АНАЛИЗ ВОЗМОЖНОСТЕЙ НЕСАНКЦИОНИРОВАННОГО СОКРЫТИЯ КОМПЬЮТЕРНОЙ ИНФОРМАЦИИ В ФАЙЛАХ ФОРМАТА OPEN DOCUMENT И РАСПОЗНАВАНИЕ ВНЕДРЕННЫХ ДАННЫХ

Ю. Н. Филиппова
(Екатеринбург, УрФУ, junfn@mail.ru)

В настоящее время предприятия активно используют обмен офисными документами посредством электронной почты и электронных носителей. В связи с этим обострилась проблема защиты компьютерной информации от утечки через пересылаемые документы.

На данный момент широкую доступность и разнообразие имеют программные продукты, позволяющие встраивать скрытую информацию в различные файлы. Кроме того, следует учитывать, что зачастую ущерб наносится не вследствие злого умысла, а из-за элементарных действий пользователей, которые случайным образом могут внести в документ избыточную информацию. Таким образом, становится очевидной актуальность совершенствования методов стеганоанализа, задача которого состоит в обнаружении факта передачи тайного сообщения.

Для успешного решения задачи обеспечения защиты от утечки информации через пересылаемые документы необходимо исследование возможных методов сокрытия информации внутри файлов и разработка алгоритма для поиска и удаления скрытой информации ограниченного доступа.

В настоящее время в России проводятся работы по переходу государственных учреждений на использование формата Open Document Format (ODF). Поэтому важным становится вопрос об обеспечении защиты компьютерной информации от утечки через пересылаемые документы формата ODF.

ODF [1] – открытый формат файлов документов для хранения и обмена редактируемыми офисными документами, в том числе текстовыми документами, электронными таблицами, рисунками,

базами данных, презентациями. Формат Open Document представляет собой упакованный в ZIP набор XML-документов и каталогов. Данный формат принят в качестве государственного стандарта (ГОСТ Р ИСО/МЭК 26300-2010 [2]) и введен в действие с 1 июня 2010 г.

Скрытие информации в файлах формата ODF методами компьютерной стеганографии возможно следующими способами:

- включением дополнительных файлов в архив;
- манипуляциями с XML-файлами;
- использованием скрытых абзацев, комментариев, не отображаемого текста и т. д.
- использованием графических объектов;
- присоединением данных к двоичным файлам до вставки в документ ODF;
- манипуляциями с внедряемыми объектами.

При разработке алгоритма защиты от распространения скрытой избыточной информации ограниченного доступа необходимо, в первую очередь, исследовать способы внедрения информации в файлы формата ODF, вызванные случайными действиями пользователей.

Компьютерный анализ файлов формата ODF на наличие скрытой избыточной информации можно разделить на два этапа:

- анализ содержимого архива;
- анализ содержимого XML-файлов.

Анализ содержимого архива

Архив документа ODF-формата содержит каталоги и файлы, отображаемые в XML-файле *manifest.xml*. Для предотвращения распространения избыточной информации необходимо проводить сравнение списка файлов, расположенных в архиве, со списком, отраженным в данном XML-файле.

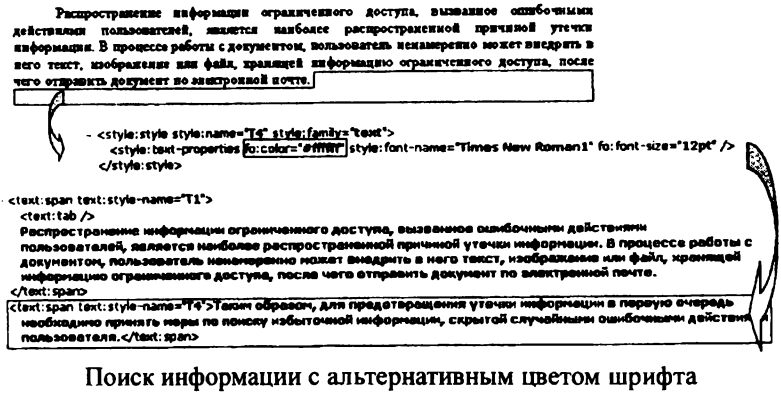
Включение дополнительных файлов в архив возможно как случайным внедрением файла, так и намеренно с редактированием списка элементов архива. Поэтому для корректной проверки на наличие внедренных файлов в архиве перед сравнением со списком *manifest.xml* необходимо проверять наличие ссылок на объекты в файле *content.xml*.

Анализ содержимого XML-файлов

Работа с документом дает пользователю возможность без особых навыков свободно форматировать документ в соответствии со своими целями. Так, пользователь может свободно задавать определенные атрибуты текстам, таблицам, графическим объектам и иным компонентам документа. Данные действия пользователя могут привести к случайному сокрытию информации, не предназначенной для распространения. Любые изменения в документе, которые вносит пользователь в документ, отражаются в XML-файле *content.xml*, в котором описывается содержимое документа и его разметка.

Если при форматировании документа пользователем устанавливаются определенные атрибуты, отличные от установленных по умолчанию, то в файле *content.xml* создаются новые стили, с указанием принятых изменений. По созданным стилям с дополнительными атрибутами можно найти ту информацию, к которой были применены данные изменения.

Например, анализ файла на наличие в документе текста с альтернативным цветом шрифта необходимо проводить путем поиска в XML-файле атрибута *fo:color* (рисунок):



Поиск информации с альтернативным цветом шрифта

Аналогичным образом проводится проверка на наличие дополнительных атрибутов, установленных для компонентов электронной таблицы, форм и отчетов баз данных, презентаций и рисунков формата ODF.

Еще одной возможностью пользователя, которой следует уделить внимание, является возможность добавления объектов. Внедренные объекты – ресурсы, за обработку которых отвечают внешние программы. Ими являются как встроенные в документы в формате Open Document формулы, графики, таблицы, так и данные, обрабатываемые внешними по отношению к офисному пакету программами, – объекты OLE.

Ресурсы в формате Open Document располагаются в отдельных папках, повторяющих в целом структуру документов ODF, при этом у них отсутствуют некоторые обязательные для обычных файлов компоненты. Данные ресурсы также позволяют пользователю при указании дополнительных атрибутов скрывать информацию. Поэтому проверка должна проводиться не только по содержимому самого документа, но и по содержимому, которое хранится внутри внедренного объекта.

Офисный пакет обращается к данным внедренного объекта только при необходимости его редактирования. Для отображения используются специальные объекты-заместители. Существует возможность принудительного удаления замещающих объектов, в этом случае при редактировании повторно создаются заместители. Эту возможность можно использовать для предотвращения подмены данных файлов.

Система защиты от утечки информации, вызванной ошибочными действиями пользователей, может быть организована с помощью разработанного приложения, реализующего поиск вложенных в каталог дополнительных файлов и отслеживание определенных атрибутов, которые позволяют скрыть информацию.

Библиографические ссылки

1. Open Document Format [Электронный ресурс]. Режим доступа: http://www.seobuilding.ru/wiki/OpenDocument.

2. ГОСТ Р ИСО/МЭК 26300-2010. Информационная технология. Формат Open Document для офисных приложений (OpenDocument) v1.0 : введ. 31.05.2011 г. [Электронный ресурс]. Режим доступа: http://protect.gost.ru/document.aspx?control=7&id=177075.