрес регистрации. Если физическое лицо проживает не по адресу регистрации, то его местоположение возможно установить по магазинам, где он совершал транзакции или по его геолокации.

Поэтому редкий просмотр сотрудником персональных данных является признаком профиля угроз «просмотр персональных данных физических лиц по заказу», но не гарантирует, что это инцидент безопасности т. к. требуется разобраться на каком временном интервале был произведен просмотр данных. Стоит отметить, что могут существовать бизнес-процессы на предприятии, требующие для обслуживания клиента просмотра персональных данных и необходимо уметь различать совершался доступ к персональным данным согласно утвержденному бизнес-процессу или просмотр был несанкционированным.

Логика работы в данном примере, следующая – сотрудник может месяц, два или больше не просматривать персональные данные, но, когда вдруг он совершает одиночный просмотр персональных данных или вдруг происходит всплеск активности по просмотру персональных данных, то такой факт может свидетельствовать о том, что сотрудник, находясь под влиянием криминальных структур, совершил просмотр персональных данных физического лица.

Если использовать временной интервал мониторинга в один день (Рис. 1, идентификационные данные сотрудников закрашены черным цветом), то на нем видны десятки и сотни просмотров данных и затруднительно сделать вывод по нарушению.

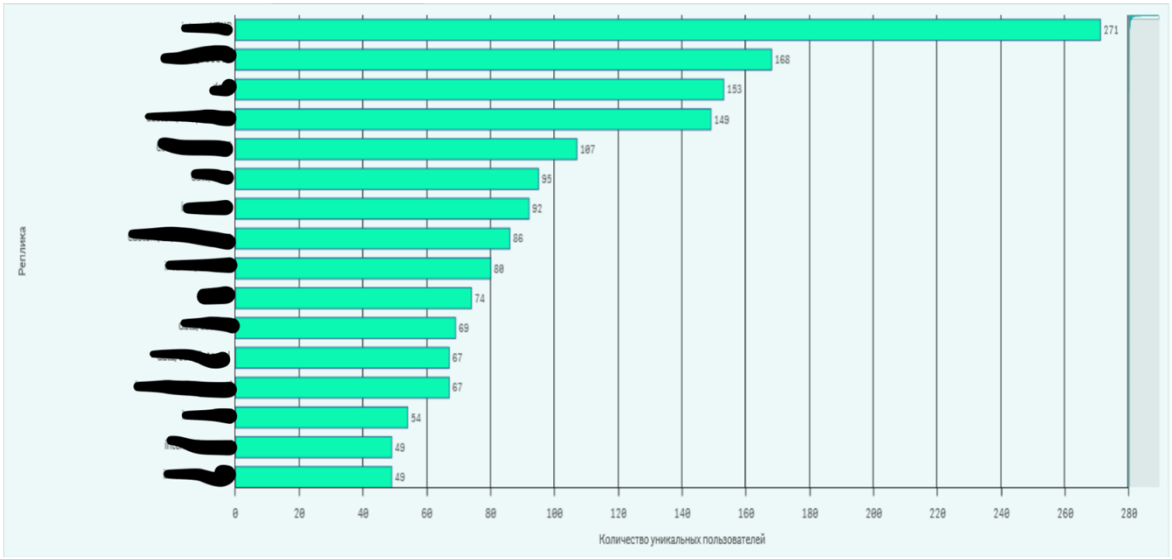

Рис.1. Факты обращения пользователей к данным (интервал - 10 мин).

Однако, если увеличить временной интервал с одного дня до 1–3 месяцев, то одиночный просмотр персональных данных будет являться признаком аномального поведения.

Почему увеличение временного интервала дает результат и подсвечивает сотрудников с аномальным поведением? Потому что не существуют бизнес-процессов, при которых просмотр персональных данных происходит один раз в три месяца (экономически не выгодно создавать бизнес-процессы, в которых что-то происходит один раз в три месяца, и обеспечивать их ресурсами). Временной интервал можно выбрать не три месяца, а например неделю, он тоже даст результат.

Пример 6 «Снижение количества ложных тревог».

В комплексе Big Data присутствуют объекты в виде реплик баз данных Oracle, MSSQL, Postgres и т. д. Реплики баз данных содержат большое количество технологических данных, которые требуются для поддержки работы исходной ИС, но не требуются для аналитических задач. Поэтому из реплик баз данных создаются витрины, в которых присутствуют только полезные данные. Например, витрина «Единый профиль сотрудника», «Депозиты», «Юридические лица Москвы», «Промышленные предприятия Урала», «Бухгалтерская книга» и т. д.

К каждой витрине, реплике базе данных, размещенной в комплексе Big Data какое-то количество пользователей имеют доступ и пользуются им. Существуют наиболее популярные витрины и базы, существуют наименее популярные (Рис. 2, идентификационные данные реплик баз данных и витрин закрашены черным цветом). Из рисунка видно, что, например к одной витрине/реплике имеют доступ 1300 сотрудников, в другой 1100 и т. д.

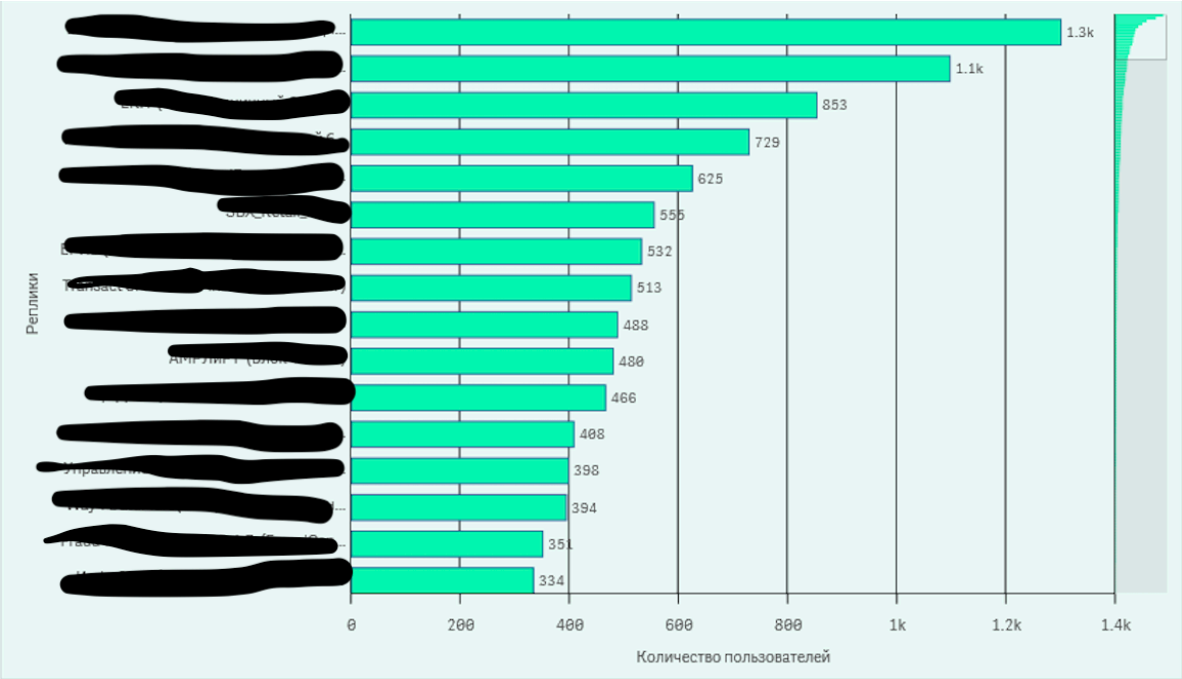

Рис.2. Распределение реплик баз данных по количеству пользователей, которые имеют туда доступ.

Особенность заключается в том, что в витрине или реплике базы данных только ~20 полей могут быть критичным с точки зрения атак инсайдера (адрес регистрации, остаток на счете, контакты, движение средств по счету и т. д.), а тысячи полей некритичные. В этой связи не имеет смысл присваивать признак сотруднику «имеет доступ к критичным данным», если он не работает с 20 критичными полями. В данном примере, если изменить параметризацию и перейти от реплики базы данных или витрины к полям, то 1300 сотрудников с признаком «имеет доступ к критичным данным» превратятся в 10–15 сотрудников, имеющим доступ к критичным полям.

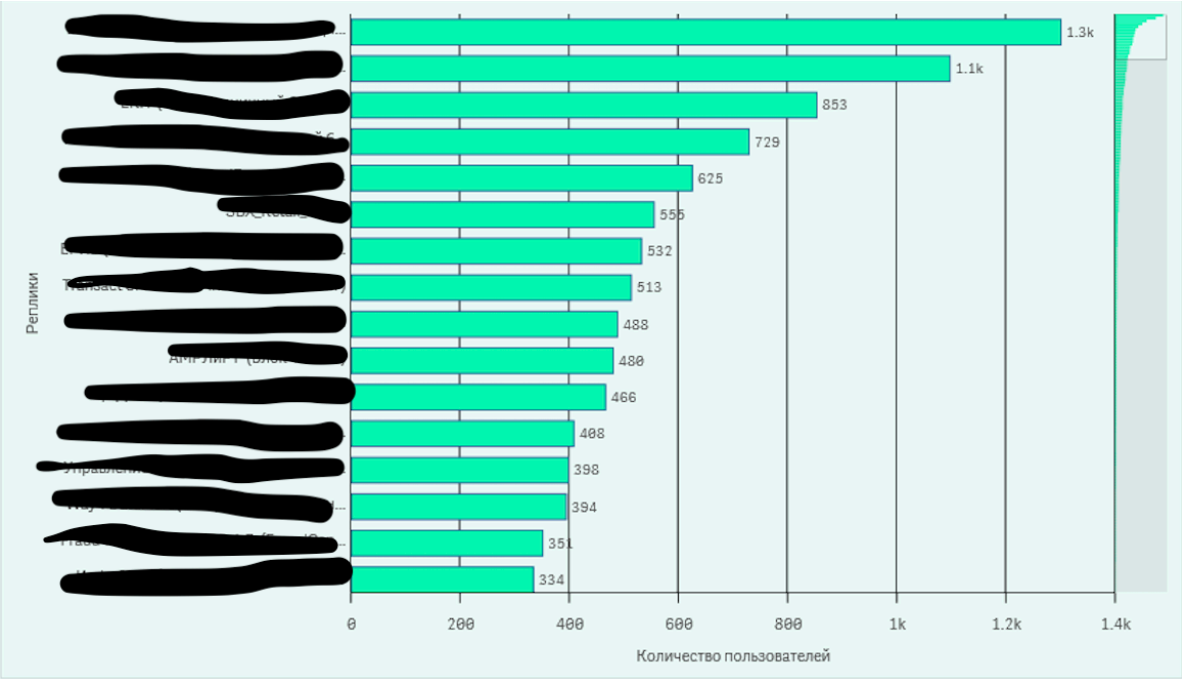

Рис.2. Распределение реплик баз данных по количеству пользователей, которые имеют туда доступ.

Итак:

Выявление устойчивых при увеличении исходных данных эмпирических причин исследуемых свойств не всегда возможно в условиях исходной параметризации. Изменение параметризации требует выполнения полноты описания известных и вновь поступающих объектов.

Возможно дополнение изменения параметризации требованием сохранения описания эмпирических причин других свойств. Поэтому необходимо строить новую параметризацию путем уточнения старой параметризации. В диссертации рассмотрены примеры построения таких уточняющих параметризаций. Возможны гибридные варианты применения изложенных методов.

В дальнейшем предполагается исследовать возможность доказательства того, что других путей уточнения параметризации нет.

2.2 Модель множества информационных пространств в задаче поиска признаков инсайдера

Инфраструктура Big Data – это расширяющиеся данные с эффектом Open, к тому же данные являются гетерогенными (различные информационные пространства). Каждое информационное пространство содержит свой тип данных. Например, в контексте задачи выявления признаков инсайдера, речь идет о данных по доступам сотрудников, информационной инфраструктуре, действиях сотрудников, кадровой структуре и т. д. Инструменты обработки данных объектно-ориентированы, поэтому могут обрабатывать тип данных, под который создавались. Задача, которую представляет и решает данный раздел – это создание метода, позволяющего работать с гетерогенными данными и выявлять компрометирующие данные в различных информационных пространствах и объединять выявленные данные в единый результат. Найденный единый результат обработки нескольких информационных пространств впоследствии уже будет представлен через пользовательский интерфейс оперативному сотруднику.

Разработка методов выявления признаков инсайдеров в организации является актуальной задачей и решению этой задачи посвящен ряд исследований [64, 65, 66, 67]. Эти исследования основаны на использовании различных статистических методов выявления аномалий в деятельности сотрудников организации. Очевидно, что хорошая идентификация инсайдера основана на объединенной картине незначительных аномалий. При этом всегда используется информация, касающаяся различных аспектов деятельности потенциальных инсайдеров. Анализ различных типов информации и формирование выводов из нее – сложная задача.

Далее предлагается подход решения этой задачи, основанный на запретах вероятностных мер [68, 69, 70]. Основная идея подхода состоит в следующем. Рассматривается конечное множество информационных пространств, в которых ищутся признаки, связанные с целями деятельности инсайдера. Накопление

информации в каждом пространстве рассматривается как вероятностный процесс. Появление даже слабых признаков вредоносной активности инсайдера интерпретируется как запрет вероятностной меры деятельности, моделирующей честного пользователя. Возникновение запретов в различных информационных пространствах можно связать дизъюнктивной формой в единое целое. Тогда значение «1» этой суммарной дизъюнктивной формы определяет наличие вредоносной деятельности инсайдера.

2.2.1 Общая математическая модель множества информационных пространств

Целью построения модели является внесение в алгоритм поиска инсайдера самой различной информации, в которой могут появиться признаки деятельности инсайдера. Необходимо отметить, что привлекаемая информация может содержать психологический поведенческий фон анализируемого лица, который усиливает малозначимые аномальные действия потенциального инсайдера.

Пусть U – контролируемый потенциальный инсайдер, $T = \{t_1, t_2, ..., t_N\}$ – цели инсайдерских атак. В ходе анализа информации об U могут рассматриваться подмножества T . Аномалии в общем фоне поведения и в характере U описываются всем множеством T . В анализе U используется m информационных пространств. Элементы каждого информационного пространства описываются своим языком и пусть $E_1, E_2, ..., E_m$ – алфавиты этих языков, а $E_1^*, E_2^*, ..., E_m^*$ – множества слов конечной длины в этих алфавитах. Языки описания информации $L_1, L_2, ..., L_m$ в информационных пространствах удовлетворяют условию $L_i \subseteq E_i^*$, $i = 1, ..., m$.

Введем понятие процесса наблюдения в информационном пространстве. Пусть время является дискретным и описывается множеством натуральных чисел. Наблюдение за U в i-м информационном пространстве к моменту времени n представляет собой конечное множество слов в языке L_i. В силу естественных технических ограничений в каждом языке L_i выделяется конечное множество

слов $X_{i,n}$, которое представляет интерес при наблюдении за U к моменту времени *n*. Обозначим через $\Sigma(X_{i,n}) = \{\sigma : \sigma \subseteq X_{i,n}\}$ множество всех подмножеств $X_{i,n}$, а через $\Sigma_j(X_{i,j}) \subseteq \Sigma(X_{i,n})$, $n \geq j$, накопленные данные к моменту времени *j*. Очевидно, что $\Sigma_1(X_{i,1}) \subseteq \Sigma_2(X_{i,2}) \subseteq \cdots$. Пусть $V_n(X_{i,n}) = \Sigma_n(X_{i,n}) \setminus \Sigma_{n-1}(X_{i,n-1})$ – дополнительная информация, полученная о *U* в момент времени *n*.

Последовательность $\{V_n(X_{i,n})\}_{n\geq 1}$ можно рассматривать как случайный процесс. Для корректного определения случайного процесса необходимо определить σ-алгебру на последовательностях $\{V_n(X_{i,n})\}_{n\geq 1}$. Пусть σ-алгебра $\mathcal{A}_i$ определяется как минимальная σ-алгебра, порожденная цилиндрическими множествами. Эта алгебра также является борелевской σ-алгеброй в тихоновском произведении дискретных топологических пространств в пространстве всех бесконечных последовательностей V_i^∞, которые могут быть траекториями рассматриваемого случайного процесса [71, 72].

Конечное множество измеримых пространств $(V_i^\infty, \mathcal{A}_i)$, $i = 1,...,m$, порождает измеримые пространства $(V^\infty, \mathcal{A})$, в котором $V^\infty = \prod_{i=1}^{m} V_i^\infty$, а σ-алгебра $\mathcal{A}$ порождена σ-алгебрами $\mathcal{A}_i$, $i = 1,...,m$, на пространстве V^∞.

Пусть поведение честного пользователя в пространстве $(V^\infty, \mathcal{A})$ описывается вероятностной мерой *P*. Как и в работах [см., например, [68],[73]] для меры *P* можно определить запреты этой вероятностной меры. Каждый запрет можно рассматривать как идентификацию инсайдера с полным описанием предыстории.

2.2.2 Модель полузапретов

Получившаяся модель является общей, но сложной для практического применения. Возможны различные способы ее упрощения. Рассмотрим один из таких способов.

Рассмотрим в отдельности каждое измеримое пространство $(V_i^{\infty}, \mathcal{A}_i)$, $i = 1,...,m$. Пусть $P^{(i)}$ проекция меры P на пространство (V_i^{∞} , $\mathcal{A}_i$). Распределение вероятностей $P^{(i)}$ характеризует поведение честного пользователя в соответствующем информационном пространстве. Для каждого информационного пространства введем упрощенную модель. Пусть $Q^{(i)}$ – некоторое распределение на пространстве $(V_i^{\infty}, \mathcal{A}_i)$, и пусть $P^{(i)}$ получается из $Q^{(i)}$ введением некоторого количества запретов [74]. Эти запреты можно интерпретировать либо как признаки некоторого подозрительного поведения, либо как прямые указания на некоторые цели атак инсайдера из множества *T*. Это означает, что можно рассматривать траектории наблюдаемого процесса в информационном пространстве несмотря на то, что появляются запреты меры $P^{(i)}$. Будем называть эти запреты меры $P^{(i)}$ в мере $Q^{(i)}$ *полузапретами*.

Определим множество бинарных переменных $x_{i,j}$, $j = \overline{1;\infty}$, следующим образом. Для $\forall n$ $x_{i,n} = 1$, когда в момент времени n появляется полузапрет меры $P^{(i)}$. В противном случае $x_{i,n} = 0$. Тогда появление полузапретов к моменту времени n характеризуется дизъюнкцией $\vee_{\{j=i\}}^{n} x_{i,j}$. Отсюда возникает возможность конструктивного определения запрета в совокупности пространств на основе полузапретов в отдельных пространствах. Например, запрет можно определить, как дизъюнкцию конъюнкций следующего вида

$$\vee_{\{1 \le i < l \le m\}} \left(\left(\vee_{\{j=i\}}^{n} x_{i,j} \right) \wedge \left(\vee_{\{j=i\}}^{n} x_{l,j} \right) \right).$$

Эта формула означает, что запрет (идентификация инсайдера) возникает при появлении не менее 2-х полузапретов в рассматриваемом множестве информационных пространств.

Ясно, что можно варьировать способ принятия решения об идентификации инсайдера, а также рассматривать уточняющие цепочки таких решений.

2.2.3 Связь целей атак инсайдера с запретами

Определение запретов основано на двух подходах. В первом подходе каждая цель из множества *T* допускает возможность построения дерева атак [см., например, [75]]. Построение деревьев атак давно используется в анализе уязвимостей. Путь к корню в дереве атак определяет последовательность шагов в каждой атаке. Кроме того, этот путь определяет цепочку событий, возможно, в нескольких информационных пространствах. Эти события в отдельных информационных пространствах характеризуются двумя способами:

- прямыми или косвенными признаками прохождения пути в дереве атак, которые могут отслеживаться функциями мониторинга безопасности;
- сопутствующими фоновыми событиями, усиливающими или смягчающими возможности реализаций событий безопасности.

С каждой вершиной или с каждым ребром дерева атак можно связать одно или несколько информационных пространств, в которых содержатся описания признаков атак и/или фона. Эти описания – суть полузапреты. Таким образом, определяется привязка целей атак к наблюдаемым данным в различных информационных пространствах.

Второй подход связан с выявлением аномалий в наблюдаемых случайных процессах в различных информационных пространствах [76]. Эти аномалии анализируются средствами интеллектуального анализа данных на предмет выявления эмпирических закономерностей и эмпирических причинно-следственных связей [77, 78]. Этот подход позволяет построить профили поведения U, из них вывести новые пути атак и новые цели атак инсайдера.

2.2.4 Примеры

Приведем примеры информационных пространств и полузапретов в них.

Пример 7. Пусть имеется описание детерминированной последовательности функций, выполняемых U. Полузапретами могут считаться отклонения от детерминированной последовательности функций в сторону доступа к ценной информации, не связанными с функциями U. Цель инсайдера – кража ценной

информации. Этот случай рассматривался в работе [70].

Пример 8. Рассмотрим связи U с другими пользователями в социальных сетях. Полузапрет определяется наличием связанных с U пользователями, имевших или имеющих отношение к криминалу. Эти полузапреты имеют значение фона. Однако они же могут указывать на цели кражи ценной информации.

Пример 9. Рассмотрим взаимодействие U с дружественными корреспондентами в социальных сетях. Из этих данных можно получить информацию о финансовых трудностях, испытываемых U. Эта информация может считаться полузапретом в негативном фоне, побуждающим U к совершению противоправных действий.

Пример 10. Полузапретом с негативным фоном можно считать внезапный интерес к дорогим материальным объектам (дорогие автомобили, дачи, квартиры и т. п.). Этот полузапрет определяется появлением аномалий в профиле U. Если U связан с доступом к ценной информации, то целью его атаки можно считать кражу ценной информации, которую с указанным полузапретом можно считать запретом.

