СӘТБАЕВ УНИВЕРСИТЕТІ SATBAYEV UNIVERSITY

Дисциплина «Этичный хакинг и противодействие взлому»

Лекция 11

Технология защиты конфиденциальной информации от внутренних угроз

Преподаватель: **Батыргалиев Асхат Болатканович, PhD, ассоц.проф. кафедры «Кибербезопасность, обработка и хранение информации»**

askhat.b.b@gmail.com

Содержание

1. Понятие технологии защиты конфиденциальной информации от внутренних угроз
2. Технологии детектирования конфиденциальной информации
3. Архивирование информации, проходящей через технические каналы утечки
4. Шифрование информации на всех точках сети. Контроль доступа к сети, приложениям и информации
5. Архитектура IPC-систем

По завершению лекции Вы будете знать:

1. Понятие технологии защиты конфиденциальной информации от внутренних угроз
2. Сведения о технологии детектирования конфиденциальной информации
3. Сведения об архивировании информации, проходящей через технические каналы утечки
4. Сведения о шифровании информации на всех точках сети и контроле доступа к сети, приложениям и информации
5. Общую архитектуру IPC-систем

Понятие технологии защиты конфиденциальной информации от внутренних угроз

Information Protection and Control (IPC) - технология защиты конфиденциальной информации от внутренних угроз. Решения класса IPC предназначены для защиты информации от внутренних угроз, предотвращения различных видов утечек информации, корпоративного шпионажа и бизнес-разведки. Термин IPC соединяет в себе две основные технологии: шифрование носителей информации на всех точках сети и контроль технических каналов утечки информации с помощью технологий Data Loss Prevention (DLP). Контроль доступа к сети, приложениям и данным является возможной третьей технологией в системах класса IPC. IPC включает в себя решения класса DLP, системы шифрования корпоративной информации и контроля доступа к ней.

Технология IPC является логическим продолжением технологии DLP и позволяет защищать данные не только от утечек по техническим каналам, то есть инсайдеров, но и от несанкционированного доступа пользователей к сети, информации, приложениям и в тех случаях, когда непосредственный носитель информации попадает в руки третьих лиц. Это позволяет не допускать утечки и в тех случаях, когда инсайдер или не имеющий легального доступа к данным человек получает доступ к непосредственному носителю информации. Например, достав жесткий диск из персонального компьютера, инсайдер не сможет прочитать имеющуюся на нем информацию. Это позволяет не допустить компрометацию конфиденциальных данных даже в случае потери, кражи или изъятия.

Задачи IPC

Основной задачей IPC-систем является предотвращение передачи конфиденциальной информации за пределы корпоративной информационной системы. Такая передача (утечка) может быть намеренной или ненамеренной. Практика показывает, что большая часть (более 75 %) утечек происходит не по злому умыслу, а из-за ошибок, невнимательности, безалаберности, небрежности работников - выявлять подобные случаи намного проще. Остальная часть связана со злым умыслом операторов и пользователей информационных систем предприятия, в частности промышленным шпионажем, конкурентной разведкой. Очевидно, что злонамеренные инсайдеры, как правило, стараются обмануть анализаторы IPC и прочие системы контроля.

Дополнительные задачи систем класса IPC:

- предотвращение передачи вовне не только конфиденциальной, но и другой нежелательной информации (обидных выражений, спама, эротики, излишних объёмов данных и т.п.);
- предотвращение передачи нежелательной информации не только изнутри наружу, но и снаружи внутрь информационной системы организации;
- предотвращение использования работниками Интернет-ресурсов и ресурсов сети в личных целях;
- защита от спама;
- защита от вирусов;
- оптимизация загрузки каналов, уменьшения нецелевого трафика;
- учет рабочего времени и присутствия на рабочем месте;
- отслеживание благонадёжности сотрудников, их политических взглядов, убеждений, сбор компромата;
- архивирование информации на случай случайного удаления или порчи оригинала;
- защита от случайного или намеренного нарушения внутренних нормативов;
- обеспечение соответствия стандартов в области ИБ и действующего Законодательства.

