грамм для ЭВМ № 2009612093 от 24.04.2009. (Федеральная служба по интеллектуальной собственности, патентам и товарным знакам). 2009.

УДК 004.056

РАЗРАБОТКА КОМПЛЕКСНОЙ ОБУЧАЕМОЙ СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИОННЫХ РЕСУРСОВ ОТ ФИШИНГОВЫХ АТАК[1]

А. В. Милошенко, Т. М. Соловьёв, Р. И. Черняк, М. В. Шумская

В настоящее время многие услуги стали доступны через Интернет. Финансовый сектор также не стал исключением. Появились различные платежные системы, интернет-кошельки и терминалы оплаты. Безопасность платежей внутри данных систем обеспечивается множеством высокотехнологичных решений, таких, как сертификаты безопасности, криптографические протоколы и др. Однако все эти решения оказываются неэффективными при применении злоумышленниками методов социальной инженерии, использующих слабости человеческого фактора.

Одним из наиболее распространенных видов такого рода атак сегодня является фишинг [1]. Фишинг (от англ. *phishing*, *password fishing* — выуживание паролей) — это вид сетевого мошенничества, целью которого является получение доступа к конфиденциальным данным пользователей обманным путём. Популярной методикой фишинга является создание поддельных веб-сайтов, внешне неотличимых от подлинных. Ущерб от преступлений, связанных с фишингом, только за 2010 г. исчисляется миллиардами долларов США. При этом, согласно статистике, количество фишинговых атак с каждым годом увеличивается примерно в полтора раза.

Остановить бурное развитие такого рода преступлений можно посредством создания комплексных систем защиты от фишинговых атак. Поскольку арсенал фишеров растёт стремительными темпами, необходимо обеспечивать обучаемость таких систем. В настоящее время не существует подобного рода решений, о чем красноречиво повествует статистика, поэтому создание комплексной обучаемой высокоэффективной системы защиты от фишинговых атак является интересным и актуальным направлением. Создание такой системы предполагает предварительное изучение характерных признаков фишинговых ресурсов и разработку на их основе методов оценивания степени опасности информационного ресурса и определения потенциально опасных ресурсов. Этому, в основном, и посвящена данная работа. На основе полученных результатов предложена схема функционирования системы защиты от фишинговых атак.

Характерные признаки фишинговых ресурсов

Анализ существующих фишинговых ресурсов позволил составить перечень их характерных признаков.

1. *Сходство графического контента*

Основная задача злоумышленников при проведении фишинговой атаки — заставить пользователя поверить в аутентичность фишингового ресурса. Наиболее простым способом сделать это является заимствование графического оформления у атакуемого сайта.

Несмотря на большое количество исследований, задача определения степени сходства изображений в настоящее время не имеет универсального и эффективного решения. Объясняется это в первую очередь тем, что понятие похожести двух изображе-

[1]Работа выполнена в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 гг. (гос. контракт № П1010).

ний неразрывно связано с особенностями человеческого восприятия и вследствие этого трудно формализуемо. Уверенно можно говорить лишь о факте полной идентичности изображений, которая определяется попиксельным сравнением или сравнением на основе значений некоторой хэш-функции.

Для определения похожести изображений, подвергаемых незначительной модификации, возможно использование методов, основанных на сравнении усредненных характеристик изображения, например метода поблочного анализа цвета или метода сравнения гистограмм.

Более сложный метод сравнения изображений — метод SURF [2]. Он базируется на сравнении ключевых точек, инвариантных как к геометрическим и фотометрическим преобразованиям, так и к изменению масштаба. Нахождение ключевых точек осуществляется с помощью прохода по пикселям изображения и поиска максимума гессиана — определителя матрицы, составленной из вторых частных производных (матрицы Гессе) функции яркости изображения. Реализация данного метода содержится во многих открытых библиотеках алгоритмов, например в библиотеке openCV.

2. *Сходство текстового контента*

Понятие похожести текстов также является нечетким, и, как и в предыдущем случае, с уверенностью можно заявлять лишь об их полной идентичности. Однако на практике при изменении текстового контента оригинального ресурса злоумышленники, как правило, используют стандартный набор преобразований, таких, как:

— вставка наборов случайных символов;
— произвольная вставка и удаление пробелов;
— замена символов одной кодировки на похожие по написанию символы другой кодировки;
— вставка ключевых слов на случайные позиции.

