РАЗРАБОТКА МЕТОДА ИДЕНТИФИКАЦИИ РАСПРОСТРАНИТЕЛЯ КОНФИДЕНЦИАЛЬНОЙ ИНФОРМАЦИИ С ПОМОЩЬЮ МЕТОДОВ ЦИФРОВОЙ СТЕГАНОГРАФИИ

Шевченко Д.Н. Email: Shevchenko644@scientifictext.ru

Шевченко Дмитрий Николаевич – студент,
кафедра защищённых систем связи, факультет инфокоммуникационных сетей и систем,
Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
г. Санкт-Петербург

Аннотация: *в данной работе рассматриваются такие понятия как конфиденциальная информация, системы DLP, цифровая стеганография и её актуальность в нынешнее время. В данной работе выделены недостатки текущих систем предотвращения утечки данных. Предложено модульное решение, обеспечивающее идентификацию распространителя (предателя, инсайдера) конфиденциальной информации. В ходе работы смоделированы методы идентификации распространителя конфиденциальной информации в пределах корпоративных сетей с использованием разных методов цифровой стеганографии к разным типам файлов.*
Ключевые слова: *конфиденциальная информация, цифровая стеганография, DLP-системы.*

DEVELOPMENT OF A METHOD FOR IDENTIFYING A DISTRIBUTOR OF CONFIDENTIAL INFORMATION USING DIGITAL STEGANOGRAPHY METHODS

Shevchenko D.N.

Shevchenko Dmitry Nikolaevich – Student,
DEPARTMENT OF SECURE COMMUNICATION SYSTEMS,
FACULTY OF INFOCOMMUNICATION NETWORKS AND SYSTEMS,
ST. PETERSBURG STATE UNIVERSITY OF TELECOMMUNICATIONS BY PROFESSOR M.A. BONCH-BRUEVICH,
ST. PETERSBURG

Abstract: *in this paper, we consider such concepts as confidential information, DLP (Data Leakage Prevention) systems, digital steganography and its relevance at the present time. In this paper, the shortcomings of current systems for preventing data leakage are highlighted. A modular solution is proposed that ensures the identification of the distributor (traitor, insider) of confidential information. In the course of work, methods for identifying the distributor of confidential information within corporate networks are simulated using different methods of digital steganography to different types of files.*
Keywords: *confidential information, digital steganography, DLP-systems.*

УДК 003.26

I. Обзор литературы

При написании данной работы были использованы, статьи в периодических изданиях, статьи конференций и интернет-источники. В статье [1] рассматриваются существующие методы идентификации распространителей информации в DLP-системах. В ней подробно объясняются основные принципы идентификации распространителей секретной информации в корпоративных сетях, а также поднимается вопрос о большой ресурсоёмкости данных методов. Основным источником, раскрывающим теоретические основы цифровой стеганографии, явилась книга [2], где рассмотрено понятие цифровой стеганографии, различные методы стеганографии для разных типов файлов и различные области применения. На основе работ [3], [4], в которых подробно рассказывается о взаимодействии сетевой составляющей DLP-систем с файлами, находящимися внутри корпоративных сетей, а также варианты скрытия информации в различные типы файлов. Для решения проблемы идентификации распространителя конфиденциальной информации используется полный уникальный идентификатор.

II. Введение

Вопрос сохранения конфиденциальности корпоративных файлов стоит очень остро в каждой компании. В современных условиях руководители организации понимают, что потеря или кража конфиденциальных данных ведет не только к прямым финансовым убыткам, но и к снижению доверия со стороны клиентов, партнеров и инвесторов. Любая утечка данных, даже отправка письма с конфиденциальными документами по ошибочному адресу, приводит к повышенному интересу со стороны регулирующих органов и СМИ.

В современных DLP-системах используются методы сетевого отслеживания, записи логов, мониторинга сетевой активности и неформатные методы, основанные на особенностях работ файловых

систем. Главное отличие данной работы и предложение заключается в том, чтобы в сам файл (какого бы типа он ни был) вкладывать информацию, о том, какой пользователь и когда создал этот документ, кому он передавал этот файл, а также, кто и когда его получил и изменял.

III. Краткий обзор DLP-систем. Уязвимости и недостатки.

DLP (Data Leakage Prevention - Предотвращение утечек данных) система — это система защиты конфиденциальных данных от внутренних угроз. Работа DLP-систем строится на перехвате, дальнейшем анализе и архивации потоков данных, пересекающих периметр в направлении внешнего контура либо циркулирующих внутри защищаемой корпоративной сети. Весь перехваченный поток данных проходит через несколько этапов работы системы предотвращения утечек информации. При обнаружении конфиденциальных данных, соответствующим выбранным критериям, срабатывает активная компонента системы, оповещающая сотрудника службы безопасности об инциденте.

На рисунке 1 представлена структурная схема работы классической DLP-системы.

