Автоматизированное обнаружение и классификация конфиденциальных данных в облачных средах

М.Э. Егоров, Д.Е. Намиот

***Аннотация*— Внедрение облачных технологий неизбежно сопровождается ростом рисков в области информационной безопасности. Одной из наиболее серьёзных проблем, с которыми сталкиваются пользователи облачных сред, является обнаружение и предотвращение утечек данных в облачной инфраструктуре. Конфиденциальные данные – это информация, которая требует защиты от несанкционированного доступа, изменения или распространения, так как она обладает высокой чувствительностью и может нанести ущерб владельцу или третьим лицам в случае утечки или злоупотребления. Эти данные могут касаться как физических лиц, так и организаций и часто регулируются нормативно-правовыми актами с целью обеспечения их безопасности и конфиденциальности. Компрометация конфиденциальной информации (персональных сведений, финансовых транзакций, интеллектуальной собственности), несанкционированный доступ к чувствительным данным могут приводить к масштабным репутационным и экономическим потерям. Согласно всем аналитическим отчётам, а также аналитическим обзорам Gartner и Forrester, число кибератак, нацеленных на облачные платформы, непрерывно растёт, а их сложность и изощрённость увеличиваются. Соответственно, вопросы идентификации конфиденциальных данных в облачных средах становятся крайне актуальными.**

***Ключевые слова*—конфиденциальные данные, облачные вычисления, персональные данные.**

I. Введение

Современная эпоха характеризуется стремительным ростом использования облачных вычислительных платформ для обработки, хранения и анализа информации. Такие сервисы, как Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), IBM Cloud и другие, предоставляют практически неограниченные возможности масштабирования ресурсов, гибкости конфигурирования и снижения временных и финансовых затрат на инфраструктуру. По данным ведущих аналитических компаний (Gartner, Forrester), объёмы данных, передаваемых, обрабатываемых и хранящихся в облаках, непрерывно возрастают, отражая глобальный тренд миграции корпоративных и государственных информационных систем в виртуализированные среды.

Однако стремительное внедрение облачных технологий неизбежно сопровождается ростом рисков в области информационной безопасности. Одной из наиболее серьёзных проблем, с которыми сталкиваются организации, является обнаружение и предотвращение утечек данных в облачной инфраструктуре. Компрометация конфиденциальной информации (персональных сведений, финансовых транзакций, интеллектуальной собственности), несанкционированный доступ к чувствительным данным или их скрытая эксфильтрация могут приводить к масштабным репутационным и экономическим потерям. Согласно отчётам ENISA [1] и OWASP [2], а также аналитическим обзорам Gartner и Forrester, число кибератак, нацеленных на облачные платформы, непрерывно растёт, а их сложность и изощрённость увеличиваются.

Особую актуальность проблема утечек приобретает в условиях мультиарендных, распределённых, масштабируемых облачных сред, для которых характерны динамическое изменение конфигураций, широкое использование микросервисной архитектуры, частая ротация виртуальных машин, контейнеров и сервисных аккаунтов. Традиционные системы Data Loss Prevention (DLP), опирающиеся преимущественно на статические сигнатуры, предопределённые политики и сравнительно стабильные инфраструктуры, оказываются недостаточно эффективными. Они не способны должным образом учитывать непрерывно меняющийся контекст, сложность топологии взаимодействий и объёмы логов и телеметрии, генерируемых облаком.

Для решения этой проблемы было предложено искать и классифицировать конфиденциальные данные в зависимости от их типа, и защищать непосредственно эту найденную чувствительную информацию. Необходимость именно в автоматизированном обнаружение и классификации конфиденциальных данных в облачных средах обусловлена рядом ключевых факторов, отражающих современные вызовы в области информационной безопасности и управления

Статья получена 22 мая 2025.
М.Э. Егоров – МГУ имени М.В. Ломоносова (email: 662366@bk.ru)
Д.Е. Намиот – МГУ имени М.В. Ломоносова (email: dnamiot@gmail.com)

данными.

Одной из ключевых проблем является управление конфиденциальностью данных в условиях, когда информация хранится на серверах, расположенных в разных юрисдикциях, и может быть доступна сторонним компаниям, в том числе облачным провайдерам. В последнее время на фоне глобальной цифровизации мира и роста количества цифровой информации и популярности облачных сред правительствами различных стран были разработаны стандарты и нормативные акты, такие как GDPR, CCPA, ISO 27001, HIPAA, PCI DSS, ФЗ-152 и другие. Конфиденциальные данные могут включать в себя личную информацию пользователей, финансовые данные, интеллектуальную собственность и другие виды чувствительной информации. Это привело к тому, что от организаций требуется внедрения эффективных механизмов защиты данных для соблюдения данных правил, иначе в случае утечек или неправильного обращения с такими данными возможны юридические, финансовые и репутационные последствия для организаций.

Еще одной сложностью является недостаточная осведомленность о том, какие данные хранятся в облаке, и их правильная классификация. В связи с тем, что обеспечение конфиденциальности и целостности данных стало основой доверия пользователей к облачным сервисам и необходимым требованием со стороны государств, соблюдение прописанных норм и обеспечение безопасности стало необходимостью. Однако осуществление этих процессов при помощи ручного анализа при работе с большим объемом данных является ресурсоемким и медленным процессом в динамичной постоянно меняющейся и растущей облачной среде. Эти факторы приводят к необходимости автоматизации этих процессов, что позволит повысить точность, скорость обработки и снизить затраты. Без эффективных методов автоматического обнаружения и классификации конфиденциальных данных организации рискуют столкнуться с проблемами, такими как утечка информации или несоответствие требованиям законодательства по защите данных.

Таким образом, исследование и внедрение автоматизированных методов обнаружения и классификации конфиденциальных данных в облачных средах являются неотъемлемой частью обеспечения информационной безопасности в условиях постоянно меняющейся цифровой среды.

Основная цель данного исследования заключается в анализе существующих методов и инструментов для автоматизированного обнаружения и классификации конфиденциальных данных в облачных средах, с выявления имеющихся недостатков и потенциальных сфер дальнейшего развития.

Среди задач можно выделить:

— Анализ существующих подходов и технологий для обнаружения конфиденциальных данных в облаке.

— Анализ моделей и алгоритмов, способных автоматически классифицировать данные в облачной среде по их степени конфиденциальности.

— Оценка эффективности предложенных методов.

— Формирование потенциальных рекомендаций по дальнейшему развитию автоматизированных систем обнаружения и классификации в существующие облачные инфраструктуры.

II. Облачные вычислительные среды: Описание сервисных моделей (IaaS, PaaS, SaaS) и их особенности

Процесс постепенной, но устойчивой миграции бизнес-процессов и информационных систем в облачные вычислительные среды стал одним из определяющих трендов современного технологического ландшафта. Облачные платформы, управляемые крупными провайдерами (такими как Amazon Web Services, Microsoft Azure, Google Cloud Platform, IBM Cloud и прочие), предлагают заказчикам широкий спектр ресурсов и сервисов, доступных по запросу, масштабируемых по мере изменения потребностей и оплачиваемых по модели «pay-as-you-go». Сущность облачных вычислений заключается в абстрагировании от физического оборудования, физического местоположения и статичных инфраструктурных моделей, что позволяет быстро адаптировать IT-ресурсы к динамично меняющимся бизнес-требованиям. Однако эти же особенности усложняют задачу контроля и мониторинга, включая своевременное обнаружение утечек данных.