В большинстве случаев производится сравнение не самих исходных текстов, а построенных на их основе подмножеств или отпечатков этих подмножеств — значений некоторых хэш-функций. В зависимости от способов построения отпечатков методы определения похожести текстов можно разделить на два класса — синтаксические и лексические. В первом случае анализируются построенные по определенным правилам последовательности слов из текста, во втором — строятся словари ключевых слов. С точки зрения эффективности и простоты реализации оптимальными являются два метода определения близости текстового контента: метод определения индекса повторяемости и метод поиска длинных предложений [3].

Сложность выявления сходства контента растёт пропорционально числу защищаемых ресурсов, что серьёзно сказывается на производительности антифишинговой системы. Проверка признаков, перечисляемых далее, не требует сравнения с оригиналом.

3. *Наличие ресурса в фишинговых базах*

Интернет-сообществом поддерживается большое количество баз опасных ресурсов. Одни пользователи добавляют ссылки на такие ресурсы, другие — подтверждают или опровергают их опасность. Окончательное решение принимается администрацией конкретной базы. Как правило, фишинговые ресурсы попадают в такие списки в течение нескольких суток. Данный признак является очень сильным, однако сама концепция ведения списка подразумевает постоянную его актуализацию, которая всегда отстаёт от деятельности злоумышленников.

4. *Использование особенностей формата URL*

Формат URL имеет много параметров, часть из которых крайне редко применяется на практике. Зачастую злоумышленники используют более полную форму, добавляя в неё редко используемые параметры, чтобы ввести пользователя в заблуждение и убедить его в подлинности фишингового ресурса.

5. *Подозрительные регистрационные данные ресурса*

К регистрационным данным можно отнести географическое положение, дату регистрации домена, имя собственника сайта или название организации-владельца. Как правило, фишинговые сайты активны в первые пять дней после их создания. В связи с этим большое значение имеет дата регистрации домена. Часто фишинговые ресурсы регистрируются в стране, отличной от той, в которой расположен оригинальный сайт, поэтому необходимо также отслеживать соответствие домена верхнего уровня реальному местонахождению сайта.

6. *Наличие ресурса на одном IP-адресе с выявленными ранее фишинговыми ресурсами*

Расположение нескольких ресурсов на одном IP-адресе является достаточно распространенной ситуацией, поэтому целесообразно использование списков IP-адресов, на которых были замечены фишинговые ресурсы. Таким образом можно идентифицировать ресурс как потенциально опасный, если он расположен на одном IP с множеством фишинговых ресурсов.

7. *Наличие изображений, содержащих в себе текст в графическом представлении*

Данный способ применяется злоумышленниками для усложнения идентификации опасного ресурса. Пользователь будет воспринимать такой объект как обычный текст, а автоматизированные антифишинговые системы — как изображения. Если контент изображений дополнительно не анализируется, система может принять неправильное решение о степени опасности ресурса.

8. *Использование «неоправданно большого» количества скриптов*

Как правило, объём исполняемого кода на странице растёт пропорционально её информативности и предоставляемой функциональности. «Неоправданно большое» количество скриптов может являться признаком недокументированных возможностей. Для определения максимально допустимого количества скриптов предлагается использовать статистистические данные из выборки в конкретной предметной области.

Все выделенные признаки имеют разную значимость, которая может варьироваться в зависимости от их комбинаций. Это говорит о необходимости разработки механизмов совокупного анализа признаков — механизмов принятия решения о степени опасности ресурса.

Методы оценивания степени опасности ресурса

Далее предлагаются алгоритмы, определяющие степень опасности информационных ресурсов с учетом описанных признаков. На вход каждого алгоритма поступает информационный ресурс, выходом является рациональное число из отрезка $[0, 1]$, которое характеризует степень опасности ресурса; наибольшей опасности соответствует единица, наименьшей — ноль.

1. *Последовательное использование алгоритмов оценки признаков*

Выполняется непосредственная проверка всех признаков. Поскольку различные алгоритмы анализа данных используют признаки разной значимости, на вход объемлющего алгоритма необходимо подать также *таблицу значимости признаков*.

Метод работает следующим образом. Если ресурс удовлетворяет некоторому признаку, то результат увеличивается на значимость этого признака. По завершении работы алгоритма результат нормируется, с тем чтобы он содержался в отрезке $[0, 1]$.