Пример 11 «Аномальные выгрузки данных из защищенного контура».

В контуре Big Data сотрудники имеют различные роли – кто-то загружает данные в контур, кто-то создает из реплик баз данных витрины, кто-то анализирует витрины и передает результаты анализа руководству, кто-то занимается интеграцией витрин в другие ИС и т. д.

Аналитикам данных часто приходится выполнять разовые задачи, например создавать отчет по юридическим лицам в разрезе региона или выполнять анализ оттока/притока клиентов в каком-либо сегменте. Как правило, аналитики работают с данным определенный срок (до 6 месяцев). После того как аналитики данных решили, поставленную перед ними задачу, они приступают к выполнению следующей задачи, очень часто на других данных. Если аналитик не использует доступ к данным более 1 месяца, такой доступ ему вероятно не нужен, но от доступа аналитик может не отказаться.

Доступы к данным могут быть не интегрированы в процесс управления доступом, в котором предусмотрен периодический аудит доступов, переподтверждение и т. д.

Процесс управления доступом к ИС работает на предприятиях два-три десятка лет и относительно неплохо отлажен, но с данными предприятия работают 5–10 лет и процесс управления доступа к данным является относительно новым. По этой причине у сотрудников с годами работы накапливаются доступы к данным. У одного сотрудника доступов может быть десятки (Рис. 3, идентификационные данные сотрудников закрашены черным цветом): «Главная бухгалтерская книга», «Нефтегазовый сектор», «Единый профиль клиента», реплика CRM и т. д.

Рис.3. Распределение пользователей по неиспользуемым более одного месяца доступов к данным.

Также сотрудник может перейти из одного подразделения, занимающееся физическими лицами, в другое подразделение, занимающееся юридическими лицами и запросить новые доступы к данным в связи с новыми задачами. Эмпирически получен порог равный 20 доступам к данным, после которого разумно присваивать признак «имеет аномально большое количество доступов к данным». Таким образом количество доступов к данным является одним информационным пространством, в котором возможно выявлять признаки аномальных доступов сотрудников к данным. Но анализ одного информационного пространства не создает инцидент безопасности.

Но если добавить другое информационное пространство – «выгрузки данных из защищенного контура», то это даст срабатывание триггера на возникновение инцидента безопасности. Стоит дополнительно пояснить, что из контура Big Data периодически требуется выгружать результаты работы такие как отчеты, таблицы, графики для передачи их руководителям или использовании в текущей деятельности подразделений. Выгрузки необходимы т. к. интеграционные связи ИС с Big Data могут быть не доделаны или их не имеет смысла делать из-за разовых задач. Анализ выгрузок на предмет наличия в них конфиденциальных данных – это отдельная задача, которая не является частью данной работы. Факт заключается в том, что выгрузки данных происходят.

Таким образом вместе с большим количеством доступов к данным, аналитик может совершать аномально большое количество файловых выгрузок из контура Big Data (Рис. 4).

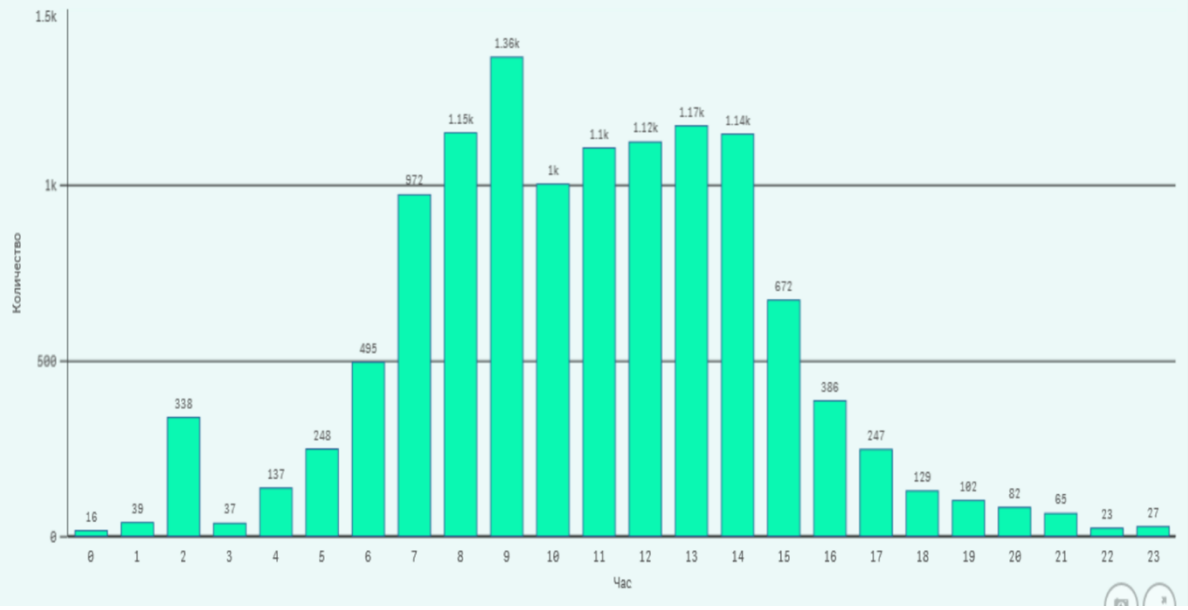

Рис.4. Пример кол-ва выгрузок файлов из Big Data на рабочее место пользователями (в месяц).

На рисунке показано количество выгрузок всеми сотрудниками, работающими в Big Data, в день. Представленный график является суммой всех графиков по каждому сотруднику и по каждому сотруднику возможно посмотреть детализацию выгрузок.

Совместив два информационных пространства «выгрузки» и «доступы» возможно вызвать реакцию на реализацию угрозы «аномальные выгрузки данных из защищенного контура». В результате анализа двух информационных пространств вскрывается аномалия, когда сотрудник долго не пользовался доступом к данным и вдруг совершает большое количество выгрузок неиспользуемых данных за периметр защищенного контура Big Data.

Стоит отметить, что информационных пространств может быть не только два. Можно соединить и другие информационные пространства такие как «офис размещения», «время работы», «доступы к данным другого подразделения», «владелец технологических учетных записей» и тоже вызвать срабатывание профиля угроз. Для реализации ПО, описанного в Главе 4 и выявляющего признаки инсайдера, были созданы десятки информационных пространств.

Итак:

В задаче поиска инсайдеров разработан подход к объединению компрометирующих данных, наблюдаемых в различных информационных пространствах. Этот подход основан на запретах и полузапретах вероятностных мер в различных информационных пространствах. С последовательностями событий, наблюдаемыми в этих информационных пространствах, связываются булевы переменные. Появление полузапретов соответствует значению «1» соответствующих булевых переменных.

Последовательности булевых переменных в различных информационных пространствах легко связываются с помощью логических выражений. Эти выражения описывают опасные тенденции, наблюдаемые в различных информационных пространствах. Выбор логических выражений (функций) может быть построен на использовании всей предыстории наблюдений за потенциальным инсайдером в соответствующих информационных пространствах или при анализе его подозрительной деятельности в заданный промежуток времени. При этом использование различных булевых функций позволяет проводить различные виды анализа на одних и тех же статистических данных.

С учетом различных целей атак инсайдера в дальнейшем можно перейти к рассмотрению k-значных функций.

Глубокие связи в каждом информационном пространстве можно исследовать с помощью методов интеллектуального анализа данных.

Выводы по Главе 2

1.Представление конечного класса объектов в форме множества характеристик (параметров) этих объектов назовем параметризацией рассматриваемого класса. Кроме идентификации объектов множествами характеристик существует задача выявления причины того, что некоторые объекты класса обладают свойством P. Для решения этой задачи, в условиях появления новых объектов, исходного множества характеристик может не хватить. В этом случае необходимо изменять параметризацию. В работе построены методы изменения начальной параметризации в задаче уточнения эмпирической причины появления свойства P при расширении исходных данных. Построенные методы продемонстрированы на практических примерах.

2. В задаче поиска инсайдеров разработан подход к объединению компрометирующих данных, наблюдаемых в различных информационных пространствах. Накопление информации в каждом пространстве рассматривается как вероятностный процесс. Рассматриваемый подход основан на запретах и полузапретах вероятностных мер в различных информационных пространствах. С последовательностями событий, наблюдаемыми в этих информационных пространствах, связываются булевы переменные. Появление полузапретов соответствует значению «1» соответствующих булевых переменных.

Последовательности булевых переменных в различных информационных пространствах легко связываются с помощью логических выражений. Эти выражения описывают опасные тенденции, наблюдаемые в различных информационных пространствах.

Глава 3. Вероятностные оценки в задачах выявления признаков инсайдера

3.1 Вероятностные оценки признаков сговора инсайдеров

В гетерогенных данных и пополняемых Big Data методы математический статистики работают неэффективно, но их можно адаптировать для решения задач. Требуется создать методы, которые позволят применять методы математической статистики на Big Data. Суть предлагаемого подхода - сегментировать Big Data таким образом, чтобы математическую статистику возможно было применить на сегментах данных. Предлагаемый подход встроен в систему выявления признаков инсайдера и позволяет сначала фильтровать большие объемы данных и затем принимать обоснованные решения с помощью методов математической статистики.

Сбор ценной информации сотрудниками, использующими личные связи, является серьезной проблемой информационной безопасности в государственных и коммерческих организациях. При этом часто нарушитель информационной безопасности сам имеет доступ к части ценной информации, но не имеет права получать ее целиком. Примером такой ситуации может являться разделение общей проблемы для решения ее в различных подразделениях организации. Тогда враждебный инсайдер (далее инсайдер) пытается узнать недостающую ценную информацию от своих друзей в других подразделениях или войти в сговор о сборе и продаже ценной информации.

В работе [79] приведен ряд вызовов, исследование которых авторы данного обзора считают важными, но трудными проблемами. Вызов 4 (см. [79]) связан с проблемой выявления инсайдеров, организованных с помощью преступного сговора.

Выявление признаков инсайдеров в течение ряда лет финансировалось американским управлением исследованиями Министерства обороны США (DARPA). Основное направление этих исследований связано с выявлением

аномалий в больших данных. В частности, в работе [80] разработан язык описания аномалий и представлена информация по его использованию.

В работах [81, 2] для выявления слабо выраженных аномалий, порожденных инсайдерами предложено использование информации из нескольких информационных пространств.

Важной проблемой является определение условий, при которых выявление признаков инсайдера возможно. Все методы в данной области строятся на основе некоторых предположений, выполнение которых необходимо для их работоспособности. Далее рассматривается проблема выявления инсайдера в банковской сфере.

В работе предполагается, что целями инсайдера являются сбор и продажа ценной информации о клиентах банка. Как правило, такой информацией являются персональные данные вместе с ценной информацией о счетах и движении денежных средств. Для того, чтобы эти данные могли дать постоянный доход инсайдеру, они должны собираться в достаточно большом объеме.

Для защиты от таких инсайдеров можно разделить информацию, например, на персональные данные и ценную информацию на счетах. При этом работать с каждым из этих блоков данных могут только различные менеджеры (сотрудники банка). Отсюда возникает задача о выявлении возможного сговора каких-то менеджеров, занимающихся либо персональными данными, либо информацией о счетах. В этом параграфе исследованы условия выявления такого сговора.

3.1.1 Формальная модель сговора инсайдеров

Для исследования задачи выявления признаков инсайдеров, использующих сговор, рассмотрим следующую модель. Обозначим через V множество клиентов, использующих сервисы банка, через U – множество менеджеров (пользователей банковской системы и реализующих информационные технологии (ИТ)). Хранилище данных, которое используют менеджеры устроено следующим образом. Все данные заносятся в прямоугольную таблицу, столбцы которой нумеруются атрибутами $A_0, A_1, \ldots$, строки содержат записи данных и результатов

действий с ними. Строки не изменяются, но при изменениях данных появляются новые строки с новыми данными. В отличие от традиционных реляционных баз данных записи не обязательно содержат значения всех атрибутов и поиск по таблице осуществляется по сложному ключу, зависящему от типа данных в строке. Однако каждая строка содержит данные о менеджере, осуществляющем обращение, идентификаторе экземпляра ИТ, которая связана с обращением, и санкционированием текущей транзакции от предыдущих шагов ИТ. Часть атрибутов $\vec{A}_1$ соответствует содержанию персональных данных, нужных для авторизации транзакций, часть атрибутов $\vec{A}_2$ соответствует ценной информации (ЦИ) (данные счетов, переводы и др.).

Политика безопасности запрещает одному менеджеру иметь доступ одновременно к $\vec{A}_1$ и к $\vec{A}_2$. Поэтому строки, соответствующие использованию атрибутов $\vec{A}_2$, обладают индексами, скрывающими данные, соответствующие атрибутам $\vec{A}_1$. Кроме того, выделены пользователи U_1, которые могут иметь доступ к информации с атрибутами $\vec{A}_1$, и пользователи U_2, не имеющие доступа к данным с атрибутами $\vec{A}_1$, но работающие с данными, соответствующие атрибутам $\vec{A}_2$.

Интересная для инсайдеров ЦИ может быть получена тогда и только тогда, когда известны значения атрибутов $\vec{A}_1$ и соответствующие им значения атрибутов $\vec{A}_2$. Для получения такой информации необходим сговор какого-либо пользователя из множества U_1 с каким-либо пользователем из множества U_2. Задача состоит в выявлении таких инсайдеров.

Пусть $u_1 \in U_1$ и $u_2 \in U_2$ образуют такую пару (u_1, u_2) инсайдеров, использующих сговор.

Сделаем дополнительные предположения о том, что пара (u_1, u_2) может так управлять потоком заявок клиентов на сервисы, что пара (u_1, u_2) появляется чаще, чем все остальные пары из множества $U_1 \times U_2$. Тогда можно пытаться

использовать методы математической статистики для выявления пары (u_1, u_2).

Однако для большого значения числа элементов в множестве $U_1 \times U_2$ выявление пары (u_1, u_2) упирается в проблему многих малых выборок [82, 83].

Пусть статистические критерии выявления (u_1, u_2) имеют ошибки $\alpha > 0$, $\beta > 0$, α - вероятность «ложной» тревоги, β - вероятность пропуска (u_1, u_2). Тогда при многих малых выборках анализ «ложных» тревог является трудоемкой задачей, а вероятность выявления пары (u_1, u_2) ограничена снизу константой (приблизительно 1/3). Поэтому для эффективного применения статистических методов необходимо снижать объем множества малых выборок. Иными словами, необходимо применять статистические методы в рамках некоторых кластеров ограниченного набора малых выборок.

В работе используется параметр, управляющий объемом множества малых выборок, и позволяющий повышать эффективность статистического анализа. Этим параметром является объем множества клиентов, в интересах которых проводятся транзакции, интересные для (u_1, u_2). Т. е. вместе с каждой парой (u, u'), $u \in U_1$, $u' \in U_2$, необходимо рассматривать параметр v со значениями в V, идентифицирующий клиента, который инициирует транзакцию. Обозначим эту тройку $(u, u')|v$.

Инсайдеров (u_1, u_2) интересуют клиенты из множества V, продажа информации о которых имеет высокую стоимость. Вместе с тем, если инсайдеров интересуют только такие клиенты (обозначим множество таких клиентов через V_1) и их транзакции, то количество троек $(u, u')|v$, где $v \in V_1$, для анализа сокращается, что позволяет повысить эффективность статистических методов.

3.1.2 Анализ выявляемости сговора инсайдеров

Пусть транзакция обрабатывается случайной парой менеджеров (u,u'), $u \in U_1$, $u' \in U_2$. Положим $|U_1| = n_1$, $|U_2| = n_2$, $|U_1 \times U_2| = n_1 \cdot n_2 = n$. Тогда вероятность появления пары (u,u') равна $\frac{1}{n}$.

Если (u_1,u_2) – инсайдеры, использующие сговор, и клиент *v* представляет для них интерес, то вероятность того, что эта пара будет обслуживать этого клиента равна

$$P\left(\left(u_1,u_2\right)|\, v\right) = p,$$

и в случае любой другой пары $(u,u') \neq \left(u_1,u_2\right)$ эта вероятность равна

$$P\left(\left(u,u'\right)|\, v\right) = (1-p)\frac{1}{n-1}.$$

С помощью методов кластеризации можно выделить множество V_1 тех клиентов, которые представляют интерес для инсайдеров, использующих сговор. Ясно, что мощность множества V_1 много меньше, чем мощность множества *V*. Как было отмечено выше, это позволит избежать эффекта малых выборок [83].

Пусть C_1 - среднее число транзакций у клиентов из множества V_1. Тогда средний объем данных для клиентов из множества V_1 равен $|V_1| \cdot C_1 = m$. Далее считаем *m* известным параметром схемы. Тогда соотношение параметров *m*, *n*, *p* определяет возможность выявления инсайдеров, использующих сговор. Инсайдеры (u_1,u_2) не могут переключить на себя весь поток клиентов из множества V_1. Поэтому для данных ξ, которые они получили, можно предложить модель биномиального распределения

$$P\left(\xi\left(u_1,u_2\right)=k\right) = \binom{m}{k} p^k (1-p)^{m-k}.$$

Тогда для любой пары (u,u') $\neq \left(u_1,u_2\right)$ распределение числа случаев предоставления сервисов описывается следующим образом

$$P\left(\xi\left(u,u'\right)=r\right) = \sum_{k=0}^{m} \binom{m}{k} p^k (1-p)^{m-k} \binom{m-k}{r} \left(\frac{1}{n-1}\right)^r \left(1-\frac{1}{n-1}\right)^{m-k-r}.$$

Поскольку числа m и n являются большими, то рассмотрение задачи выявляемости признаков инсайдеров (u_1, u_2) в множестве V_1 будет вестись в терминах асимптотических распределений вероятностей в схеме серий, т.е. в предположении, что $m \to \infty, n \to \infty, p \to 0$. Задача состоит в поиске значений вероятности p, при которых вероятность выявления инсайдеров (u_1, u_2) стремится к 1.

В практическом плане интересен случай, когда $m \to \infty, n \to \infty$ и $\alpha = \frac{m}{n}$ удовлетворяет условию $\frac{\alpha}{\ln n} \to 0$. Это условие соответствует преобладанию числа предоставления сервисов для клиентов из множества V_1 над числом пар менеджеров $n_1 \cdot n_2 = n$, которые проводят это обслуживание. В каком-то смысле множитель $\ln n$ можно интерпретировать, как количество обслуживаний на одну пару менеджеров, не являющихся инсайдерами (u_1, u_2), использующими сговор.

Рассмотрим случайную величину $\eta_{(n)}$, равную максимальному числу обслуживаний клиентов из множества V_1 по всем парам менеджеров (u, u'). Обозначим

$$\alpha = \frac{m}{n},\ p_k = \frac{\alpha^k}{k!} e^{-\alpha},\ k = 0, 1, \ldots$$

В работе [84] приведена следующая теорема.

Теорема. *Если* $m, n \to \infty$, $\frac{\alpha}{\ln n} \to 0$ *и* $r = r(\alpha, n)$ *выбрано так, что* $r > \alpha$ *и* $np_r \to \lambda$, *где* λ *– положительная постоянная, то*

$$P\{\eta_{(n)} = r - 1\} \to e^{-\lambda},\ P\{\eta_{(n)} = r\} \to 1 - e^{-\lambda}.$$

Рассмотрим простой пример применения данной теоремы. Пусть

$$n \frac{\alpha^3}{6} e^{-\alpha} \to \lambda > 0,$$

т. е. $r = r(\alpha, n) = 3$. Тогда максимальное число обслуживаний клиентов из множества V_1 по всем парам менеджеров $(u,u') \neq (u_1,u_2)$ равно 3 с вероятностью, стремящейся к 1.

Превышение числа появлений пары (u_1,u_2), начиная с 4, однозначно идентифицирует эту пару инсайдеров с вероятностью, стремящейся к 1. Вероятность того, что пара (u_1,u_2) в определенной выше схеме появится не больше 3-х раз равна

$$(1-p)^m + mp(1-p)^{m-1} + \frac{m(m-1)}{2} p^2 (1-p)^{m-2} + \frac{m(m-1)(m-2)}{6} p^3 (1-p)^{m-3}.$$

Отсюда следует, что при $p \le \frac{\ln m}{m}$ пара инсайдеров (u_1,u_2), использующих сговор, выявляется с вероятностью, стремящейся к 1.

Пример 12 «Нарушения в порядке предоставления доступа к данным (потенциальный сговор)».

Как правило любой доступ к данным (и не только к данным) сопровождается согласованной заявкой на доступ или автоматическим предоставлением доступа по ролевой модели подразделения, что тоже отмечается в учетных системах. К сожалению, бывают случаи, когда у сотрудника имеется доступ к данным, а заявки на доступ отсутствуют (Рис. 5, идентификационные данные сотрудников закрашены черным цветом).