Контроль каналов утечки информации (DLP)

Технология DLP в IPC поддерживает контроль следующих технических каналов утечки конфиденциальной информации:

- корпоративная электронная почта,
- веб-почта,
- социальные сети и блоги,
- файлообменные сети,
- форумы и другие интернет-ресурсы, в том числе выполненные на AJAX-технологии,
- средства мгновенного обмена сообщениями (ICQ, Mail.Ru Агент, Skype, AOL AIM, Google Talk, Yahoo Messenger, MSN Messenger и прочее),
- p2p-клиенты,
- периферийные устройства (USB, LPT, COM, WiFi, Bluetooth и прочее),
- локальные и сетевые принтеры.

Технологии DLP в IPC поддерживают контроль в том числе следующих протоколов обмена данными:

- FTP;
- FTP-over-HTTP;
- FTPS;
- HTTP;
- HTTPS (SSL);
- NNTP;
- POP3;
- SMTP.

Технологии детектирования конфиденциальной информации

Сигнатуры

Самый простой метод контроля - поиск в потоке данных некоторой последовательности символов. Иногда запрещенную последовательность символов называют «стоп-выражением», но в более общем случае она может быть представлена не словом, а произвольным набором символов, например, определенной меткой. Если система настроена только на одно слово, то результат её работы - определение 100% совпадения, т.е. метод можно отнести к детерминистским. Однако чаще поиск определенной последовательности символов все же применяют при анализе текста. В подавляющем большинстве случаев сигнатурные системы настроены на поиск нескольких слов и частоту встречаемости терминов.

К достоинствам метода можно отнести простоту пополнения словаря запрещенных терминов и очевидность принципа работы, а также то, что это самый верный способ, если необходимо найти соответствие слова или выражения на 100%. **Недостатки** же становятся очевидными после начала промышленного использования при поиске утечек и настройке правил фильтрации. Большинство производителей DLP-систем работают для Западных рынков, а английский язык очень «сигнатурен» - формы слов чаще всего образуются с помощью предлогов без изменения самого слова. В русском языке все гораздо сложнее, так как у нас есть приставки, окончания, суффиксы. Для примера можно взять слово «ключ», которое может означать как «ключ шифрования», «ключ от квартиры», «родник», «ключ или PIN-код от кредитной карты», так и множество других значений. В русском языке из корня «ключ» можно образовать несколько десятков различных слов. Это означает, что если на Западе специалисту по защите информации от инсайдеров достаточно ввести одно слово, в русскоязычном регионе специалисту придется вводить пару десятков слов и затем еще изменять их в шести различных кодировках. Реальное применение этого метода требует наличие лингвиста или команды лингвистов как на этапе внедрения, так и в процессе эксплуатации и обновления базы. Несомненным недостатком является и то, что «сигнатуры» неустойчивы к примитивному кодированию, например, заменой символов на похожие по начертанию.

Технологии детектирования конфиденциальной информации

«Цифровые отпечатки» (Digital Fingerprints или DG)

В данной технологии используются различного рода хеш-функции образцов конфиденциальных документов. Общий сценарий действия такой: набирается база образцов конфиденциальных документов. Суть работы DG довольно проста и часто этим и привлекает: DLP/IPC-системе передается некий стандартный документ-шаблон, из него создается цифровой отпечаток и записывается в базу данных DF. Далее в правилах контентной фильтрации настраивается процентное соответствие шаблону из базы. Например, если настроить 75% соответствие «цифровому отпечатку» договору поставки, то при контентной фильтрации DLP обнаружит практически все договоры этой формы. Иногда, к этой технологии относят и системы вроде «Антиплагиата», однако последняя работает только с текстовой информацией, в то время как технология «цифровых отпечатков», в зависимости от реализации, может работать и различным медийным контентом и применяться для защиты авторских прав и препятствию случайному или намеренному нарушению законов и нормативов информационной безопасности.

