Международный союз электросвязи

МСЭ-Т

СЕКТОР СТАНДАРТИЗАЦИИ
ЭЛЕКТРОСВЯЗИ МСЭ

X.1235

(01/2022)

СЕРИЯ X: СЕТИ ПЕРЕДАЧИ ДАННЫХ,
ВЗАИМОСВЯЗЬ ОТКРЫТЫХ СИСТЕМ
И БЕЗОПАСНОСТЬ

Безопасность киберпространства –
Противодействие спаму

Технологии противодействия спуфингу веб-сайтов для организаций электросвязи

Рекомендация МСЭ-Т X.1235

РЕКОМЕНДАЦИИ МСЭ-Т СЕРИИ X

СЕТИ ПЕРЕДАЧИ ДАННЫХ, ВЗАИМОСВЯЗЬ ОТКРЫТЫХ СИСТЕМ И БЕЗОПАСНОСТЬ

СЕТИ ПЕРЕДАЧИ ДАННЫХ ОБЩЕГО ПОЛЬЗОВАНИЯ	X.1–X.199
ВЗАИМОСВЯЗЬ ОТКРЫТЫХ СИСТЕМ	X.200–X.299
ВЗАИМОДЕЙСТВИЕ МЕЖДУ СЕТЯМИ	X.300–X.399
СИСТЕМЫ ОБРАБОТКИ СООБЩЕНИЙ	X.400–X.499
СПРАВОЧНИК	X.500–X.599
ОРГАНИЗАЦИЯ СЕТИ ВОС И СИСТЕМНЫЕ АСПЕКТЫ	X.600–X.699
УПРАВЛЕНИЕ В ВОС	X.700–X.799
БЕЗОПАСНОСТЬ	X.800–X.849
ПРИЛОЖЕНИЯ ВОС	X.850–X.899
ОТКРЫТАЯ РАСПРЕДЕЛЕННАЯ ОБРАБОТКА	X.900–X.999
БЕЗОПАСНОСТЬ ИНФОРМАЦИИ И СЕТЕЙ	
Общие аспекты безопасности	X.1000–X.1029
Безопасность сетей	X.1030–X.1049
Управление безопасностью	X.1050–X.1069
Телебиометрия	X.1080–X.1099
БЕЗОПАСНЫЕ ПРИЛОЖЕНИЯ И УСЛУГИ (1)	
Безопасность многоадресной передачи	X.1100–X.1109
Безопасность домашних сетей	X.1110–X.1119
Безопасность подвижной связи	X.1120–X.1139
Безопасность веб-среды (1)	X.1140–X.1149
Безопасность приложений (1)	X.1150–X.1159
Безопасность одноранговых сетей	X.1160–X.1169
Безопасность сетевой идентификации	X.1170–X.1179
Безопасность IPTV	X.1180–X.1199
БЕЗОПАСНОСТЬ КИБЕРПРОСТРАНСТВА	
Кибербезопасность	X.1200–X.1229
Противодействие спаму	**X.1230–X.1249**
Управление определением идентичности	X.1250–X.1279
БЕЗОПАСНЫЕ ПРИЛОЖЕНИЯ И УСЛУГИ (2)	
Связь в чрезвычайных ситуациях	X.1300–X.1309
Безопасность повсеместных сенсорных сетей	X.1310–X.1319
Безопасность умных электросетей	X.1330–X.1339
Сертифицированная электронная почта	X.1340–X.1349
Безопасность интернета вещей (IoT)	X.1350–X.1369
Безопасность интеллектуальных транспортных систем (ИТС)	X.1370–X.1399
Безопасность технологии распределенного реестра (DLT)	X.1400–X.1429
Безопасность приложений (2)	X.1450–X.1459
Безопасность веб-среды (2)	X.1470–X.1489
ОБМЕН ИНФОРМАЦИЕЙ, КАСАЮЩЕЙСЯ КИБЕРБЕЗОПАСНОСТИ	
Обзор кибербезопасности	X.1500–X.1519
Обмен информацией об уязвимости/состоянии	X.1520–X.1539
Обмен информацией о событии/инциденте/эвристических правилах	X.1540–X.1549
Обмен информацией о политике	X.1550–X.1559
Эвристические правила и запрос информации	X.1560–X.1569
Идентификация и обнаружение	X.1570–X.1579
Гарантированный обмен	X.1580–X.1589
Киберзащита	X.1590–X.1599
БЕЗОПАСНОСТЬ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ	
Обзор безопасности облачных вычислений	X.1600–X.1601
Проектирование безопасности облачных вычислений	X.1602–X.1639
Передовой опыт и руководящие указания в области облачных вычислений	X.1640–X.1659
Обеспечение безопасности облачных вычислений	X.1660–X.1679
Другие вопросы безопасности облачных вычислений	X.1680–X.1699
КВАНТОВАЯ СВЯЗЬ	
Терминология	X.1700–X.1701
Квантовый генератор случайных чисел	X.1702–X.1709
Структура безопасности QKDN	X.1710–X.1711
Проектирование безопасности QKDN	X.1712–X.1719
Методы обеспечения безопасности QKDN	X.1720–X.1729
БЕЗОПАСНОСТЬ ДАННЫХ	
Безопасность больших данных	X.1750–X.1759
Защита данных	X.1770–X.1789
БЕЗОПАСНОСТЬ СЕТЕЙ IMT-2020	X.1800–X.1819

Для получения более подробной информации просьба обращаться к перечню Рекомендаций МСЭ-Т.