- Ключевым элементом понимания облачных сред являются сервисные модели, определяющие уровень абстракции и степень ответственности между провайдером и пользователем. Наибольшее распространение получили три классические модели: IaaS (Infrastructure as a Service),
- PaaS (Platform as a Service) и
- SaaS (Software as a Service).

A. Инфраструктура как услуга (IaaS)

Модель IaaS [3] предоставляет клиенту доступ к базовым вычислительным ресурсам: виртуальным машинам, сетям, хранилищам, балансировщикам нагрузки и прочим низкоуровневым компонентам. Пользователь может разворачивать собственные операционные системы, устанавливать необходимое программное обеспечение, конфигурировать сети и системы безопасности, практически так же, как в традиционном дата-центре, но без необходимости владения физической инфраструктурой. Примерами IaaS-сервисов могут служить Amazon EC2, Azure Virtual Machines, Google Compute Engine.

Основное преимущество IaaS — гибкость и полный контроль над базовыми ресурсами. Однако из этого следует и повышенная ответственность клиента за правильную настройку систем, мониторинг и безопасность. В контексте утечек данных IaaS-среды характеризуются тем, что пользователь сам решает, как построить архитектуру защиты: от выбора ОС и

межсетевых экранов до шифрования хранилищ. При этом высокая степень свободы сочетается с риском неверных конфигураций, ошибок доступа и пробелов в управлении секретами. Из-за этого платформа IaaS может становиться почвой для скрытых утечек, если не будут применены надлежащие контрмеры.

B. Платформа как услуга (PaaS)

Модель PaaS [4] предоставляет клиенту готовую платформу для разработки и эксплуатации приложений. Пользователю платформы не нужно администрировать системное программное обеспечение, организовывать бесперебойную работу дата-центра и системных программных компонент. Эти элементы находятся в сфере ответственности вендора (провайдера платформы). Пользователи платформы фокусируются на разработке приложений, и они отвечают за их устойчивую работу, безопасность, права доступа к API и т. д. Примерами PaaS-сервисов могут служить Amazon Web Services (AWS) Elastic Beanstalk, Google App Engine, Red Hat OpenShift, Microsoft Azure.

Преимущества PaaS заключаются в следующем. PaaS может снизить затраты на разработку, устраняя необходимость в покупке инфраструктуры и оборудования. PaaS позволяет быстрее развертывать и масштабировать приложения, сокращая время выхода на рынок. PaaS предоставляет гибкую платформу, которая может адаптироваться к меняющимся потребностям бизнеса и технологическим требованиям и облегчает совместную работу разработчиков, предоставляя общую среду и инструменты.

C. Программное обеспечение как услуга (SaaS)

Модель SaaS [5] представляет собой наиболее высокий уровень абстракции: конечный пользователь получает полностью готовое приложение, доступное по сети, без необходимости управлять инфраструктурой, базами данных или платформенными компонентами. Примером SaaS могут служить такие решения, как Office 365, Google Workspace, Salesforce и другие бизнес-приложения.

В SaaS-модели практически вся ответственность за инфраструктуру, обновления, исправления и большую часть мер безопасности лежит на провайдере.

С точки зрения утечек данных SaaS может казаться более защищённой моделью, поскольку клиент часто не имеет прямого доступа к низкоуровневым ресурсам и, следовательно, риски неправильной конфигурации снижаются. Однако, сохраняется множество каналов для утечки: неправомерное использование функционала приложения инсайдерами, неправильно настроенные разрешения внутри тенантов (мультиарендных экземпляров), уязвимости в бизнес-логике самого SaaSпродукта. Пользователи, например, могут по неосторожности предоставлять доступ к документам или данным внешним лицам, перенаправлять конфиденциальные сведения через интеграции с другими сервисами или использовать слабые аутентификационные механизмы. Кроме того, мультиарендность в SaaS означает, что один и тот же экземпляр программного обеспечения обслуживает множество клиентов (тенантов), разделяя инфраструктуру и ресурсы на логическом уровне. Утечка данных может произойти, если разделение данных между арендаторами будет нарушено или будет обнаружена уязвимость, позволяющая одному клиенту получить доступ к данным другого.

D. Мультиарендность как особенность облака

Мультиарендность (multi-tenancy) – одна из фундаментальных характеристик большинства облачных сервисов [6]. В отличие от классических корпоративных инфраструктур, где ресурсы часто выделены под нужды одной организации, в облаке множество клиентов (арендаторов) пользуются одними и теми же физическими серверами, сетевыми устройствами, системами хранения и сервисными компонентами. Разграничение доступа достигается средствами виртуализации, контейнеризации, политиками и контролем идентификационных данных. Это даёт экономию средств, повышает эффективность использования ресурсов провайдером, но одновременно осложняет логику обеспечения безопасности.

Для обнаружения утечек в мультиарендных средах необходимо детально отслеживать, какие данные, когда и куда перемещаются, а также как политики безопасности применяются к отдельным арендаторам и их приложениям. Сбой в настройках доступа, ошибка в коде общего сервиса или пропуск в логике разграничения может позволить одному арендатору просмотреть или извлечь данные другого, что крайне опасно и создаёт широкий спектр потенциальных инцидентов.

Более того, мультиарендность усложняет аудит и корреляцию событий, поскольку логи и метрики генерируются большим числом субъектов, чьи активности переплетены в общем пуле ресурсов.

E. Географическое распределение данных и ресурсов

Современные облачные провайдеры предлагают клиентам возможность размещать приложения и данные в различных географических регионах и зонах доступности, расположенных по всему миру. Это повышает отказоустойчивость, снижает задержки доступа для пользователей из разных стран и обеспечивает непрерывность бизнеса даже при сбоях в одной локации.

Однако географическое распределение данных создает дополнительные сложности для обнаружения утечек. Применяемые нормативно-правовые требования (например, местные законы о хранении персональных данных), политика репликации и резервирования, кросс-региональные передачи информации – всё это может затруднять централизованный мониторинг. Когда данные фрагментированы и хранятся на разных континентах, расширенный анализ логов и взаимодействий становится нетривиальным.

Злоумышленники могут использовать эту фрагментацию, маскируя эксфильтрацию через различные региональные узлы.

F. Динамичность ресурсов и эластичность масштабирования

Одним из ключевых преимуществ облачных вычислений является возможность динамически изменять объёмы используемых ресурсов в ответ на колебания нагрузки. Приложения в облаке могут автоматически масштабироваться: увеличивать число виртуальных машин или контейнеров в периоды пикового трафика и уменьшать их по мере снижения нагрузки. Это обеспечивает экономию и высокую производительность, однако усложняет задачу постоянного мониторинга данных.

Когда компоненты инфраструктуры появляются и исчезают буквально за секунды, а сервисные аккаунты и токены доступа генерируются динамически, привычные статические модели обнаружения инцидентов, основанные на стабильных сетевых топологиях и предсказуемых паттернах, теряют эффективность. Вчерашняя конфигурация не гарантирует, что сегодня всё устроено так же. Динамические инфраструктуры требуют адаптивных, самонастраиваемых систем обнаружения утечек, способных быстро приспосабливаться к изменениям состава ресурсов, топологий и политик.

G. Влияние сервисных моделей на задачи обнаружения утечек

Соотношение между IaaS, PaaS и SaaS-моделями и задачами обнаружения утечек может быть представлено сопоставлением ответственности и возможностей контроля. На уровне IaaS у клиента больше ручного контроля и инструментов низкоуровневого мониторинга, но и гораздо больше рисков неправильной конфигурации. На уровне SaaS у клиента меньше возможностей влиять на внутреннее устройство системы, но он может столкнуться с проблемами, связанными с ограниченной наблюдаемостью и сложностью расследования инцидентов, поскольку большинство управленческих функций скрыто за абстракцией провайдера.

