УДК. 004.056 DOI: 10.34708/GSTOU.2022.18.98.012

РОЛЬ DLP-СИСТЕМ В ЗАЩИТЕ ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ НАУЧНЫХ УЧРЕЖДЕНИЙ

© Гацаев Б.Б., Зархматова М.Ш., Хашумов И.У., Абдулаев И.Х.
ГГНТУ им. акад. М.Д. Миллионщикова, г. Грозный

Невозможно представить образование будущего без конфиденциальности и конфиденциальных данных, которые необходимы для большинства научных учреждений. Именно поэтому защита этих данных требует внимания со стороны высшего руководства учреждения, администраторов и IT-менеджеров. Утечка данных оказывает негативное влияние на организации. Традиционные подходы к безопасности, такие как межсетевые экраны, не могут защитить данные от утечки. Системы предотвращения утечки данных (DLP) — это решения, которые защищают конфиденциальные данные от попадания в недоверенные руки. Эта статья является результатом изучения DLP-систем в качестве способа защиты интеллектуальной собственности научных учреждений, в которой также будет проводиться сравнение с другими системами безопасности и защиты данных.

Ключевые слова: *Предотвращение потери данных (DLP), утечка данных, сетевая уязвимость, безопасность данных университетов, снижение риска, устойчивость системы.*

DLP-технологии в университетах и других учреждениях. Широкое использование информационных технологий и облачных вычислений увеличило угрозы утечки, манипулирования и распространения данных. В то время как мир все больше полагается на облачные вычисления, проблемы безопасности и конфиденциальности данных, хранящихся и совместно используемых в распределенной сети, также возрастают и риски утечки информации. Многие предприятия были жертвами кражи, потери, утечки, манипулирования конфиденциальными бизнес-данными, и такая прямая и косвенная потеря конфиденциальных данных физическим или юридическим лицом представляет большую угрозу деловой репутации и доверию человека. Из причин утечки и потери данных наиболее часто сообщаемой причиной является случайное и небрежное поведение сотрудников при взаимодействии с общими данными или файлами, такими как конфиденциальная информация, деловые документы, деловые предложения, финансовые отчеты, контракты, интеллектуальная собственность и другая личная информация корпорации, которая передается через скрытый канал [1]. В таких случаях, когда общие файлы перемещаются по скрытому каналу, вероятность несанкционированного доступа к данным увеличивается, поскольку скрытые каналы формируются через общие ресурсы и обходят обычные механизмы защиты данных законного пользователя, что приводит к непреднамеренной утечке данных. Сообщается, что 90% утечек корпоративных данных можно было бы предотвратить за счет усовершенствования стратегий защиты данных и предотвращения их утечек. Обманные и мошеннические атаки, инициированные с использованием учетных данных законного инсайдера, являются основными причинами нарушения внутренней безопасности, за которыми следует утечка информации в результате инсайдерской атаки. Таким образом, очевидно, что информационная безопасность является первоочередной темой в области IT, целью которой является обеспечение безопасной вычислительной среды для индивидуальных и корпоративных пользователей [2].

В этой статье представлен критический обзор распределенной модели предотвращения утечки данных. В обзоре критических исследований были проанализированы распределенные системы обнаружения и предотвращения вторжений на основе DLP-систем с точки зрения их конструкции, возможностей и недостатков. В статье также представлен обзор исследований, в которых предложена модель распространения для обнаружения и предотвращения утечек

данных на основе мобильных агентов. Обсуждение различных методов проводится для анализа и сравнения этих систем.

Предотвращение потери данных (DLP). Утечка данных относится к несанкционированной передаче данных внутри организации внешнему адресату или получателю. Типы утечек данных обычно включают конфиденциальную информацию, интеллектуальную собственность, пользовательские и другие различные данные.

Рис. 1. Краткое изложение технических подходов к предотвращению и обнаружению утечки данных.

Учитывая нынешние строгие нормативные и правовые требования к защите интеллектуальной и личной информации, организации, включая университеты, вложили много времени и ресурсов в защиту своей информации от потенциального несанкционированного доступа и раскрытия конфиденциальной информации. Многие исследователи разработали различные контрмеры для борьбы с проблемами утечки данных, которые в совокупности известны как решения для предотвращения утечки данных (DLP). DLP можно определить, как любые системы или инструменты, которые идентифицируют, отслеживают и защищают данные [1]. Примеры таких данных включают:

1) Данные в движении — данные, передаваемые по проводам или по беспроводной сети.

