Защита от утечки данных: комплексная терапия

Александр Кузнецов, *руководитель отдела безопасности информационных систем НТЦ "Вулкан"*

В последние годы неуклонно растет интерес организаций к применению систем класса Data Loss Prevention (DLP). Производители DLP-решений стараются предложить потребителю "серебряную пулю" для решения задач обеспечения безопасности информации от инсайдеров и случайных утечек, тесня классические средства защиты информации. Можно ли, обеспечивая безопасность критичных данных, рассчитывать только на DLP-систему? Об этом мы поговорим далее.

Говорят, что ...

DLP – панацея от всех бед!

DLP защищает от всех типов угроз.

DLP не дает ложных срабатываний.

DLP – решение "из коробки".

DLP работает само по себе.

Технические аспекты

Сразу оговорюсь, что в статье не затрагиваются правовые вопросы применения данных систем, возможности их сертификации по требованиям безопасности информации, а также организационные моменты, связанные с кропотливой подготовительной работой при их внедрении. Я хотел бы поделиться с вами практическим опытом нашей команды и показать узкие места, которые требуют особого внимания как при выборе и эксплуатации DLP-решений, так и при их интеграции с другими средствами защиты информации.

Утечка данных по сети

Одним из ключевых каналов утечки данных является сеть. На сегодняшний день практически все компьютеры на постоянной или временной основе имеют подключение к сетям передачи данных, в том числе и к сети Интернет. Большинство производителей DLP-решений заявляют о наличии в составе их продукта средств анализа и контроля данных, передаваемых по сети. Но зачастую при ближайшем рассмотрении данный функционал выглядит весьма и весьма ограниченным.

Начнем с того, что в большинстве исполнений DLP-систем анализ сетевого трафика осуществляют программно-аппаратные комплексы (ПАК), которые должны найти свое место в ЛВС организации, работая в одном из двух режимов:

- активный (подключение "в разрыв");
- пассивный (на SPAN-порту).

У каждого из данных режимов есть свои достоинства и недостатки. Установка ПАК "в разрыв" увеличивает время обмена данными и добавляет новую точку отказа в ЛВС, но при этом позволяет предотвращать утечки данных. Пассивный режим, организованный за счет репликации сетевого трафика, не позволяет активно предотвращать инциденты, а используется в большинстве своем для "разбора полетов". В активном режиме ПАК должен устанавливаться в точке выхода трафика из ЛВС в стороннюю сеть, в пассивном – в точках сбора трафика. А если таких точек несколько, то и количество устройств возрастает. Стоит отметить, что архитектура некоторых DLP-систем предусматривает наличие не одного, а нескольких ПАК в каждой такой точке, так как задачи консолидации данных, анализа почтового трафика или Web-трафика могут быть разнесены на отдельные специализированные устройства, что создает дополнительные трудности при внедрении. Необходимо добавить, что приобрести сетевой компонент в виде виртуальной машины зачастую нельзя, и владельцам защищаемой информации приходится выискивать место и мощности для обслуживания нескольких новых "железок". Организациям с филиальной структурой или имеющим множество точек подключения к различным ведомственным или партнерским сетям требуется быть готовыми к значительному расширению своего парка технических средств.

Говоря о распределенных сетях, нельзя забывать о практике применения в них протоколов IPSec. Увы, в данном случае ни один вариант развертывания ПАК не позволит проводить анализ зашифрованного трафика.

Вообще говоря, практически любой зашифрованный трафик является камнем преткновения для DLP-систем, за исключением случаев использования протоколов на базе SSL, например FTPS или HTTPS. Остановимся на них подробнее.

Чтобы "разобрать" данный трафик, ПАК должен, по сути, реализовать атаку "человек посередине". Сам по себе ПАК сделать этого не может, поэтому требуется его интеграция с прокси-сервером, поддерживающим протокол ICAP, например Squid, ProxySG или др. (см. рис. 1).

Но далеко не в каждой ЛВС есть такой прокси-сервер, а те компании, у которых его нет, могут быть не готовы внедрить прикладного посредника по ряду причин – начиная от устоявшейся сетевой инфраструктуры и заканчивая строгой политикой унификации используемого в составе ЛВС оборудования. Таким образом, "разбор" популярного протокола, используемого в том числе для доступа к сервисам электронной почты (один из ключевых каналов утечки данных), становится невозможен.

К "неразбираемым" еще стоит отнести протоколы Skype. Многие DLP-решения позиционируют готовность к анализу данного канала утечки. На самом же деле контроль осуществляется на более ранней стадии, на клиенте до отправки данных по сети (в основном за счет

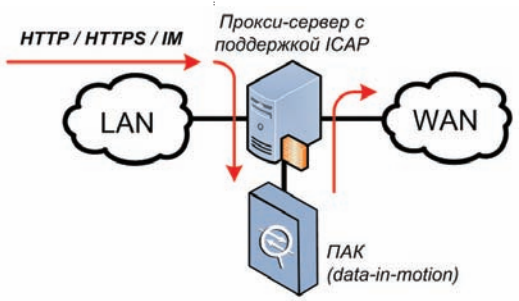

Рис. 1. Схема "разбора" DLP-решением протоколов на базе SSL

контроля буфера обмена, то есть контролируются только вложения). Стоит отметить, что иногда это характерно еще для ряда каналов утечки: электронная почта, IM и т.п.

Форматы документов

Как показывает опыт внедрения DLP-решений, владельцев защищаемой информации в первую очередь интересует контроль данных на хосте.

