Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра теоретических основ компьютерной безопасности и криптографии

Система предотвращения утечки конфиденциальной информации

АВТОРЕФЕРАТ

дипломной работы

студента 6 курса 631 группы

специальности 10.05.01 Компьютерная безопасность

факультета компьютерных наук и информационных технологий

Беспалова Кирилла Игоревича

Научный руководитель

ассистент _____ Н.Н. Бондарев

31.12.2016 г.

Заведующий кафедрой

профессор, к.ф.-м.н. _____ В.Н. Салий

31.12.2016 г.

Саратов 2017

ВВЕДЕНИЕ

На сегодняшний день главным источником угроз информационной безопасности любой компании или предприятия являются действия собственных сотрудников, которые могут привести к утечке конфиденциальной информации. Причем действия сотрудников, повлекшие за собой утечку конфиденциальной информации, необязательно носят умышленный характер. Чаще всего это происходит по неосторожности, как следствие отсутствия контроля за перемещением конфиденциальной информации внутри компании.

Необходимость в контроле действий пользователей, направленных на передачу конфиденциальной информации, привела к появлению на рынке так называемых DLP-систем – систем предотвращения утечек конфиденциальной информации. В таких системах подход, с помощью которого осуществляется этот контроль, основывается на попытке перехвата и анализа исходящих информационных потоков в режиме реального времени. Однако, в тоже время это является и основным недостатком таких систем. На практике они имеют поддержку контроля достаточно ограниченного списка каналов и протоколов передачи данных. В связи с этим они не могут предотвращать утечку конфиденциальной информации путем перехвата всех возможных исходящих информационных потоков.

Целью данной дипломной работы является разработка и реализация альтернативного подхода, основанного на создании специальной изолированной среды для пользователя, единственным каналом передачи данных из которой являлась бы сама DLP-система. Причем должна быть исключена любая возможность утечки конфиденциальной информации за пределы этой среды.

Для достижения поставленной цели необходимо решить следующие задачи:

- изучить подход к DLP в системах, присутствующих сегодня на рынке;

- выявить его недостатки;
- разработать альтернативный подход, лишенный выявленных недостатков;
- реализовать разработанный подход в виде программных модулей, выполняющих функции DLP-системы.

Дипломная работа состоит из введения, 4 разделов, заключения, списка использованных источников и 3 приложений. Общий объем работы – 81 страница, из них 44 страницы – основное содержание, включая 22 рисунка, 1 таблицу, список использованных источников из 14 наименований.

КРАТКОЕ СОДЕРЖАНИЕ

Во введении формулируется цель дипломной работы: реализация системы предотвращения утечки конфиденциальной информации на основе альтернативного подхода к контролю исходящих информационных потоков, лишенного тех недостатков, которыми обладают аналогичные системы, представленные сегодня на рынке.

В разделе 1 «Классический подход к DLP» приводится описание подхода, который используется в большинстве современных DLP-системах для контроля исходящих информационных потоков. Дается описание основных компонентов таких систем, а также функций, которые они выполняют. В конце раздела перечисляются выявленные недостатки такого подхода.

Раздел 2 «Изоляционный подход к DLP» посвящен описанию предлагаемого в дипломной работе подхода к DLP, который позиционируется как альтернатива классическому подходу и лишен его недостатков. В разделе вводятся все необходимые понятия и определения, такие как: открытая среда, изолированная среда, посредник. Дается описание возможностей пользователя в системе при работе в каждой из сред, определяются роль и функции посредника, а также возможности взаимодействия с ним пользователей.

В разделе 3 «Пользовательское взаимодействие с DLP-системой» отражается процесс взаимодействия пользователя с разработанной на основе предложенного подхода DLP-системой на рабочих станциях под управлением ОС Windows. Демонстрируется графический интерфейс изолированной среды, особенности работы в ней, а также графические интерфейсы меню «управление доступом» и «извлечение файлов» изолированных контейнеров – специальных файлов, в которых DLP-система хранит содержимое изолированной среды.

Раздел 4 «Архитектура DLP-системы» содержит описание деталей реализации разработанной системы и состоит из 5 подразделов.

В подразделе 4.1 «Основные модули и компоненты» перечисляются программные модули, из которых состоит разработанная DLP-система. Дается

описание их роли в системе, а также функций и возможностей, которыми они обладают.

В подразделе 4.2 «Аутентификация пользователей» содержится описание механизма аутентификации DLP-системой пользователей в среде операционной системы Windows.

В подразделе 4.3 «Структура изолированного контейнера» описывается структура изолированных контейнеров – зашифрованных zip-архивов, в которых DLP-система архивирует файлы, доступные в открытом виде пользователю при работе в изолированной среде.