Рекомендация МСЭ-Т X.1235

Технологии противодействия спуфингу веб-сайтов для организаций электросвязи

Резюме

Подделка (спуфинг) веб-сайтов – серьезная угроза для организаций электросвязи, особенно для операторов. Операторам электросвязи рекомендуется применять технологии противодействия спуфингу веб-сайтов для защиты своих клиентов и сохранения своей репутации и доходов. В Рекомендации МСЭ-Т X.1235 анализируются основные способы спуфинга веб-сайтов и предлагаются технологии для выявления поддельных веб-сайтов. Эти меры можно рассматривать как руководящие принципы защиты веб-сайтов от спуфинга для организаций электросвязи. Аналогичный подход можно применять для защиты от спуфинга любых веб-сайтов, включая сайты банков, страховых компаний, интернет-магазинов и т. д.

Хронологическая справка

Издание	Рекомендация	Утверждено	Исследовательская комиссия	Уникальный идентификатор*
1.0	МСЭ-Т X.1235	07.01.2022 г.	17-я	11.1002/1000/14797

Ключевые слова

Контрмеры, спуфинг веб-сайтов.

* Для получения доступа к Рекомендации наберите в адресном поле вашего браузера URL http://handle.itu.int/, после которого укажите уникальный идентификатор Рекомендации.
Например: http://handle.itu.int/11.1002/1000/11830-en.

ПРЕДИСЛОВИЕ

Международный союз электросвязи (МСЭ) является специализированным учреждением Организации Объединенных Наций в области электросвязи и информационно-коммуникационных технологий (ИКТ). Сектор стандартизации электросвязи МСЭ (МСЭ-Т) – постоянный орган МСЭ. МСЭ-Т отвечает за изучение технических, эксплуатационных и тарифных вопросов и за выпуск Рекомендаций по ним в целях стандартизации электросвязи на всемирной основе.

На Всемирной ассамблее по стандартизации электросвязи (ВАСЭ), которая проводится каждые четыре года, определяются темы для изучения исследовательскими комиссиями МСЭ-Т, которые, в свою очередь, вырабатывают Рекомендации по этим темам.

Утверждение Рекомендаций МСЭ-Т осуществляется в соответствии с процедурой, изложенной в Резолюции 1 ВАСЭ.

В некоторых областях информационных технологий, которые входят в компетенцию МСЭ-Т, необходимые стандарты разрабатываются на основе сотрудничества с ИСО и МЭК.

ПРИМЕЧАНИЕ

В настоящей Рекомендации термин "администрация" используется для краткости и обозначает как администрацию электросвязи, так и признанную эксплуатационную организацию.

Соблюдение положений данной Рекомендации осуществляется на добровольной основе. Однако данная Рекомендация может содержать некоторые обязательные положения (например, для обеспечения функциональной совместимости или возможности применения), и в таком случае соблюдение Рекомендации достигается при выполнении всех указанных положений. Для выражения требований используются слова "следует", "должен" (shall) или некоторые другие обязывающие выражения, такие как "обязан" (must), а также их отрицательные формы. Употребление таких слов не означает, что от какой-либо стороны требуется соблюдение положений данной Рекомендации.

ПРАВА ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

МСЭ обращает внимание на вероятность того, что практическое применение или выполнение настоящей Рекомендации может включать использование заявленного права интеллектуальной собственности. МСЭ не занимает какую бы то ни было позицию относительно подтверждения, действительности или применимости заявленных прав интеллектуальной собственности независимо от того, доказываются ли такие права членами МСЭ или другими сторонами, не относящимися к процессу разработки Рекомендации.

На момент утверждения настоящей Рекомендации МСЭ не получил извещения об интеллектуальной собственности, защищенной патентами, которые могут потребоваться для выполнения настоящей Рекомендации. Однако те, кто будет применять Рекомендацию, должны иметь в виду, что вышесказанное может не отражать самую последнюю информацию, и поэтому им настоятельно рекомендуется обращаться к патентной базе данных БСЭ по адресу http://www.itu.int/ITU-T/ipr/.

СОДЕРЖАНИЕ

Введение

В последние годы спуфинг веб-сайтов является одним из основных методов мошенничества в интернете. Мошенники обычно выбирают веб-сайты известных организаций или компаний для сбора учетных данных посетителей или распространения вредоносного ПО. В результате как посетители, так и операторы электросвязи терпят финансовые потери, а операторы электросвязи – еще и ущерб, нанесенный деловой репутации.

Поскольку веб-сайты операторов электросвязи стали для клиентов важнейшим порталом, где они могут запрашивать все виды услуг и подписываться на них, во всем мире наблюдаются непрерывные попытки мошенников подделать веб-сайты операторов электросвязи, чтобы обмануть клиентов. В настоящей Рекомендации проводится тщательный анализ спуфинга веб-сайтов и даются рекомендации по ряду контрмер для организованного противодействия.