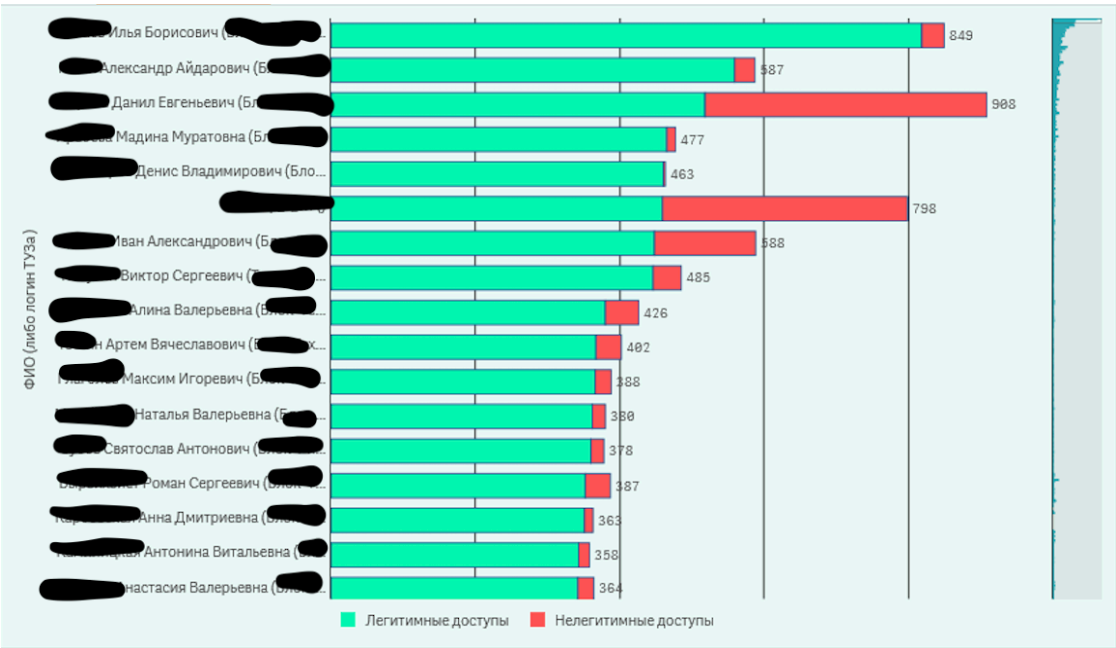

Рис. 5. Распределение сотрудников, имеющих признаки нарушения в порядке получения доступа к данным.

Нарушения в порядке предоставления доступа могут случаться, когда сотрудник создал условия, при которых ИТ-администратор предоставил ему доступ без заявок. ИТ-администратора могли ввести в морально сложную ситуацию. Морально сложные ситуации возникают в отчетный период, когда ключевые вехи знаковых проектов находятся под риском или ИТ-администратор имеет неформальные связи аналитиком данных, кому требуется доступ к данным или ИТ-администратор зависит от аналитика данных в других задачах, успех которых влияет на его квартальную или годовую премию. Также ИТ-администратор может быть молодым недавно нанятым сотрудником, который еще не успел усвоить порядок работы и на него оказывают давление другие подразделения одновременно с просьбой предоставить доступ к данным.

Также нарушения в порядке предоставления доступов к данным возможны по техническим причинам. Могли быть произведены работы по предоставлению массового доступа группе сотрудников в связи с переводом их в другое подразделение или в связи и переименованием подразделений, которое повлекло сбой в доступах. Массовые доступы могут делаться по спискам (например, в форме Excel). Очевидно, что списки Excel составляются вручную сотрудниками. А если подразделения территориально распределены и состоят из сотен

сотрудников, то составление корректного списка Excel является сложной задачей, требующей много времени и внимания. Соответственно ввиду большой загруженности сотрудники могут халатно подходить к задаче составления списков доступа. Поэтому в списках доступа могут быть опечатки, ошибки, однофамильцы из других подразделений, сотрудники, которым доступ к данным не нужен и т.д. Чем больше у сотрудника доступов с нарушениями, тем вероятнее его деятельность необходимо рассматривать детальнее и разумно присвоить признак «большое кол-во доступов к данным, полученное с нарушением установленного порядка (потенциальный сговор)».

Пример 13 «Несанкционированная передача данных в командном пространстве (потенциальный сговор)».

В контуре Big Data работа с данными может производиться единолично сотрудником в его каталоге/пространстве, а может производиться группой сотрудников в групповом пространстве. Сотрудники работают в группах т. к. задачи анализа данных часто требуют различных компетенций. Группа сотрудников работает в своем командном пространстве (файловый каталог). Если у сотрудника имеется доступ к командному пространству, это означает, что у него имеется доступ ко всем данным командного пространства, а это могут быть терабайты данных или десятки фрагментов витрин или реплик баз данных. На большом предприятии команд может быть десятки или сотни.

У любого командного пространства (как и у любого информационного ресурса) имеется владелец, который разрешает или не разрешает доступ в свое командное пространство. Как правило команды состоят из 2–10 человек.

Но бывают и аномалии в структуре команд. Если в командном пространстве работают одновременно больше 30 человек (Рис. 6, идентификаторы команд закрашены черным цветом), то высока вероятность, что владелец командного пространства лично не знает пользователей и механически согласует все заявки на доступ к своему командному пространству.

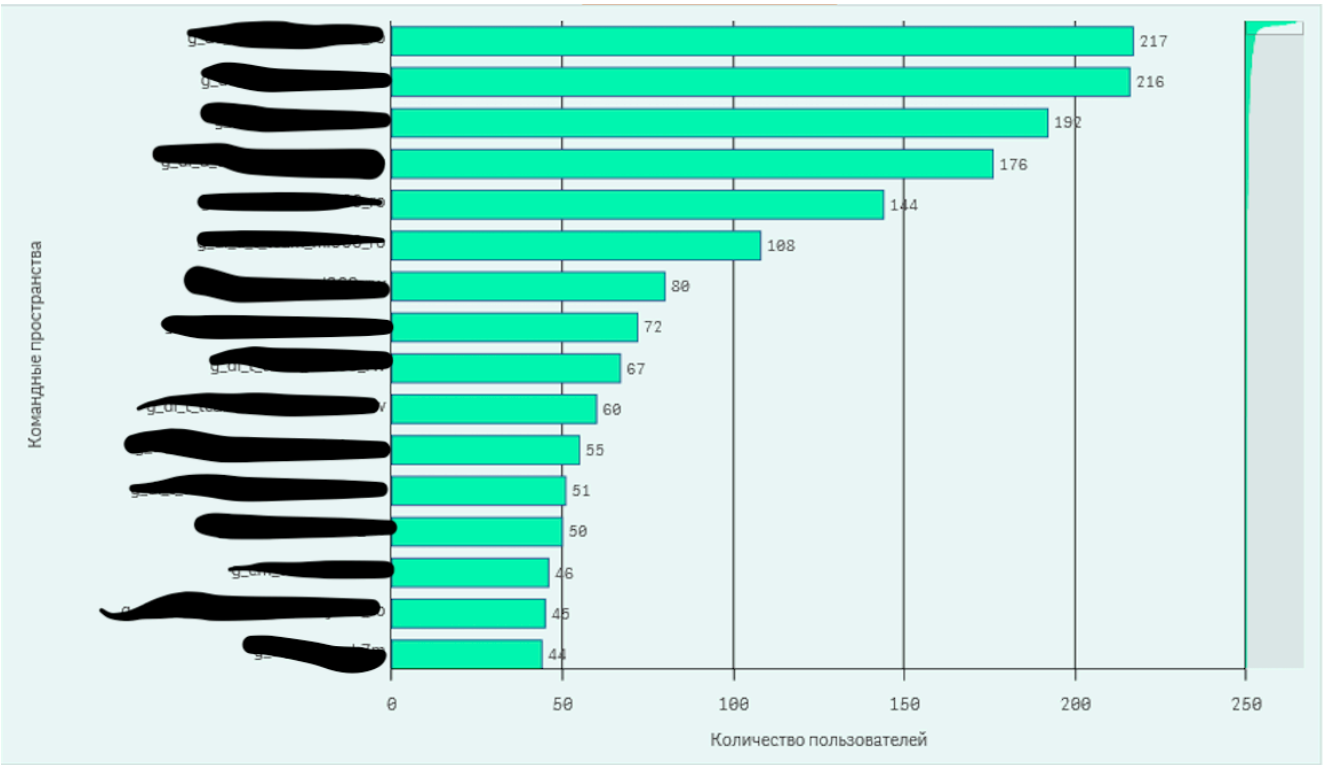

Рис.6. Распределение командных пространств (каталогов) по количеству пользователей, которые имеют туда доступ.

Возможны случаи, когда владелец командного пространства делегировал право предоставлять доступ своему коллеге и коллега не анализирует каждую заявку на доступ к командному пространству, а механически согласует т. к. воспринимает данную функцию как дополнительную нагрузку, отвлекающую от его основных задач.

Также возможен случай так называемого «теневого ИТ (shadow IT)», когда вместо того, чтобы получать доступ к каждой витрине или реплике базы данных в установленном порядке, владелец командного пространства копирует витрины и реплики в командное пространство, где сам уже предоставляет доступ кому пожелает.

Таким образом чем больше сотрудников находится в командном пространстве, тем вероятнее, что всех сотрудников в командном пространстве необходимо сканировать детальнее, а владельцу командного пространства присвоить признак «предоставление доступов к данным по сговору».

Итак:

Показано, как сочетать методы кластеризации и методы математической статистики для выявления инсайдеров, использующих сговор. А именно, для эффективного применения математической статистики на множестве разнородных данных, которые состоят из множества малых выборок, необходимо приближенное выполнение соотношения между параметрами схемы.

Однако согласование параметров возможно, когда существует возможность управления объемами данных. Для такого управления можно использовать методы кластеризации.

В рассмотренной задаче можно выделить систему вложенных кластеров данных, повышающих шансы выявления инсайдеров, использующих сговор. Выделение таких кластеров позволило доказать пренебрежимо малую вероятность превышения построенного порога для числа случайных повторений встречаемости пар менеджеров, не являющихся инсайдерами, использующими сговор.

Вместе с тем, при полученной отсюда оценке вероятности встречаемости инсайдеров, использующими сговор, их идентификация происходит с вероятностью, стремящейся к 1.

3.2. О вероятностных оценках достоверности эмпирических выводов

Обнаружить враждебные действия инсайдера намного сложнее, чем вредоносное воздействие хакера, потому что инсайдеры знакомы с системой безопасности организации и имеют доступ к ее части, а также количество вредоносных действий, которые можно зафиксировать, мало по сравнению с количеством легальных действий, совершаемых пользователями в процессе работы с информационной системой (ИС) организации. Разработанные алгоритмы поиска инсайдеров часто не способны выявить новые методы атак инсайдера, а лучшие результаты показывают на шаблонных входных данных.

Необходимо отметить, что как правило поиск инсайдеров связан с поиском

аномалий в различных информационных пространствах, где могут появляться признаки инсайдеров. Например, министерство обороны США провело исследования под названием Anomaly Detection at Multiple Scales (ADAMS) [85], основная цель которого разработать и внедрить технологию классификации и обнаружения аномалий в больших объемах данных с целью поиска следов инсайдеров.

В настоящее время множество статей посвящено проблеме поиска инсайдеров. Отметим некоторые работы, имеющие связь с исследованиями данной работы.

Пример корреляционного подхода в поиске инсайдера можно найти в работе [86]. В этой работе рассматриваются описания связей между событиями и инсайдерами. В работе [87] прямо ставится проблема верификации вывода о результатах поиска инсайдера в нескольких информационных пространствах, но методы верификации не рассматриваются. Вероятностные методы рассмотрены в [88], где показано, как использовать скрытые Марковские модели для оценок нормального поведения пользователей ИС. Обзор [89] рассматривает различные подходы к анализу угроз, порождаемых инсайдерами.

Приведем несколько российских работ, посвященных проблеме поиска инсайдеров. В статьях [81, 3, 2] построены методы работы с несколькими информационными пространствами в задачах поиска инсайдера и поиска аномалий, связанных с инсайдером.

Срабатывания системы выявления признаков инсайдеров могут быть вызваны как случайными, так и закономерными аномалиями. Может получиться, что аномалия вызвана спонтанным действием сотрудника или опечаткой в запросе к данным (например, вместо select id=5, написал select id=6). Очевидно, что необходимы методы, способные оценивать вероятность возникновения аномалии. Это позволит ранжировать аномалии по вероятности, и оперативный работник сможет обрабатывать аномалии с наименьшей вероятностью случайного возникновения. Например, если вероятность случайной аномалии крайне мала, значит это закономерность и оперативному работнику стоит реагировать. Задача

данного раздела – разработать метод, позволяющий произвести оценку вероятности аномалии с целью приоритизировать обработку аномалий. В этом параграфе рассматривается задача верификации найденных признаков инсайдеров вероятностными методами.

Пусть задано множество пользователей ИС, содержащей профессиональную информацию и ценную информацию (в том числе – дорогостоящую). Допустим, что среди пользователей есть один инсайдер. Относительно его выявления возможны различные частные постановки задач.

A. Если произошло компьютерное преступление, в котором виновен инсайдер, то надо провести расследование и определить личность его в множестве пользователей, собрать доказательную базу.

B. Если есть множество конкретных должностных лиц, то надо определить есть ли в этом множестве инсайдер.

Признаком возможных действий инсайдера являются аномалии в технологических и информационных процессах. Тогда надо выявлять аномалии различных типов и на основе найденных аномалий определить, существует ли инсайдер.

Перечислим некоторые подходы к выявлению искомой информации.

1. Подход поиска связи какого-нибудь пользователя *u* и свойства, которое наблюдается в некоторых событиях, связанных с возможным появлением инсайдера. Например, появление у клиентов банка больших сумм на счету сопровождается дорогими покупками пользователя – служащего банка. Гипотеза состоит в том, что этот пользователь продает информацию о появлениях указанного свойства и получив деньги, тратит на заранее запланированные покупки или участие в азартных играх. Связь пользователя с появлением данного свойства связано с задачами A и B.

2. В ходе расследования компьютерного преступления в банке выяснилось, что преступления подобного типа уже совершались и проходили по определённой схеме. Если пользователь участвует в схеме, то должны появляться другие элементы схемы при проявлении действия из

этой схемы. Тогда подход состоит в выявлении действия схемы.

3. На фоне стандартной деятельности пользователей появились вкрапления нестандартных действий пользователя, которые следует рассматривать как аномалию, если эти действия удается заметить.

4. Одним из важных свойств, используемых для выявления инсайдеров, являются противоречия, например размер зарплаты и размеры расходов, круг профессиональных задач и широта интересов и др.

5. Подходом, близким по эффективности, но отличающимся от метода поиска противоречий, является поиск некоторых характеристик пользователя, порождающих желание стать инсайдером. Например, знакомства с криминальными элементами, жадность, игромания и др. являются такими характеристиками.

Выше перечислены некоторые свойства, но этот набор свойств не является полным. Во всех случаях возможно построение вероятностной модели и проверки согласия с ней. Но надо заметить, что выявленные свойства, которые назовем эмпирическими закономерности, могут проистекать от случайного стечения обстоятельств. Поэтому надо иметь методы, позволяющие фильтровать возможности случайных сочетаний свойств, которые можно принять за искомые закономерности. Отметим, что чем сложнее эти методы, тем меньше они применимы на практике. С другой стороны, упрощение моделей и методов может привести к потере реальных закономерностей. Выскажем гипотезу, что пропущенные или отброшенные кандидаты на закономерности в случае реального инсайдера должны повторяться и при запоминании отброшенных кандидатов могут получить уверенное подтверждение при вторичном анализе.

Параграф посвящен построению простых методов фильтрации кандидатов на закономерности в подходах 1-го типа или фильтрации аномалий.

3.2.1 Модель корреляции важных событий и действий пользователей

Пусть $U = \{u_1, u_2, ..., u_m\}$ - пользователи ИС (сотрудники организации),

$V = \{v_1, v_2, ..., v_n\}$ - клиенты, на которых работают пользователи и эти множества не пересекаются.

Простейшие модели и логику работы с ними построим для подходов 1-го типа.

Пусть C – события, связанные с появлением и доступностью ценной информации. Предположим, что события C возникают в соответствии с пуассоновским процессом с параметром λ. Траты (большие) произвольного пользователя u из множества U также происходят в соответствии с пуассоновским процессом с параметром μ. Все процессы независимы. Введем параметр T – время возможного появления зависимости пользовательских трат и появления ценной информации. Вероятность того, что данный пользователь в промежуток времени T не производил больших трат, равна $e^{-\mu T}$.

Рассмотрим возможность появления трат в фиксированный промежуток времени T, связанный с фиксированным событием C. Вероятность, что все пользователи в промежуток T не производили больших трат, равна $e^{-\mu Tm}$. То есть можно окружить каждое появление события C отрезком длины T, начинающимся с момента появления события C. По предположению события C появляются в соответствии с пуассоновским процессом и можно выбрать промежуток времени T таким образом, чтобы отрезки длины T не пересекались с вероятностью, близкой к 1 (для этого надо рассматривать события С в процессе с $\lambda << \mu$). Вероятность того, что ни в одном промежутке времени при появлении k событий C не было больших трат, не превосходит $ke^{-\mu Tm}$.

Если μTm большая величина, то вероятность непопадания больших трат на какой-нибудь отрезок длины T вокруг событий C мала. Тогда появление таких событий указывает на множество подозрительных пользователей. Можно выписать распределение числа таких пользователей, но вопрос, когда их будет мало или не будет вовсе с большой вероятностью в данном случае более интересен. То есть поставим вопрос о выборе параметров целесообразности поиска. При каких соотношениях параметров асимптотически в результате

корреляционной проверки проявляется только инсайдер? Для решения этой задачи используем параметр k – число появлений события C.

Предположим, что все пользователи являются честными. Пусть φ - случайная величина, равная числу появлений события C за время наблюдения t за поведением и тратами пользователей из множества $U = \{u_1, u_2, ..., u_m\}$. События C появляются в соответствии с законом простого пуассоновского процесса:

$$P(\varphi = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Предположим, что выброшены все пересечения промежутков T после событий C. Распределение Пуассона при больших k можно аппроксимировать нормальным распределением. Отсюда получим:

$$P(\varphi = \lambda t(1 + o(1)) = 1 + o(1).$$

Тогда время t наблюдения для получения k событий C приближенно равно $\frac{k}{\lambda}$.

Для определения необходимых значений k и определения невозможного события для всех честных пользователей воспользуемся следующими оценками.

Вероятность появления большой траты в данный промежуток T для фиксированного пользователя u равна

$$p = 1 - e^{-\mu T}$$

По предположению промежутки T не пересекаются. Тогда появление большой траты в каждый промежуток T для данного пользователя u равна p^k. Отсюда среднее число таких пользователей равно mp^k. При $T = const$ и $k = \frac{\ln m + \ln \ln m}{e^{-\mu T}}$ получим оценку среднего числа таких пользователей $\frac{1}{\ln m}$.

В этих условиях проявление инсайдера с вероятностью, близкой к 1, возможно при появлении пользователя, который делает большие траты после каждого появления события C не менее k раз. Отметим, что время наблюдения оценивается величиной

$$t = \frac{k}{\lambda}(1 + o(1)).$$

Такие параметры поиска значительно лучше, чем при ожидании полного отсутствия трат.

3.2.2 Модель машинного обучения

В предыдущем разделе показано, что предложенный метод работает, но требует выполнения двух условий:

- наблюдение за всеми тратами;
- появление трат после появления событий *C*.

Наблюдения за всеми большими тратами возможно, например, когда в игровом клубе находится инсайдер, который докладывает в службу безопасности обо всех сотрудниках организации, играющих на крупные суммы. В таком случае в предыдущем параграфе найдены условия однозначного выявления именно этого типа инсайдеров.

Рассмотрим схожую задачу поиска зависимости больших трат и появления ценной информации в базе данных организации в несколько других предположениях. В рассматриваемом случае возможность выявления инсайдера основана на сокращении числа подозреваемых пользователей (2-ой подход).

Пусть за время *t* в множестве *U* выделено подмножество U_1 пользователей с очень большими единовременным тратами (покупка машины, квартиры, дачи и др.). Под эти покупки возможно взяты кредиты или займы у родственников или знакомых. Ясно, что объем множества U_1 значительно меньше, чем объем множества *U*. Воспользуемся этим обстоятельством. Каждый участник множества U_1 имеет основания использовать свое служебное положение для дополнительного заработка. Соберем информацию о доступах пользователей из U_1 за время *t* к ценной информации. Пусть $C(1, \varphi)$, $C(2, \varphi)$,…, $C(k(\varphi), \varphi)$ случаи получения ценной информации пользователем φ из U_1 за время *t* (за счет расширения множества запросов в базе данных и других способов).

Сравним полученные данные с доступами к ценной информации у каждого из пользователей в множестве $U \setminus U_1$ (машинное обучение). Пусть $\nu(i)$ - число пользователей, которые *i* раз получили данные за время *t*, которые можно считать ценной информацией. Рассмотрим относительные частоты чисел $\nu(i)$ к мощности множества $U \setminus U_1$. Эти числа можно рассматривать как оценки вероятностей P, что данный (честный) пользователь *i* раз получит ценную информацию за время *t*. Тогда при маленьких значениях $P(k(\varphi))$ для φ из U_1 и небольших множествах U_1 получаем оценку вероятности, что пользователь φ злоупотреблял своим служебным положением и может быть инсайдером, который продает ценную информацию.

Пример 14 «Сотрудник имеет доступ не к своим данным».

Аналитиков данных нанимают, чтобы выполнять задачи собственного подразделения, как правило, на основе данных собственного подразделения. Конечно, бывают исключения, когда создаются подразделения, анализирующие данные всего предприятия (сводят баланс, рассчитывают операционные риски за все предприятие в целом и т. д.), но таким подразделениям, как правило не нужны персональные данные, а нужна агрегированная информация.

Задачи аналитиков данных могут быть разные, например, выполнить расчеты оттоков/притоков клиентов, разработать витрины премирования сотрудников. Также, как привило, руководители не нанимают в свой штат сотрудников, чтобы решать задачи другого подразделения (бывают исключения, когда сотрудника оформляют в соседнее подразделение т. к. в целевом подразделение закончились ставки).

Таким образом, если у сотрудника большое количество доступов к данным другого подразделения, то это является косвенным признаком того, сотрудник выполняет не свою функцию (Рис. 7, идентификаторы сотрудников закрашены черным цветом).