2) Используемые данные — данные в конечных точках сети, такие как информация, хранящаяся на переносных жестких дисках.

3) Данные в состоянии покоя — данные, которые находятся в файловой системе, базах данных и других методах хранения.

Некоторые из широко доступных систем DLP:

1) Сетевой DLP. Сетевой DLP предназначен для обнаружения любых случаев утечки, связанных с данными в движении, путем определения того, передаются ли определенные важные файлы данных по сетям. Этот тип устройств DLP обычно подключается к оборудованию, такому как маршрутизаторы, коммутаторы, поддерживает несколько протоколов, таких как HTTP, FTP, P2P и SMTP.

2) Endpoint DLP: продукты Endpoint DLP представляют собой агенты или программное обеспечение, которые обычно находятся на терминалах конечных пользователей, таких как мобильные устройства и ноутбуки. Обычно Endpoint DLP используется для предотвращения

хранения конфиденциальной информации пользователями на съемных носителях, таких как USB-накопители и диски CD/ROM, а также для защиты от несанкционированной передачи конфиденциальной информации, когда пользователь не подключен к общедоступной точке доступа WI-FI. Программное обеспечение Endpoint DLP также может использовать шифрование диска, что предотвращает несанкционированный доступ к информации на потерянном или украденном ноутбуке.

3) Встроенная защита от потери данных. Встроенная защита от потери данных внедряется в определенные приложения для эффективного отслеживания потоков данных, определения ключевых слов или связанных шаблонов, относящихся к конфиденциальной информации, и блокирования любых подозрительных попыток утечки данных. Например, сканирование и отклонение исходящих сообщений электронной почты на наличие важных ключевых слов или вложений, ограничение печати электронных документов, защищенных авторским правом.

Некоторые из ключевых преимуществ DLP: предотвращение утечки данных, снижение стоимости расследования и ущерба репутации организаций, облегчение раннего обнаружения и смягчение рисков, а также повышение уровня комфорта пользователей.

DLP — это решение многогранной проблемы, которая касается того, что несет ответственность за потерю данных и что следует делать, чтобы избежать этой ситуации. В прошлом было представлено множество методов обнаружения и предотвращения утечек данных, которые так или иначе обеспечивают решение проблемы или каким-либо образом не являются абсолютным решением.

Обнаружение утечки данных в организациях. Цифровые водяные знаки уже давно используются для аутентификации и обеспечения целостности несущего сигнала. Кроме того, они также используются для идентификации личности владельца, таким образом, они определяют источник подделки или утечки данных и предотвращают незаконные манипуляции с данными. По этой причине следы водяных знаков используются для обнаружения манипуляций с данными и для нахождения утечки информации. Но проблема с техникой водяных знаков заключается в том, что когда водяной знак встраивается в исходный код, исходный код не остается в своем первоначальном виде, а видоизменяется. С другой стороны, если исходник нельзя модифицировать, в него нельзя добавить водяной знак. SELinux и ColoredLinux — это два метода, основанные на водяных знаках, которые используются для обнаружения утечки данных. ColoredLinux является преемником и формируется за счет модификаций SELinux. ColoredLinux генерирует слепые водяные знаки на основе тегов, предоставленных файлу (разрешение на доступ к файлу) файловой системой. Таким образом, при доступе к файлу он проверяет тег с водяным знаком и соответственно предоставляет доступ. Таким образом, несанкционированный доступ исключается, если тег не соответствует водяному знаку файла, в противном случае SELinux обеспечивает контроль доступа для пользователя, обращающегося к файлу. Такие механизмы хорошо работают для закрытых систем, где одна система не подключена к внешним машинам, а модифицированная операционная система выполняется на каждой системе. Однако, это плохо работает для открытой системы, где система подключена к нескольким системам с нецветными (модифицированными) операционными системами. В таких случаях внутренние атаки могут происходить по скрытым каналам, оставляя файлы измененными и незамеченными. Был предложен новый гибридный метод с использованием мобильных агентов и схемы цветных водяных знаков, который определяет тесную связь между защитной меткой и слепым водяным знаком и эффективно противостоит атакам по внутренним и скрытым каналам.