Таким образом, выбор модели облачного сервиса неизбежно отражается на способах детектирования утечек данных, применяемых инструментах и архитектуре мониторинга. Не существует единого универсального решения: каждая модель диктует собственные требования к политикам безопасности и контролю данных.

III. Конфиденциальные данные

Конфиденциальные данные – это информация, которая требует защиты от несанкционированного доступа, изменения или распространения, так как она обладает высокой чувствительностью и может нанести ущерб владельцу или третьим лицам в случае утечки или злоупотребления. Эти данные могут касаться как физических лиц, так и организаций и часто регулируются нормативно-правовыми актами с целью обеспечения их безопасности и конфиденциальности. Понятие «конфиденциальность информации» раскрывается в законе N 149-ФЗ, а именно – в ст. 2. Информация ограниченного доступа – сведения, которые нельзя передавать без согласия их владельца.

A. Типы конфиденциальных данных

Глобально всю конфиденциальную информацию можно разделить на следующие типы:

Персональные данные – информация, относящаяся к определенному или определяемому физическому лицу, позволяющая установить его личность. К таким данным относятся фамилия, имя, отчество, дата и место рождения, паспортные данные, контактная информация и другие сведения. В Российской Федерации обработка персональных данных регулируется Федеральным законом № 152-ФЗ «О персональных данных».

Финансовая информация – сведения, связанные с финансовым состоянием и операциями физических и юридических лиц. Это могут быть данные о доходах, расходах, банковских счетах, кредитной истории, налоговых декларациях и другие финансовые показатели. Защита финансовой информации важна для предотвращения мошенничества, кражи личности и других финансовых преступлений.

Коммерческая тайна – сведения, которые имеют коммерческую ценность для организации и не являются общедоступными. К ним могут относиться данные о бизнес-планах, маркетинговых стратегиях, ценовой политике, клиентской базе, поставщиках и другие внутренние документы. Защита коммерческой тайны способствует сохранению конкурентных преимуществ и финансовой стабильности компании.

Другие типы – медицинские данные (HIPAA): истории болезней, результаты анализов; государственная тайна: военные секреты, дипломатическая переписка; данные IoT-устройств: геолокация, метрики с датчиков; и другие данные.

В РФ конфиденциальная информация бывает следующих типов [9]: государственная тайна, коммерческая тайна, персональные данные, налоговая тайна, банковская тайна, врачебная тайна, служебная тайна, коммерческая тайна и другие виды тайн.

В рамках обработки конфиденциальной информации ее можно классифицировать на разные группы в зависимости от необходимости:

1) Тема принадлежности: военные, политические, экономические, персональные, финансовые, медицинские, коммерческие и другие.

2) По степени конфиденциальности: публичная, для внутреннего пользования, секретная, совершенно секретная, особой важности.

3) По формату: структурированные (базы данных) и неструктурированные (текст, изображения).

B. Нормативные требования к защите конфиденциальных данных

Нормативные требования к защите конфиденциальных данных варьируются в зависимости от страны и региона. Эти требования определяют, как компании и организации должны обрабатывать, хранить и защищать данные физических и юридических лиц. Отличия в законодательстве и нормативных актах обусловлены культурными, правовыми и экономическими различиями между странами, а также различиями в уровнях развития технологий. В России основными нормативными актами являются:

— Федеральный закон № 152-ФЗ «О персональных данных». Регулирует обработку персональных данных, устанавливает требования к их защите и определяет права субъектов данных. Согласно этому закону, обработка персональных данных в облачных сервисах возможна при условии, что дата-центр находится на территории России, а передача данных за рубеж осуществляется в соответствии с установленными требованиями.

— Гражданский кодекс Российской Федерации. Содержит положения о коммерческой тайне, определяя режим ее защиты и ответственность за разглашение.

— Стандарты серии ISO/IEC 27000. Предлагают международные лучшие практики в области управления информационной безопасностью, включая защиту конфиденциальных данных. Соблюдение этих стандартов помогает организациям обеспечивать системный подход к защите информации.

Ниже приведены наиболее важные и известные регламенты и законы других стран.

GDPR (Общий регламент по защите данных – General Data Protection Regulation) — это один из самых строгих нормативных актов в мире, регулирующий защиту персональных данных в Европейском Союзе. Он распространяется на любые организации, которые обрабатывают данные граждан ЕС, независимо от их местонахождения. GDPR устанавливает требования к сбору, хранению, обработке и передаче персональных данных.

CCPA (Закон о конфиденциальности потребителей Калифорнии – California Consumer Privacy Act) — закон, регулирующий защиту персональных данных в Калифорнии, США. Он предоставляет жителям Калифорнии право знать, какие личные данные собирает компания, право на доступ к этим данным, право на их удаление и право на отказ от продажи их личных данных третьим лицам. CCPA ориентирован на защиту прав потребителей и накладывает на организации обязательства по уведомлению пользователей о целях использования их данных.

HIPAA (Закон о переносимости и подотчетности медицинского страхования – Health Insurance Portability and Accountability Act) регулирует защиту медицинских данных в США. Он устанавливает требования к конфиденциальности и безопасности медицинской информации, требуя от организаций, работающих с медицинскими данными (например, больниц и страховых компаний), соблюдать строгие меры безопасности для защиты этих данных от утечек и несанкционированного доступа.

Закон о защите персональных данных Китая (Personal Information Protection Law, PIPL), вступивший в силу 1 ноября 2021 года, является основным нормативным актом, регулирующим защиту личных данных в Китае. Это первый закон в Китае, который регулирует обработку персональных данных на уровне национального законодательства и во многом аналогичен европейскому GDPR.

В таблице I представлена сводная информация по критериям защиты данных

Критерий	ЕС	США	Россия	Китай
Локализация данных	Не требуется	Не требуется	Обязательна	Обязательна
Согласие на сбор	Обязательно	Частично (зависит от штата)	Обязательно	Обязательно
Штрафы	До 4% оборота	До $9,4 млн (например, Clearview AI)	До 18 млн руб.	Не указаны, но жёсткий контроль государства
Трансграничная передача	Разрешена в страны с адекватной защитой	Ограничена для госорганов	Только в утверждённые страны	Запрещена без разрешения

IV Методы обнаружения и классификации

К традиционным методам обнаружения и классификации конфиденциальных данных относится метод сопоставления шаблонов, который включает в себя идентификацию конфиденциальных данных путем распознавания конкретных шаблонов или последовательностей, связанных с типами информации.

Шаблоны могут быть определены с помощью

регулярных выражений (regex) или настраиваемых алгоритмов, которые сопоставляют определенные последовательности в данных. Эти шаблоны представляют собой общие формы или структуры конфиденциальной информации, такие как номера кредитных карт, телефонные номера, электронные адреса и другие аналогичные виды информации.

Процесс сопоставления шаблонов происходит в несколько этапов. Во-первых, система обнаружения сканирует данные (например, текстовые файлы, документы, электронные письма, базы данных) для выявления последовательностей символов, которые соответствуют заранее определенным шаблонам. Когда последовательность соответствует шаблону, система помечает ее как конфиденциальную информацию. На основе найденного шаблона система назначает данным определенный класс.

