ИДЕНТИФИКАЦИЯ УСТРОЙСТВА С ПОМОЩЬЮ ОТПЕЧАТКА ФАЙЛОВОЙ СИСТЕМЫ

IDENTIFICATION OF THE DEVICE USING THE FINGERPRINT OF A FILE SYSTEM

V. Abashin
S. Makeev

Summary. In given article authors shine a problem of identifying devices by the fingerprint of the file system, i.e., by the structure of information and metadata of the file system.

Keywords: identification; devices; template; user activity; file system fingerprint.

Абашин Валерий Геннадьевич
К.т.н., доцент, Московский Государственный Лингвистический Университет (Москва, Россия)
valeriy@abashin.ru
Макеев Сергей Александрович
Старший преподаватель, Московский Государственный Лингвистический Университет (Москва, Россия)
mak3y1984@gmail.com

Аннотация. В данной статье авторы описывают возможность идентификации устройств по отпечатку файловой системы, т.е. по структуре информации и метаданным файловой системы носителей информации.

Ключевые слова: идентификация; устройства; эталон; активность пользователя; отпечаток файловой системы.

Введение

Каждая ЭВМ, использующая операционную систему, имеет уникальную комбинацию данных о файловой системе (ФС). Несмотря на различия в ФС разных типов, они используют одни базовые принципы, и состоят из двух базовых компонент: хранимой информации и служебной информации самой файловой системы.

Главное преимущество использования ФС для целей идентификации в отличие от служебной информации об аппаратном или программном обеспечении в её доступности для большинства пользователей ЭВМ. Все пользователи производят обращения к ФС в том или ином виде. Наименее значимыми правами являются права на чтения информации ФС, которые делегируются пользователю чаще всего. Прав на чтение достаточно, чтобы получить доступ к информации о ФС. Также использование ФС позволяет использовать для идентификации существующую информацию, а не создавать дополнительную.

Полученные данные пользовательских обращений можно использовать для решения задач по обеспечению информационной безопасности, в частности, для выявления мошеннических операций в системах дистанционного банковского обслуживания [1], идентификации пользователя в сети Интернет [2].

Постановка задачи

Информация, хранимая в файловых системах, является отражением окружающего мира и так же, как окружающий мир, постоянно изменяется. Часть информации в файловой системе остается неизменной достаточно продолжительное время, чтобы её можно было использовать как идентифицирующий признак. Для того, чтобы выделить информацию ФС, пригодную для создания идентифицирующего признака, необходимо определить причины изменения ФС. К ним относятся:

- состояние и функционирование аппаратного обеспечения;
- состояние и функционирование программного обеспечения;
- активность (действия) пользователя ЭВМ.

В конечном итоге, наибольшее влияние на изменение ФС оказывают социальные, биологические ритмы человека, а также его психологические особенности. Например, обновление программного обеспечения, вызывающее значительные изменения в ФС, обычно согласуются с ритмичностью разработки программного обеспечения, которая, в свою очередь, связана с годовым, квартальным и месячными интервалами. Изменение аппаратной конфигурации, за исключением брака, связано с рассчитанной производителем продолжительности работы устройства. На практике, замена устройства производится до потери им работоспособности и согласуется с амортизацией, учитываемой в бухгалтерской деятельности. Её основным циклом является годовая бухгалтерская отчетность. Деятельность пользователя ЭВМ напрямую зависит от суточных и недельных ритмов и психологических особенностей человека.

Для идентификации необходим эталонный отпечаток ФС, с которым будет сравниваться предъявляемый

отпечаток, а также предъявляемый отпечаток, сформированный по запросу пользователя, программы или аппаратного обеспечения. Причем, в связи с тем, что ФС постоянно подвержена изменениям, необходимо иметь возможность учитывать изменения в её отдельных элементах, т.е. иметь некоторое дискретное представление о ФС.