В другом исследовании сообщается, что идентификация источника утечки данных осуществляется с помощью технологии происхождения данных, которая идентифицирует законных владельцев данных (найденных в базе данных) и, таким образом, идентифицирует источники утечки данных. Был предложен еще один метод, использующий вставку поддельных данных для повышения вероятности выявления утечки данных. Небольшой модификацией методики является добавление вероятности вины пользователя, что улучшает

алгоритм распространения файлов. Одной из основных причин утечки данных является участие человека, т.е. умысел злонамеренных работников. Сообщается, что 11% сотрудников продавали бизнес-информацию с целью получения прибыли посредством несанкционированного доступа. По этой причине были представлены две модели определения утечки данных: модель наблюдателя и модель вины. Модель наблюдателя анализирует несанкционированный доступ корпоративного сотрудника, а модель вины использует вероятность вины для анализа того, передает ли подозреваемый пользователь этот файл какой-либо третьей стороне. Эти две модели предлагаются для расчета вероятности участия пользователя в утечке данных путем наблюдения пересечения между утечкой данных и доступом к данным подозреваемого пользователя.

Перехват утекающих данных. Несмотря на то, что существуют десятки алгоритмов для выявления утечки данных, приоритетным является выбор методов, которые максимально смягчают ситуации, приводящие к проблеме утечки данных. Критерии релевантности атак на данные и описание инцидентов с потерей данных определены в системах, в которых построено понятие предотвращения утечки данных и информации (DLP & ILP) в области информационной безопасности. Для защиты организационной информации от внутренних атак раскрытия информации представлена модель управления доверием. Также существует модель надежности для предотвращения утечки данных, в которой учитывается надежность пользователя, которая рассчитывается на основе истории пользователя, а затем его преднамеренное поведение утечки рассчитывается на основе его надежности и истории доступа к файлам. Перекрытие между ними указывает на преднамеренное поведение пользователя при утечке. Также дистрибутив оценивает уязвимость платформы пользователя. На основе этой коллективной оценки рисков распространитель принимает решение о том, следует ли распространять файлы среди этих пользователей.

На рис. 2 представлена гибридная платформа, которая сочетает в себе мобильные агенты и DLP для обнаружения и предотвращения утечек данных путем разработки этой платформы в файловой системе уровня ядра операционной системы. Он предназначен для преодоления слабых мест, возникающих при отслеживании действий по утечке данных. Этот предлагаемый механизм основан на файловой подсистеме уровня ядра, которая ранее была доступна в стандартной операционной системе для выявления потенциальных утечек конфиденциальных данных. Эта гибридная структура сокращает административную работу и предоставляет возможность изменять и добавлять функции обнаружения потери данных, а также объединять их в модули с использованием управляющего агента. Сообщается, что этот подход дает наименьший процент ложноположительных результатов из-за его способности идентифицировать неизвестные атаки.

Предлагаемый метод эффективно идентифицирует потенциальные источники утечки как со стороны скрытого канала, так и со стороны инсайдеров.

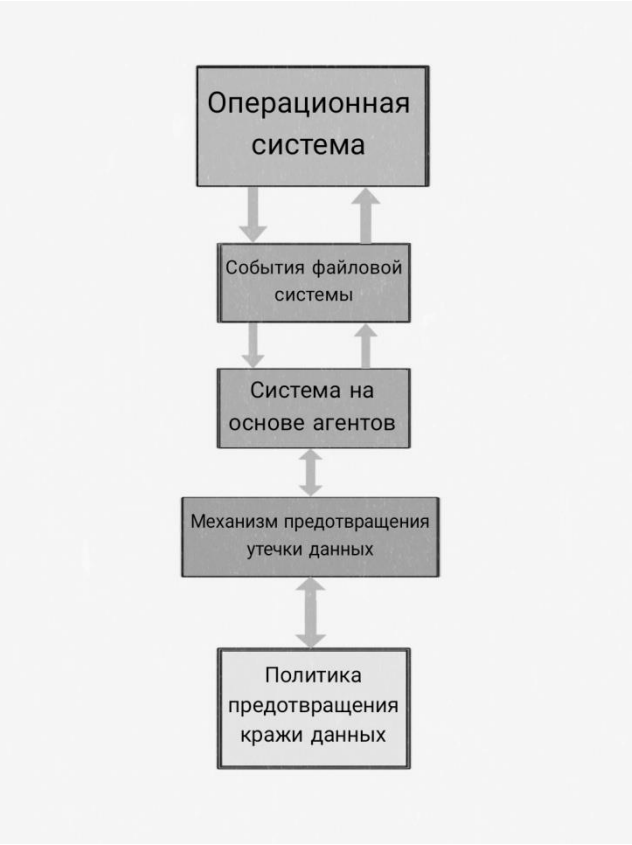

Рис. 2. Архитектура обнаружения и предотвращения утечки информации на основе мобильных агентов и DLP