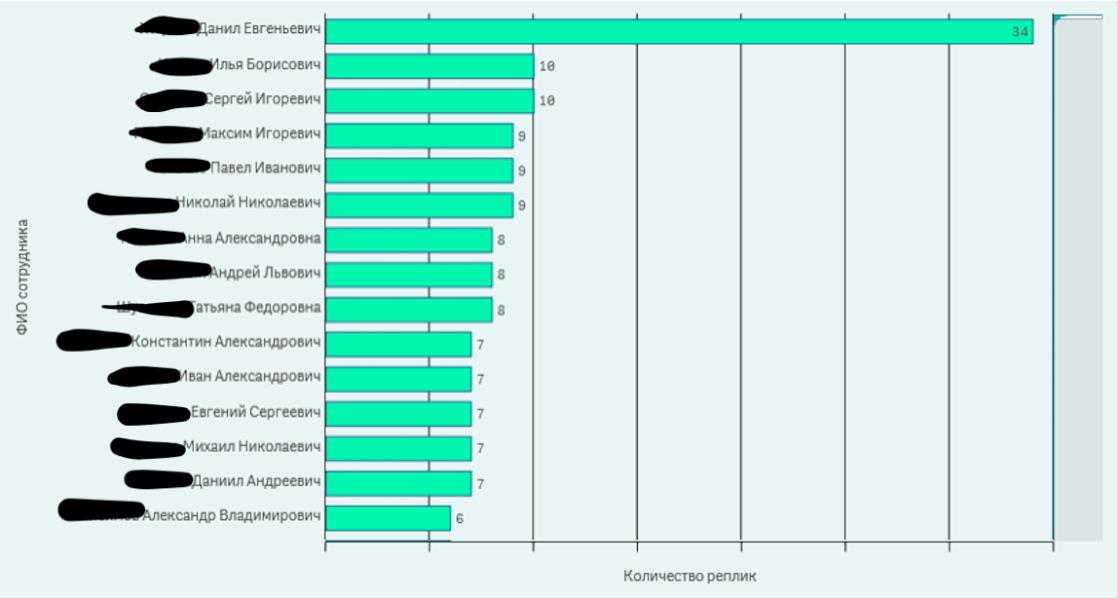

Рис.7. Распределение сотрудников, которые имеют доступ к данным соседних подразделений, а не к данным своих подразделений.

Имея доступы к данным соседних подразделений, может свидетельствовать о том, что сотрудник одновременно оформлен на предприятии и в дочернюю компанию, занимающую анализом данным. Оформление сотрудника в штат предприятия упрощает процедуры получения доступа к данным. Сотрудник может несанкционированно работать с данными без ведома владельца данных в интересах дочерней компании.

Также сотрудник имеет большинство доступов к данным другого подразделения может свидетельствам о том, что сотрудник перешел из одного подразделение в другое и забыл отозвать старые доступы. Ложная тревога.

Так или иначе сотрудник, имеющий доступы к данным соседнего подразделения вероятнее будет иметь признак «Сотрудник имеет доступ не к своим данным». Эмпирически получен порог равный 80% (80% доступов к данным чужих подразделений, 20% доступов к данным своего подразделения), при котором разумно проставлять данный признак.

Пример 15 «Работа в Big Data в удаленном режиме».

Как правило предприятия не разрешают работать сотрудникам в Big Data в удаленном режиме т. к. среда удаленной работы сотрудника не контролируемая.

Неконтролируемая среда удаленной работы означает, что у предприятия отсутствуют средства способные установить действительно ли, сотрудник работает в Big Data или сотрудник передал кому-то доступ, или находится под влиянием третьих лиц.

Витрины возможно псевдообезличивать (удалять/заменять прямые идентификаторы такие как номер паспорта, номер машины, номер телефона и т.д. на псевдомины т. е. поля, которые однозначно определяют физическое лицо заменяются на какую-либо последовательность символов). Псевдообезличенные витрины не содержат персональных данных физических лиц и некоторые предприятия разрешают в них работать удаленно.

Однако в реальной среде существуют сотни или тысячи витрин и случаются ошибки в признании витрины псевдообезличенной, например сотрудники невнимательно проверили витрину или механизм, проверяющий витрину на предмет наличия в ней персональных данных дал сбой, или к псевдообезличенной витрине добавили несколько полей, содержащих персональные данные и не запустили процедуру проверки.

В этой связи удаленная работа в комплексе Big Data несет повышенный риск (Рис. 8, идентификаторы сотрудника закрашены черным цветом).

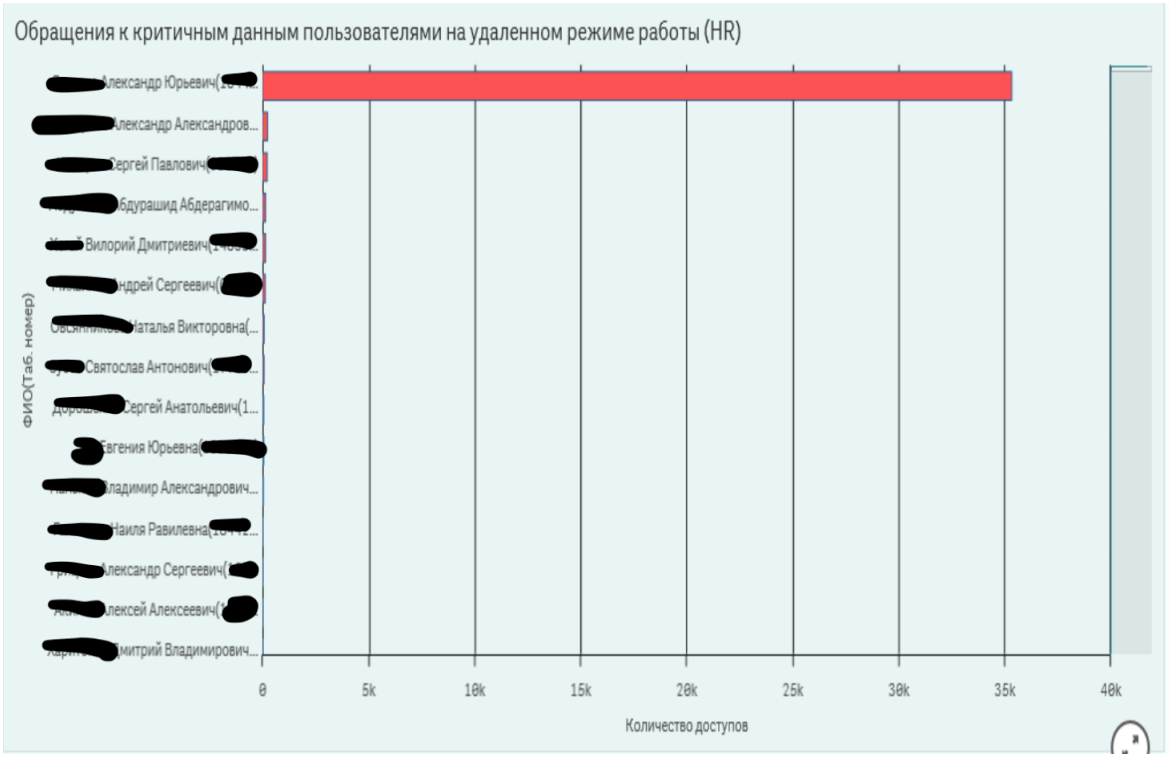

Рис.8. Факты обращения пользователя к данным, в случаях, когда пользователь находился не на территории офисов компании.

Возможно разработать проверяющий механизм, который позволит вычислить сотрудников, находящихся в удаленном режиме работы (согласно подписанным документам, хранящимися в кадровой базе и группами удаленного доступа, настроенными в корпоративном каталоге типа Active Directory) и просматривающих критичные данные. Таким сотрудникам можно присвоить признак «Работа в Big Data в удаленном режиме».

Теоретически ИТ-инфраструктура не должна позволять получать удаленный доступ к критическим данным. Однако, в условиях, когда в ИТ-инфраструктуре сотни тысяч или миллионы объектов управления (серверов, сетевого оборудования и т. д.) и в сутки происходит сотни или тысячи изменений в ИТ-инфраструктуре, то случаются ошибки. Поэтому разумно встроить независимый проверяющий механизм в контур Big Data (Рис. 8), который объективно показывает кто их сотрудников просматривал критичные данные в удаленном режиме.

Итак:

Рассмотренные в этом параграфе методы выявления инсайдеров являются простейшими примерами задач поиска и анализа причинности (идентификации каузальных оснований) [3]. В самом деле, мы предполагаем наличие появления события C как причину появления больших трат пользователя, торговца ценной информацией. То есть предполагаем, что продажа информации о событиях C влечет возможность осуществления больших трат. Поскольку сделанное предположение может оказаться ложным, то необходимо найти критерии фильтрации ложных каузальных оснований. Наоборот, должны быть подтверждения правильности предполагаемых каузальных оснований. Для решения этих задач подходят вероятностно-статистические методы [91]. Однако применения этих методов явно отличаются от традиционных методов математической статистики [92]. Основная идея применения вероятностных методов – это построение событий, вероятности которых либо равны 0, либо при выбранных соотношениях параметров асимптотически стремятся к 0. При равенстве 0 мы имеем дело с запретами вероятностных мер в конечных вероятностных пространствах [93], при стремлении вероятностей к 0 имеем дело с состоятельными процедурами статистических решений [94] (чаще всего в схеме серий, что отличает эти подходы от классической математической статистики). Но такие подходы соответствуют задаче фильтрации кандидатов на каузальные основания. Возникает вопрос об истинности предположений о рассматриваемых вероятностных мерах и точности решений. Здесь также видим отличия традиционных задач математической статистики от фильтрации каузальных оснований [81, 95]. Эти отличия состоят в избыточности поиска каузальных оснований, которое состоит в том, что как правило можно рассматривать несколько информационных пространств [2] и использовать для анализа различные каузальные основания различных исходных характеристик.

Выводы по Главе 3

1. Исследованы подходы к выявлению враждебных инсайдеров организации, использующих сговор. Проблема выявления организованной группы нарушителей

информационной безопасности является одной из самых сложных задач обеспечения безопасности организации.

Исходное множество данных для анализа состоит из множества малых выборок, описывающих функционал информационных технологий организации. Это множество можно считать большими данными. Для сокращения объема исходных данных использован метод кластеризации. Это позволило эффективно использовать методы математической статистики, т. е. выявить малые выборки, несущие информацию о признаках инсайдеров. Сложность задачи заключалась в том, чтобы как можно меньше потерять искомых малых выборок. Найдены условия, когда в схеме серий вероятность выявления признаков инсайдеров, использующих сговор, стремится к 1.

2. Исследованы возможности использования различных подходов к описанию диагностики действий инсайдеров при анализе больших эмпирических данных. В задачах этого типа необходимо установить (спрогнозировать, диагностировать, и др.) наличие или отсутствие целевых свойств у каких-либо пользователей из заданного множества.

Оценка правильности правдоподобных рассуждений проверяется на основе оценок вероятностей случайного появления найденных закономерностей в простейших вероятностных моделях.

Рассмотренные в статье примеры показывают, при каких соотношениях параметров возможно эффективное выявление причинно-следственных связей между событиями, с помощью которых можно выявлять инсайдеров.

Указаны два способа управления соотношениями между параметрами, позволяющие получать содержательную информацию. Первый способ основан на разделении периода наблюдений на промежутки, в течение которых искомая связь может проявиться. Второй способ связан со способами сокращения множества пользователей, которые потенциально могут стать инсайдерами. Т. е. речь идет о формировании кластеров, в которых вероятностные оценки становятся работоспособными. Искомые соотношения между параметрами для поиска зависимостей можно определять с помощью предельных теорем в схеме серий.

Глава 4. Сбор и анализ информации из различных источников в условиях Big Data

Потребность в разработке средств «навигации» в Big Data очевидным образом ассоциируется с потребностью в разработке эффективных поисковых технологий, которые позволяли бы результативно обрабатывать большие объемы не только собственно исходных «сырых» данных, но и формируемой на их основе аналитики – витрин данных, графической информации, dashboard'ов и др. Критичными здесь оказываются потребности в обработке больших объемов постоянно изменяющейся, то есть пополняемой новыми сведениями информации, оставаясь при этом в рамках жестких временных ограничений анализа данных и поддержки принятия решений. Обычно поисковые сервисы предоставляются их пользователям в рамках трех «архитектурных» моделей:

- облачные услуги (SaaS),
- готовое ПО, включающее в себя библиотеки, пользовательские интерфейсы и т. д.,
- библиотеки «базовых» программных инструментов, реализующие основные функции поисковых систем.

Примеры широко используемых коммерческих корпоративных поисковых систем – это, в частности, Algolia[1], IBM Watson Discovery[2], Yext[3], Swiftype[4], SearchUnify[5] и др. Популярные корпоративные поисковые системы с открытым исходным кодом – это, в частности, ElasticSearch[6], Solr[7], Sphinx[8] и др.. Среди отечественных

[1] The flexible AI-powered Search & Discovery platform - https://www.algolia.com
[2] IBM Watson Discovery - https://www.ibm.com/cloud/watson-discovery
[3] Power your website with the world's best search - https://www.yext.com
[4] A powerful search experience for your website—without the learning curve - https://swiftype.com
[5] SearchUnify for Sales & Customer Service - https://www.searchunify.com
[6] ELASTIC: Search more, spend less - https://www.elastic.co
[7] Solr is the popular, blazing-fast, open source enterprise search platform built on Apache Lucene™ - https://lucene.apache.org/solr/
[8] Introduction to Search with SPHINX - http://sphinxsearch.com

коммерческих корпоративных поисковых систем, по-видимому, наиболее известными являются Спутник (Ростелеком)[1] и 1С[2].

Критически значимыми характеристиками корпоративных поисковых систем считаются:

- скорость индексирования первичных данных (быстрота переработки поисковиком входных «сырых» данных для занесения в свой внутренний поисковый аппарат – системы поисковых индексов, классификаторы и т.п. Обычно этот параметр оценивается в мегабайтах чистого\«сырого» входного текста в секунду);
- скорость переиндексации, то есть реконструкции поискового инструментария - обновления индексов или создание новых с приходом новой входной информации. При этом может поддерживаться как инкрементальное индексирование, так и полная перестройка – переформирование – индекса, могут использоваться и дополнительные индексы – в т. ч. так называемый дельта-индекс, в который включается только новая информация;
- поддерживаемые API (поисковое ядро необходимо связывать с приложениями: у приложений могут быть библиотеки, работающие с API поисковика);
- взаимосвязь размеров базы и скорости поиска, так как некоторые поисковики перестают отвечать на запросы при индексах, содержащих более 50 млн записей;
- поддерживаемые типы входных документов (возможности индексации различных типов источников - СУБД, файловых хранилищ и т. п.).

Обычно такие характеристики обеспечиваются использованием специализированных алгоритмов поддержки поиска. Примерам таких инструментов обработки данных могут, в частности, быть средства обеспечения

- ранжирования и сортировки результатов поиска;

[1] Корпоративный поиск «Спутник» - https://www.sputnik.ru/searchbox
[2] Архитектура платформы 1С-Предприятие: глобальный поиск. - https://v8.1c.ru/platforma/globalnyy-poisk/

- лексической нормализации, в том число – *стемминга*[1], удаления стоп-слов[2], удаление «шума»[3], *лемматизации*[4] и др.;
- так называемого *бустинга,* то есть сервисов смыслового анализа текстов – выделения ключевых слов, словосочетаний и фраз; идентификации релевантных определенной теме сущностей – персон, объектов, понятий и отношений.

Среди наиболее известных библиотек, реализующих поисковые технологии:

- Apache Lucene (ElasticSearch, Solr, MongoDB Atlas Search, Datafari, CrateDB)[5]
- Apache Lucy[6]
- FTS, Tsearch2, RUM, GIN, OpenFTS, GIST (Postgres)[7]
- Sphinx/Manticore[8]
- Indri (Lemur)[9]
- Fulltext (MySQL)[10]
- Terrier[11]
- Manatee[12]
- iSearch Library (ArangoSearch)[13]
- Lunar[14]
- Xapian[15]

Очевидно, что инструменты поддержки поиска позволяют отслеживать и анализировать действия и намерения пользователя, которые, например, можно

---

[1] Stemming - усечение слов (например “trouble“, “troubled“, “troubling“ конвертируется в “trouble”)
[2] Удаление таких слов, как, например, “I”, “me”, “etc”, “must”, “you” и т.п.
[3] Удаление функционально схожих со стоп-словами технических слов (например, “#”, “tag”, “url” и т.п.)
[4] Lemmatization – ситуации, когда усеченное алгоритмом стемминга слово восстанавливают до формы естественного языка (например, “troubl” = “trouble” и т.п.)
[5] Welcome to Apache Lucene - https://lucene.apache.org
[6] The Apache Software Foundation - https://lucy.apache.org
[7] Do you need a Full-Text Search in PostgreSQL? - https://www.postgresql.eu/events/pgconfeu2018/sessions/session/2116/slides/137/pgconf.eu-2018-fts.pdf
[8] Open-source database for search applications - https://manticoresearch.com
[9] INDRI: Language modeling meets inference networks - https://www.lemurproject.org/indri/
[10] MySQL: Full-Text Search Functions - https://dev.mysql.com/doc/refman/8.0/en/fulltext-search.html
[11] Welcome to the Terrier IR Platform - http://terrier.org
[12] NLP-Center: NoSketch Engine - https://nlp.fi.muni.cz/trac/noske
[13] ArangoDB: Powerful Search Included - https://www.arangodb.com/full-text-search-engine/
[14] LUNR: Search made simple - https://lunrjs.com
[15] Xapian: Open Source Search Engine Library - https://xapian.org

сопоставлять с различными сведениями управленческого характера – текущими с целями и задачами его производственной деятельности, должностными полномочиями по доступу к тем или иным информационным ресурсам. К сожалению, на текущий момент рынок не предлагает надежных коммерческих систем поддержки поиска и идентификации признаков вредоносной инсайдерской активности, способных обеспечить результативное применение в крупных отечественных финансовых структурах.

4.1 ИТ-среда анализа данных и поддержки принятия решений

Разработаны требования к системно-техническому решению и архитектуре системы поиска признаков вредоносной инсайдерской активности. Комплекс информационной системы (Рис. 9*)* — это комплекс накапливания и обработки Big Data, охватывающий примерно 2000 серверов. Данный комплекс Big Data (в силу специфики решаемых на его базе бизнес-задач) является динамической структурой, конфигурация которой варьируется в части увеличения или же, наоборот, уменьшения примерно на 10 серверов ежесуточно. В этом комплексе одновременно работают несколько тысяч специализированных сотрудников-аналитиков данных, в задачи которых входит обеспечение полного цикла анализа данных и принятия соответствующих решений (подготовка данных, разработка моделей и т. д.). Требуемых для их работы (так называемых *полезных*) данных - порядка десятка петабайт при общем объеме накапливаемой и обрабатываемой информации в несколько раз больше. *Полезные* данные представлены как более чем сотня баз данных и несколько сотен аналитических продуктов (витрин данных).

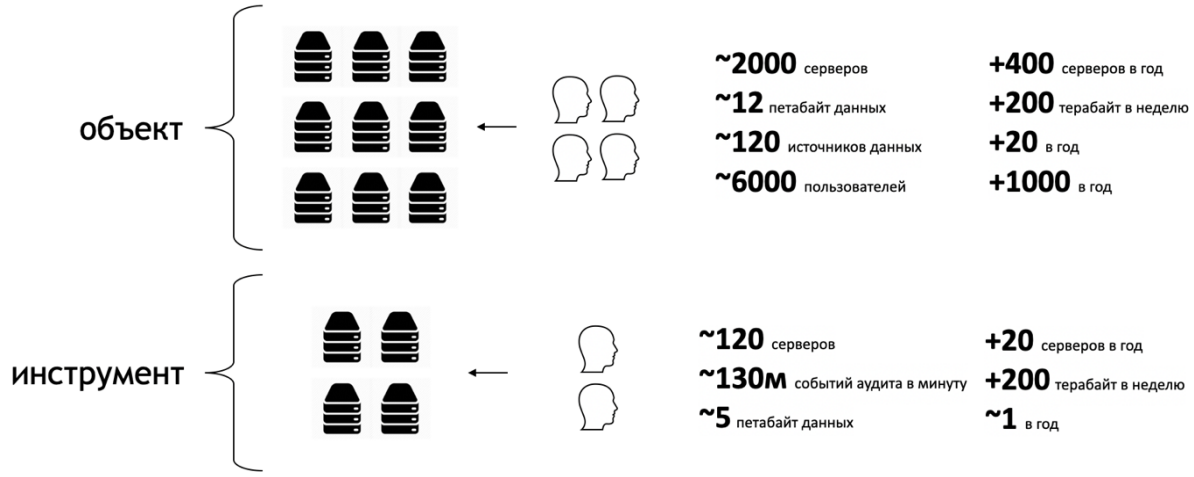

Рис.9. Объект поиска признаков инсайдеров (основные характеристики объекта).

Разработанные методика и программный инструментарий выявления признаков инсайдера — это комплекс меньшего, однако также вполне внушительного размера (примерно 6% от общего парка серверов Big Data). Данный инструментальный комплекс защиты порождает несколько десятков терабайт так называемых «сырых» данных в сутки, которые необходимо обрабатывать, фильтровать, приводить к нормализованному виду, индексировать, чтобы постоянно обеспечивать требуемые ограничения процессно-реального времени для обновления\актуализации параметров системы мониторинга и защиты основного комплекса Big Data.