Делается предположение, чем меньше времени просуществовали данные, тем менее они устойчивы. Несмотря на то, что скопированный для годового хранения архив тоже имеет сначала небольшое время хранения, выдвигается предположение, при регулярном использовании накопителя информации, большая часть элементов ФС остаются неизменными.

Изменения в ФС накапливаются с течением времени. Для повышения эффективности использования эталонного отпечатка ФС предлагается создавать его заново при каждом использовании ФС как ключа для идентификации. Другим способом повышения стабильности эталонного отпечатка ФС является использование наборов признаков, относящихся к различным ритмам: технологическим, социальным, биологическим. Таким образом, в случае утери значащего признака в связи со значительным обновлением программного обеспечения ЭВМ, будут сохранены признаки, связанные с социальными и биологическими ритмами.

Для создания отпечатков ФС необходимо определить наборы признаков, по которым возможно сформировать эталонный отпечаток и провести идентификацию ФС.

Формализованное решение задачи

Под отпечатком ФС понимается информация о наборе идентифицирующих признаков ФС. В источнике [3] предлагается создание идентифицирующего признака ФС с использованием имени элемента ФС, информации о местоположении, размере, дате создания и дате редактирования. Уточним набор данных идентифицирующего признака следующим образом:

- полный путь к файлу;
- размер файла;
- время последнего доступа к файлу;
- время последней модификации;
- время последнего изменения состояния файла.

Представим описание идентифицирующего признака на языке С11.

```
struct FS_ELEMENT {
char full_name[500]; //Полный путь к файлу
int filesize; //Размер файла
time_t last_access; //Время последнего доступа
time_t last_modification; //Время последней модификации
time_t last_change; //Время последнего изменения
};
```

В результате анализа ритмов источников изменения информации ФС были выделены следующие классы интервалов изменений:

- до суток;
- от суток до недели;
- от недели до месяца;
- от месяца до полугода;
- от полугода до года;
- от года и более.

В связи с требованием к эталонному отпечатку ФС оставаться стабильным на период до месяца, ритмы с тактом до месяца отбрасываются. Пригодными к использованию являются такты:

- от месяца до полугода;
- от полугода до года;
- от года и более.

Считаем, что минимальный эталонный отпечаток ФС состоит из трех идентифицирующих признаков, каждый из которых соответствует одному из трех используемых ритмов.

Представим в виде текстового описания обобщённый алгоритм создания эталонного отпечатка ФС. Необходимо выполнить рекурсивный обход всех каталогов, начиная с указанного корневым и найти объекты ФС, соответствующие одному из трех ритмов, пригодных для создания эталонного отпечатка ФС. Далее, сформировав из них идентифицирующие признаки, заполнить соответствующие поля эталона. Сам эталонный отпечаток ФС будет представлять из себя массив структур.

```
struct FS_ELEMENT etalon[3];
```

Обобщенный алгоритм идентификации заключается в определении наличия объекта ФС с указанными характеристиками, подсчет количества совпавших и несовпавших идентифицирующих признаков, итоговом принятии решения об успешности идентификации.

Для практического применения алгоритма создания эталонного отпечатка ФС требуется учесть следующие ситуации:

- эталонный отпечаток ФС не сформирован из-за отсутствия элементов ФС;
- эталонный отпечаток ФС не может быть сформирован полностью;

- создание эталонного отпечатка ФС занимает слишком продолжительное время из-за большого объема информации в ФС.

Учитывая практические ограничения при разработке функции создания эталонного отпечатка ФС, определим её входные и выходные данные.

Входными данными являются:
- корневой каталог, относительно которого будет построен эталонный отпечаток ФС;
- количество идентифицирующих признаков для каждого ритма изменения данных в ФС;
- количество секунд на создание эталона или количество перебираемых элементов ФС, которые будут использованы для построения эталонного отпечатка ФС.

Выходными данными являются:
- массив идентифицирующих признаков как эталонный отпечаток ФС;
- степень заполнения эталонного отпечатка ФС.