Одним из ключевых критериев выбора DLP-системы, обеспечивающей контроль данных, обрабатываемых на АРМ и серверах, является перечень форматов файлов, которые система способна анализировать. Заметим, что не все производители озаботились наличием такого перечня, и это зачастую затрудняет подбор наиболее оптимального продукта.

С высокой степенью вероятности мы найдем в данных перечнях форматы документов PDF и TXT.

PDF на сегодняшний день особенно популярен за счет того, что позволяет защитить информацию от редактирования. При этом содержание PDF-файлов может быть либо векторное, либо растровое. Первый тип содержания DLP-система способна анализировать, а вот второй – нет, поскольку данные воспринимаются как изображение. Как следствие, прекрасно читаемые человеческим глазом конфиденциальные документы могут бесследно "уплыть" за пределы организации, если только DLP-система не использует для контроля таких файлов цифровые отпечатки. Ожидается, что в ближайшем будущем появятся технологии автоматического распознавания теста. Но на сегодняшний день с учетом значительных объемов данных и требования "защита не должна тормозить бизнес", обеспечить корректное распознавание и анализ такого рода документов "на лету" практически невозможно.

Если морально к трудностям с PDF многие изначально готовы, то нюансы работы DLP-систем с TXT-файлами ряд специалистов застают врасплох. На сегодняшний день TXT является достаточно часто используемым текстовым форматом. При сохранении в нем текста можно использовать различные кодировки (ANSI, Unicode и др.). Как раз в них все и дело. Большинство DLP-систем не могут анализировать данные в кодировке ANSI, то есть злоумышленник может сохранить конфиденциальный документ в данной кодировке и отправить его любым способом, и DLP-система пропустит такую транзакцию (см. рис. 2).

Пароль на службе зла

Практически каждый из нас пользовался встроенными механизмами парольной защиты файлов MS Office или ZIP/RAR-архивов. Это самый дешевый способ обеспечить конфиденциальность. Для контекстного анализа защищенных таким образом данных необходимо знать пароль либо его подобрать. На сегодняшний день существуют единичные DLP-решения, которые используют подбор паролей по словарям и способны-таки добраться до содержимого. Но большинство DLP-систем просто блокирует передачу зашифрованных или защищенных паролем файлов. Безусловно, это создает сложности при санкционированном обмене конфиденциальными данными. При этом формирование исключений не всегда является возможным.

При анализе данного вопроса мы решили "копнуть глубже" и проверить проведение контекстного анализа цепочки архивов, на дне которой находиться архив или файл, защищенный паролем (см. рис. 3).

Исследование показало, что некоторые DLP-системы не справляются с решением подобной многоярусной задачи, пропуская файлы с конфиденциальными сведениями.

Комплексный подход

За рамками статьи остался еще добрый десяток способов, с помощью которых можно попытаться обмануть DLP-систему. Безусловно, этот класс решений является важным шагом на пути развития систем обеспечения безопасности информации, но пока они не могут единолично решить все проблемы в этой области и стать волшебной таблеткой, избавляющей от всех ИБ-болезней.

Хочу отметить, что особый кумулятивный эффект достигается за счет применения совместно с DLP-системами таких решений, как Digital Rights Management (DRM) и Security Information and Event Management (SIEM). В таком симбиозе системы взаимно дополняют друг друга. Компоненты DLP-системы осуществляют поиск и классификацию защищаемой информации по установленным критериям. В свою очередь, DRM-решение накладывает политики доступа на выявленные файлы, что ограничивает набор операций с ними. А SIEM формирует "единое окно" для администратора безопасности, в котором сводятся данные о выявленных файлах, подлежащих защите, попытках доступа к ним, а также увязывается (коррелируется) технологическая информация, поступающая от ОС, СУБД, сетевого оборудования и других источников, формируя полную картину состояния ИБ в организации.

Применение данного набора решений позволит снизить или как минимум отследить возникновение нештатных для DLP-решений ситуаций, описанных выше.

Только комплексный подход к защите информации позволит максимально приблизиться к ожидаемому результату и снизить вероятность утечки информации до приемлемого уровня. В нашей команде именно комплексному подходу отдается наибольший приоритет, и в этом залог успешного удовлетворения потребностей заказчиков в надежной, умной и управляемой защите. ●

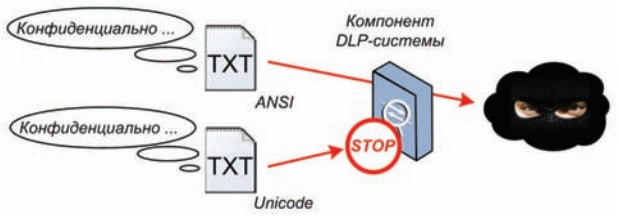

Рис. 2. Задача с кодировками файлов

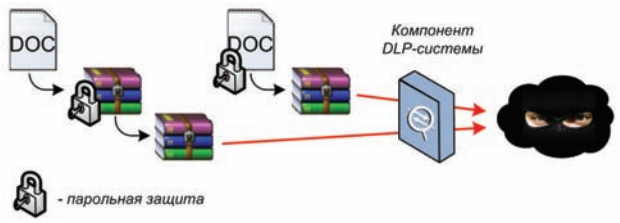

Рис. 3. Задача с паролями

На самом деле ...

DLP – "лекарство" для внутреннего нарушителя.

DLP нейтрализует (парирует) угрозы хищения, утечек и/или потерь.

Частота False Positive и False Negative зависит от качества настроек.

DLP требует профессионального сервиса при внедрении.

DLP требует сопровождения и мониторинга.