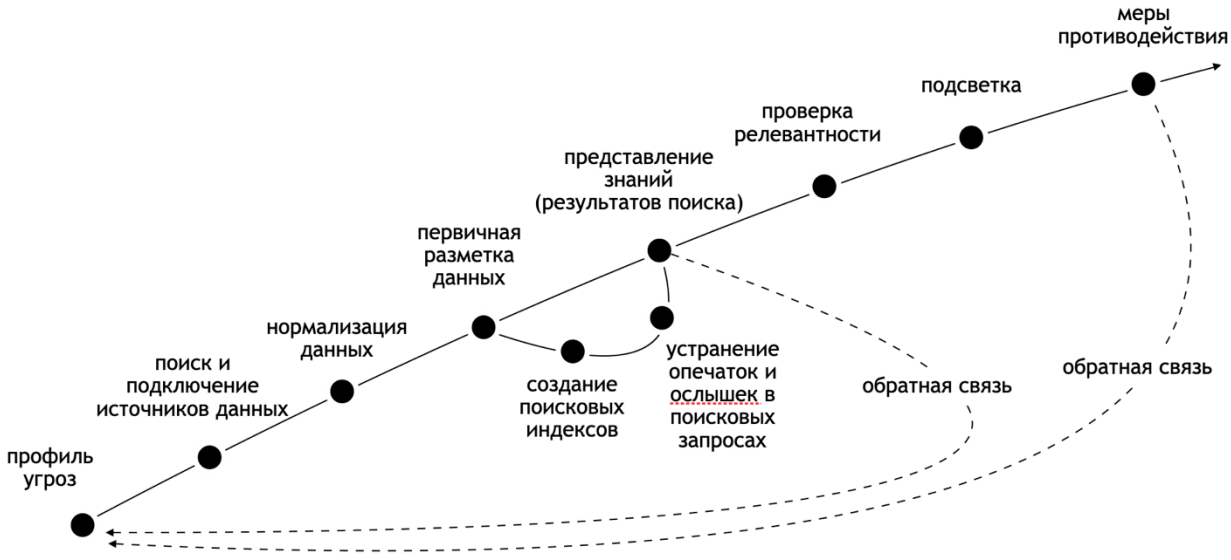

Рис.10. Основные этапа жизненного цикла обработки данных

Как объект управления разработанный комплекс инструментов защиты характеризуется специфичным жизненным циклом (ЖЦ) анализа данных и поддержки принятия управленческих решений. На разных этапах этого ЖЦ разные типы данных обрабатываются соответствующими проблемно-ориентированными инструментами (Рис 10). Особенности «шагов» ЖЦ отражаются в соответствующих требованиях к обрабатываемым данным и задействованным программным средствам на каждом из таких этапов.

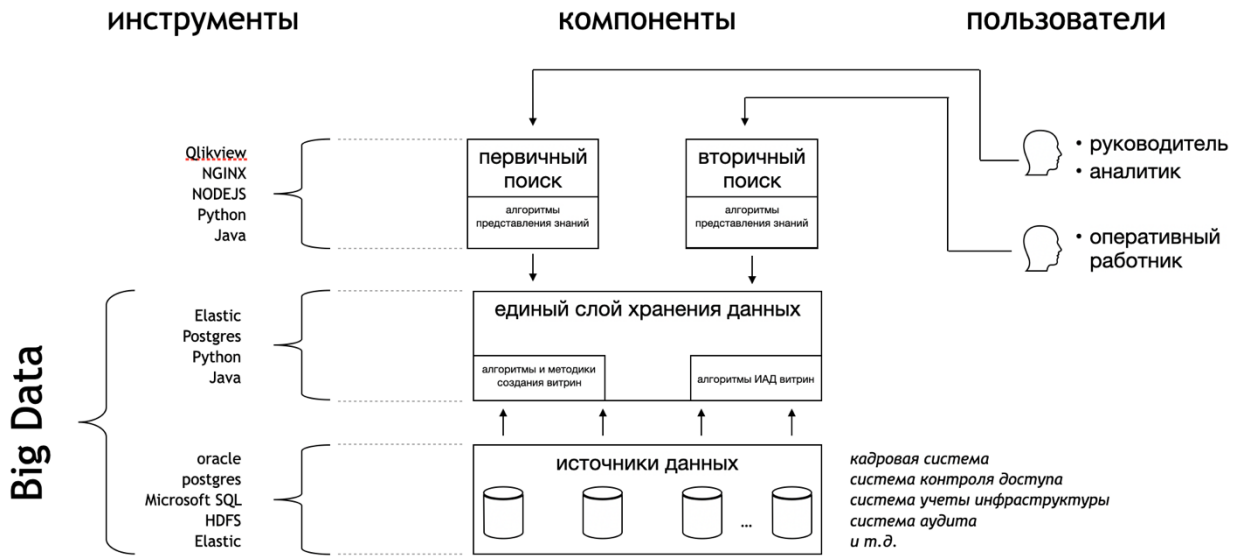

Рис.11. Архитектура комплекса средств анализа данных мониторинга

Представляемый комплекс средств защиты от инсайдеров имеет следующую архитектуру (см. Рис.11). Базовые компоненты Комплекса — это единый слой хранения (Big Data) и аналитическое ядро, содержащее программно-технические инструменты анализа данных и поддержки принятия решений. Использованные виды программных инструментов - типовые. Это - базы данных, серверы приложений, среды исполнения программного кода - Python, Java. Созданный инструментарий защиты имеет два различных типа интерфейсов, ориентированных на задачи формализованного представления знаний для так называемых первичного и вторичный поиска. Первичный поиск характеризуется классом задач, где аналитик выявляет в исходных «сырых» данных те, которые релевантны отдельно накапливаемым знаниям из текущего профиля угроз, аккумулирующего в себе уже накопленный опыт экспертов служб безопасности об

особенностях наблюдавшихся ранее инсайдерских активностей. В первичном поиске определяются необходимые характеристики в «сырых» данных (поля, идентификаторы и т.п.), а далее с использованием специальных средств машинного обучения, опираясь на прецеденты ранее идентифицированных инсайдерских угроз, выделяются из текущих «сырых» данных вся та информация, которая далее будет использована для поддержки текущей работы служб безопасности (мониторинга основного комплекса Big Data на предмет идентификации в его текущей операционной работе тех или иных аномальных активностей а также организации противодействия противоправным действиям инсайдерского характера) и обеспечения оперативной отчетности.

Вторичный поиск обеспечивает оперативное отображение и ответы на релевантные целям мониторинга безопасности запросы (это - своего рода «локальный Яндекс»). Здесь оперативный сотрудник может ввести необходимые идентификационные данные (ФИО, табельный номер или источник данных и др.) и посмотреть детальный профиль соответствующего сотрудника или подразделения, например, штатный профиль доступов данного сотрудника к ресурсам защищаемого комплекса Big Data в соотнесении с текущими характеристиками, полученными в результате работы алгоритмов машинного обучения.

4.2 Проблемы, потребовавшие решения при разработке системы защиты от инсайдерских действий

При разработке обсуждаемого подхода был идентифицирован ряд типичных для работы с Big Data барьеров (см. Рис. 12), для преодоления которых пришлось разрабатывать проблемно-ориентированные результативные решения. Такие барьеры можно объединить в следующие четыре однородные группы:

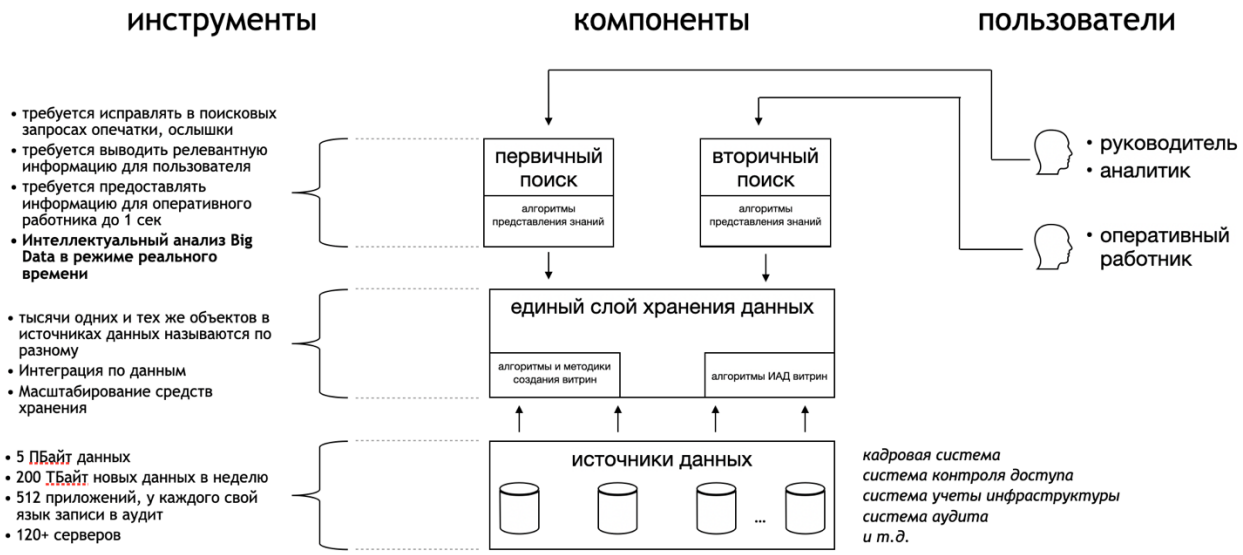

Рис.12. Процедурные барьеры

1. Ограничения по времени анализа данных, постоянно пополняемые большие данные (эффекты *Big* и *Open)*.
2. Интеграция данных, извлекаемых из различных источников. Интеграция данных, отбираемых из различных источников «сырых» первичных данных представляет собою нетривиальную задачу: необходимо в режиме ограниченного времени отбирать релевантную целям мониторинга информацию из огромного перечня объектов (ресурсов), характеризуемых своими собственными типами представления данных (именами полей и доменов, именами и значениями атрибутов и т. п.). Для преодоления таких барьеров был предложен и реализован в виде программных инструментов ряд проблемно-ориентированных эвристик, отражающих зарекомендовавшую себя на практике «логику» оперирования с разнородными данными – «склеивания» согласуемых данных, используемую профильными экспертами службы безопасности при поиске инсайдерских активностей.
3. Нормализация обрабатываемых данных и сокращение объемов перечней объектов-примитивов за счет элиминации объектов-дубликатов. Так, например, пользователи сервисов вторичного поиска при работе со средствами диалогового интерфейса допускают различного рода неточности и\или ошибки в именовании искомого объекта. Именно это обстоятельство

потребовало разработки соответствующих средств автоматической идентификации и коррекции ошибок (клавиатурных ошибок, опечаток, «ослышек» и т. п.).

4. Ресурсные ограничения. Характеристики разрабатываемого комплекса должны удовлетворять следующему ограничению. Стоимость инструментального комплекса защиты от инсайдерских активностей не должна превышать 10% процентов от стоимости собственно всего комплекса информационной системы.

Для достижения необходимых характеристик были решены следующие технические проблемы.

1. При обработке исходных «сырых» данных на первом этапе их фильтрации «стартовые» несколько десятков терабайт характеристик анализируемых событий удалось «ужать» до 600 Гб (1,5 млрд записей об анализируемых активностях).

2. Далее на втором этапе фильтрации данных эти 600 Гб «ужали» до 2 гигабайт (3 млн записей об активностях).

3. При этом удалось добиться, чтобы обеспечивающие вторичный поиск индексы обновлялись в режиме имеющихся ограничений процессно-реального времени, а время отклика на запрос не превышало 10 сек. (на выделенном для этого программно-техническом комплексе).

Алгоритм фильтрации лог-файлов.

Разработан алгоритм фильтрации логов. Логика работы алгоритма, следующая:

1) Каждое приложение комплекса Big Data при его запуске настраивается вручную записывать события в централизованном хранилище (Elastic Search). В централизованном хранилище лог-файлов логи фильтруются по типу источника данных HDFS (поле «source» равно «/var/log/hadoop-hdfs/hdfs-audit.log»).

   Поскольку любые приложения в Big Data обращаются к данным через слой HDFS, то таким фильтром гарантируется, что ни одно обращение пользователя к данным не пропускается. Таким образом после первого

этапа фильтрации логов, **десятки терабайт** логов ужимаются до **сотен гигабайт,** что тоже «много». Первый этап фильтрации логов «грубый» и фильтрует только файлы и не анализирует их содержимое. В этой связи был создан второй этап.

2) На втором этапе фильтрации логов анализируется содержимое каждого оставшегося лог-файла отдельно. Цель второго этапа фильтрации – оставить в HDFS логах только факты обращения к данным и удалить все остальное, например технологическую информацию. Гарантия того, что алгоритм не удалит ни один из фактов обращения к данным достигается за счет административно-технической меры, которая обязывает все файлы, хранимые в HDFS размещать в корневом каталоге «data». Отсюда фильтр - src=«/data» AND NOT «part-» AND «cmd=getfileinfo». Второй этап сокращает сотни гигабайт до десятков гигабайт.
3) Третий этап фильтрации логов удаляет дубли внутри лог-файлов. Дубли возникают из-за особенностей работы HDFS и особенностей работы пользователей. Например, пользователь обратился к данным, а HDFS записал в лог несколько одинаковых или очень похожих строк, но отличающихся меткой времени в несколько миллисекунд. Или пользователь два раза выполнил одно и тоже или похожее обращение к данным. Поэтому необходимо удалить повторы. Повторы удаляются следующий образом – все обращения пользователя в интервале 10 минут с одинаковой учетной записью (login) и репликой данных считаются как одно обращение. Интервал 10 минут выбран экспериментально т. к. позволяется превращаться тысячи технологических записей в одну.
4) Четвертый этап аналогичный третьему, только анализируется кортеж (“учетная запись”,” реплика”, “таблица в реплике”, “метка времени”).

В результате работы алгоритма фильтрации логов получается, что на каждом этапе фильтрации количество строк в лог-файлах уменьшается на порядок (Таб. 1),

тем самым уменьшается объем поиска. Причем алгоритм фильтрации не отбрасывает ни одной полезной строки.

Дата	**Кол-во событий в логах**	**1 ЭТАП ОБРАБОТКИ:** Кол-во hdfs-audit.log логов (строк)	**2 ЭТАП ОБРАБОТКИ:** Кол-во логово после применения фильтров "src=/data" AND "cmd=getfileinfo" (elastic filter) (строк)	**3 ЭТАП ОБРАБОТКИ:** Кол-во логов после принменения фильтра по тексту "part-" (elastic filter) (строк)	**4 ЭТАП ОБРАБОТКИ:** Кол-во событий в логах
12.04.2021	7 500 000 000	1 400 000 000	560 000 000	280 000 000	4 140 000
13.04.2021	7 500 000 000	1 400 000 000	560 000 000	280 000 000	644 390
14.04.2021	7 500 000 000	1 400 000 000	560 000 000	280 000 000	1 121 000
15.04.2021	7 500 000 000	1 450 000 000	580 000 000	290 000 000	1 370 000
16.04.2021	7 500 000 000	1 450 000 000	580 000 000	290 000 000	599 000
19.04.2021	7 600 000 000	1 450 000 000	580 000 000	290 000 000	6 410 000
20.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	2 460 000
21.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 340 000
22.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 470 000
23.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 280 000
24.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	803 300
25.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	2 360 000
26.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 950 000
27.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 540 000
28.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 530 000
29.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	882 800
30.04.2021	7 600 000 000	1 500 000 000	600 000 000	300 000 000	1 280 000
10.05.2021	7 700 000 000	1 600 000 000	640 000 000	320 000 000	6 610 000
11.05.2021	7 700 000 000	1 600 000 000	640 000 000	320 000 000	2 840 000
12.05.2021	7 700 000 000	1 600 000 000	640 000 000	320 000 000	423 300

Таб.1. Эффект алгоритма фильтрации данных (отфильтрованы избыточные данные)

Разное количество отфильтрованных строк на 4 этапе объясняется неравномерностью работы сотрудников в комплексе Big Data.

Одной из критически значимых целей такой фильтрации данных было сокращение количества актуальных для поисковой обработки записей. Так, например, известно, что используемое в целом ряде задач поиска промышленное ПО Elastic Search перестает в штатном режиме отвечать на запросы при размерах индекса более 50 млн записей.

4.3 Методика анализа данных и поддержки принятия решений

Разработка методики идентификации признаков инсайдерской активности стартует с формирования актуальной модели угроз. Модель угроз формализуется в виде Профиля Угроз (ПУ), представляющего собою постоянно поддерживаемый в актуальном состоянии перечень так называемых Типовых Сценариев (ТС). Каждый из Типовых Сценариев порождается обобщением опыта оперативных

сотрудников, вовлеченных в расследования конкретных случаев мошенничества (инсайдерских активностей). Опыт оперативных сотрудников сперва фиксируется в виде текстового описания (см. Рис. 13), которое далее преобразуется в машиночитаемый формализованный вид.

- Сотрудник выполнил “точечный” запрос к базе, в которой 100 млн записей
- Сотрудник, работающий в одном подразделении, имеет 80% доступ к данным другого подразделения
- Сотрудник, работающий с данными, не посещает офис более 1 дня в неделю
- Сотрудник, имеющей те же доступы, что и его коллеги, физически размещается в другом офисе
- Сотрудник, имеющей одновременно доступ в аналитическую систему и транзакционную систему.

Рис. 13. Текстовые описания типовых сценариев

При этом задействовано промежуточное представление знаний о каждом из Типовых Сценариев в виде фрейма (см. пример на Рис. 14). Для описания данных в слотах подобных фреймов предусмотрены иерархии типов данных от булевских значений признаков – Да\Нет, до графов параметров и отношений между такими параметрами с пометками на вершинах и ребрах, а также текстовых комментариев, например, в виде Binary Large Objects - BLOB.

Подобные иерархии типов данных могут быть задействованы в случае необходимости получения более тонкой «дифференциации» состояний НОРМА\АНОМАЛИЯ использованием более детального представления знаний об анализируемых инцидентах. Простейший вариант представления знаний в Типовых Сценариях Профиля Угроз – использование булевских значений Да\Нет, позволяющих описать каждый такой фрейм в виде множества, характеризующих именно его признаков. В свою очередь множество всех используемых при описании текущего ПУ признаков определяет битовую строку, соответствующими единицами которой кодируется каждый из соответствующих Типовых Сценариев в Профиле Угроз (см. пример на Рис.14.). Обработка машиночитаемого описания фреймов, представленных в виде именно битовых строк, дает возможность

получить существенный выигрыш в производительности при анализе текущих данных (т. к. позволяет организовать сравнение текущей ситуации с описаниями ТС средствами одной вычислительной макрооперации).

Рис. 14. Текстовые знаний об угрозах в виде фреймов

Текущий (актуальный на данный момент) Профиль Угроз – динамически изменяемая во времени (пополняемая с учетом постоянно накапливаемого опыта оперативных действий – Рис. 16) конструкция. В таком ПУ могут находиться десятки или сотни Типовых Сценариев. Ниже представлены некоторые примеры текстовых версий ТС:

- Сотрудник X выполнил “точечный” запрос к базе Y, в которой 100 млн записей
- Сотрудник Z, работающий в одном подразделении X, имеет 80% доступ к данным другого подразделения Y.
- Сотрудник X, работающий с данными, не посещает офис более 1 дня в неделю
- Сотрудник X, имеющей те же доступы, что и его коллеги из офиса Y, физически размещается в другом офисе Z.
- Сотрудник X, имеющей одновременно доступ в *аналитическую* систему Y и *транзакционную* систему Z.

Пример 16. Ранее были описаны примеры, когда у аналитика имеется доступ к данным, полученный в обход установленного порядка. Но ниже описан немного другой пример (Рис. 15) - аналитик данных реально просматривает данные, доступ к которым получен в обход установленного порядка. Это является почти гарантированным признаком, т. е. вызовет срабатывание ТС.

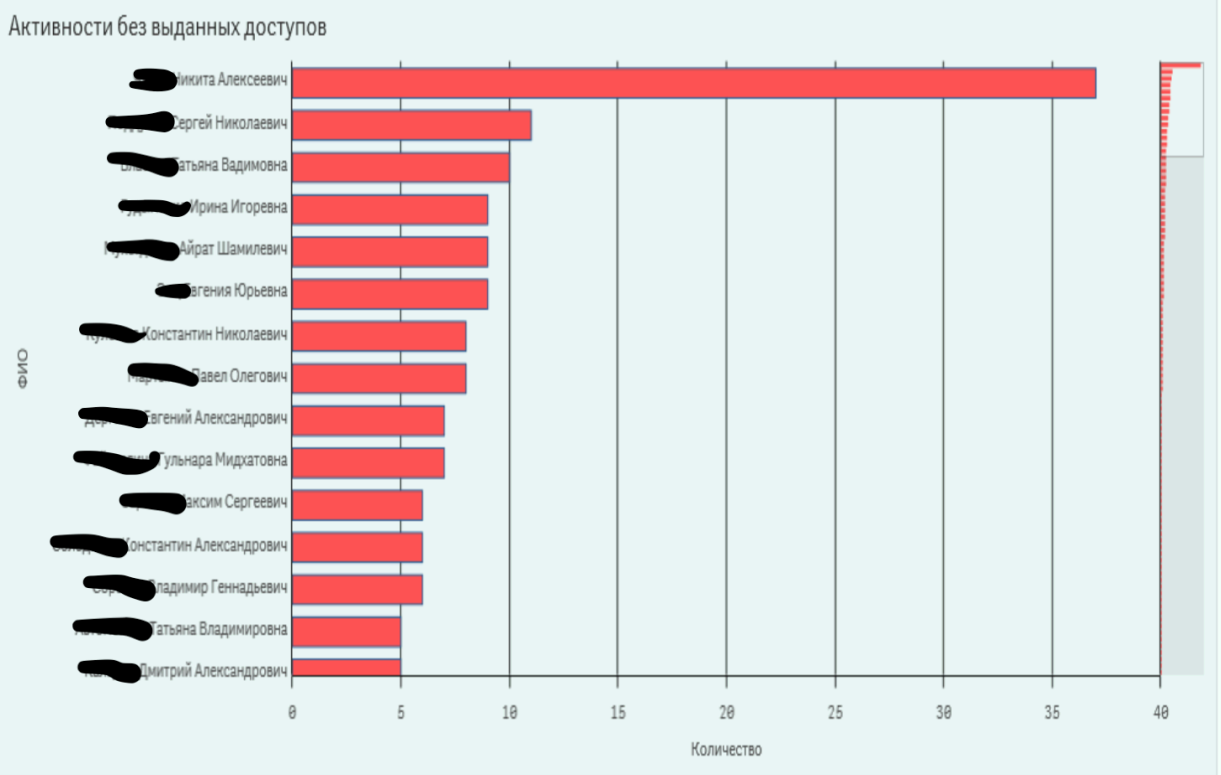

Рис.15. Факты обращения пользователя к данным без оформленного доступа к данным (период 1 месяц).