Рекомендация МСЭ-Т X.1235

Технологии противодействия спуфингу веб-сайтов для организаций электросвязи

1 Сфера применения

В настоящей Рекомендации организациям электросвязи предлагаются технологии для своевременного выявления спуфинга своих веб-сайтов и их защиты от подделки. Сначала в ней дается системный анализ параметров и функций спуфинга, а затем предлагаются передовые методы применения сетевых технологий противодействия спуфингу веб-сайтов с дополнением в виде пользовательских технологий.

Соответствие настоящей Рекомендации не следует рассматривать в качестве какого-либо доказательства заявленного соблюдения любых национальных или региональных законов, норм или политики. Описанные в настоящей Рекомендации технические, организационные и процедурные средства ни в коей мере не гарантируют создания какого-либо уровня безопасности, который может налагаться на определенную корреспонденцию в соответствии с конкретным национальным или региональным законом, нормой или политикой.

Аналогичный подход можно применять для защиты от спуфинга любых веб-сайтов, включая веб-сайты банков, страховых компаний, интернет-магазинов и т. д.

2 Справочные документы

Указанные ниже Рекомендации МСЭ-Т и другие справочные документы содержат положения, которые путем ссылок на них в данном тексте составляют положения настоящей Рекомендации. На момент публикации указанные издания были действующими. Все Рекомендации и другие справочные документы могут подвергаться пересмотру; поэтому всем пользователям данной Рекомендации предлагается изучить возможность применения последнего издания Рекомендаций и других справочных документов, перечисленных ниже. Перечень действующих на настоящий момент Рекомендаций МСЭ-Т регулярно публикуется. Ссылка на документ, приведенный в настоящей Рекомендации, не придает ему как отдельному документу статус Рекомендации.

Отсутствуют.

3 Определения

3.1 Термины, определенные в других документах

Отсутствуют.

3.2 Термины, определенные в настоящей Рекомендации

В настоящей Рекомендации определены следующие термины.

3.2.1 сверточная нейронная сеть (convolutional neural network): Алгоритм глубокого обучения, который принимает входное изображение, присваивает значения различным его аспектам/объектам и может отличать одно изображение от другого.

3.2.2 сети с долгой краткосрочной памятью (long short-term memory networks): Тип рекуррентных нейронных сетей, способных к изучению порядка зависимости в задачах прогнозирования последовательностей.

3.2.3 рекуррентная нейронная сеть (recurrent neural network): Класс нейронных сетей, которые позволяют использовать предыдущие выходные данные в качестве входных данных при наличии скрытых состояний.

3.2.4 масштабно-инвариантное преобразование признаков (scale-invariant feature transform (SIFT)): В технологии компьютерного зрения алгоритм выявления признаков для обнаружения и описания локальных признаков изображения, который составляет описательный идентификатор, инвариантный к сдвигам, поворотам и преобразованиям масштабирования в области изображения и устойчивый к умеренным преобразованиям перспективы и вариациям освещения. На практике описательный идентификатор SIFT оказался чрезвычайно полезным для сопоставления изображений и распознавания объектов в реальных условиях.

3.2.5 быстрообнаруживаемые надежные признаки (speeded up robust features): В технологии компьютерного зрения детектор локальных признаков и описательный идентификатор для обнаружения и описания локальных признаков изображения, который работает в несколько раз быстрее и более устойчив к различным преобразованиям изображения, чем масштабно-инвариантное преобразование признаков (SIFT).

3.2.6 поддельный веб-сайт (spoofed website): Веб-сайт, созданный с помощью спуфинга (см. пункт 3.2.8).

3.2.7 метод опорных векторов (support vector machine): Модель машинного обучения с учителем, которая решает задачи классификации по двум группам с заданными наборами помеченных обучающих данных для каждой категории; они позволяют классифицировать новые наборы.

3.2.8 спуфинг веб-сайта (website spoofing): Набор злонамеренных действий в целях имитации веб-сайта, хорошо известного широкой публике или группе людей; затем фальшивый веб-сайт используется для злоупотребления доверием посетителей и совершения злонамеренных/ противоправных действий, таких как мошенничество, нарушение конфиденциальности и т. д.

4 Сокращения и акронимы

В настоящей Рекомендации используются следующие сокращения и акронимы.

AI	Artificial Intelligence	ИИ	Искусственный интеллект
DNS	Domain Name Service		Служба доменных имен
GUI	Graphical User Interface		Графический интерфейс пользователя
LSTM	Long Short-Term Memory		Долгая краткосрочная память
OCR	Optical Character Recognition		Оптическое распознавание символов
PII	Personally Identifiable Information		Информация, позволяющая установить личность
QR	Quick Response		Немедленное реагирование
SIFT	Scale-Invariant Feature Transform		Масштабно-инвариантное преобразование признаков
SURF	Speeded Up Robust Features		Быстрообнаруживаемые надежные признаки
SVM	Support Vector Machine		Метод опорных векторов
URL	Uniform Resource Locator		Унифицированный указатель ресурса

5 Соглашения

В настоящей Рекомендации используются следующие соглашения по терминологии:

Ключевое слово "**следует**" или "**должен**" (should) означает требование, которое рекомендуется, но не является абсолютно необходимым; таким образом, для заявления о соответствии это требование не является обязательным;

Ключевое слово "**может**" (may) означает необязательное требование, которое допустимо, но не имеет рекомендательного значения;

В тексте настоящей Рекомендации иногда встречается слово "**может**" (can); в этом случае его следует понимать как **"имеется возможность"**.