К достоинствам технологии можно отнести простоту добавления новых шаблонов, довольно высокую степень детектирования и прозрачность алгоритма технологии для сотрудников подразделений по защите информации. Специалистам СБ и ИБ не надо думать о «стоп-выражениях» и прочей лингвистике, тратить много времени на анализ потенциально опасных словоформ и вбивать их в базу, тратить ресурсы на внедрение и поддержку лингвистической базы. Основным недостатком, который на первый взгляд неочевиден и скрыт за «патентованными технологиями», является то, что, несмотря на всю простоту и фактическое отсутствие лингвистических методов, необходимо постоянно обновлять базу данных «цифровых отпечатков». И если в случае с «сигнатурами», такой метод не требует постоянного обновления базы словами, то он требует обновления базы «цифровых отпечатков». **К недостаткам** можно отнести то, что фактически от «*дополнения базы словами*» поддержка DLP в эффективном состоянии переходит «*поиск и индексирование новых и измененных файлов*», что является более сложной задачей, даже если это делается DLP-системой полуавтоматически.

Технологии детектирования конфиденциальной информации

«Метки»

Суть метода заключается в расстановке специальных «меток» внутри файлов, содержащих конфиденциальную информацию. С одной стороны, такой метод дает стабильные и максимально точные сведения для DLP-системы, с другой стороны требуется много довольно сильных изменений в инфраструктуре сети. Несмотря на явное **достоинство** «меток» - качество детектирования, есть множество существенных **недостатков**: от необходимости значительной перестройки инфраструктуры внутри сети до введения множества новых правил и форматов файлов для пользователей. Фактически внедрение такой технологии превращается во внедрение упрощенной системы документооборота.

Регулярные выражения

Поиск по регулярным выражениям («маскам) является также давно известным способом детектирования необходимого содержимого, однако в DLP стал применяться относительно недавно. Часто этот метод называют «текстовыми идентификаторами». Регулярные выражения позволяют находить совпадения по форме данных, в нем нельзя точно указать точное значение данных, в отличие от «сигнатур». Поиск по «маскам» позволяет IPC-системе обеспечивать соответствие требованиям все более популярного стандарта PCI DSS, разработанного международными платежными системами Visa и MasterCard для финансовых организаций.

К достоинствам технологии относится то, что они позволяют детектировать специфичный для каждой организации тип контента, начиная от кредитных карт и заканчивая названиям схем оборудования, специфичных для каждой компании. Кроме того, формы основных конфиденциальных данных меняются крайне редко, поэтому их поддержка практически не будет требовать временных ресурсов. **К недостаткам** можно отнести их ограниченную сферу применения в рамках IPC-систем, так как найти с помощью них можно только конфиденциальную информацию лишь определенной формы. Регулярные выражения не могут применяться независимо от других технологий, однако могут эффективно дополнять их возможности.

Технологии детектирования конфиденциальной информации

Лингвистические методы (морфология, стемминг)

Самым распространенным на сегодняшний день методом анализа в IPC-системах является лингвистический анализ текста («контентная фильтрация»). В лингвистических методах есть свои отпечатки, базирующиеся на статистике; например, берется документ, считаются пятьдесят самых употребляемых слов, затем выбирается по 10 самых употребляемых из них в каждом абзаце. Такой «словарь» представляет собой практически уникальную характеристику текста и позволяет находить в «клонах» значащие цитаты. Разбор всех тонкостей лингвистического анализа не входит в рамки этой статьи, однако необходимо заметить ширину возможностей данной технологии в рамках IPC-систем.

К достоинствам можно отнести то, что в морфологии и других лингвистических методах высокая степень эффективности, сравнимая с сигнатурами, при намного меньших трудозатратах на внедрение и поддержку (снижение трудозатрат на 95% по отношению к «сигнатурам»). При этом в случае с использованием лингвистических методов детектирования нет необходимости отслеживать появление новых документов и направлять их на анализ в IPC-систему, так как эффективность лингвистических методов определения конфиденциальной информации не зависит от количества конфиденциальных документов, частоты их появления и производительности системы фильтрации содержимого. **Недостатки** также довольно очевидны, первый из них - зависимость от языка - если организация представлена в нескольких странах, базы конфиденциальных слов и выражений придется создавать отдельно для каждого языка и страны, учитывая всю специфику. При этом обычная эффективность такого метода составит в среднем 85%. Если привлекать профессиональных лингвистов, то эффективность может возрасти до 95% - больше может обеспечить лишь ручная проверка или «сигнатуры», однако по отношению эффективности и трудозатрат равных лингвистическим методам пока не имеется.