В формализованном описании текущий ПУ может быть описан (см. Рис.16) как матрица, строки которой (при использовании булевского варианта представления знаний от ТС в соответствующих фреймах) соответствуют задействованным при описании угроз параметрам/признакам, а каждый из столбцов этой матрицы представляет соответствующий ТС.

динамически изменяемая конструкция

	признаки	угрозы								
		ТС1	ТС2	ТС3	ТС4	ТС5	ТС6	ТС7	ТС8	ТС9	...
1	пол мужской	1	1	0	1	0	1	1	1	0	...
2	возраст менее 30 лет	1	0	1	1	0	0	1	0	1	...
3	время работы в организации менее 6 месяцев	1	0	0	1	0	0	1	0	0	...
4	признак смены должности	0	0	0	0	1	1	0	0	0	...
5	кол-во доступов к данным больше 20	1	0	0	0	0	1	1	0	0	...
6	% доступов к данным других подразделений больше 80	1	0	0	0	0	0	1	0	0	...
7	адрес офис не совпадает с адресом офиса коллег	0	0	0	1	0	1	0	0	0	...
8	80% переписки с коллегами другого подразделения	1	0	0	0	0	0	1	0	0	...
9	4 дня в месяц не посещает офисы компании	0	0	1	0	0	0	0	0	1	...
10	имеется доступ к данным двух несвязанных друг с другом подразделений	1	0	0	0	0	1	1	0	0	...
11	имеется доступ хотя бы к одной активной системе	0	0	0	0	1	1	0	0	0	...
12	ранее работал во фронт-офисе	1	0	1	0	0	1	1	0	0	...
13	сотрудник делает более 30 точечных SQL-запросов в сутки	0	1	0	0	0	0	0	1	0	...
14	объем файлов в личном каталоге хранилища больше 10Gb	1	1	0	0	0	1	1	1	0	...
...	...	...	...	...	...	...	...	...	...	...	...

расширение

битовая строка

Рис.16. Формализация описания Профиля Угроз

Сравнивая элементы (ячейки) этой матрицы с характеристиками (профилем значений признаков) доступа к защищаемым ресурсам комплекса Dig Data, актуальными в данный момент для конкретного мониторимого сотрудника, можно оценить весомость угроз несанкционированных активностей этого сотрудника – релевантность его текущего поведения каким-либо известным угрозам из текущего ПУ[1] (Рис. 17). Однако, проведение таких сравнений «лобовым» методом «грубой силы» оказывается чрезвычайно ресурсоемким (см. выше Раздел «*ИТ-среда* …»). Таким образом, востребованными оказываются любые результативные приемы, подходы и методы сокращения объемов перебора при формировании «диагностических» заключений по каждому из мониторимых сотрудников.

[1] В т.ч. - использование статистических средств анализа рисков при идентификации аномалий ([25-29] и др.)

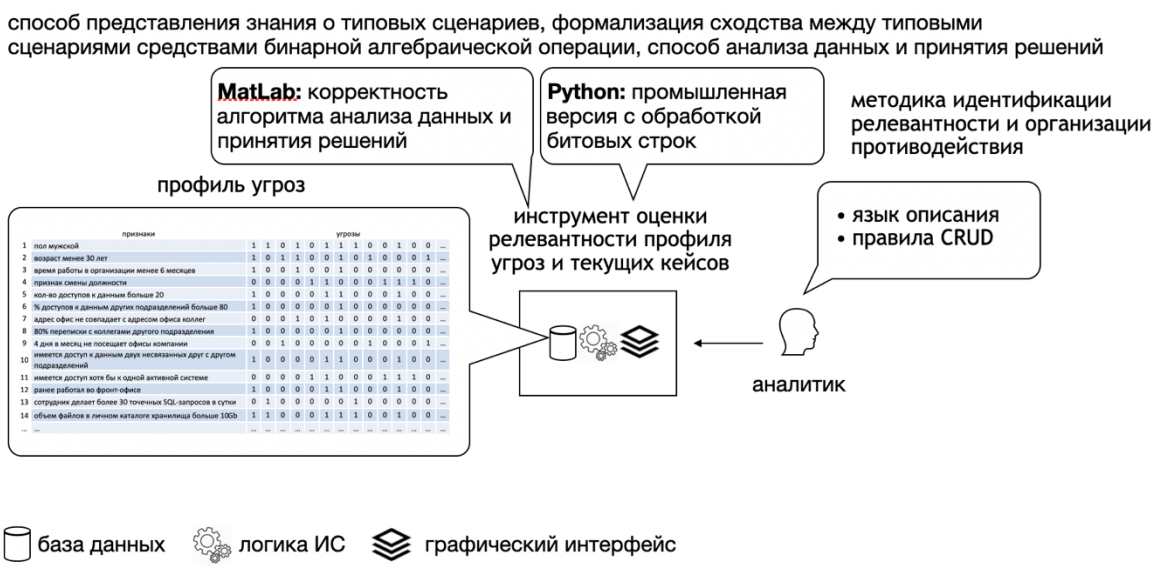

Рис. 17. Задачи анализ Профиля Угроз

Наряду с уже представленными выше инструментами нормализации и фильтрации исходных «сырых» данных существенный выигрыш в объемах необходимых вычислений позволяет получить процедурное уточнение идеи (эвристики) учета сходств в описаниях ТС. Действительно, при оценке релевантности текущего профиля доступов конкретного сотрудника к защищаемым информационным ресурсам представляется вполне естественным начать такие проверки с наиболее общих для всех актуальных ТС множеств признаков, переходя далее ко все менее и менее общим, завершая весь процесс сравнением с собственно каждым из имеющихся ТС.

Говоря формально (см., в частности, работы [96] и др.), определив бинарную алгебраическую операцию сходства описаний ТС [96, 97 и др.] можно построить диаграмму взаимной вложимости множеств признаков, задействованных в описаниях ТС. А далее (один раз сформировав такую диаграмму) проверять релевантность текущего анализируемого профиля доступов конкретного сотрудника имеющемуся Профилю Угроз, начиная со сравнения его элементов с элементами нижнего «этажа» (минимальных по вложению подмножеств признаков, одновременно актуальных для нескольких ТС) и далее двигаясь лишь по релевантным цепочкам частичного порядка этой диаграммы к ее верхнему «этажу» (подмножеств максимальных по числу актуальных общих признаков), а от него – к релевантным данной ситуации описаниям ТС (см. Рис.18).

Формализованное описание переборных задач, возникающих при формировании диаграммы сходств описаний Типовых Сценариев текущего Профиля Угроз представлено ниже.

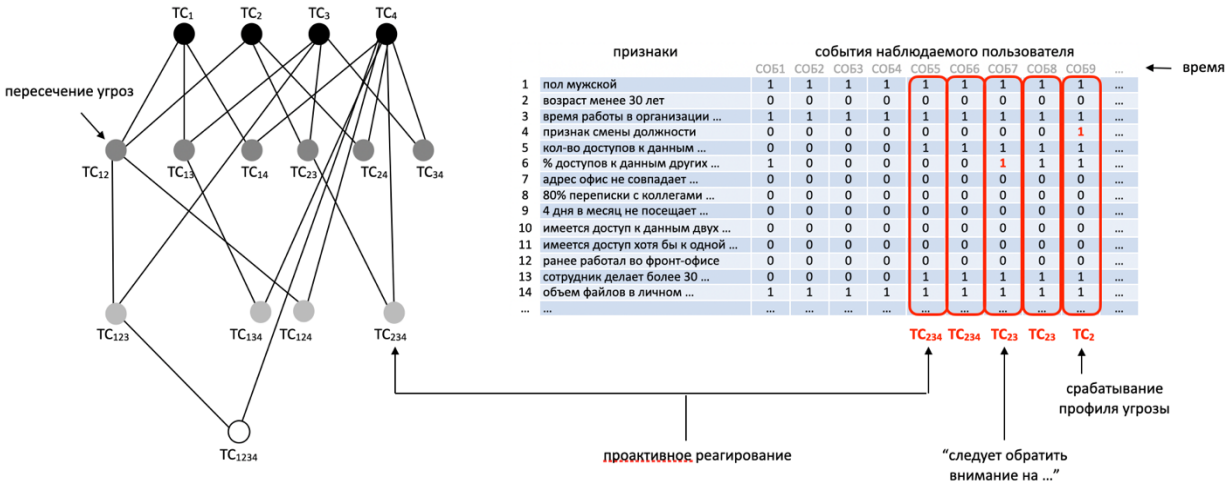

Рис.18. Диаграмма сходств Типовых Сценариев из Профиля Угроз

Подобная тактика первоочередного использования наиболее общих для имеющихся описаний ТС множеств признаков и последующего движения лишь вдоль актуальных цепочек частичного порядка в один раз построенной диаграмме позволяет не только существенным образом сократить необходимые объемы вычислений при проверке релевантности текущей профиля доступов конкретного сотрудника и актуального ПУ, но и в *проактивном* режиме подсказать офицеру безопасности в текущем конкретном случае наиболее опасные варианты дальнейшего развития событий («подсвечивая» соответствующие цепочки частичного порядка на диаграмме сходств описаний ТС, двигаясь с ее нижнего «этажа» вверх к релевантным этому конкретному профилю доступов описаниям Типовым Сценариям).

В случае булевского представления данных об имеющихся ТС каждый такой Типовой Сценарий характеризуется как битовая строка. Таким образом получаем возможность вычислять сходства описаний ТС, используя стандартные для многих современных системных программных сред макро-операции с битовыми строками. Это позволяет работать с имеющимися Big Data достаточно быстро и эффективно (формируя результат сходства описаний ТС средствами соответствующей машинной макро-операции).

Идею оценки релевантности текущего профиля доступов конкретного сотрудника Типовым Сценариям актуального Профиля Угроз иллюстрирует Рис. 19. Для выполнения такой оценки достаточно выявления общих частей описаний объекта мониторинга и Типовых Сценариев (в том числе - с учетом ранее рассчитанных риск-индикаторов идентификации аномалий на наличие признаков инсайдерской активности – см. подробнее работы [61, 2, 3, 4, 5] и др.).

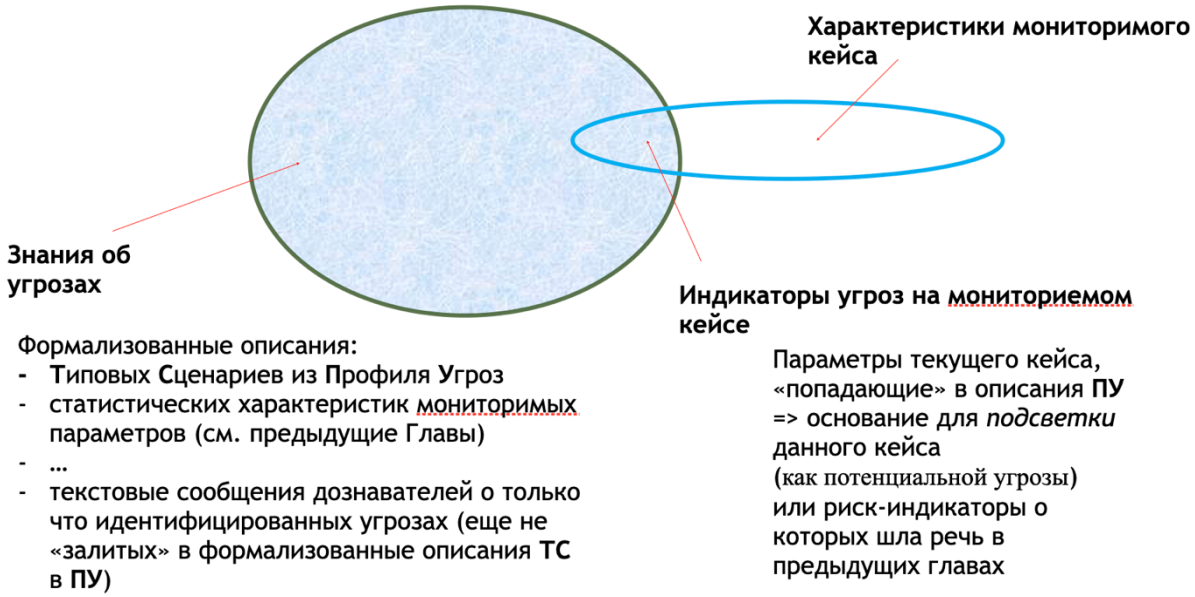

Рис.19. Отношение релевантности *текущая ситуация ~ Профиль Угроз*

Особого внимания требует учет того обстоятельства, что Профиль Угроз — это динамически изменяемая конструкция, которая может быть в любое время модифицирована аналитиками, обрабатывающими накапливаемый опыт идентификации и противодействия вредоносным активностям. При этом следует учитывать, что управление перебором (уход от сравнения «всего» в описании анализируемого профиля доступов со «всем» в ПУ при обсуждаемой оценке их релевантности) в рассматриваемом контексте Big Data оказывается неприемлемо ресурсоемкой тактикой анализа данных и поддержки принятия решений. А ведь при этом необходимо провести *исчерпывающий* анализ совпадений фрагментов текущего профиля доступов каждого конкретного сотрудника по всем ТС актуального Профиля Угроз, что требует обработки данных о нескольких миллионах пользователей в сутки.

В результате данной работы был реализован анализ нескольких Профилей Угроз. Например, в одном из реализованных Профилей Угроз требовалось выявлять технологические учетные записи (TUZ), под которыми работают не процессы, а сотрудники (PUZ), тем самым скрывая свои следы обращения к данным (“маскировщики”) – см. Рис. 20.

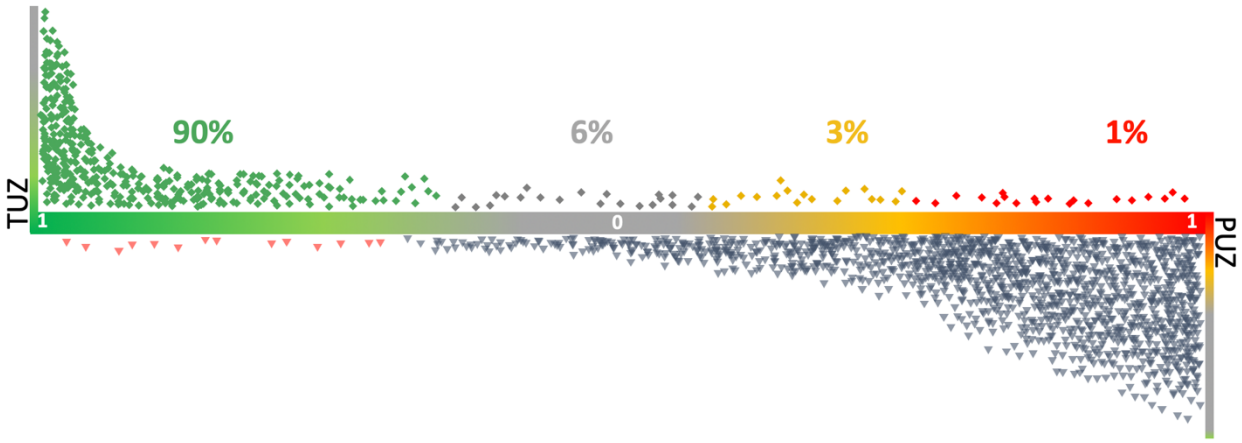

Рис.20. Результаты работы Профиля Угроз по выявлению технологических учетных записей, которые используют сотрудники.

Для создания данного Профиля Угроз были использованы 1316 признаков (часть признаков приведены в Таб. 2), характеризующую учетную запись. В результате анализа Профиля Угроз были обнаружены 102 технологические учетные записи (TUZ), которая использовались сотрудниками для доступа к данным, а не процессами.

Atribute Name (Фича)	Role(Роль)	Type (Тип данных)
CNT_numSources	Feature	Integer
kadmin	Feature	Integer
TTL_actsPerHour	Feature	Integer
mean_actsPerHour	Feature	Integer
hdfs	Feature	Integer
Max*(daysActDiff_NEXT)	Feature	Integer
hive	Feature	Integer
CNT_dateActs	Feature	Integer
Max*(daysActDiff_PREV)	Feature	Integer
host_7_IPS	Feature	Integer
lst_actHour	Feature	Integer
host_8_IPS	Feature	Integer
fst_actHour	Feature	Integer
host_9_IPS	Feature	Integer
CNT_actsHour	Feature	Integer
host_10_IPS	Feature	Integer
yarn	Feature	Integer

Таб. 2. Основные признаки Профиля Угроз «маскировщики».

Итак, в предлагаемом подходе (при использовании сходств ТС из актуального ПУ при оценке опасности действий конкретного сотрудника при его доступе к

защищаемым от инсайдерских активностей информационным ресурсам) можно существенным образом оптимизировать (по сравнению с тактикой «грубой силы», предусматривающей сравнения «всего» со «всем») объемы необходимых вычислений. Для этого следует один раз сформировать диаграмму сходств и последовательно вести проверки ее пересечений ее фрагментов с теми «релевантными» анализируемому профилю доступов конкретного сотрудника элементами диаграммы сходств ТС, которые размещены на ее цепочках частичного порядка. В ситуации, когда речь идет о миллиардах событий и о сотнях ТС, таким способом можно сформировать значительный выигрыш в скорости принятия финальных решений. Необходимость подобной оптимизации мотивируется достаточно естественным образом: угроз защищаемому комплексу Big Data со временем становится все больше, и это требует эффективного управления имеющимися вычислительными ресурсами.

Дополнительный аргумент в пользу предлагаемого подхода – возможность организовать *проактивный* мониторинг негативного развития «аномальных» ситуаций, подсказывая конкретному сотруднику безопасности наиболее опасные варианты изменения отслеживаемой им конкретной ситуации (вдоль релевантных ей цепочек частичного порядка на диаграмме сходств Типовых Сценариев).

4.4 Программный инструментарий реализации предложенной методики

Как уже отмечалось выше, Профиль Угроз — это динамически изменяемая конструкция, предполагающая возможность модификации в соответствии со вновь накапливаемыми эмпирическими данными о поведении объектов мониторинга, а также опытом противодействия (как успешного, так и нерезультативного) идентифицированным вредоносным активностям. Поддержка изменений в «архитектуре» актуального ПУ потребовала разработки соответствующих программных инструментов экономного реинжиниринга структуры диаграммы сходств описаний ТС. Показано (см., например, [97] и др.), что при порождении диаграммы сходств ТС в общем случае приходится иметь

дело с объектом, размеры которого растут экспоненциально быстро при линейном росте размеров множества ТС. Таким образом, актуальной оказалась задача оптимизации перебора вариантов (локальных сходств описаний ТС) при формировании диаграммы сходств ТС. Для этого был разработан специальный программный инструмент со встроенными алгоритмом экономной организации генерации локальных сходств описаний ТС. Следуя подходу Rapid Application Development сперва в инструментальной среде Matlab, был проведена отладка и проверка корректности этого алгоритма анализа данных и принятия решений, а далее на Python была реализована его промышленная версия (использующая возможности экономной обработки битовых строк).

Вместе со специально разработанным проблемно-ориентированным графическим редактором (обеспечивающим аналитикам возможности формировать новые ТС и поддерживать ПУ в актуальном состоянии) эта промышленная Python-версия генератора диаграммы сходств описаний ТС образует ядро программного инструментария представления и обработки знаний о вредоносных инсайдерских активностях.

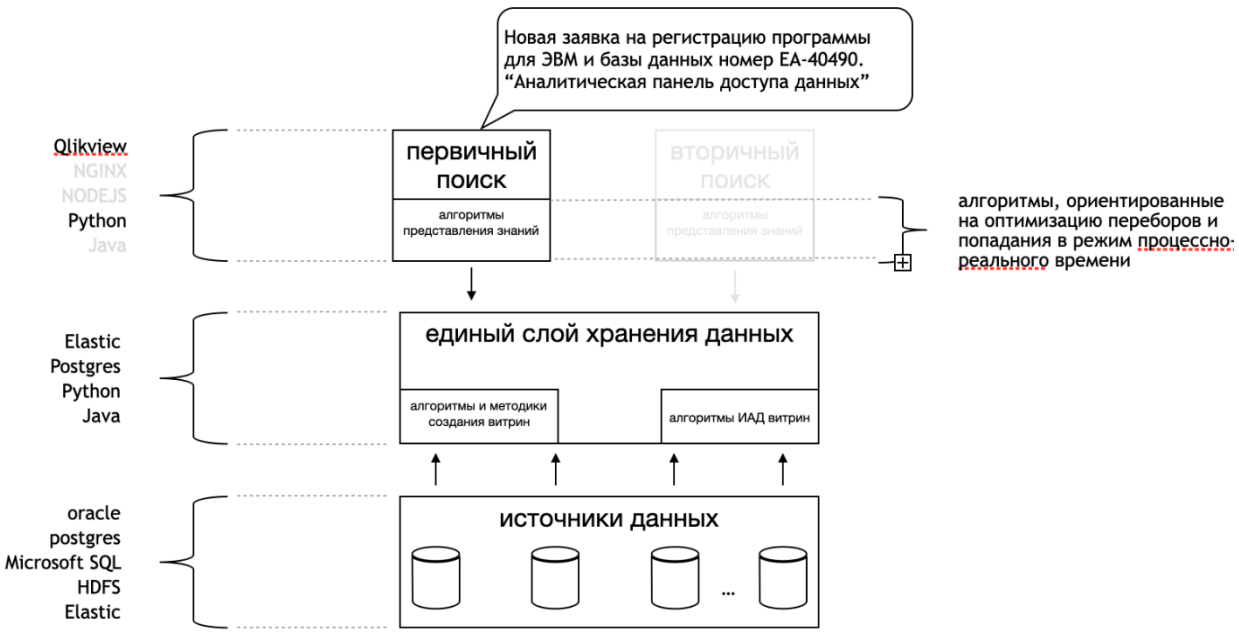

Рис. 21. Первичный поиск в «сырых» данных.