6 Что такое спуфинг веб-сайтов

Подделка (спуфинг) веб-сайта – это ряд злонамеренных действий по имитации веб-сайта, известного широкой публике или группе людей. Поддельные веб-сайты собирают учетные данные посетителей или распространяют вредоносное ПО, преследуя злонамеренные/незаконные цели, такие как мошенничество, нарушение конфиденциальности и т. д.

Спуфинг веб-сайтов может нанести серьезный ущерб по нескольким причинам. Например, он может приводить к финансовым потерям как для клиентов, так и для операторов, наносить ущерб репутации операторов и т. д. Более того, собранная информация о пользователях может применяться в злонамеренных целях в течение длительного времени, даже когда поддельный веб-сайт исчезнет. Поэтому важно, чтобы организации электросвязи, особенно операторы, внедряли полезные технологии для противодействия спуфингу веб-сайтов.

6.1 Сценарий спуфинга веб-сайта

Типичный сценарий создания поддельного веб-сайта для достижения неправомерной цели состоит из трех этапов, как показано на рисунке 6-1.

– **Первый этап**. Создание веб-сайта
  - Шаг первый. Разработка и эксплуатация поддельного веб-сайта, идентичного или похожего на некий известный веб-сайт.

– **Второй этап**. Распространение унифицированного указателя ресурса (URL)
  - Шаг второй. Распространение адреса поддельного веб-сайта с помощью различных методов, таких как электронная почта, услуга передачи коротких сообщений, услуги мгновенного обмена сообщениями и т. д.

– **Третий этап**. Сбор учетных данных или распространение вредоносного ПО
  - Шаг третий. Посетителей обманным путем вынуждают ввести учетные данные или загрузить вредоносное ПО. Получив доступ к поддельному веб-сайту, потерпевший может ошибочно полагать, что он подлинный, и войти со своими учетными данными или загрузить вредоносное программное обеспечение.
  - Шаг четвертый. Получение учетных данных посетителя и подготовка следующего акта мошенничества или другой злонамеренной деятельности.

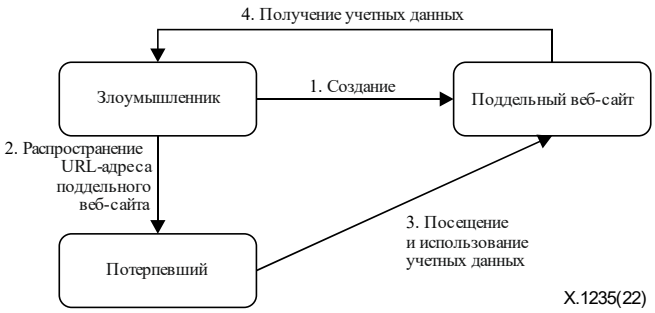


Рисунок 6-1 – Четыре этапа использования поддельного веб-сайта для сбора учетных данных пользователей

6.2 Характеристики спуфинга веб-сайтов

6.2.1 Характеристики поддельных веб-сайтов

На поддельных веб-сайтах используются одни и те же методы, чтобы заставить посетителей принять их за подлинные. Все они имеют несколько общих характеристик:

– визуальное сходство – поддельный веб-сайт может быть полной копией подлинного, включая его визуальный контент и интерактивную логику;

– заимствование визуальных элементов – подложный веб-сайт может заимствовать некоторые важные визуальные элементы у подлинного веб-сайта;

– ссылки на известные веб-сайты – подложный веб-сайт может содержать ссылки на другие известные веб-сайты, чтобы имитировать веб-сайт авторизованного партнера, родственный веб-сайт или другую версию того же веб-сайта;

– похожее доменное имя – доменное имя поддельного веб-сайта очень похоже на подлинное;

– скрытое доменное имя – иногда злоумышленники используют службы сокращения URL-адресов или коды немедленного реагирования (QR), чтобы скрыть реальное доменное имя веб-сайта при распространении ссылки;

– дублирование доменных имен – для поддельного веб-сайта может быть зарезервирован целый набор доменных имен. Некоторые из них публикуются только один раз, чтобы продлить жизнь такого поддельного веб-сайта;

– обфусцированный исходный код – исходный код подложного веб-сайта может быть обфусцирован (запутан), чтобы обмануть программы, производящие сканирование в целях безопасности;

– ввод учетных данных – поддельные веб-сайты могут выманивать у посетителей учетные данные;

– перенаправление на известный сайт – во время простоя поддельный веб-сайт может перенаправлять посетителей на известные веб-сайты (например, google.com) во избежание автоматического сканирования на предмет мошенничества в интернете;

– поддельная адресная строка – некоторые поддельные веб-сайты специально предназначены для мобильных веб-браузеров. Злоупотребляя функцией графического интерфейса пользователя (GUI) мобильного браузера, веб-сайт может разместить поддельную копию адресной строки мобильного браузера в верхней части экрана мобильного телефона и закрепить эту копию.