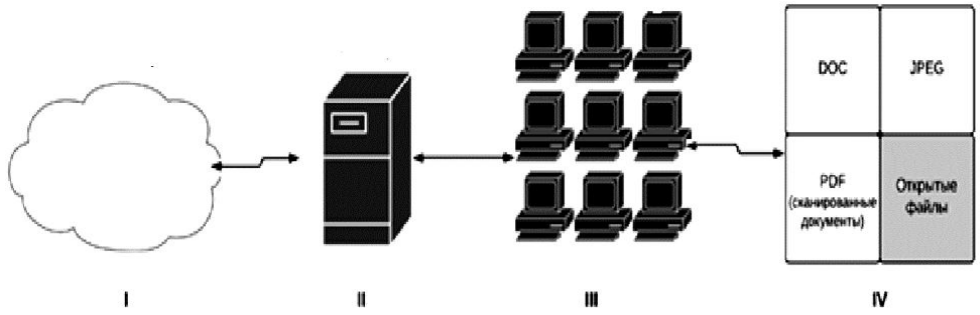

Рис. 1. Структурная схема работы классической DLP-системы

Пояснения к рисунку:

I- Интернет, всемирная сеть, утечка информации, в которую секретных документов недопустима.

II- DLP-Server главный центр DLP системы, который проводит мониторинг файлов, отправляющихся за внешний контур.

III- Персональные компьютеры работников, каждому присвоен свой уникальный идентификационный номер.

IV- Общее хранилище файлов компании.

В «классической» DLP-системе идентификация сотрудника, который передаёт конфиденциальный файл за пределы внутреннего контура, основана на сетевых технологиях. Это значит, что передача файлов (через сервер DLP-системы) логируется (ведётся запись заголовков пакетов передачи). И если срабатывает сигнализация, системный администратор или офицер безопасности просматривают пакеты, на которые разбивался этот файл. Далее идентифицируется IP-адрес (а соответственно и конкретная рабочая станция), с которого была произведена попытка передачи того или иного конфиденциального файла.

Но данный метод не может обеспечить полной уверенности в том, что распространитель данной информации – это владелец данного ПК.

Возможны ситуации, когда инсайдер (предатель, traitor) может злонамеренно воспользоваться персональным компьютером другого работника.

Эту проблему возможно решить другими методами, но основное предложение данной работы решает её гораздо проще.

IV. Полный уникальный идентификатор.

Для удобства работы с различными методами вложения данных, было принято решение создать систему уникальных идентификаторов каждого сотрудника.

10001110110100110100110010101101000 11

Первый символ указывает на степень конфиденциальности файлы. Это поле бинарно.

Символы (2-6) указывают на авторство. Это поле можно расширить до любого другого количества символов.

Символы (7-37) 31 символ, указывают на время. Для определения точного времени создания того или иного файла в данной работе предлагается использовать Timestamp в формате UNIX-время.

V. Реализация методов цифровой стеганографии для различных типов файлов.

В основном, в документообороте компаний участвуют основные три типа файлов:

- Изображения (цветные и чёрно-белые) – jpeg
- Текстовые документы – doc, txt
- Сканированные документы – pdf

Для изображений предлагается использовать стеганографический метод F5. Алгоритм F5 вкладывает информацию в случайные КДКП, выбираемые с помощью ПСП, генерируемой на основе пароля, заданного пользователем.

Для текстовых документов используется метол модуляции хвостовых пробелов. Метод хвостовых пробелов предполагает дописывание в конце каждой строки файла-контейнера одного пробела, в случае кодирования единичного бита стегосообщения. Если нужно закодировать нулевой бит, пробел в конце строки не дописывается.

Для вложения информации в сканированные документы, и в целом, в файлы формата PDF существует стегосистема, маскирующая вложения дополнительной информации шумами сканера. Основная идея данного метода заключается в том, чтобы отсканировать напечатанный документ и внести в него секретную информацию имитируя шумы сканера. Имитация осуществляется за счёт того, что изменению могут быть подвергнуты только пиксели, находящиеся на границе черного с белым.

VI. Заключение.

Используя методы цифровой стеганографии можно значительно упростить процесс идентификации распространителя конфиденциальной информации.

В настоящее время число утечек информации в различных компаниях увеличивается с каждым днем, и будет продолжать расти. Соответственно в актуальность DLP-систем в современном мире повышается. Помимо определения самого факта утечки данных, важен процесс идентификации сотрудника-предателя. А так как стандартные методы, используемые в DLP-системах, не могут обеспечить стопроцентную уверенность в идентификации предателя, то актуальность применения методов цифровой стеганографии тоже будет расти. В рамках данной работы был проведен анализ научно-технической литературы, результатом которого является выявление уязвимостей «классических» DLP-системы.

Разработка, описанная в данной работе, не представляет из себя полноценную замену DLP-системам. По сути, это дополнительный модуль, встраиваемый в корпоративную сеть, чтобы обеспечить процесс расследования инцидентов безопасности, а именно, процесс идентификации предателя.

Список литературы / References

1. *Silowash G., Cappelli D., Moore A., Trzeciak R., Shimeall T., Flynn L.* Common Sense Guide to Mitigating Insider Threats. 4th edition. (CMU/SEI-2012-TR-012) Software Engineering Institute. (2012).
2. *Коржик В.И., Небаева К.А., Герлинг Е.Ю., Догиль П.С., Федянин И.А*. Монография “Цифровая стеганография и цифровые водяные знаки. Часть 1. Цифровая стеганография. СПбГУТ. - СПБ., 2016. 226 с.
3. *Mamta Jain.* A Review on Data Leakage Prevention using Image Steganography, Department of Computer Science and Engineering, Mody University of Science and Technology ISSN: 2319-7323.
4. *Jessica Fridrich.* Steganography in Digital Media: Principles, Algorithms, and Applications, 1st Edition, 2009.