В подразделе 4.4 «Устройство изолированной среды» содержится описание всех механизмов и процессов, связанных с созданием и работой изолированной среды. В начале подраздела дается описание этапов извлечения содержимого изолированного контейнера и его размещения в изолированной среде, а также роли в этом процессе защищенной файловой системы EFS семейства ОС Windows. Далее в подразделе рассказывается о возможности ОС Windows запуска программ и сервисов от имени других пользователей, а также о том, как эта возможность используется DLP-системой для ограничения прав доступа пользователя в изолированной среде. Завершает подраздел описание механизма перехвата системных вызовов Windows API, который используется DLP-системой для блокировки доступа пользователей в изолированной среде к буферу обмена, принтерам, локальной и глобальной сети, а также снятию снимков экрана.

В подразделе 4.5 «Модуль классификации информации» описывается модуль классификации информации и методы, которые используются в нем для классификации извлекаемой по запросу пользователя в открытую среду информации.

В заключении содержатся выводы по результатам работы.

В приложении А «Модуль dlp-server» приведен код серверной части разработанной DLP-системы.

В приложении Б «Модуль dlp-client» приведен код клиентской части разработанной DLP-системы.

В приложении В «Модули classifier и classifier-builder» приведен код модулей, которые используются для проведения классификации информации.

ЗАКЛЮЧЕНИЕ

В ходе работы было изучено понятие систем предотвращения утечки конфиденциальной информации, а также цели и задачи, которые стоят перед ними. Был рассмотрен классический подход к построению таких систем, основанный на попытке перехвата и контролирования в режиме реального времени всех исходящих информационных потоков. Обозначены недостатки такого подхода.

В качестве альтернативы был предложен и реализован подход, основанный на разделении пользовательского взаимодействия с системой на две среды – открытую и изолированную.

В изолированной среде пользователь имеет возможность работы с защищаемой информацией, однако доступа к каналам её передачи внутри этой среды у него нет. Фактически, единственным каналом передачи данных здесь является сама DLP-система, которая осуществляет контроль за содержимым этой среды и не позволяет пользователям извлекать конфиденциальную информацию из нее в открытую среду.

В открытой среде наоборот, пользователь никак не ограничивается в доступе к каналам передачи данных. Однако подразумевается, что в ней у пользователя отсутствуют средства работы с информацией, такие, как например, текстовые редакторы. Поэтому он вынужден осуществлять взаимодействие с информацией только в изолированной среде, где такая возможность присутствует.

Вся информация, с которой пользователь осуществлял работу в изолированной среде, предоставляется ему DLP-системой и в открытой среде. Однако в зашифрованном виде и без знания ключа. Это позволяет сотрудникам компании обмениваться зашифрованной информацией через каналы передачи данных, доступные в открытой среде. При этом доступ к переданной информации возможен только под контролем DLP-системы и в изолированной среде.

Ключом шифрования, а также возможностями передачи данных между средами обладает только посредник – специальная служба DLP-системы, с которой пользователь осуществляет взаимодействие и посредством которой реализуется функционал по работе со средами и содержимым зашифрованных файлов. Кроме того, посредник осуществляет аутентификацию и авторизацию пользователей, а также реализует контроль над действиями пользователя в изолированной среде.

В данном подходе у пользователей присутствует возможность извлечения в открытую среду информации в незашифрованном виде через посредника, однако только в том случае, если она не является конфиденциальной, и её владелец разрешил такую операцию. Таким образом реализуется механизм предотвращения утечки конфиденциальной информации.

Выше описанный подход был реализован в виде нескольких программных модулей клиент-серверной архитектуры, устанавливаемой на рабочих станциях под управлением ОС Windows, где клиентом выступает пользователь, а в качестве сервера выступает служба-посредник.

Открытая среда в данном случае представляет собой обычную среду операционной системы Windows, в которой пользователю запрещается доступ администратором безопасности к программам и средствам по работе с файлами и документами, принуждая пользователя тем самым осуществлять работу с ними в изолированной среде, где эта возможность присутствует.

Изолированная среда была реализована путем запуска службой посредника в активной сессии проводника Windows, в котором пользователь осуществляет необходимую работу с файлами и программами от имени специального пользователя, максимально ограниченного в правах в системе. Права доступа предварительно настраиваются администратором безопасности с помощью механизма групповых политик. Блокировка в изолированной среде выполнения программами таких операций, как запись в буфер обмена, снятие

снимков экрана, передача данных по локальной или глобальной сети, печать документов на принтерах, была реализована путем перехвата вызовов функций интерфейса Windows API, которые соответствуют этим операциям.