6.2.2 Характеристики соответствующих видов деятельности

Могут применяться многочисленные способы заманить или обмануть людей, так чтобы они посещали поддельные веб-сайты и взаимодействовали с ними. Ниже перечислены некоторые характеристики или варианты этих действий:

– распространение сообщений, чтобы как можно больше людей узнало о существовании поддельного веб-сайта;

– взлом или порча службы доменных имен (DNS), так что посетители подлинного веб-сайта будут перенаправляться на поддельный;

– обещание тех или иных выгод (купонов, подарков и т. п.), чтобы побудить пользователей переходить по ссылкам на поддельные веб-сайты;

– установка на телефон потерпевших вредоносного ПО, заменяющего запрос к подлинному сайту запросами к поддельному;

– обращение к пользователям от имени службы клиентской поддержки, чтобы навязать им ссылки на поддельные веб-сайты.

6.3 Последствия

Поддельный веб-сайт может принести пользователям и владельцам веб-сайтов множество неприятностей. Возможные последствия перечислены ниже.

– Материальные потери пользователей – могут быть похищены учетные данные и информация, позволяющая установить личность (PII) пользователя. Мошенники могут использовать эту информацию для дальнейшего мошенничества. Пользователей могут побудить загрузить трояны, программы криптомайнинга и другие вредоносные программы. Контроль над терминалами пользователей может обеспечить возможность для кражи финансовых счетов, распространения спама и заражения других пользователей. Эти действия могут привести к потере заряда батареи или выходу терминалов из строя, а также к огромным прямым финансовым потерям.

– Материальные потери операторов – обычные транзакции между пользователями и операторами будут нарушены и ограничены даже в долгосрочной перспективе.

– Репутационные потери – будет подорвана репутация подлинных веб-сайтов и даже утрачено общественное доверие к оператору; это приведет к большому количеству жалоб клиентов и долгосрочной негативной оценке.

7 Контрмеры

Меры противодействия спуфингу веб-сайтов можно разделить на две категории – выявление и защита. Первые выявляют поддельные веб-сайты с других веб-сайтов; см. пункт 7.1. Вторые защищают пользователей от посещения известных поддельных веб-сайтов или от обмана с их стороны; см. пункт 7.2.

Организациям электросвязи необязательно использовать все рекомендованные в этом разделе технологии одновременно. Они должны гибко применять соответствующие технологии, описанные в этом разделе, действуя на основе собранных данных, законодательства, требований абонентов и т. д.

7.1 Выявление

Рекомендуется рассмотреть следующие контрмеры в соответствии с характеристиками поддельных веб-сайтов, как указано в пункте 6.2.1.

7.1.1 Сравнение схожих доменных имен

Операторам рекомендуется вести список доменных имен известных веб-сайтов, которые подвергаются подделке. Если доменное имя очень похоже на доменное имя из списка (но не совпадает с ним), это может быть доменное имя поддельного веб-сайта. Если доменное имя сокращено службой сокращения URL-адресов, то перед сравнением его следует восстановить до исходного URL-адреса.

Существует множество методов определения подобия двух доменных имен, в том числе: расстояние преобразования, оценка сходства по Жаккару, наибольшая общая подпоследовательность, преобразование по визуальному сходству и т. д.

- **Метод на основе расстояния преобразования**. Для двух доменных имен A и B вычисляется минимальное количество правок, необходимых для преобразования из A в B. Чем меньше расстояние преобразования, тем больше сходство.
- **Метод оценки сходства по Жаккару**. Размер пересечения двух наборов символов доменного имени делится на размер их объединения. Чем больше отношение, тем больше сходство.
- **Метод наибольшей общей подпоследовательности**. Рассчитывается длина наибольшей общей подпоследовательности двух доменных имен. Чем больше длина, тем больше сходство.

  ПРИМЕЧАНИЕ. – Наибольшая общая подпоследовательность (longest common subsequence) – это самая длинная подпоследовательность, общая для двух данных последовательностей, причем элементы подпоследовательности не обязательно должны занимать последовательные позиции в исходных последовательностях. Общая подпоследовательность должна быть строго возрастающей последовательностью индексов двух данных последовательностей. Например, если заданы две последовательности – "abcde" и "akcve", – то наибольшей общей подпоследовательностью двух этих последовательностей будет "ace".
- **Метод преобразования по визуальному сходству**. Перед сравнением заменяются визуально схожие символы. Например, можно "0" заменить на "o"; "1" на "i" и т. д. Так, доменное имя "z00.com" можно преобразовать в "zoo.com". Этот метод может повысить эффективность сравнения. Однако при определенных обстоятельствах преобразование по визуальному сходству может снизить его эффективность. Поэтому целесообразно выполнять сравнение сходства до и после преобразования.

7.1.2 Обнаружение официальных логотипов

Изображения или другие визуальные элементы на неизвестном веб-сайте рекомендуется сравнивать с официальными логотипами. Официальные логотипы включают, помимо прочего, бренд компании или организации, рекламный дизайн и другие знаковые элементы. Метод обнаружения официальных логотипов включает следующие шаги.