Ограничения в существующих моделях обнаружения и предотвращения утечки данных. Несмотря на то, что была предложена дюжина решений для обнаружения и предотвращения утечки данных, не было найдено решения, обеспечивающего полную защиту от проблемы. Поэтому в существующих решениях обнаруживаются некоторые ограничения. Одним из выявленных недостатков является отсутствие общей основы для обнаружения и предотвращения утечки данных, поэтому было предложено множество решений, но ни одно из них не имеет общего рабочего процесса для базового понимания проблемы. Еще одним ограничением является отсутствие практичной и функциональной основы распределенной системы для решения проблемы обнаружения и предотвращения утечки данных. Кроме того, существующие предлагаемые решения были смоделированы в операционной системе Linux, в то время как большая часть IT-специалистов состоит из людей, использующих операционную систему Windows. Кроме того, наблюдается ограниченное использование методов машинного обучения, т.е., только метод опорных векторов (SVM). Наконец, также отмечается, что в некоторых случаях не предоставляется поддержка для сценариев предупреждений об утечке данных.

В большинстве рассмотренных идей показано, что для распределенной системы не было реализовано практической/функциональной системы обнаружения и предотвращения утечки информации. Также в них отсутствует разработка общей основы для обнаружения и предотвращения утечки данных на различных операционных системах. Большинство работ, выполняемых в DLP, моделируются в операционных системах Linux (вместо ОС Windows с наибольшим количеством пользователей) [3].

Таким образом, это исследование мотивировано необходимостью преодоления этих ограничений посредством новой интеграции системы Mobile-Agent с системой обнаружения и предотвращения утечки информации, построенной на файловой системе уровня ядра операционной системы; как способ использования их соответствующих уникальных

преимуществ для улучшения и устранения слабых мест в задаче мониторинга деятельности системы против утечек информации.

Рекомендуемые решения. Обработка данных во многом зависит от категории, к которой она принадлежит с точки зрения использования данных, их остановки или движения. Поскольку организации содержат объемные данные, которые трудно классифицировать вручную по разным причинам, это создает серьезную проблему для обработки данных и предотвращения их утечки. Многие корпоративные организации уже давно пытаются решить проблему несанкционированного доступа, потери и утечки данных, поскольку для любой организации крайне важно сохранить свое конкурентное преимущество и поддерживать постоянные отношения со своими клиентами. Общей стратегией организаций является использование механизма защиты от потери или утечки данных (DLP), который представляет собой набор технологий для мониторинга доступа и передачи конфиденциальных данных и защиты их от любого несанкционированного доступа [5].

Можно сказать, что существует четыре основных категории подходов, которые используются для обнаружения и предотвращения утечки данных. Специальные механизмы на основе DLP предотвращают несанкционированный доступ к данным, а доступ и передача данных возможны только для законных пользователей. Типичные методы решений DLP включают статистические методы на основе машинного обучения, сопоставление с образцом, методы на основе ключевых слов. Другой распространенный и простейший метод проверки несанкционированного доступа заключается в использовании механизмов контроля доступа и алгоритмов шифрования, которые предотвращают атаки изнутри или снаружи среды [6]. Некоторые традиционные механизмы обнаружения и предотвращения утечки данных включают брандмауэр, процедуры и процессы обнаружения вторжений, а также антивирусы. Стандартные процедуры не работают в сложной среде и часто менее надежны в некоторых сценариях, и по этой причине другие методы, такие как методы приманки, а также методы машинного обучения лучше подходят для таких задач. Несмотря на множество решений, самая большая проблема заключается в том, что потеря данных может произойти по разным причинам, например, через сеть или устройства хранения информации, во время обмена файлами P2P, через интернет или при передаче информации по электронной почте. Таким образом, предотвращение потери данных является многогранной проблемой, и DLP может стать готовым к использованию решением, но требует больших усилий для создания серьезного решения, основанного на тщательной подготовке и постоянном обслуживании [4].