Для функции идентификации входными данными являются:
- эталонный отпечаток ФС;
- корневой каталог, для которого проводится идентификация;
- минимальное значение процента совпадений идентифицирующих признаков между эталонным отпечатком ФС и переданным корневым каталогом.

Выходными данными функции идентификации являются:
- процент совпадений между эталонным отпечатком ФС и переданным корневым каталогом;
- ответ да/нет.

Экспериментальное исследование

Для подтверждения выработанных гипотез и разработанных алгоритмов использовалась информация ФС 12 ПЭВМ. Перед развертыванием рабочих мест данные на накопителях информации ПЭВМ были клонированы. На ПЭВМ использовались операционные системы Microsoft Windows 7 Professional и дистрибутив Альт-Линукс. Работа с ПЭВМ велась каждый рабочий день на протяжении календарного года.

Информация для анализа была собрана через год после начала эксплуатации ЭВМ. Для сбора и анализа собранной информации о ФС было разработано соответствующее задачам программное обеспечение [4]. В результате обработки собранных данных не было найдено ни одного совпадающего значащего признака среди всех 12 ПЭВМ, т.е. выбранный набор характеристик уникален для каждого элемента ФС, даже при начальном клонировании информации на накопителях информации. Также было выявлено, что в ФС с большим количеством объектов (более 1000 объектов), обычно в течение одного рабочего дня изменяется менее 1% ФС.

Для проверки работоспособности алгоритмов создания эталонного отпечатка ФС и алгоритма идентификации разработано консольное программное обеспечение, которое запускается в трех режимах:
- выдача справочного сообщения;
- создание эталонного отпечатка ФС;
- идентификация ФС по созданному ранее эталонному отпечатку ФС.

Выводы

Результатами исследования являются:
- определение набора данных идентифицирующего признака на основе анализа предметной области;
- выработанные требования к создаваемому эталонному отпечатку ФС на основе предложенных идентифицирующих признаков;
- разработанный обобщенный алгоритм создания эталонного отпечатка ФС;
- разработанный обобщенный алгоритм идентификации с использованием эталонного отпечатка ФС.

С целью практического подтверждения научных результатов проведено экспериментальное исследование, в рамках которого разработан ряд программ для сбора и обработки собранных данных, а также для создания эталонного отпечатка ФС и идентификации с его использованием.

Ограничением применения разработанных алгоритмов является их низкая эффективность в условиях информационного противоборства, т.к. пользователю достаточно удалить все объекты ФС с накопителя информации, чтобы каталог нельзя было идентифицировать.

К направлениям дальнейшего исследования относятся:
- оценка пригодности произвольного каталога для целей идентификации по отпечатку ФС;
- математическое обоснование выбора формулы для определения расстояния от центра эталона до идентифицируемого отпечатка;
- выделение психологических особенностей пользователя ПЭВМ.

ЛИТЕРАТУРА

1. Слипенчук П. В. Алгоритм извлечения характерных признаков из данных пользовательских активностей // Вопросы кибербезопасности. 2019. № 1(29). С. 53–58.
2. Eckersley P. How Unique Is Your Web Browser? // Electronic Frontier Foundation.— 2014. URL: https://panopticlick.eff.org/static/browser-uniqueness.pdf (дата обращения: 29.10.2019).
3. Абашин В. Г. Идентификация пользователя и идентификация устройства в Интернет / Текст: электронный // VI Международная научно-техническая конференция «Информационные технологии в науке, образовании и производстве (ИТНОП — 2014)».— 2014. URL: http://irsit.ru/files/article/503.pdf (дата обращения: 29.10.2019).
4. Свид. 2018663874 Российская Федерация. Свидетельство об официальной регистрации программы для ЭВМ. ФС идентификация / В. Г. Абашин; заявитель и правообладатель В. Г. Абашин (RU).— № 2018661648; заявл. 16.10.2018; опубл. 06.11.2018, Реестр программ для ЭВМ.— 1 с.

Журнал «Современная наука: актуальные проблемы теории и практики»

Московский Государственный Лингвистический Университет