- Перед проведением обнаружения официальные логотипы следует собрать и хранить в базе данных логотипов.

- Визуальные элементы или снимки экрана неизвестной веб-страницы загружаются и сравниваются с логотипами из базы данных логотипов. Сходство между двумя изображениями можно вычислить с помощью двух популярных дескрипторов – масштабно-инвариантного преобразования признаков (SIFT) и быстрообнаруживаемых надежных признаков (SURF).
- Если исходный код веб-страницы обфусцирован, загрузка визуальных элементов может быть затруднена. Тогда можно использовать какую-либо систему автоматического тестирования веб-приложений на основе браузера, чтобы сделать снимок экрана веб-страницы. Перед сравнением логотипов их можно вырезать со снимка экрана с помощью модели обнаружения логотипов на основе ИИ. Эту модель обнаружения логотипов на основе ИИ можно эффективно обучить, используя базы данных логотипов с открытым исходным кодом.

7.1.3 Обнаружение обфускации кода

Поддельные веб-сайты могут обфусцировать вредоносный или подделывающий код, чтобы затруднить анализ. Обфусцированный код сильно отличается от обычного. Существует множество методов обнаружения обфусцированного кода, включая сверточные и рекуррентные классификаторы на основе нейронных сетей и т. п.

- Классификаторы на основе сверточной нейронной сети могут автоматически извлекать n-граммные признаки нормального и обфусцированного исходного кода. N-грамма – это непрерывная последовательность из n слов, составленная из заданного исходного кода. Чтобы уловить все отличительные признаки для классификации, можно установить количество слов, равное пяти или более.
- Классификаторы на основе рекуррентных нейронных сетей в основном используются в сетях с долгой краткосрочной памятью (LSTM) или трансформерных сетях. Такие классификаторы обрабатывают исходный код как последовательность символов, автоматически идентифицируют шаблоны последовательности между нормальным и обфусцированным кодом, а затем определяют, обфусцирован ли этот код.

7.1.4 Обнаружение ввода учетных данных

Вводимые учетные данные обычно присутствуют в исходном коде веб-страниц в виде "форм ввода". Как правило, в исходном коде формы ввода имеется атрибут password (этот атрибут гарантирует, что когда пользователь вводит его, пароль не отображается). Информацию об этом атрибуте можно легко найти с помощью инструментов анализа регулярных выражений или инструментов анализа веб-страниц. Если исходный код веб-сайта обфусцирован, можно использовать метод оптического распознавания символов (OCR) для обнаружения вводимых учетных данных на снимках экрана веб-сайта. Метод OCR сначала выделит на снимке экрана веб-страницы текстовые области, а затем преобразует эти области в текст. Если преобразованные тексты содержат некоторые ключевые слова, такие как password, то на веб-странице обнаруживается форма ввода учетных данных.

7.1.5 Сторонняя служба безопасности

Операторам рекомендуется использовать услуги сторонних служб безопасности для получения атрибутов и статистики веб-сайта, включая трафик веб-сайта, сведения о его репутации, плагины, сведения о регистрации доменного имени, сведения о сертификации и другие сведения, относящиеся к безопасности. Эти сведения могут быть использованы в качестве дополнительной информации для идентификации поддельных веб-сайтов.

7.1.6 Анализ ссылок

Поддельный веб-сайт может копировать гиперссылки с официального веб-сайта для использования на своих поддельных веб-страницах визуальных элементов официального веб-сайта. Такие действия можно обнаружить с помощью анализа ссылок. Анализ ссылок предусматривает сбор ссылок на визуальные элементы на официальном веб-сайте и их сравнение со ссылками на анализируемом веб-сайте. Если на обоих веб-сайтах есть одинаковые гиперссылки или таких гиперссылок много, этот веб-сайт может быть поддельным.

7.1.7 Окончательная идентификация на основе оценки репутации

Чтобы объединить все результаты анализа или обнаружения с применением рекомендованных выше контрмер и принять обоснованное решение о том, является ли веб-сайт поддельным, операторам рекомендуется использовать модель принятия решений на основе искусственного интеллекта (ИИ). Модель принятия решений на основе ИИ после этапа обучения автоматически определяет вес каждой из контрмер.