Регулярные выражения – это инструмент для определения шаблонов, которые соответствуют определенным типам данных. Используя заранее определенные шаблоны для информации, регулярные выражения могут быстро и эффективно сканировать большие объемы данных с целью выявления чувствительной информации.

Среди вариантов, основанных на этом методе, можно выделить применение эвристического анализ. Он включает использование правил шаблонов, которые ищут общие атрибуты чувствительной. информации, такие как длина, структура или известные префиксы. Это используется для выявления чувствительных данных, даже если точный шаблон неизвестен.

Сопоставление шаблонов является широко используемым и эффективным методом для обнаружения и классификации чувствительной информации, особенно когда речь идет о структурированных данных или хорошо известных типах чувствительной информации. Несмотря на свою эффективность в многих сценариях, этот метод имеет ограничения, особенно при работе с более сложными, неструктурированными или новыми типами чувствительных данных. Для преодоления этих ограничений многие организации комбинируют сопоставление шаблонов с другими методами, вроде эвристического анализа, чтобы улучшить точность и адаптивность обнаружения.

A. Анализ метаданных

Обнаружение и классификация конфиденциальной информации на основе анализа метаданных – это подход, при котором чувствительная информация определяется и классифицируется на основе метаданных, а не содержания самих данных. Метаданные – это данные, которые описываются другими данными, вроде, даты создания, автора, типа файла, размера файла, права доступа и прочего. Этот метод использует тот факт, что конфиденциальная информация часто имеет определенные шаблоны метаданных или связанных с ней атрибутами. Путем анализа этих метаданных можно обнаружить и классифицировать чувствительную информацию. Это особенно полезно в средах, где полноценный анализ содержания может быть слишком ресурсоемким или в тех случаях, когда существуют опасения по поводу конфиденциальности данных при их прямом сканировании. После сбора метаданных, их анализ включает в себя сканирование и интерпретацию метаданных файла для выявления потенциально чувствительной или регулируемой информации. Система может обнаружить файлы или документы, вероятно содержащие чувствительную информацию, анализируя их метаданные. Например, PDF-документ с конфиденциальной информацией может иметь метаданные, указывающие на ограниченный доступ, или может быть создан автором, занимающим чувствительную должность. После того как чувствительные метаданные обнаружены, система может классифицировать данные на основе заранее определенных правил. Данный метод для обнаружения и классификации конфиденциальной информации, особенно эффективен в средах с большими объемами данных, вроде облачных, где сканирование содержания может быть непрактичным. Анализ метаданных позволяет эффективно обнаруживать потенциальные риски безопасности, обеспечивать соблюдение нормативных требований и внедрять политику надлежащего обращения с данными, при этом минимизируя нагрузку на ресурсы. Однако важно отметить, что анализ метаданных работает лучше всего в сочетании с другими методами обнаружения, такими как анализ содержания или машинное обучение, чтобы обеспечить комплексный подход к управлению информацией.

B. Алгоритмы машинного и глубокого обучения

Методы, вроде сопоставления шаблонов или анализа метаданных, могут быть ограничены, особенно когда речь идет о сложных или динамичных наборах данных. Алгоритмы машинного обучения, в свою очередь, предоставляют возможность более эффективно выявлять и классифицировать чувствительные данные, обучаясь на паттернах данных и принимая решения на основе этих усвоенных закономерностей. Этот метод предоставляет большую гибкость и точность в выявлении чувствительной информации, даже если явные шаблоны не всегда легко различимы. Подходы, основанные на машинном обучении, для обнаружения и классификации конфиденциальной информации включают обучение моделей на размеченных данных, чтобы они научились 10 определять и классифицировать различные типы чувствительных данных. При этом зачастую данный метод используется в совокупности с другими, описанными ранее. Так, CASES [10] использует метод на основе анализа контекста для обнаружения конфиденциальных данных в структурированных наборах данных. В ней интегрированы преобразователи глубокого обучения, в частности BERT, с традиционными системами, основанными на правилах, для определения взаимосвязей внутри столбцов данных и между ними.

Этот гибридный подход повышает точность обнаружения конфиденциальных данных в структурированных форматах.

В статье [11] описывается методику обнаружения чувствительной информации в текстовых файлах с помощью скрытых марковских моделей (HMM) и машин опорных векторов (SVM). В этой работе анализируются текстовые данные, используя HMM для моделирования последовательностей и SVM для классификации, с целью выявления и защиты конфиденциальной информации.

Статья [12] представляет новый подход к классификации данных с целью обеспечения их конфиденциальности в облачных средах. В работе предлагается улучшенная версия алгоритма kближайших соседей, названная Training dataset Filtration-kNN (TsF-kNN), которая снижает вычислительную сложность по сравнению с традиционными методами, путем первоначального анализа метаданных файла с целью определения специальной модели, обученной на отдельном наборе данных, для последующего процесса классификации.

В статье [13] предлагается метод обнаружения конфиденциальной информации в неструктурированных текстах с использованием сверточных нейронных сетей (CNN). Авторы отмечают, что традиционные методы, такие как сопоставление чувствительных слов и рекуррентные нейронные сети (RNN), могут быть недостаточно точными и эффективными для сложных паттернов чувствительной информации. Вместо этого они предлагают использовать Text-CNN, который обеспечивает высокую точность обнаружения и сокращает время обучения модели. Экспериментальные результаты подтверждают эффективность предложенного подхода.

Статья [14] предлагает подход к автоматическому обнаружению неструктурированной конфиденциальной информации в текстах с учетом контекста и зависимостью от данных. Авторы разработали модель, использующую глубокие нейронные сети для анализа текста с учетом контекста. В отличие от традиционных методов, таких как основанные на правилах или простые алгоритмы машинного обучения, их подход позволяет учитывать сложные зависимости между словами и фразами в тексте, что особенно важно при обработке неструктурированных данных, где контекст может существенно изменять значение информации.

В статье [15], модель Sherlock ставит перед собой задачу проанализировать взаимосвязи между ячейками и представить более глубокое понимание контекста. Для достижения этой цели Sherlock использует архитектуру глубокой нейронной сети ,которая, помимо использования информации из текущей ячейки для формирования ее характеристик, учитывает контекст, принимая во внимание все остальные ячейки в том же столбце посредством модифицированных векторов абзацев, сгенерированных заранее. Этот метод предполагает учет только обобщенной информации по столбцу, к которому относится текущая ячейка. В дополнение к векторам абзацев, Sherlock также использует статистические данные текущего столбца, такие как распределение символов и средняя длина ячеек. Sato [16] развивает теорию из статьи [15], добавив модель предсказания тем, действующую для всей таблицы. Следовательно, Sato также использует другие столбцы базы данных для расширения контекста ячейки. Хотя и Sato, и Sherlock учитывают контекст таблицы, они делают это пассивным способом: контекст создается заблаговременно, и модели не позволяют активно исследовать значения иных ячеек при формировании представления текущей ячейки.

C. Сравнительный анализ методов

Сравнительный обзор всех, рассмотренных ранее методов, с целью выявления наиболее оптимальных решений для различных ситуаций приведен в таблице II.