Вывод. Технологический прогресс и преобладающее использование облачных и распределенных вычислений увеличили проблемы безопасности, связанные с данными и IT, а угрозы, связанные с утечкой данных в такой среде, также многократно увеличились. В этом обзоре исследований критически проанализированы решения, предложенные за последнее десятилетие для обнаружения и предотвращения утечки данных. Цель различных исследований была описательной, и многие исследования не предоставили существенного и эффективного решения для обнаружения и предотвращения утечки данных для распределенных систем. Хотя различные предлагаемые решения и улучшали тот или иной аспект для предотвращения утечки данных в различных научных организациях, а также для выявления вредоносных источников и пользователей, данные решения не являются полными и абсолютными. Однако, некоторые предложенные решения лишены теории, хотя в некоторых сценариях они показали себя хорошо. Поэтому делается вывод, что системы DLP требуют значительных исследований для разработки более совершенных систем, чтобы они окончательно закрепились как средство защиты данных в различных научных и других организациях. Кроме того, они требуют постоянных исследований и обслуживания до тех пор, пока не будет разработана приемлемая и всеобъемлющая структура.

ЛИТЕРАТУРА

1. Крылова, С. Л. DLP - эффективная защита электронного документооборота / С. Л. Крылова, Н. А. Сипатров // Материалы XXI научно-практической конференции молодых ученых, аспирантов и студентов Национального исследовательского Мордовского государственного университета им. Н.П. Огарёва: В 3-х частях, Саранск, 22–29 мая 2017 года / Составитель А.В. Столяров. Ответственный за выпуск П.В. Сенин. Том Часть 1. – Саранск: Национальный исследовательский Мордовский государственный университет им. Н.П. Огарёва, 2017. – С. 447-453. URL: https://elibrary.ru/item.asp?id=30719942

2. Гречанная А.Ю., Тастенов А.Д. DLP-системы и их роль в защите от утечек конфиденциальной информации // Наука и техника Казахстана. 2015. №3-4. URL: https://cyberleninka.ru/article/n/dlp-sistemy-i-ih-rol-v-zaschite-ot-utechek-konfidentsialnoy-informatsii (дата обращения: 11.11.2022).

3. Зарубин А.В., Смирнов М.Б., Харитонов С.В., Денисов Д.В. Основные драйверы и тенденции развития DLP-систем в Российской Федерации // Прикладная информатика. 2020. Т. 15. – № 3(87). С. 75-90. URL: https://elibrary.ru/item.asp?id=43063644

4. Копырулина, О.А. Защита конфиденциальной информации на предприятиипосредством DLP - технологии / О. А. Копырулина, Е. В. Устюжанин // Вестник современных исследований. 2018. № 10.1(25). С. 367-369. https://elibrary.ru/item.asp?id=36310918

5. Минцаев М. Ш., Алисултанова Э. Д., Усамов И. Р. Технологии машинного обучения в современной образовательной среде // Вестник ГГНТУ. Гуманитарные и социально-экономические науки. Том: XVIII, номер: 3 (29), Грозный 2022

6. Алисултанова Э. Д., Магомадова А. Р., Мусаева М. С-А. Системы мониторинга тенденции развития дистанционного обучения// Вестник ГГНТУ. Гуманитарные и социально-экономические науки. Том: XVII, номер: 2 (24), Грозный 2021

THE ROLE OF DLP SYSTEMS IN THE PROTECTION OF SCIENTIFIC INSTITUTIONS INTELLECTUAL PROPERTY

© B.B. Gatsaev, M.Sh. Zarkhmatova, I.U. Khashumov, I. Kh. Abdulaev
GSTOU named after acad. M.D. Millionshchikov, Grozny

It is impossible to imagine the education of the future without confidentiality and confidential data, which are necessary for most scientific institutions. That is why the protection of this data requires attention from the top management of the institution, administrators and IT managers. Data leakage has a negative impact on organizations. Traditional security approaches such as firewalls cannot protect data from leakage. Data Leak Prevention (DLP) systems are solutions that protect sensitive data from falling into untrusted hands. This article is the result of an examination of DLP systems as a way to protect the intellectual property of scientific institutions, which will also be compared with other security and data protection systems.

Keywords: *Data Loss Prevention (DLP), data leakage, network vulnerability, university data security, risk mitigation, system resilience.*