Архитектура разработанной системы мониторинга и противодействия инсайдерским угрозам предусматривает реализацию двух типов поиска – первичного (выделение релевантной информации из первичных «сырых» данных

– см. Рис. 21) и вторичного (быстрый поиск в уже отобранных релевантных данных для подготовки управленческой отчетности, а также для информационного сопровождения оперативной деятельности сотрудников службы безопасности – Рис. 22).

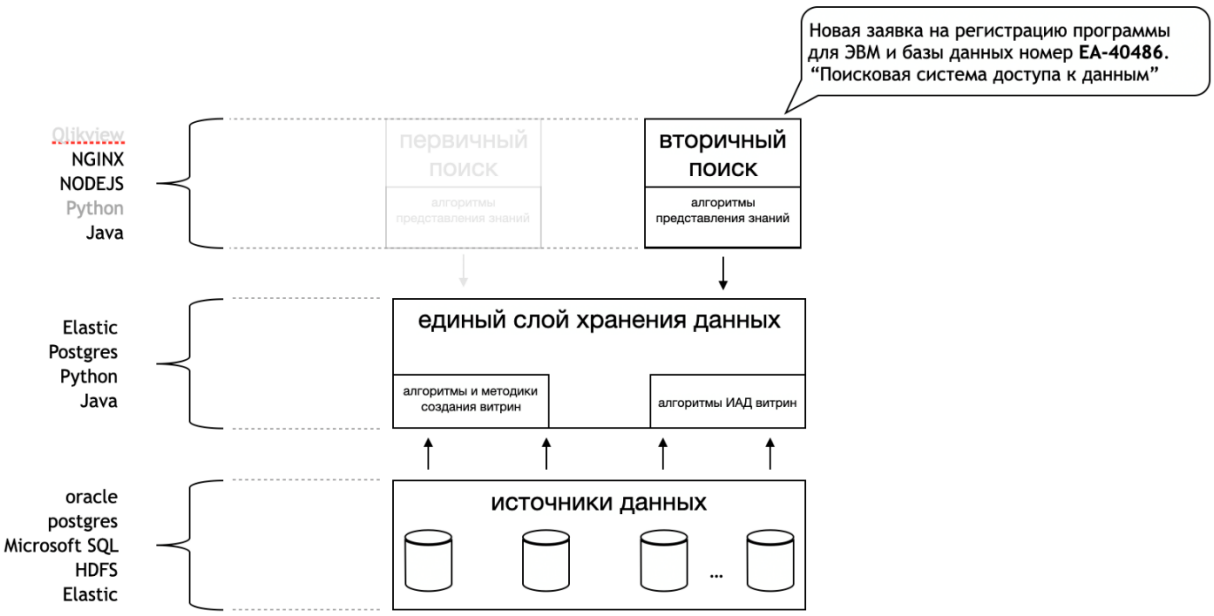

Рис. 22. Вторичный поиск.

Для поддержки этих двух классов информационных сервисов разработаны соответствующие пользовательские интерфейсы. Так в первичном поиске интерфейс помогает аналитику отбирать все те сведения, что должны быть «подсвечены» в последующей работе как релевантные знания об угрозах. Исходные данные для машинного обучения (как описания прецедентов, формируемые на базе анализа инцидентов безопасности, проанализированных экспертами) вводятся в систему первичного поиска через соответствующие интерфейсы. Сюда “удобным” образом подключены нормализованные данные из единого слоя хранения и имеются алгоритмы, ориентированные на оптимизацию перебора вариантов (которая необходима для соблюдения ограничений режима процессно-реального времени анализа данных и поддержки принятия соответствующих управленческих решений).

Вторичный поиск обеспечивается отдельной поисковой системой, пользовательский интерфейс которой поддерживает работу с текстовым полем для ввода запросов и кнопкой “искать”. Цель вторичного поиска - оперативно

предоставлять информацию, включающую результаты работы алгоритмов машинного обучения, в простом и понятном виде для сотрудников, не имеющих продвинутых ИТ-навыков. Важнейший эффект, обеспечиваемый средствами вторичного поиска, — это ускорение работы оперативных сотрудников (не являются техническими экспертами), занятых мониторингом и противодействием вредоносным инсайдерским активностям.

4.5 Программный инструментарий нормализации данных

Особого внимания заслуживают возможности специально разработанных лингвистических программных сервисов, которые в автоматическом режиме поддерживают процесс нормализации анализируемых данных (Рис. 23).

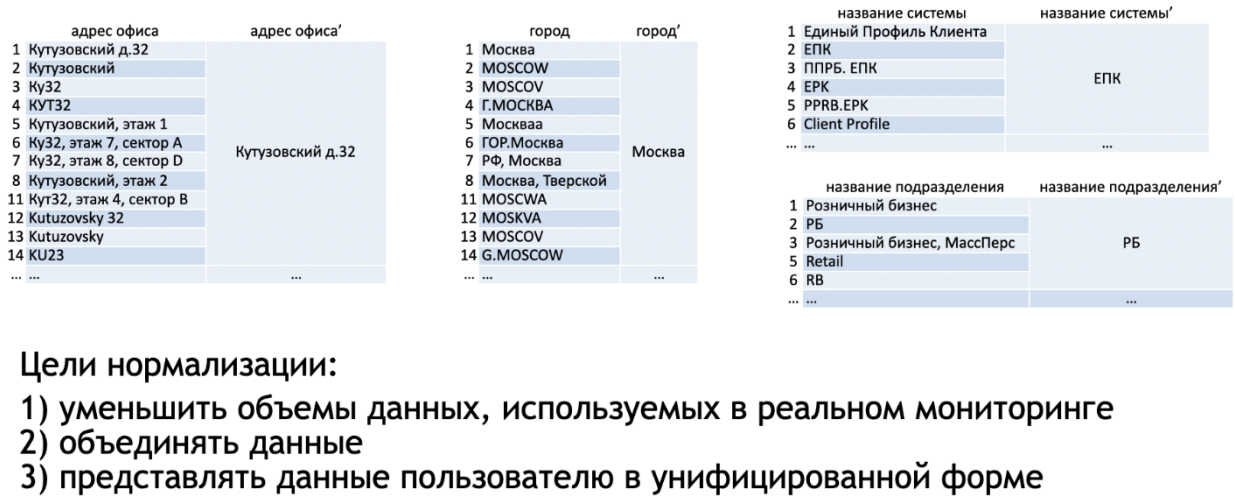

Рис. 23. Проблема нормализации данных.

Так в различных источниках поля и значения таких полей данных, как правило, называются различным образом. Так, в частности, одно и тоже наименование места или города в различных базах данных может иметь разные называния (см., например: город Москва может иметь десятки различных написаний, в т.ч. - “MOSCOW”, “G. MOSKVA”, “MOSKVA”, “ГОРОД МОСКВА”, “МОСКВА, МОСКОВСКАЯ ОБЛАСТЬ” и т.д.). Именно по этой причине анализируемые данные необходимо нормализовать. При этом проводимая нормализация преследует три базовые цели:

1. уменьшить объемы данных, используемых в реальном мониторинге

2. объединять данные из разных источников,
3. представлять данные пользователю в унифицированной форме.

Результаты работы алгоритмов нормализации данных:

Например, в одной из 30 разработанных витрин – ненормализованных названий доступов к данным – 167000 (уникальных – 3400), а нормализованных (официальных) названий доступов – уже 114000 (уникальных – 212). Нормализация в среднем снизила потребность в месте хранения на 40% (114000/167000).

Ниже приведен пример работы алгоритма нормализации данных, где показаны как различные наименования объекта доступа преобразуются в единое наименование.

Ненормализованный вид	Нормализованный вид
g_dc_d_internal_saphcm_bkp_ro	SAP HCM
g_dc_d_internal_saphcm_delta_ro	SAP HCM
g_dc_d_internal_saphcm_dv_ro	SAP HCM
g_dc_d_internal_saphcm_qa_ro	SAP HCM
g_dc_d_internal_saphcm_src_ro	SAP HCM
g_dc_d_internal_saphcm_stg_ro	SAP HCM
…	…

Таб. 3. Пример исполнения нормализации данных.

В инфраструктуру вторичного поиска встроены специально разработанные программные инструменты, реализующие алгоритмы корректировки опечаток и «ослышек» (Рис. 24). Исправление «ослышек» необходимо, например, в ситуациях, когда оперативный работник узнал фамилию из сообщения в телефоне и не знает, как именно пишется эта фамилия. Он вбивает в строку поиска то, что услышал, и алгоритм корректирует результаты ввода. Для решения таких задач коррекции был разработан собственный фонетический алгоритм[1], реализующий два этапа: фонетическое редуцирование и механизм (правила) так называемого

[1] Собственный алгоритм основан на алгоритме Metaphone. Алгоритм Metaphone допускал ошибки корректировке имен и фамилий, поэтому его пришлось изменить в части стандартных окончаний.

оглушения (Рис. 25). Для устранения опечаток ввода запроса были использованы обыкновенные триграммы (алгоритм измерения дистанции между эталонным названием и опечаткой - Рис. 26). Тестирование разработанных программных инструментов эмпирическим путем подтвердило корректность работы алгоритмов на базе в несколько сотен тысяч сотрудников.

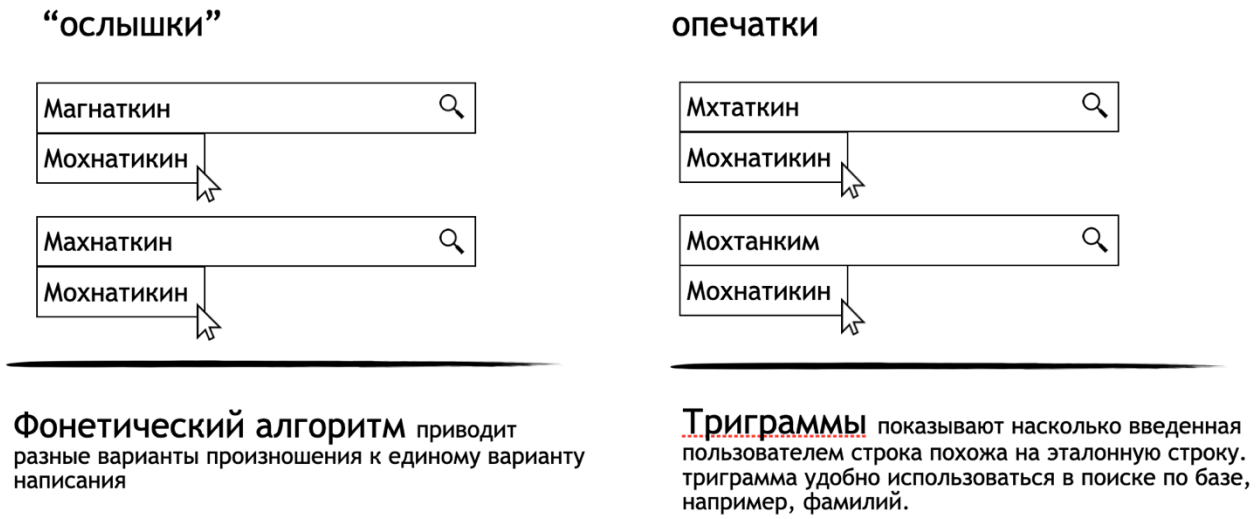

Рис. 24. Проблема опечаток и «ослышек».

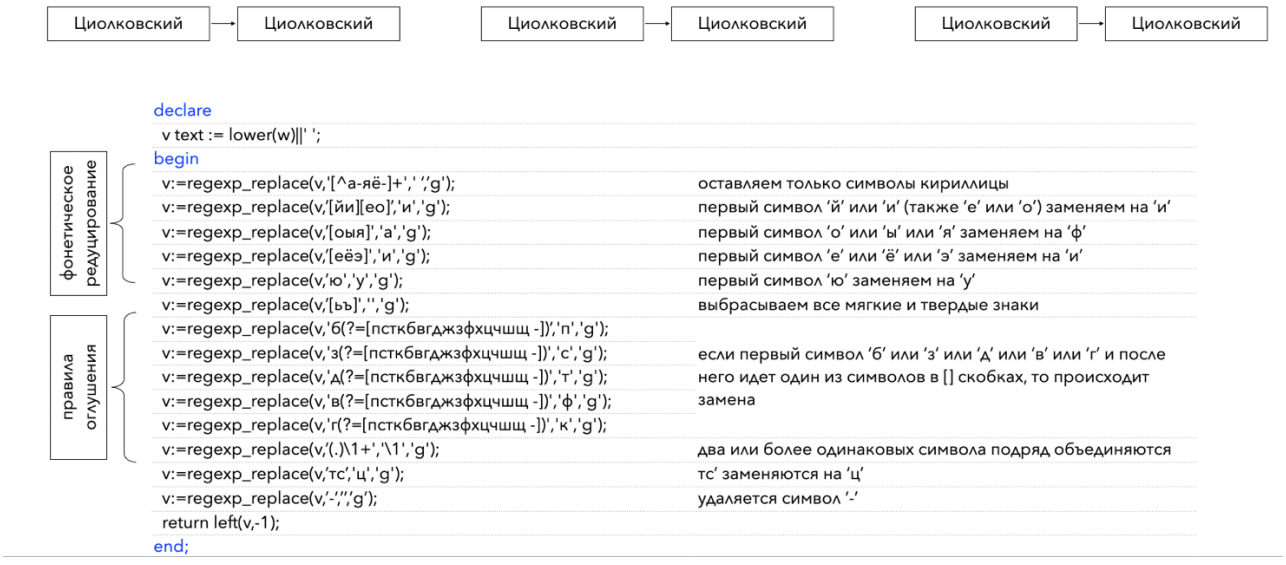

Рис. 25. Алгоритмы обработки опечаток и «ослышек».

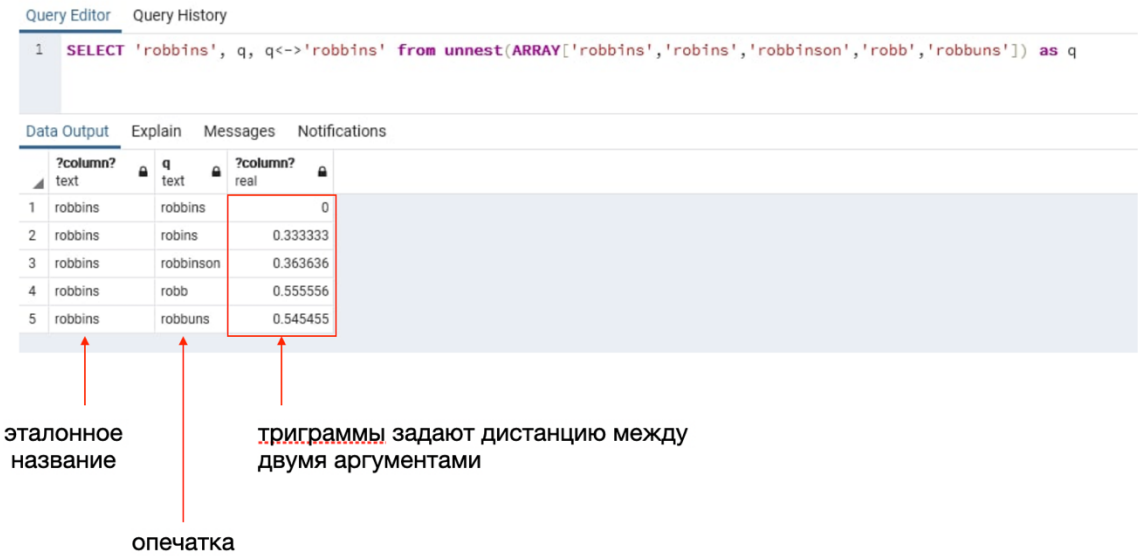

Рис. 26. Алгоритм сравнения триграмм

Были разработаны регламентыи также программные инструменты поддержки изменений и развития обсуждаемой системы защиты от вредоносных инсайдерских активностей:

- средства для поддержки реорганизации (расширения и модификации) Профиля Угроз с учетом динамически накапливаемого опыта. Инструментальные средства поддержки таких реорганизаций (методики\регламенты, программные инструменты анализа данных и визуализации результатов);
- средства для поддержки реорганизации (расширения и модификации) поискового аппарата (поисковых индексов, классификационных систем и т. п.) для поддержания эффективности вторичного поиска в динамически изменяемой информационной среде.

4.6 Основные результаты Главы 4

Основные результаты представленных исследований и разработок можно суммировать следующим образом:

1. Разработана методика поиска и ИАД релевантных заданной цели поиска признаков инсайдера среди сотрудников в условиях данных мониторинга с

эффектами «Big» и «Open» в условиях жестких ограничений времени для поддержки принятия решений. Сформировано научное обоснование[7] корректности и эффективности задействованных при реализации этого программного комплекса математических моделей и алгоритмов интеллектуального анализа больших данных. Состав методики:

1.1. Методы и алгоритмы:

1.1.1. ИАД, обеспечивающие эффективный первичный поиск в "сырых" исходных данных

1.1.2. механизмы представления знаний об угрозах;

1.1.3. механизмы ускорения вычислений поиска в текущих данных фрагментов, которые релевантны компонентам описания профиля угроз - “подсветка", например, диаграмма сходств типовых сценариев угроз;

1.1.4. механизмы управления ресурсами при мониторинге "подсвеченных" ситуаций, например, направленное продвижение по цепочкам частичного порядка в диаграмме сходств от наиболее "простых" комбинаций признаков\параметров "вверх" к описаниям собственно ТС);

1.1.5. механизмы статистического анализа для поиска аномалий в поведении объектов мониторинга.

1.2. Методы оценки качества:

1.2.1. Методы оценки качества\надежности формируемых статистическими средствами заключений о классификации аномалий в поведении объектов мониторинга, дающие дополнительные основания для принятия решений о приоритетности отработки соответствующих "подсвеченных" ситуаций;

1.3. Рекомендации:

1.3.1. рекомендации для дознавателей СБ по использованию соответствующих инструментов ИАД в их профильной деятельности;

2. Разработан программный комплекс, реализующий методику. Состав программного комплекса:

2.1. разработан набор сервисных программных инструментов, поддерживающих нормализацию данных как в первичном, так и во вторичном поиске;

2.2. разработаны оригинальные программные инструменты формирования и реконструкции диаграммы сходств ТС;

2.3. разработаны проблемно-ориентированные средства имитационного моделирования для оценки ряда эффектов и поддержки принятия управленческих решений;

2.4. предложен вариант интеграции вновь разработанных программных инструментов ИАД с уже имеющимися в организации промышленными программными инструментами обработки данных;

3. Проведена демонстрация работоспособности программного комплекса (апробация):

3.1. получено (экспериментальным путем - в процессе апробации программных инструментов ИАД[1] в деятельности крупной коммерческой организации) подтверждение работоспособности и результативности разработанных методики и реализующего ее программного комплекса

3.2. Получено свидетельство о государственной регистрации программы для ЭВМ № 2021614494 «Аналитическая панель доступов к данным», дата государственной регистрации 25.03.2021.

3.3. Свидетельство о государственной регистрации программы для ЭВМ № 2021613506 «Поисковая система доступа к данным», дата государственной регистрации 19.04.2021.

4.7 Выводы и рекомендации

1. Анализ Big Data методами «brute-force» - бесперспективная задача,
2. В задачах анализа Big Data ключевым достижением являются алгоритмы нормализации и фильтрации данных. Как только решены задачи нормализации и фильтрации данных, создание целевых алгоритмов

[1] И это зафиксировано соответствующими свидетельствами о регистрации РИД.

становится относительно «простой» задачей, выполняемой на структурированных и понятных данных меньшего объема.

3. Из алгоритмов нормализации и фильтрации данных, алгоритмы фильтрации наиболее сложные т. к. обрабатывают большие потоки гетерогенных данных.

Основные результаты диссертации

1. Определены условия, при которых возможен поиск вкраплений признаков враждебного инсайдера в Big Data.
2. Разработаны и применены методы анализа гетерогенных данных. Ранние работы анализировали один тип данных.
3. Определены условия, при которых возможно применять методы математической статистики при анализе Big Data.
4. Разработан метод работы с противоречиями при выявлении аномалии в поведении сотрудников, позволяющий подтвердить или опровергнуть выявленную аномалию.
5. Разработан метод, позволяющий определять является ли аномалия в поведении сотрудников случайным событием или закономерностью.
6. Создано системно-техническое решение (методика, программная реализация методики и обоснование), способное выявлять признаки враждебных действий сотрудников к комплексу Big Data, несмотря большие объемы данных и ограничение по времени.