Основное достоинство данного подхода — простота реализации; недостатки — необходимость проверки всех признаков и негибкий учет значимости каждого из них.

2. *Использование булевой функции*

При определении степени опасности информационного ресурса некоторый набор признаков может оказаться достаточным для отнесения ресурса к разряду опасных. Например, если доменное имя сайта содержится в фишинговых базах, нет необходимости проверять остальные признаки — такой сайт априори представляет опасность.

Для работы алгоритма необходимо задать булеву функцию $f : \{0, 1\}^n \to \{0, 1\}$, принимающую значение 1 на тех и только тех наборах значений признаков информационного ресурса, которые гарантируют опасность последнего. Выделение достаточных подмножеств признаков (определение функции f) относится к предвычислениям, а сам алгоритм состоит в ее однократном вычислении на наборе аргументов, соответствующем признакам проверяемого ресурса.

Недостатком алгоритма является высокая зависимость от таблицы значимости признаков. Кроме того, процесс определения достаточных условий не автоматизирован, поэтому на данном этапе возможны ошибки разного рода.

3. *Использование нейронных сетей*

Одним из способов отказа от таблицы значимости признаков является использование статистических методов. Пусть значение функции $f : [0, 1]^n \to \{0, 1\}$ характеризует опасность (1) или безопасность (0) ресурса. Аргументы функции — вещественные переменные со значениями в отрезке $[0, 1]$, определенными алгоритмами оценивания признаков.

Функция f неизвестна, и вычислить ее непосредственно невозможно. Поведение функции может быть промоделировано с помощью нейронной сети. Вычисление функции f происходит в два этапа:

1) обучение нейронной сети;
2) вычисление значения функции.

При этом обучение может производиться параллельно, не оказывая существенного влияния на производительность самого алгоритма. В качестве обучающей выборки могут быть использованы базы фишинговых ресурсов.

Основными преимуществами данного способа являются гибкость при оценивании значимости признаков и высокая адекватность действительности. Последнее достигается путем использования в алгоритме статистических методов.

4. *Использование методов линейной регрессии*

Обобщим предыдущую модель на непрерывный случай, полагая, что степень опасности ресурса оценивается как вещественное число в отрезке $[0, 1]$ и определяется функцией $f : [0, 1]^n \to [0, 1]$.

Одним из статистических способов вычисления неизвестной вещественнозначной функции f является метод линейной регрессии. Он позволяет построить функцию g — линейную аппроксимацию f в виде $g = \sum\limits_{i=1}^{n} w_i x_i$. Для того чтобы оценить степень опасности того или иного ресурса, достаточно вычислить значение функции g от соответствующих ему аргументов.

Для построения приближенной функции используются заданные значения опасностей некоторых информационных ресурсов. При неизвестной функции определения

степени опасности не вполне понятно, как получить достаточно большое количество значений этой функции на заданных значениях аргументов. Эта проблема решается с помощью экспертных оценок.

Механизмы выбора потенциально опасных ресурсов

Поскольку наибольший ущерб фишинговые ресурсы наносят в первые сутки своего существования, для обеспечения эффективной защиты от них необходимо проводить непрерывный мониторинг информационных ресурсов. Однако анализировать все ресурсы сети Интернет не представляется возможным, поэтому следует сконцентрироваться на тех, которые с наибольшей вероятностью могут оказаться опасными.

При разработке фишинговых сайтов злоумышленники ориентируются на то, что пользователь либо допустит ошибку при наборе доменного имени, либо перейдёт по ссылке, визуально похожей на доверенный ресурс. Согласно этому, были выделены несколько принципов построения доменных имён, представляющих потенциальную опасность по отношению к заданному.

1. *Доменные имена с незначительными изменениями*

Данный принцип включает в себя следующие методы генерации: опечатки (замена символа другим символом, расположенным на соседней клавише); выпадение гласной между согласными; приписывание к началу доменного имени комбинации «www»; вставка дефиса; вставка символа вследствие нажатия нужной клавиши и соседней; удвоение символов доменного имени; замена символов таким образом, что полученное имя схоже по произношению с исходным; добавление в конце доменного имени, написанного латиницей и оканчивающегося на гласную, символов «s», «r» или «rs»; перестановка букв.