Таблица II Сравнительный анализ подходов

Критерий	Signature-based [17]	Rule-based [18]	Анализ контекста [19]	Анализ метаданных [20]	Машинное обучение
Эффективность против известных угроз	Высокая	Высокая	Средняя	Низкая	Высокая
Эффективность против новых угроз	Низкая	Низкая	Средняя	Низкая	Высокая
Адаптивность	Низкая	Низкая	Средняя	Низкая	Высокая
Точность	Высокая (для известных)	Высокая	Средняя	Низкая	Зависит от данных
Ресурсоемкость	Низкая	Средняя	Средняя	Низкая	Высокая
Устойчивость к обфускации	Низкая	Средняя	Высокая	Низкая	Высокая
Интеграция с облаком	Легкая	Средняя	Сложная	Легкая	Сложная

На основе таблицы можно сделать следующие выводы:
1) Традиционные методы уступают в тех или иных областях, делая систему потенциально уязвимой. Однако, стоит выделить метод, основанный на анализе контекста, который может являться достаточным при определенных обстоятельствах.
2) Метод на основе анализа метаданных кажется весьма бесполезным, но его основная в простоте реализации и скорости работы. Он был создан для работы в высокоскоростных средах, где требуется слабая защита, а на данный момент является вспомогательным методом, используемым в гибридных реализациях. Например, как часть системы с машинным обучением.
3) Методы на основе машинного и глубокого обучения, при правильной настройке и обучении на достаточном наборе данных, показывают наилучшие показатели, но они также являются самыми ресурсоемкими и сложными в реализации.

Также рассмотрим результативность различных моделей на основе алгоритмов машинного и глубокого обучения, приведенных ранее – таблица III.

Таблица III: Сравнительный анализ моделей на основе алгоритмов машинного и глубокого обучения

Модель	Accuracy	Precision	F1-score	База данных	Особенности
BERT [10]	-	0. 9707	0.9832	DeSSI	Контекстно-ориентированный подход к структурированным конф-м данным
NLP [10]	-	0. 935	0.8672	DeSSI	Контекстно-ориентированный подход к структурированным конф-м данным
Сопоставление ключевых слов (signature-based) [13]	0.73	-	-	Собственная бд из 14 тыс. текстовых документов	Плохие показатели
Традиционные методы МО [13]	0.87	-	-	Собственная бд из 14 тыс. текстовых документов	Сложность определения признаков и опасность влияния человеческого фактора
RNN [13]	0.9424	-	-	Собственная бд из 14 тыс. текстовых документов	Требует большого времени для обучения
Text-CNN [13]	0.9517	-	-	Собственная бд из 14 тыс. текстовых документов	Положительный отзыв авторов
Линейный SVM [14]	0.65	0.62	0.65	Tweet датасет (текст)	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой
Логистическая регрессия с TF-IDF [14]	0.7	0.63	0.73	Tweet	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой конфиденциальной информации
Дерево решений с TF-IDF [14]	0.9	0.92	0.88	Составная из MIDV, DIQA, SVT (изображения)	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой конфиденциальной информации
Text-CNN [14]	0.8	0.81	0.75	Tweet	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой конфиденциальной информации
Text-CNN [14]	0.86	0.85	0.87	Составная	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой конфиденциальной информации
LSTM [14]	0.95	0.95	0.93	Tweet	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой конфиденциальной информации
LSTM [14]	0.75	0.62	0.69	Составная	Направлен на автоматическое обнаружение неструктурированной контекстно-зависимой конфиденциальной информации
Нейронная сеть (Sherlock) [15]	-	-	0.89	VizNet	Учитывает контекст информации, хранящейся в виде информации
Нейронная сеть + LDA [16] (Sato)	-	-	0.9	VizNet	Развивает идеи модели из статьи [15] Sherlock

На основе таблицы можно сделать следующие выводы:

1) Традиционные методы машинного обучения в связи с необходимостью анализировать неструктурированные данные (текст, изображения и т.п.). В случае анализа структурированной информации они могут быть эффективнее других методов.
2) Методы, в основе которых лежат нейронные сети, показывают хорошие результаты, при достаточно полном наборе информации для обучения, а также требуют меньше ресурсов чем LLM (Large Language Model).
3) Методы на основе LLM показали лучшие результаты,

но их разработка и эксплуатация требуют больших ресурсов. Также нужно изучить их влияние на скорость работы системы.

V Проблемы существующих методов

Обеспечение конфиденциальности данных в облачных средах является ключевым аспектом, и автоматические системы, которые обнаруживают и классифицируют конфиденциальные данные, играют важную роль в защите личной информации и соблюдении нормативных требований. Однако при разработке и внедрении таких систем для обнаружения и классификации конфиденциальных данных в облачных средах существует несколько проблем и требований, которые необходимо учитывать.

Одна из проблем обусловлена масштабируемостью и сложностью облачных сред. Облачные среды динамичны, данные могут быстро увеличиваться в объеме и изменяться. Это затрудняет использование статичных моделей обнаружения, которые не успевают адаптироваться к изменяющимся данным. В облачных средах различные пользователи делят ресурсы, что означает, что автоматическим системам необходимо различать данные одного арендатора и данные других арендаторов. Ошибки классификации могут привести к утечке данных. При этом, большая часть данных в облаке является неструктурированной (например, текстовые документы, электронные письма, изображения), что усложняет точное обнаружение и классификацию конфиденциальной информации. Также облачные сервисы динамически регулируют ресурсы в зависимости от нагрузки.

Системы автоматического обнаружения должны адаптироваться к изменениям в рабочей нагрузке и вычислительных мощностях, что может привести к задержкам или снижению точности. В связи с этим, система должна быть способна масштабироваться в зависимости от изменений объема данных и вычислительных мощностей без потери точности. Алгоритмы должны эффективно работать с большими и разнообразными наборами данных. Системы должны обеспечивать возможность обнаружения и классификации конфиденциальных данных в реальном или близком к реальному времени, когда данные загружаются, обрабатываются или передаются. Другая проблема – видимости данных и мониторинга. Поставщики облачных услуг часто управляют инфраструктурой, и организации имеют ограниченную видимость того, как и где хранятся, обрабатываются и передаются их данные. Данные в облачных средах часто распределены по нескольким географическим регионам и системам. Без единого представления о данных становится сложнее последовательно обнаруживать и управлять конфиденциальной информацией. Это приводит к необходимости системы автоматического обнаружения обеспечивать детализированное управление доступом и разрешениями, чтобы защитить конфиденциальные данные в облаке и гарантировать, что только авторизованные пользователи и процессы имеют к ним доступ. Также особенности структуры облачных сред приводят к возникновению проблемы точности и ложных срабатываний, либо пропусков конфиденциальной информации. Одна из основных проблем систем автоматического обнаружения и классификации заключается в риске ложных срабатываний, когда не конфиденциальные данные ошибочно классифицируются как конфиденциальные. Напротив, ложные отрицания происходят, когда конфиденциальная информация не классифицируется должным образом, что может привести к утечке данных или несоответствию нормативным требованиям. При этом классификация данных как конфиденциальных или нет зависит от контекста их использования. Модель, которая не учитывает контекст, может ошибочно классифицировать данные. Для покрытия этой проблемы системы должны использовать более сложные алгоритмы машинного обучения (например, глубокое обучение, обработка естественного языка), которые могут более точно понимать контекст данных и минимизировать ложные срабатывания и пропуски, а также позволять организациям задавать пользовательские правила и классификацию, соответствующие их конкретным требованиям, с учетом разных типов конфиденциальных данных и контекста их использования. Две последние проблемы свойственным любым информационным системам – соответствие количества затраченных ресурсов и эффективности системы и сохранение безопасности и целостности данных, обрабатываемых системой.

Система должна быть спроектирована так, чтобы эффективно использовать вычислительные ресурсы, а также масштабироваться динамически, чтобы минимизировать затраты, используя модель оплаты за использование, сбалансированную с производительностью. При этом конфиденциальные данные должны обрабатываться в защищенных средах, с шифрованием, как при передаче, так и при хранении. Системы также должны использовать методы проверки целостности для обеспечения защиты от подделок данных.