СПИСОК ЛИТЕРАТУРЫ

1. Смирнов Д. В., Грушо А.А., Забежайло М.И., Тимонина Е. Е. Система сбора и анализа информации из различных источников в условиях Big Data // International Journal of Open Information Technologies, 2021. V. 9. № 4. Pp. 64-74. http://injoit.org/index.php/j1/article/view/1099 (ВАК- 05.13.19)
2. Грушо А.А., Забежайло М.И., Смирнов Д.В., Тимонина Е.Е. Модель множества информационных пространств в задаче поиска инсайдера // Информатика и ее применения, 2017, том 11, № 4, с. 65-69. (Scopus, ВАК - 05.13.19.)
3. Грушо А.А., Грушо Н.А., Забежайло М.И., Смирнов Д.В., Тимонина Е.Е. Параметризация в прикладных задачах поиска эмпирических причин // Информатика и ее применения, ИПИ РАН (М.), 2018, том 12, № 3, с. 62-66 (Scopus, ВАК -05.13.19.)
4. Грушо А.А., Забежайло М.И., Смирнов Д.В., Тимонина Е.Е., С.Я. Шоргин. Методы математической статистики в задаче поиска инсайдера // Информатика и ее применения, 2020. Т. 14. Вып. 3. С. 71-75 (Scopus, ВАК - 05.13.19.)
5. Грушо А.А., Забежайло М.И., Смирнов Д.В., Тимонина Е.Е. О вероятностных оценках достоверности эмпирических выводов // Информатика и ее применения, 2020. Т. 14. Вып. 4. С. 3-8. (Scopus, ВАК - 05.13.19.)
6. Смирнов Д. В., Об одной методике проблемно-ориентированного анализа Big Data в режи-ме процессно-реального времени // International Journal of Open Information Technologies, 2021. V. 9. № 4. Pp. 64-74. http://injoit.org/index.php/j1/article/view/1099 (ВАК- 05.13.19)

7. Chinchani R., Ha D., Iyer A., Ngo H.Q., and Upadhyaya S. Insider threat assessment: Model, analysis and tool // In Network Security. – Boston: Springer, 2010. – P. 143-174.
8. Garfinkel R., Gopal R., Goes P. Privacy protection of binary confidential data against deterministic, stochastic, and insider threat // Management Science – 2002. – Vol. 48(6). – P. 749-764.
9. Sinclair S., Smith S.W., Preventative directions for insider threat mitigation via access control // In Insider Attack and Cyber Security. – Springer, 2008. – P. 165-194.
10. Probst C.W., Hunker J., Bishop M., Gollmann D. Summary-Countering Insider Threats // In Countering Insider Threats (Dagstuhl Seminar). – Germany: Leibniz-Zentrum fuer Informatik, 2008.
11. Greitzer F.L., Frincke D.A., Zabriskie M. Social/ethical issues in predictive insider threat monitoring. Information Assurance and Security Ethics // In Complex Systems: Interdisciplinary Perspectives. – 2010. – P.132-161.
12. Bishop M., 2005. Position: Insider is relative // Workshop on New Security Paradigms – 2002. – ACM. – P. 77-78.
13. Bishop M., Engle S., Peisert S., Whalen S., Gates C. 2009a. Case studies of an insider framework // In Hawaii Int. Conference on System Sciences. – IEEE, 2009a – P. 1-10.
14. Predd J., Pfleeger S.L., Hunker J., Bulford C., Insiders behaving badly // IEEE Security & Privacy 6, 4 (2008), P.66-70. C. W. Probst, R. R. Hansen, and F. Nielson. 2006. Where can an insider attack? In Int. Workshop on Formal Aspects in Security and Trust. Springer. – P. 127-142.
15. Salem M.B., Hershkop S., Stolfo S. J. 2008. A Survey of Insider Attack Detection Research // In Insider Attack and Cyber Security. – US: Springer, 2009. – P. 69-90.
16. Magklaras G., Furnell S. Insider threat prediction tool: Evaluating the probability of IT misuse // Computers & Security. – 2002. – Vol. 21(1). – P. 62-73.

17. Jabbour G., Menascé D. Stopping the insider threat: the case for implementing autonomic defense mechanisms in computing systems // In Int. Conference of Information Security and Privacy. – 2009a.
18. Probst C. W., Hunker J. The risk of risk analysis and its relation to the economics of insider threats // In Economics of information security and privacy. – Springer, 2010. – P. 279-299.
19. Parveen P., Weger Z. R., Thuraisingham B., Hamlen K., Khan L. Supervised learning for insider threat detection using stream mining // In Int. Conference on Tools with Artificial Intelligence. – IEEE, 2011b. – P. 1032-1039.
20. Azaria A., Richardson A., Kraus S., Subrahmanian V.S. Behavioral Analysis of Insider Threat: A Survey and Bootstrapped Prediction // In Imbalanced Data. Transactions on Computational Social Systems. – 2014. – No.1(2). – P. 135-155.
21. Hunker J., Probst C.W. Insiders and Insider Threats: An Overview of Definitions and Mitigation Techniques // Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications. – 2011. – Vol. 2, No. 1. – P. 4-27.
22. Ophoff J., Jensen A., Sanderson-Smith J., Porter M., Johnston K. A Descriptive Literature Review and Classification of Insider Threat Research. Technical Report. – 2014.
23. Bertacchini M., Fierens P. A survey on masquerader detection approaches // In Congreso Iberoamericano de Seguridad Informática, Universidad de la Repúblicade Uruguay. – 2008. – P. 46-60.
24. Gheyas I.A., Abdallah A.E. Detection and prediction of insider threats to cybersecurity: a systematic literature review and meta-analysis // Big Data Analytics – 2016. – Vol.1, Iss.1.
25. Sanzgiri A., Dasgupta D. Classification of Insider Threat Detection Techniques // In Annual Cyber and Information Security Research Conference. – 2016. – ACM, 25.
26. Khan M.I., Foley S.N. Detecting anomalous behavior in DBMS logs. // In F. Cuppens, N. Cuppens, J.-L. Lanet, and A. Legay, editors, Risks and Security of Internet and Systems - 11th International Conference, CRiSIS 2016, Rosco,

France, September 5-7, 2016. – Revised Selected Papers. – Lecture Notes in Computer Science, – Springer, 2016. – Vol. 10158. – P.147-152.

27. Hussain S. R., Sallam A. M., Detanom E. B. Detecting anomalous database transactions by insiders // In Proceedings of the 5th ACM Conference on Data and Application Security and Privacy. – CODASPY ’15, New York, NY, USA, 2015. ACM. – P.25-35.
28. Sallam A., Fadolalkarim D., Bertino E., Xiao Q. Data and syntax centric anomaly detection for relational databases // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, – 2016. – Vol.6(6). – P. 231-239.
29. Mathew S., Petropoulos M., Ngo H.Q., Upadhyaya Sh. A data centric approach to insider attack detection in database systems // Proceedings of the 13th International Conference on Recent Advances in Intrusion Detection, RAID’10. – Berlin, Heidelberg: Springer-Verlag, 2010. – P.382-401
30. Alizadeh M., Peters S., Etalle S., Zannone N. Behavior analysis in the medical sector: Theory and practice // In Proceedings of the 33rd Annual ACM Symposium on Applied Computing, SAC ’18, New York, NY, USA, 2018. ACM, – P. 1637-1646.
31. Kul G., Luong D., Xie T., Coonan P., Chandola V., Kennedy O., Upadhyaya Sh. Ettu: Analyzing query intents in corporate databases // In Proceedings of the 25th International Conference Companion on World Wide Web, WWW’16 Companion, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee. – P. 463-466
32. Low W.L., Lee J., Teoh P. DIDAFIT: detecting intrusions in databases through fingerprinting transactions // ICEIS 2002, Proceedings of the 4st International Conference on Enterprise Information Systems. – CiudadReal, Spain, April 2-6, 2002. – P.121-128.
33. Lee S.Y., Low W.L., Wong P.Y. Learning fingerprints for a database intrusion detection system // In Proceedings of the 7th European Symposium on Research in Computer Security, ESORICS ’02, London, UK, 2002. – Springer Verlag, 2002. – P.264-280

34. Kamra A., Terzi E., Bertino E. Detecting anomalous access patterns in relational databases // The VLDB Journal. – 17(5):1063–1077. – August 2008.

35. Kemmererand R. A., Vigna G. Intrusion detection: a brief history and overview. Computer, 35(4):27–30. – Apr 2002.

36. Sallam A., Bertino E., Hussain S. R., Landers D., Lefler R. M., Steiner D. Dbsafe: an anomaly detection system to protect databases from exfiltration attempts // IEEE Systems Journal. – 2015. – PP (99):1–11

37. Chung C.Y., Gertz M., Levitt K.N. DEMIDS: A misuse detection system for database systems // In M. E. van Biene-Hershey and L. Strous, editors, Integrity and Internal Control in Information Systems, IFIP TC11 Working Group 11.5, Third Working Conference on Integrity and Internal Control in Information Systems: Strategic Views on the Need for Control. – Amsterdam, The Netherlands, November 18-19, 1999, Vol.165 of IFIP Conference Proceedings. – P.159-178. Kluwer.

38. Bertino E., Terzi E., Kamra A., Vakali A. Intrusion detection in rbac-administered databases // In 21st Annual Computer Security Applications Conference (ACSAC'05), Dec 2005, pages 10 pp.–182,.

39. Khan M. I., O'Sullivan B., Foley S.N. Towards modelling insider's behavior as rare behavior to detect malicious rdbms access // In 2018 IEEE International Conference on Big Data (Big Data). – 2018. – P. 3094-3099.

40. Khan M. I., O'Sullivan B., Foley S. N. A semantic approach to frequency-based anomaly detection of insider access in database management systems. // In N. Cuppens, F. Cuppens, J.-L. Lanet, A. Legay, and J. Garcia-Alfaro, editors, Risks and Security of Internet and Systems. – Cham: Springer International Publishing, 2018. – P.18-28.

41. Lazarevic A., Kumar V., Srivastava J. Intrusion Detection: A Survey. – US, Boston, MA: Springer, 2005. – P. 19-78.

42. Gama J., Žliobaitė I., Bifet A., Pechenizkiy M. & Bouchachia A. (2014) A survey on concept drift adaptation. ACM computing surveys (CSUR) 46, p. 44.

43. Salehi M., and Rashidi L. (2018) A survey on anomaly detection in evolving data. ACM SIGKDD Explorations Newsletter 20, – P. 13–23. – URL: https: //doi.org/10.1145/3229329.3229332.

44. Goldstein M., Uchida S. (2016) A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PloS one 11, p. e0152173. – URL: https://dx.doi.org/10.1371%2Fjournal.pone. 0152173.

45. Ertöz L., Eilertson E., Lazarevic A., Tan P.N., Kumar V., Srivastava J., Dokas P. Chapter 3 the minds-minnesota intrusion detection system.

46. Schölkopf B., Platt J.C., Shawe-Taylor J., Smola A.J. & Williamson R.C. (2001) Estimating the support of a high-dimensional distribution. Neural Computation 13, pp. 1443–1471. – URL: https://doi.org/10.1162/ 089976601750264965.

47. Liu F. T., Ting K. M., Zhou Z. H. (2012) Isolation-based anomaly detection. ACM Transactions on Knowledge Discovery from Data (TKDD) 6, P.3.

48. Doshi-Velez F., Kim B. (2017) Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

49. Das S., Islam M.R., Jayakodi N.K., Doppa J.R. (2019) Active anomaly detection via ensembles: Insights, algorithms, and interpretability. arXiv:1901.08930 [Online; accessed 20-Aug-2019].

50. Görnitz N., Kloft M., Rieck K. & Brefeld U. Toward supervised anomaly detection // Journal of Artificial Intelligence Research 46. – 2013. – P.235-262.

51. Ribeiro M.T., SinghS., Guestrin C. (2016) Model-agnostic interpretability of machine learning. arXiv preprint arXiv:1606.05386.

52. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead // Nature Machine Intelligence 1, – 2019. – P. 206-215.

53. Brdiczka O. et al. Proactive Insider Threat Detection through Graph Learning and Psychological Context // In Security and Privacy Workshops (SPW) – IEEE Symposium on. – 2012. – P.142-149.

54. Nurse J. R. C. et al. Understanding Insider Threat: A Framework for Characterising Attacks // In Security and Privacy Workshops (SPW). – IEEE, 2014. – P. 214-228.
55. Nance K., Marty R. Identifying and Visualizing the Malicious Insider Threat Using Bipartite Graphs // In System Sciences (HICSS), 2011 44th Hawaii International Conference on. – 2011. – P. 1-9.
56. *У. Росс Эшби*. Конструкция мозга / Пер. с англ. – М.: Иностранная литература, 1962. 397 с. (W. Ross Ashby. Design for a Brain. – N.Y.: Wiley, 1954. 260 p.).
57. ДСМ-метод автоматического порождения гипотез: Логические и эпистемологические основания / Сост. О.М. Аншаков, Е.Ф. Фабрикантова; Под общ. ред. О.М. Аншакова. – М.: Книжный дом «ЛИБРОКОМ», 2009. – 432 с.
58. Милль Д.С. Система логики силлогической и индуктивной: Изложение принципов доказательства в связи с методами научного исследования. Пер. с англ. / Предисл. и прил. В.К. Финна. Изд. 5-е испр. и доп. – М.: ЛЕНАНД, 2011. 832 с. (*Mill J.S.* A System of Logic Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation. – 1st ed. – London: John W. Parker, 1843. 622 p.).
59. *Финн В. К.* Искусственный интеллект: Методология, применения, философия. – М.: КРАСАНД, 2011. – 448 с.
60. *Грушо А. А., Применко Э. А., Тимонина Е. Е.* Теоретические основы компьютерной безопасности: учеб. пособие для студентов высш. учеб. заведений. – М.: Издательский центр «Академия», 2009. – 272 с.
61. *Грушо А. А., Забежайло М. И., Смирнов Д. В., Тимонина Е. Е.* О комплексной аутентификации // Системы и средства информатики. – 2017. – Т. 27. № 3. – С. 3–10.
62. *Грушо А. А., Забежайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О.,* Тимонина *Е.Е.* Классификация ошибочных состояний в распределенных вычислительных системах и источники их возникновения //

Системы и средства информатики. – 2017. – Т. 27. № 2. – С. 30–41.

63. *Грушо А. А., Забежайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О., Сенчило В. В., Судариков И. В., Тимонина Е. Е.* Об анализе ошибочных состояний в распределенных вычислительных системах // Системы и средства информатики, 2018. – Т. 28. № 1. – С. 99–109.
64. Anomaly Detection at Multiple Scales (ADAMS). General Services Administration. 2010-10-22. Retrieved 2011-12-05.
65. https://www.fbo.gov/download/2f6/2f6289e99a0c04942bbd89ccf242fb4c/DARPA-BAA-11-04_ADAMS.pdf
66. Yu R., He X., Liu Y. GLAD: Group Anomaly Detection in Social Media Analysis // arXiv.org e-Print Archive, 7 Oct 2014. arXiv:1410.1940.
67. Senator T., Bader D. et al. Detecting Insider Threats in a Real Corporate Database of Computer Usage Activity // Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. – New York, NY, USA: ACM, 2013. – P. 1393-1401.
68. Grusho A., Grusho N., Timonina E. Quality of tests defined by bans // Proceedings of the 16th Applied Stochastic Models and Data Analysis International Conference (ASMDA2015). – ISAST: Piraeus, Greece, 2015. – P. 289–295.
69. Grusho A., Grusho N., Timonina E. Modelling for ensuring information security of the distributed information systems // Proceedings of 31th European Conference on Modelling and Simulation (ECMS 2017). – Digitaldruck Pirrot GmbHP Dudweiler, Germany, 2017. – P. 656–660.
70. Мартьянов Е.А., Возможность выявления инсайдера статистическими методами // Системы и средства информатики. – 2017. – Т. 27. № 2. – С. 41–47.
71. Бурбаки Н. Общая топология. Основные структуры / Пер. с франц. – М.: Наука, 1968. 272 с. (Bourbaki N. Topologie Générale. Chapitre 1: Structures topologiques. Chapitre 2: Structures uniformes. – Paris: Hermann, 1940. 129 p.
72. Прохоров Ю. В., Розанов Ю. А. Теория вероятностей. – М.: Наука, 1993. –

496 с.

73. Grusho A., Grusho N., Timonina E. Consistent sequences of tests defined by bans // Springer proceedings in mathematics & statistics. Vol. 31: Optimization Theory, Decision Making, and Operation Research Applications. – New York, Heidelberg, Dordrecht, London: Springer-Verlag, 2013. – P. 281-291.
74. Грушо А. А., Грушо Н. А., Тимонина Е. Е. Включение новых запретов в случайные последовательности // Информатика и ее применения. – 2014. – Том. 8. № 4. – С. 48–54.
75. O. Sheyner, J. Haines, S. Jha, R. Lippmann and J. M. Wing. Automated generation and analysis of attack graphs // In Proceedings 2002 IEEE Symposium on Security and Privacy. – 2002. – P. 273-284.
76. Grusho A., Grusho N., Timonina E. Detection of anomalies in non-numerical data // Proceedings of the 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops. – IEEE Piscataway, NJ, USA, 2016. – P. 273-276.
77. Grusho A. Data mining and information security // Lecture Notes in Computer Science. – 2017. – Vol. 10446. – P. 28-33.
78. А. А. Грушо, Н. А. Грушо, М. И. Забежайло, Е. Е. Тимонина. Интеграция статистических и детерминистских методов анализа информационной безопасности // Информатика и ее применения. – 2016. – Том. 10. № 3. – С. 19-25.
79. Gheyas Iffat, Abdallah Ali. Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis // Big Data Analytics. – 2016. Vol. 1. – P. 1-29. – doi: 10.1186/s41044-016-0006-0
80. Memory Alex, Goldberg Henry G., Senator Ted E. Context-Aware Insider Threat Detection // Activity Context-Aware System Architectures: Papers from the AAAI 2013 Workshop, 2013. – P. 44-47. – URL: https://pdfs.semanticscholar.org/04aa/e6d97900ba62e90b07ac682fb7bd8c2e1029.pdf
81. Grusho A., Grusho N., Timonina E. Method of several information spaces for

identification of anomalies // Intelligent Distributed Computing XIII. IDC 2019 / Eds. I. Kotenko, C. Badica, V. Desnitsky, El Baz D., M. Ivanovic. – Studies in Computational Intelligence. – 2020. – Vol. 868. – P. 515-520. doi: 10.1007/978-3-030-32258-8_60.

82. Axelsson S. The Base-Rate Fallacy and the Difficulty of Intrusion Detection // ACM Transactions on Information and System Security. – 2000. – Vol. 3. No. 3. – P. 186–205.
83. Grusho A., Grusho N., and Timonina E. The bans in finite probability spaces and the problem of small samples // Distributed Computer and Communication Networks. DCCN 2019 / Eds. Vishnevskiy V., Samouylov K., Kozyrev D. — Lecture notes in computer science ser. – Springer, 2019. – Vol. 11965. – P. 578-590.
84. Колчин В. Ф., Севастьянов Б. А., Чистяков В. П. // Случайные размещения. – М.: Наука, 1976. – 224 с.
85. Anomaly Detection at Multiple Scales (ADAMS). – 2011. – URL: https://info.publicintelligence.net/DARPAADAMS.pdf.
86. Memory Alex, Goldberg Henry G., Senator Ted E. Context-Aware Insider Threat Detection // Activity Context-Aware System Architectures: Papers from the AAAI 2013 Workshop. – 2013. – P. 44-47.
87. Ruttenberg B. et al. Probabilistic Modeling of Insider Threat Detection Systems // In: Liu P., Mauw S., Stolen K. (eds) Graphical Models for Security. GraMSec 2017. Lecture Notes in Computer Science. – Vol. 10744. – Cham: Springer, , 2018. – P. 91-98. – URL: https://doi.org/10.1007/978-3-319-74860-3_6
88. Rashid Tabish, Agrafiotis Ioannis, Nurse Jason R. C. A New Take on Detecting Insider Threats: Exploring the Use of Hidden Markov Models // MIST '16: Proceedings of the 8th ACM CCS International Workshop on Managing Insider Security Threats. – October 2016. – P. 47-56
89. Gheyas Iffat A., Abdallah Ali E. Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis // Big Data Analytics. – 2016. – Vol. 1, No. 6. – P. 1-29.

90. Грушо А.А., Забежайло М.И., Тимонина Е.Е. О каузальной репрезентативности обучающих выборок прецедентов в задачах диагностического типа // Информатика и ее применения. –2020. – Т. 14. № 1. – С. 80-86.
91. Грушо А.А., Грушо Н.А., Тимонина Е.Е. Методы выявления "слабых" признаков нарушений информационной безопасности // Информатика и ее применения. – 2019. – Т. 13. № 3. – С. 3-8.
92. Ширяев А.Н. Вероятность. – В 2-х кн. – 3-е изд., перераб. и доп. – М.: МЦНМО, 2004. – 521 с.
93. Грушо А.А., Тимонина Е.Е. Запреты в дискретных вероятностно-статистических задачах // Дискретная математика. – 2011. – Том. 23. № 2. – С. 53-58.
94. Grusho A., Timonina E., Kniazev A. Detection of illegal information flow // Lecture Notes in Computer Science. – 2005. – Vol. 3685 LNCS. – P. 235-244.
95. Забежайло М.И., Трунин Ю.Ю. К проблеме доказательности медицинского диагноза: интеллектуальный анализ данных о пациентах в выборках ограниченного размера // Научно-техническая информация. – Сер 2, 2019. №12. – С. 12-18.
96. Забежайло М.И. О некоторых возможностях управления перебором в ДСМ-методе // Искусственный интеллект и принятие решений. – 2014. – Часть I: № 1. – С.95-110. – Часть II: № 3. – С. 3-21.
97. Забежайло М.И. О некоторых оценках сложности вычислений в ДСМ-рассуждениях // Искусственный интеллект и принятие решений. – 2015. – Часть I: №1, С.3-17, Часть II: №2. – С. 3-17.
98. Кон П. М. Универсальная алгебра. – М.: Мир, 1968. 359 с.