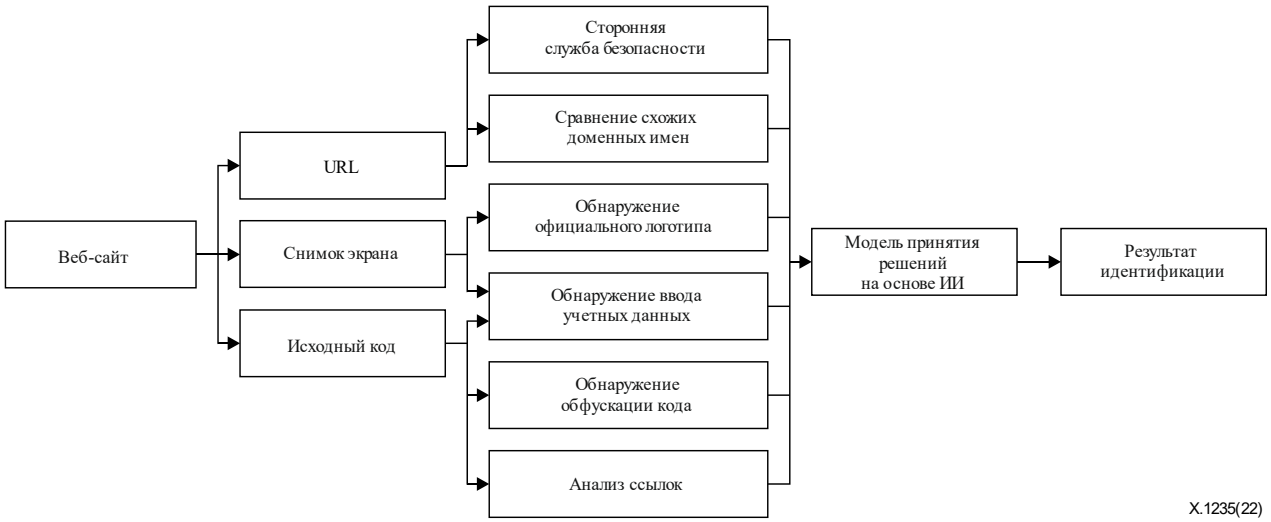


Рисунок 7-1 – Окончательная идентификация на основе оценки репутации

Как показано на рисунке 7-1, веб-сайт анализируется по его URL-адресу, снимкам экрана и исходному коду. По URL-адресу веб-страницы рассчитывается балльная оценка сходства доменных имен и проверяется результат сторонних служб безопасности. Снимки экрана используются для обнаружения и анализа официальных логотипов и ввода учетных данных. По исходному коду обнаруживается ввод учетных данных и выполняется анализ ссылок. Чтобы определить, является ли веб-сайт поддельным, модель принятия решений на основе ИИ должна учитывать все эти аспекты.

Для обучения этой модели принятия решений на основе ИИ следует собрать в качестве образцов все официальные веб-сайты, которые требуется защитить, известные поддельные веб-сайты и другие обычные веб-сайты. Все образцы (то есть собранные веб-сайты) оцениваются с точки зрения контрмер, рекомендованных выше, как показано на рисунке 7-1, и каждому образцу присваивается оценочный вектор. Все эти оценочные векторы и типы (поддельный – не поддельный) всех образцов можно использовать в качестве обучающих данных для обучения классификатора для формирования модели решений на основе ИИ. Классификаторы могут быть классификаторами на основе метода опорных векторов (SVM) или глубоких нейронных сетей.

Во избежание ошибок идентификации будет полезно, если поддельные веб-сайты, идентифицированные моделью принятия решений на основе ИИ, проверят эксперты.

7.2 Защита

Операторам рекомендуется защитить пользователей, приняв с их разрешения следующие контрмеры.

- Предупреждение пользователей. Если пользователь собирается посетить известный поддельный веб-сайт, оператору рекомендуется задержать запрос, перенаправить его на страницу с предупреждением, чтобы предупредить о риске подделки, и попросить пользователя подтвердить запрос.
- Создание черного списка. Оператору рекомендуется создать черный список поддельных веб-сайтов, чтобы блокировать все запросы к веб-сайтам из этого черного списка. Черный список может использоваться на шлюзах в сетях оператора, на DNS-серверах оператора или на других соответствующих сетевых объектах.
- Предотвращение распространения. Оператору рекомендуется блокировать сообщения о заторах, содержащие ссылку на поддельные веб-сайты.

- Информирование пользователей о способах самозащиты от подозрительных поддельных веб-сайтов. Периодически напоминайте пользователям о рисках и характеристиках поддельных веб-сайтов и предлагайте им практические рекомендации по предотвращению посещения поддельных веб-сайтов. В число таких рекомендаций входят:
  - не нажимать на странные ссылки в электронных письмах или сообщениях;
  - внимательно изучить информацию о доменном имени в адресной строке и сравнить ее с официальным сайтом;
  - использовать функцию безопасности веб-браузеров для проверки подлинности веб-сайтов.
- Защита пользователей путем выявления действующих доменных имен тех же поддельных веб-сайтов. Набор доменных имен поддельного веб-сайта можно использовать для прогнозирования других действующих доменных имен этого веб-сайта. Если содержание веб-сайтов под новыми действующими доменными именами сходно с содержанием известных поддельных веб-сайтов, их можно сразу рассматривать как те же поддельные веб-сайты и принять рекомендованные выше меры защиты.

8 Механизм

Для противодействия поддельным веб-сайтам рекомендуется использовать систематический механизм, сочетающий в себе все контрмеры, указанные в разделе 7.