VI Потенциальные направления развития

Автоматическое обнаружение и классификация конфиденциальных данных в облачных средах имеют критическое значение для обеспечения безопасности данных, конфиденциальности и соблюдения нормативных требований. Несмотря на значительный прогресс, существуют области, где текущие системы могут быть улучшены для более эффективного управления динамичной и сложной природой облачных данных. Ниже представлены ключевые области, в которых есть возможности для улучшений в процессах автоматического обнаружения и классификации.

A. Повышение точности и снижение ложных срабатываний/пропусков.

Ключевым направлением является повышение

точности детектирования чувствительных данных при одновременном сокращении ложных срабатываний и пропусков. Традиционно используются регулярные выражения и шаблоны для поиска совпадений с PII/PHI, но они часто не учитывают контекст и генерируют множество ложных тревог. В ответ появилась практика объединения нескольких методов. Например, точное сравнение данных позволяет детерминировано сверять записи с заранее известным списком конфиденциальных значений, что практически устраняет ложные срабатывания по этим данным. В то же время EDM склонен давать пропуски при даже незначительных изменениях формата (например, отсутствие дефисов в номере).

Современные системы добавляют машинное обучение и NLP: продвинутые языковые модели (BERT, GPT и др.) лучше распознают контекст и семантику, что повышает точность классификации. Например, Google Cloud DLP (Sensitive Data Protection) [20] позволяет настраивать порог вероятности и ограничивать набор используемых детекторов: документ рекомендуется начинать со «суженной» выборки и исключений, чтобы идентифицировать избыточные совпадения и скорректировать правила, что снижает число ложных находок. В продуктах также реализованы механизмы отсева шума. Так, AWS Macie [21] ввел функцию *allow list* (список разрешённых шаблонов), позволяющий исключить из анализа известные значения или форматы, не требующие действий. Это позволяет отфильтровать верные, но не релевантные совпадения и сфокусироваться на действительно критичных находках. Комбинация таких списков с интеллектуальными моделями повышает общую точность: чередование детерминированного подхода (EDM) для строгого соответствия и ML-моделей для понимания контекста даёт более точные результаты.

Существующие решения: AWS Macie, Google Cloud DLP (Sensitive Data Protection), Microsoft Purview [22] и другие предлагают встроенные «идентификаторы» конфиденциальных типов данных, настройки порогов и списки исключений. Они используют шаблоны, эвристики и модели ML для поиска. Практический пример – Macie анализирует выборку объектов и генерирует находки по чувствительным данным, а пользователи могут уточнять шаблоны и исключения, минимизируя ложные тревоги. В Google DLP можно настроить «исключения» и приоритеты сканирования, чтобы избежать широких совпадений.

B. Вызовы и ограничения: несмотря на прогресс, проблемы сохраняются.

Регулярные выражения и шаблоны по-прежнему нередко дают ложные срабатывания на похожие, но не релевантные данные. Модели ML чувствительны к качеству обучающей выборки и могут неправильно обобщать, что требует тщательной подготовки данных. Точный подбор обученных моделей для разных языков остаётся сложной задачей. Кроме того, баланс между чувствительностью и точностью часто требует ручной настройки и итераций: избыточное снижение ложных срабатываний может увеличить пропуски реальной конфиденциальной информации.

2. Преодоление сложности многопользовательских и мультиоблачных сред. В крупных организациях возникает потребность защищать данные в десятках и сотнях учетных записей и облачных сервисов одновременно. Фрагментация среды (множество регионов и платформ) создаёт сложности единого мониторинга и классификации чувствительной информации. Разные облачные провайдеры имеют свои API и модели безопасности, что затрудняет централизованное управление. Существующие решения: Microsoft Purview имеет список «sensitive information types» для глобальных законодательных требований (например, GDPR, CCPA) и может применять их ко всем источникам данных. Google DLP (через Security Command Center) позволяет подключить AWS, предоставить ей права на чтение и запустить профилирование данных.

C. Вызовы и ограничения: основные проблемы

Основные проблемы – это несоответствие метаданных и политик между провайдерами, дублирование усилий и сложность нормализации результатов. Классификация, выполненная в AWS Macie, может отличаться по семантике от классификации в Google DLP, поэтому без центральной системы сложно интерпретировать общую картину. Настройка соединений и прав доступа для мультиоблачных сканеров требует значительных усилий. Помимо технических сложностей, неизменной становится проблема согласования политик безопасности и приоритетов между разными командами и облаками. Например, настройки Macie в AWS и настройки DLP в GCP или Azure надо приводить к единому стандарту, а это непросто. Несмотря на доступные инструменты, часто требуется сторонняя платформа DSPM или SIEM, чтобы централизовать результаты сканирования и обеспечить единообразное управление обнаружением чувствительных данных.

D. Обнаружение и классификация в реальном времени без значительного замедления процессов работы облачных сред.

Критическим требованием является своевременное выявление конфиденциальных данных без серьёзного ухудшения производительности рабочих нагрузок. Однако постоянное сканирование всех данных при срабатывании может существенно замедлить системы. К примеру, попытка запускать DLP-сканирование на каждом документе во всех облачных сервисах может привести к резкому снижению производительности. Существующие решения: многие продукты поддерживают сочетание реактивных и превентивных методов. Например, Macie может генерировать находки только по незащищённым объектам, а задачи запускаются по расписанию. Для потоковых данных используются другие сервисы. В целом данная проблема часто решается через интеграцию с системами управления событиями и автоматическими ответами.

E. Вызовы и ограничения: главная трудность

Главная трудность – это баланс между полнотой сканирования и скоростью обработки. Полное, сканирование всего объема в реальном времени приведёт к перегрузкам и задержкам в работе системы. Поэтому приходится жертвовать либо частотой проверок, либо охватом: например, использовать частичную выборку, каскадное обнаружение (сначала метаданных, затем контента) или сосредоточиться на критичных ресурсах. Кроме того, задержка в обработке или ошибки обнаружения могут приводить к пропускам в защите. Отдельно следует отметить, что аналитика потоковых данных пока недостаточно развита в рамках DLP-продуктов и часто требует специализированных решений.

F. Обработка неструктурированных данных.

Большая часть конфиденциальной информации хранится в неструктурированных форматах: текстовых документах, презентациях, скан-копиях, изображениях, а также аудио- и видеозаписях. Эффективное обнаружение в таких источниках требует специализированных методов. Современные DLP-системы работают с широким спектром форматов: текст, Office-документы, PDF, но содержат также возможности OCR для изображений.

Существующие решения: Google DLP и Microsoft Purview имеют встроенные OCR: они умеют распознавать текст на изображениях и встраивать его в анализ. Macie и аналогичные сервисы обрабатывают документы как потоки текста. Некоторые коммерческие решения используют машинное обучение для извлечения смысловых признаков из неструктурированных файлов. В Microsoft Purview в состав коннекторов входят средства извлечения метаданных и текста из файлов разных форматов. Вызовы и ограничения: обработка неструктурированной информации наиболее сложная проблема. OCR может ошибаться при низком качестве изображения, нестандартных шрифтах или использовании разных языков. Аудио-транскрипция неточно передаёт акценты и жаргон. Таким образом, точность обнаружения персональной информации в таких данных ниже, а сами операции существенно медленнее и более затратное. Кроме того, многие алгоритмы ML обучены на табличных данных или обработанном тексте и хуже справляются с шумами.