Шифрование данных было реализовано средствами штатных криптографических сервисов ОС Windows, таких, как Data Protection API и Encryption File System.

Возможность извлечения данных в открытую среду была реализована путем использования посредником разработанного в рамках работы модуля классификации, в основе работы которого применяются методы машинного обучения, а также статические методы классификации, такие как базы стоп-слов, регулярных выражений и базы хэш-значений конфиденциальных файлов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Windows Data Protection API [Электронный ресурс] // MSDN Developer Network [Электронный ресурс] : [сайт]. URL: https://msdn.microsoft.com/en-us/library/ms995355.aspx (дата обращения 06.11.2016). Загл. с экрана. Яз.англ.

2 PKCS #5: Password-Based Cryptography Specification Version 2.0 [Электронный ресурс] // The Internet Engineering Task Force [Электронный ресурс] : [сайт]. URL: https://tools.ietf.org/html/rfc2898 (дата обращения 20.11.2016). Загл. с экрана. Яз.англ.

3 Windows Sysinternals [Электронный ресурс] // Resources and Tools for IT Professionals TechNet [Электронный ресурс] : [сайт]. URL: https://technet.microsoft.com/en-us/sysinternals/bb545021.aspx (дата обращения 23.11.2016). Загл. с экрана. Яз.англ.

4 Security Tool for the Encrypting File System [Электронный ресурс] // Microsoft Support [Электронный ресурс] : [сайт]. URL: https://support.microsoft.com/ru-ru/kb/298009 (дата обращения 15.11.2016). Загл. с экрана. Яз.англ.

5 Group Policy Management Console [Электронный ресурс] // Resources and Tools for IT Professionals TechNet [Электронный ресурс] : [сайт]. URL: https://technet.microsoft.com/en-us/library/cc753298(v=ws.11).aspx (дата обращения 17.11.2016) Загл. с экрана. Яз.англ.

6 Windows DLL Injection Basics [Электронный ресурс] // Open Security Research [Электронный ресурс] : [сайт]. URL: http://blog.opensecurityresearch.com/2013/01/windows-dll-injection-basics.html (дата обращения 22.09.2016). Загл. с экрана. Яз.англ.

7 Understanding the Import Address Table [Электронный ресурс] // Sandsprite Tools Web application auditing – custom activex controls [Электронный ресурс] : [сайт]. URL: http://sandsprite.com/CodeStuff/Understanding_impo

rts.html (дата обращения 12.11.2016). Загл. с экрана. Яз.англ.

8 Deviare API Hook Overview [Электронный ресурс] // Nektra – Custom Software Development Company [Электронный ресурс] : [сайт]. URL: http://www.nektra.com/products/deviare-api-hook-windows (дата обращения 12.09.2016). Загл. с экрана. Яз.англ.

9 Support Vector Machines [Электронный ресурс] // Scikit-learn – Machine Learning in Python [Электронный ресурс] : [сайт]. URL: http://scikit-learn.org/stable/modules/svm.html (дата обращения 22.11.2016). Загл. с экрана. Яз.англ.

10 Textract documentation [Электронный ресурс] // Textract library documentation [Электронный ресурс] : [сайт]. URL: https://textract.readthedocs.io/en/stable/ (дата обращения 16.09.2016). Загл. с экрана. Яз.англ.

11 Zecurion Zdiscovery (Discovery) [Электронный ресурс] // Zecurion Protection as a Priority [Электронный ресурс] : [сайт]. URL: http://zecurion.ru/products/zdiscovery/description/benefits/ (дата обращения 16.12.2016). Загл. с экрана. Яз.рус.

12 Finding Important Words in Text Using TF-IDF [Электронный ресурс] // Me Steven Loria [Электронный ресурс] : [сайт]. URL: http://stevenloria.com/finding-important-words-in-a-document-using-tf-idf (дата обращения 07.09. 2016). Загл. с экрана. Яз.англ.

13 Scikit-learn Machine Learning in Python [Электронный ресурс] // Scikit-learn Machine Learning in Python [Электронный ресурс] : [сайт]. URL: http://scikit-learn.org/ (дата обращения 12.09.2016). Загл. с экрана. Яз.англ.

14 Training and Tuning SVC: R vs Python [Электронный ресурс] // Joel Carlson blog [Электронный ресурс] : [сайт]. URL http://joelcarlson.me/2016/05/14/RvsPython-GridSearch (дата обращения 8.09.2016). Загл. с экрана. Яз.англ.