2. *Транслитерация кириллического доменного имени*

Под транслитерацией домена понимается замена кириллических символов на один или несколько латинских символов. Можно выделить три принципа транслитерации: транскрипция (преобразование на основе схожести соответствующих звуков), преобразование типа «один символ — одна буква» и преобразование на основе графической схожести символов. При генерации доменных имён могут использоваться любые из принципов транслитерации, а также их комбинации.

3. *Перевод на английский язык лексем заданного доменного имени*

В данном случае сгенерированные доменные имена содержат лексемы, являющиеся либо переводом, либо транслитерацией (если невозможно осуществить перевод) лексем заданного домена.

4. *Доменные имена, полученные на основе использования сервисов поисковых систем по поиску URL*

В первые дни существования опасные ресурсы имеют высокую популярность, поэтому при использовании запросов, осуществляющих поиск заданных лексем в URL, такие ресурсы могут оказаться в числе первых, выданных поисковой системой.

5. *Доменные имена, содержащие комбинации лексем*

Принцип генерации доменов, содержащих комбинации лексем, включает в себя следующие варианты написания: слитное написание, использование дефиса, а также их комбинации.

Для последующего мониторинга с целью определения наличия ресурса и оценки степени его опасности каждое сгенерированное имя должно быть расширено каждым возможным для него постфиксом, идентифицирующим доменную зону. Кроме этого, предлагается сгенерировать группу доменных имён, используя уже расширенные домены. Из каждого доменного имени вида *доменное_имя.доменная_зона1* генериру-

ются всевозможные имена вида *доменное_имя|доменная_зона2.доменная_зона1*, где операция | — операция конкатенации, *доменная зона2* принимает все значения из списка возможных доменных зон для доменного имени *доменное_имя*. Такая мера целесообразна, поскольку при разработке фишинговых ресурсов злоумышленники часто используют «двойные» доменные зоны.

Система защиты от фишинговых атак

На основе полученных результатов может быть создана полноценная система защиты от фишинговых атак. Предлагается следующая схема функционирования такой системы защиты для каждого информационного ресурса:

1) генерация списка потенциально опасных доменных имен;
2) получение списка зарегистрированных и доступных потенциально опасных информационных ресурсов;
3) определение степени опасности каждого из потенциально опасных информационных ресурсов;
4) пополнение антифишинговых баз вновь обнаруженными опасными ресурсами.

Для обеспечения высокой степени защиты ресурса необходимо регулярно повторять последовательность действий 2 — 4 и проводить непрерывное обучение системы.

ЛИТЕРАТУРА

1. *Lininger R. and Vines D.* Phishing: Cutting the Identity Theft Line. Wiley, 2005. 334 p.
2. *Bay H., Ess A., and Tuytelaars T.* SURF: Speeded Up Robust Features // Computer Vision and Image Understanding (CVIU). 2008. V. 110. No. 3. P. 346–359.
3. *Зеленков Ю. Г., Сегалович И. В.* Сравнительный анализ методов определения нечетких дубликатов для Web-документов // Труды 9 Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» — RCDL'2007. Переславль-Залесский, Россия, 2007. С. 166–174.

УДК 004.94

ПОДХОДЫ К РАЗРАБОТКЕ ДИСКРЕЦИОННОЙ ДП-МОДЕЛИ СОВРЕМЕННЫХ ЗАЩИЩЕННЫХ ОПЕРАЦИОННЫХ СИСТЕМ[1]

В. Г. Проскурин

Рассматривается дискреционная ДП-модель, развивающая и конкретизирующая существующие ДП-модели для случая, когда моделируемая компьютерная система является защищенной операционной системой (ОС). Далее будем называть предлагаемую модель ЗОС ДП-моделью. Она основывается на базовой ДП-модели [1], ФПАС ДП-модели [2] и ФС ДП-модели [3], а также отчасти на модели мандатной политики целостности информации Биба [1]. Будем использовать систему обозначений [1].

Предлагаемая модель характеризуется следующими основными отличиями от ранее разработанных ДП-моделей.

1. В модель вводится новый элемент — множество учетных записей пользователей U. Права доступа к сущностям назначаются не напрямую субъектам, как в других ДП-моделях, а учетным записям пользователей, которые делегируют права доступа субъектам, выполняющимся от имени соответствующей учетной записи. Другими словами, множество разрешенных прав доступа R определяется как подмножество множества $U \times (E \cup U) \times R_r$. Задается отображение $user : S \to U$, определяющее для каждого

[1]Работа выполнена при поддержке Министерства образования и науки Российской Федерации.