G. Повышение прозрачности и объяснимости системы.

С ростом применения ML-моделей в классификации данных становится актуальной задача обеспечения их прозрачности. «Черный ящик» алгоритмов усложняет доверие к системе и аудит её решений. В критичных областях, особенно когда речь идёт о персональных данных, важна возможность объяснить, почему конкретный объект был отнесён к конфиденциальной информации. Регуляторы GDPR, HIPAA и др. фактически требуют такой прозрачности: XAI-модели (Explainable AI) позволяют предоставлять понятные «объяснения» выводов алгоритмов, что необходимо для аудита и для соблюдения прав субъектов данных. Как указывают эксперты, чёрный ящик ИИ считается «юридической проблемой», особенно в контексте GDPR/HIPAA, где «требуется прозрачность» в автоматизированных решениях по защите данных. Существующие решения: многие DLP-системы по умолчанию снабжены логами и контекстной информацией: AWS Macie в отчетах указывает найденные фрагменты и тип информации, Google DLP возвращает координаты в документе, а Microsoft Purview хранит «метки чувствительности» с описанием причин. Для ML-моделей применяются техники пост-объяснения (LIME, SHAP и др.), но такие возможности пока редки в коммерческих продуктах. В целом отрасль движется к внедрению XAI-механизмов: либо через использование более простых моделей (решающие деревья вместо нейросетей), либо через генерацию «отчетов об обосновании» для каждой находки. Вызовы и ограничения: открытые технологии XAI всё ещё активно развиваются. С одной стороны, добавление объяснений может ухудшить производительность или точность (например, замена нейросети на дерево решений), а с другой – сами ML-модели часто слишком сложны для полного понимания. Ещё одна проблема – отсутствие единых стандартов: какие именно объяснения нужны аудиторам или пользователям, – поэтому производители лишь начинают экспериментировать. Тем не менее требования регуляторов и внутренний контроль над данными всё больше требуют прозрачности, что оправдывает инвестирование в объяснимые системы.

H. Стоимость и управление ресурсами.

Разработка более эффективных алгоритмов управления ресурсами, которые могут динамично выделять вычислительные мощности в зависимости от рабочей нагрузки и сложности задач классификации данных. Существующие решения: некоторые платформы оптимизируют вычисления через архитектурные решения. Например, использование «small language model» (SLM) – облегчённую языковую модель, работающую на CPU – вместо тяжёлых LLM. Такая модель даёт приближение представления документов и классифицирует их примерно за 200 мс (один файл) и ˜40 мс (AI Mesh). Это позволяет экономить ресурсы и ускорять сканирование больших библиотек файлов. Также применяются квоты на сканирование (как в Macie), а Cloud DLP опирается на облачные вычисления «по требованию», что даёт эластичность. Вызовы и ограничения: основные сложности – это выбор компромисса между стоимостью и полнотой защиты. Слишком дорогое общее сканирование может неприемлемо тормозить бизнес-процессы. Частые сканирования с высокой точностью требуют больших затрат процессорного времени, особенно при использовании глубоких нейросетей или OCR.

I. Улучшение соблюдения нормативных требований и суверенитета данных.

Внедрение автоматической проверки соответствия

нормативным требованиям в процессе классификации данных поможет гарантировать, что конфиденциальная информация обрабатывается в соответствии с актуальными законами без вмешательства человека. Усиление систем для мониторинга и соблюдения политик местоположения данных обеспечит более строгий контроль над тем, где хранятся и обрабатываются конфиденциальные данные в облачных средах. Существующие решения: инструменты классификации обычно интегрируются с механизмами шифрования и меток соответствия. Кроме того, облачные платформы сертифицированы по стандартам безопасности. Наконец, отчёты о сканировании и метаданные обеспечивают доказательства аудиту – какие данные где и как были защищены. Главная сложность здесь – разноплановость нормативных требований. Каждая юрисдикция предъявляет свои специфические правила к тому, какие данные считать личными и как их обрабатывать. Обновления законодательства требуют быстро адаптировать классификаторы (добавлять новые шаблоны или модели), что может отставать от реального времени.

VII Заключение.

Текущие решения для автоматизированного обнаружения и классификации чувствительных данных активно развиваются, но сложность современных облачных ландшафтов диктует новые требования. Мы видим явный тренд к сочетанию традиционных методов с ML/NLP. Однако на каждом из рассмотренных этапов – точность, мультиоблачность, детектирование в реальном времени, обработка «больших данных, прозрачность и стоимость – остаются значительные вызовы. Во-первых, повышение точности требует более продвинутых моделей и обхода эффектов «черного ящика». Во-вторых, задачи мультиоблака приводят к необходимости централизации данных и унификации политик, что является сложной задачей. В-третьих, баланс между полнотой сканирования и задержками процессов требует новых архитектур. Четвёртое, сложные форматы (изображения, аудио) расширяют распознавание, но повышают требования к вычислениям и обучению. Пятое, объяснимость и прозрачность становятся решающим аспектом для доверия пользователей и регуляторов: системы должны давать понятные обоснования своих находок, иначе их эффективность под вопросом. Таким образом, необходимо усиленное внимание к разработке объяснимых механизмов классификации. Рост юридических требований и потребности бизнеса в аудите защиты данных делают объяснимость приоритетом. Улучшение видимости критериев срабатывания позволит быстрее выявлять ошибки, настраивать алгоритмы и обосновывать перед контролирующими органами, что организация действительно понимает и контролирует конфиденциальные данные. В целом можно заключить, что автоматизация обнаружения конфиденциальных данных в облаках при всей своей эффективности должна быть подкреплена механизмами прозрачности и объяснимости. Это повысит доверие к системе, упростит настройку и аудит, а также поможет соблюсти всё более строгие требования информационной безопасности и суверенитета данных.

Благодарности

Работа выполнена в рамках развития магистерской программы Кибербезопасность [23] (совместно с ПАО Сбербанк) на факультете ВМК МГУ. Работа продолжает серию публикаций, начатых статьей [24].