Технологии детектирования конфиденциальной информации

Ручное детектирование («Карантин»)

Ручная проверка конфиденциальной информации иногда называется «Карантином». Любая информация, которая попадает под правила ручной проверки, например, в ней встречается слово «ключ», попадает в консоль специалиста информационной безопасности. Последний по очереди вручную просматривает такую информацию и принимает решение о пропуске, блокировке или задержке данных. Если данные блокируются или задерживаются, отправителю посылается соответствующее сообщение. Несомненным **достоинством** такого метода можно считать наибольшую эффективность. Однако, такой метод в реальном бизнесе применим лишь для ограниченного объема данных, так как требуется большого количества человеческих ресурсов, так как для качественного анализа всей информации, выходящий за пределы компании, количество сотрудников информационной безопасности должно примерно совпадать с количеством остальных офисных сотрудников. А это невозможно даже в силовых и военных структурах. Реальное применение для такого метода — анализ данных выбранных сотрудников, где требуется более тонкая работа, чем автоматический поиск по шаблонам, «цифровых отпечатков» или совпадений со словами из базы.

Архивирование информации, проходящей через технические каналы утечки

Обязательной компонентой IPC является архив, который ведется для выбранных потоков информации (пакетов, сообщений). Вся информация о действиях сотрудников хранится в одной и нескольких связанных базах данных. Лидирующие IPC-системы позволяют архивировать все каналы утечки, которые они могут контролировать. В архиве IPC хранятся копии закачанных в интернет документов и текста, электронных писем, распечатанных документов и файлов, записанных на периферийные устройства. В любой момент администратор ИБ может получить доступ к любому документу или тексту в архиве, используя лингвистический поиск информации по единому архиву (или всем распределенным архивам единовременно). Любое письмо при необходимости можно посмотреть или переслать, а любой закачанный в Интернет, записанный на внешнее устройство или распечатанный файл или документ просмотреть или скопировать. Это позволяет проводить ретроспективный анализ возможных утечек и, в ряде случаев, соответствовать регулирующим деятельность документам.

Шифрование информации на всех точках сети Контроль доступа к сети, приложениям и информации

Обязательной компонентой IPC является архив, который ведется для выбранных потоков информации (пакетов, сообщений). Вся информация о действиях сотрудников хранится в одной и нескольких связанных базах данных. Лидирующие IPC-системы позволяют архивировать все каналы утечки, которые они могут контролировать. В архиве IPC хранятся копии закачанных в интернет документов и текста, электронных писем, распечатанных документов и файлов, записанных на периферийные устройства. В любой момент администратор ИБ может получить доступ к любому документу или тексту в архиве, используя лингвистический поиск информации по единому архиву (или всем распределенным архивам единовременно). Любое письмо при необходимости можно посмотреть или переслать, а любой закачанный в Интернет, записанный на внешнее устройство или распечатанный файл или документ просмотреть или скопировать. Это позволяет проводить ретроспективный анализ возможных утечек и, в ряде случаев, соответствовать регулирующим деятельность документам.

Двухфакторная аутентификация - это реализация контроля доступа, представляющая собой идентификацию пользователя на основе того, что он знает и того, чем он владеет. Наиболее распространенная форма аутентификации часто - это обычные пароли, которые пользователь держит у себя в памяти. Пароли создают слабую защиту, так как они могут быть легко раскрыты или подобраны. Политика безопасности, основанная на одних паролях, делает организацию уязвимой, поэтому в IPC применяется двухфакторная аутентификация с использованием распространенных USB-токенов.

Архитектура IPC-систем

IPC-системы обладают агентами на всех ключевых точках сети: сервера, хранилища, шлюзы, ПК (десктопы и ноутбуки), периферийные и сетевые пользовательские устройства. Технологии IPC реализованы для Windows, Linux, Sun Solaris, Novell. Поддерживается взаимодействие с Microsoft Active Directory, Novell eDirectory и другими LDAP. Большинство компонент может эффективно работать в рабочих группах.