Библиография

[1] ENISA (European Union Agency for Cybersecurity). "Cloud Security: Key Recommendations."European Union Agency for Cybersecurity, 2018.
[2] OWASP. "OWASP Cloud-Native Application Security Top 10."OWASP Foundation, 2021.
[3] Bhardwaj, Sushil, Leena Jain, and Sandeep Jain. "Cloud computing: A study of infrastructure as a service (IAAS)."International Journal of engineering and information Technology 2.1 (2010): 60-63.
[4] Boniface, Michael, et al. "Platform-as-a-service architecture for real-time quality of service management in clouds."2010 fifth nternational conference on internet and web applications and services. IEEE, 2010.
[5] Tsai, WeiTek, XiaoYing Bai, and Yu Huang. "Software-as-a-service (SaaS): perspectives and challenges."Science China Information Sciences 57 (2014): 1-15.
[6] Hussein, Mohamed K., Mohamed H. Mousa, and Mohamed A. Alqarni. "A placement architecture for a container as a service (CaaS) in a cloud environment."Journal of Cloud Computing 8 (2019): 1-15.
[7] Abdurachman, Edi, Ford Lumban Gaol, and Benfano Soewito. "Survey on threats and risks in the cloud computing environment."Procedia Computer Science 161 (2019): 1325-1332.
[8] Alshammari, Abdulaziz, et al. "Security threats and challenges in cloud computing."2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud). IEEE, 2017.
[9] Консультант плюс. "Перечень нормативных актов, относящих сведения к категории ограниченного доступа".
[10] Ku˘zina, Vjeko, et al. "CASSED: context-based approach for structured sensitive data detection."Expert systems with applications 223 (2023): 119924.
[11] Huang, Ziyi. "Sensitive Information Detection Using HMM&SVM."Proceedings of the 2021 3rd International Conference on Intelligent Medicine and Image Processing. 2021.
[12] Ali, Munwar, and Low Tang Jung. "Confidentiality based file attributes and data classification using tsf-knn."2015 5th International Conference on IT Convergence and Security (ICITCS). IEEE, 2015.
[13] Xu, Guosheng, et al. "Detecting sensitive information of unstructured text using convolutional neural network."2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). IEEE, 2019.
[14] Ahmed, Hadeer, et al. "Automated detection of unstructured context-dependent sensitive information using deep learning."Internet of Things 16 (2021): 100444.
[15] Hulsebos, Madelon, et al. "Sherlock: A deep learning approach to semantic data type detection."Proceedings of the 25th ACM SIGKDD International Conference on knowledge discovery & data mining. 2019.
[16] Zhang, Dan, et al. "Sato: Contextual semantic type detection in tables."arXiv preprint arXiv:1911.06311 (2019).
[17] Masdari, Mohammad, and Hemn Khezri. "A survey and taxonomy of the fuzzy signature-based intrusion detection systems."Applied Soft Computing 92 (2020): 106301.
[18] Qin, Biao, et al. "A rule-based classification algorithm for uncertain data."2009 IEEE 25th international conference on data engineering. IEEE, 2009.
[19] [Phoha, Shashi, et al. "Context-aware dynamic data-driven pattern classification." Procedia Computer Science 29 (2014): 1324-1333.
[20] Cloud Data Loss Prevention https://cloud.google.com/security/products/dlp?hl=en Retrieved: May, 2025
[21] AWS Macie https://aws.amazon.com/macie/ Retrieved: May, 2025
[22] Microsoft Purview https://www.microsoft.com/en-us/security/business/microsoft-purview Retrieved: May, 2025

[23] Сухомлин, Владимир Александрович. "Концепция и основные характеристики магистерской программы" Кибербезопасность" факультета ВМК МГУ." International Journal of Open Information Technologies 11.7 (2023): 143-148.
[24] Намиот, Д. Е., Е. А. Ильюшин, and И. В. Чижов. "Искусственный интеллект и кибербезопасность." International Journal of Open Information Technologies 10.9 (2022): 135-147.

Automated Detection and Classification of Sensitive Data in Cloud Environments

Maxim Egorov, Dmitry Namiot

***Abstract* - The introduction of cloud technologies is inevitably accompanied by an increase in information security risks. One of the most serious problems faced by cloud users is the detection and prevention of data leaks in the cloud infrastructure. Confidential data is information that requires protection from unauthorized access, modification or distribution, as it is highly sensitive and can cause damage to the owner or third parties in the event of leakage or abuse. This data may concern both individuals and organizations and is often regulated by legal acts in order to ensure their security and confidentiality. Compromise of confidential information (personal information, financial transactions, intellectual property), unauthorized access to sensitive data can lead to large-scale reputational and economic losses. According to all analytical reports, as well as analytical reviews by Gartner and Forrester, the number of cyberattacks targeting cloud platforms is constantly growing, and their complexity and sophistication are increasing. Accordingly, the issues of identifying confidential data in cloud environments are becoming extremely relevant.**

***Keywords* - confidential data, cloud computing, personal data.**

REFERENCES

[1] ENISA (European Union Agency for Cybersecurity). "Cloud Security: Key Recommendations."European Union Agency for Cybersecurity, 2018.
[2] OWASP. "OWASP Cloud-Native Application Security Top 10."OWASP Foundation, 2021.
[3] Bhardwaj, Sushil, Leena Jain, and Sandeep Jain. "Cloud computing: A study of infrastructure as a service (IAAS)."International Journal of engineering and information Technology 2.1 (2010): 60-63.
[4] Boniface, Michael, et al. "Platform-as-a-service architecture for real-time quality of service management in clouds."2010 fifth nternational conference on internet and web applications and services. IEEE, 2010.
[5] Tsai, WeiTek, Xiao Ying Bai, and Yu Huang. "Software-as-a-service (SaaS): perspectives and challenges."Science China Information Sciences 57 (2014): 1-15.
[6] Hussein, Mohamed K., Mohamed H. Mousa, and Mohamed A. Alqarni. "A placement architecture for a container as a service (CaaS) in a cloud environment."Journal of Cloud Computing 8 (2019): 1-15.
[7] Abdurachman, Edi, Ford Lumban Gaol, and Benfano Soewito. "Survey on threats and risks in the cloud computing environment."Procedia Computer Science 161 (2019): 1325-1332.
[8] Alshammari, Abdulaziz, et al. "Security threats and challenges in cloud computing."2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud). IEEE, 2017.
[9] Konsul'tant pljus. "Perechen' normativnyh aktov, otnosjashhih svedenija k kategorii ogranichennogo dostupa".
[10] Kuˇzina, Vjeko, et al. "CASSED: context-based approach for structured sensitive data detection."Expert systems with applications 223 (2023): 119924.
[11] Huang, Ziyi. "Sensitive Information Detection Using HMM&SVM."Proceedings of the 2021 3rd International Conference on Intelligent Medicine and Image Processing. 2021.
[12] Ali, Munwar, and Low Tang Jung. "Confidentiality based file attributes and data classification using tsf-knn."2015 5th International Conference on IT Convergence and Security (ICITCS). IEEE, 2015.
[13] Xu, Guosheng, et al. "Detecting sensitive information of unstructured text using convolutional neural network."2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). IEEE, 2019.
[14] Ahmed, Hadeer, et al. "Automated detection of unstructured context-dependent sensitive information using deep learning."Internet of Things 16 (2021): 100444.
[15] Hulsebos, Madelon, et al. "Sherlock: A deep learning approach to semantic data type detection."Proceedings of the 25th ACM SIGKDD International Conference on knowledge discovery & data mining. 2019.
[16] Zhang, Dan, et al. "Sato: Contextual semantic type detection in tables."arXiv preprint arXiv:1911.06311 (2019).
[17] Masdari, Mohammad, and Hemn Khezri. "A survey and taxonomy of the fuzzy signature-based intrusion detection systems."Applied Soft Computing 92 (2020): 106301.
[18] Qin, Biao, et al. "A rule-based classification algorithm for uncertain data."2009 IEEE 25th international conference on data engineering. IEEE, 2009.
[19] [Phoha, Shashi, et al. "Context-aware dynamic data-driven pattern classification." Procedia Computer Science 29 (2014): 1324-1333.
[20] Cloud Data Loss Prevention https://cloud.google.com/security/products/dlp?hl=en Retrieved: May, 2025
[21] AWS Macie https://aws.amazon.com/macie/ Retrieved: May, 2025
[22] Microsoft Purview https://www.microsoft.com/en-us/security/business/microsoft-purview Retrieved: May, 2025
[23] Suhomlin, Vladimir Aleksandrovich. "Koncepcija i osnovnye harakteristiki magisterskoj programmy" Kiberbezopasnost'" fakul'teta VMK MGU." International Journal of Open Information Technologies 11.7 (2023): 143-148.
[24] Namiot, D. E., E. A. Il'jushin, and I. V. Chizhov. "Iskusstvennyj intellekt i kiberbezopasnost'." International Journal of Open Information Technologies 10.9 (2022): 135-147.