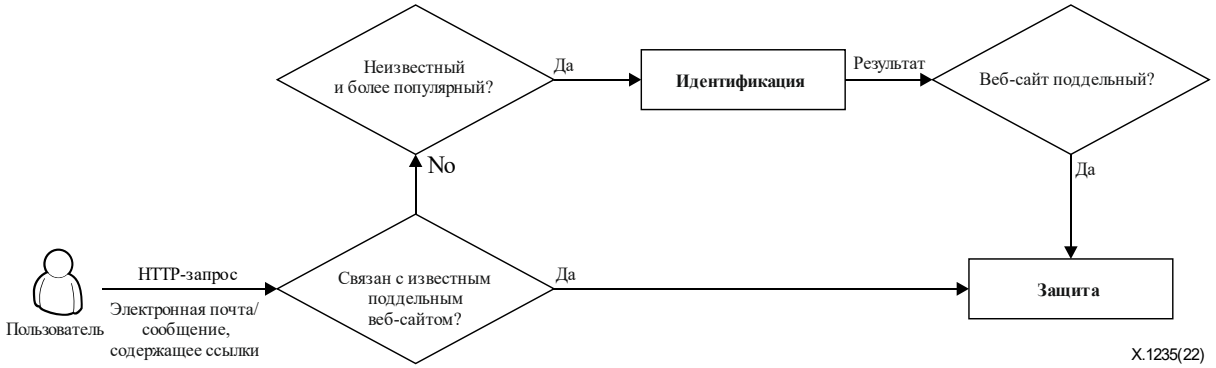


Рисунок 8-1 – Механизм противодействия поддельным веб-сайтам

1) Как показано на рисунке 8-1, URL-адреса могут находиться в HTTP-запросах пользователей, электронных письмах или мгновенных сообщениях пользователей (которыми они обмениваются с третьей стороной или которые поступают с их разрешения).
2) Если URL-адрес указывает на известный веб-сайт, следует принять защитные контрмеры для обработки запроса или сообщения.
3) Если URL-адрес указывает на популярный неизвестный веб-сайт, следует принять контрмеры по идентификации, чтобы определить тип веб-сайта.
4) Если результат контрмер по идентификации указывает на поддельный веб-сайт, следует принять защитные контрмеры, чтобы предотвратить посещение веб-сайта пользователем.

Дополнение I

Пример механизма противодействия спуфингу веб-сайтов

(Данное Дополнение не является неотъемлемой частью настоящей Рекомендации.)

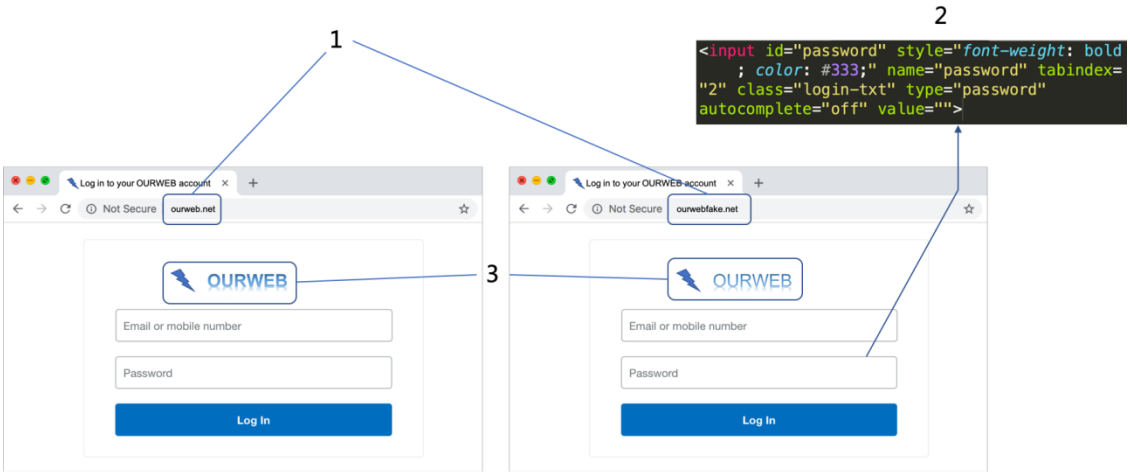


Рисунок I.1 – Пример официального веб-сайта и поддельного веб-сайта

Предположим, что существует известный веб-сайт OURWEB с доменным именем ourweb.net. Хакер пытается создать поддельный веб-сайт с доменным именем ourwebfake.net. Предположим, что доменное имя поддельного веб-сайта неизвестно оператору, тогда при большом количестве запросов пользователей к поддельному веб-сайту этот веб-сайт становится популярным неизвестным веб-сайтом. В этом случае веб-сайт следует обработать с применением контрмер по идентификации. Как показано на рисунке I.1, возможны три типа контрмер для выявления поддельного веб-сайта.

1) Доменные имена двух сайтов схожи. Предположим, что показатель сходства, соответствующий доле общих символов, равен 0,8.

2) На поддельном веб-сайте имеется форма ввода учетных данных, которую можно обнаружить в исходном коде. Показатель ввода учетных данных равен 1.

3) Логотип поддельного сайта аналогичен логотипу официального сайта. Предположим, что показатель сходства, который измеряет долю общих визуальных описательных идентификаторов, таких как описательный идентификатор SIFT, составляет 0,9.

4) Предположим, что результаты других контрмер равны нулю.

Тогда можно составить оценочный вектор [0,8; 0,9; 0; 1; 0; 0]. Первое значение в векторе – это показатель сходства доменных имен. Второе значение – это показатель сходства логотипа. Четвертое значение – показатель ввода учетных данных. Для получения результата этот вектор можно ввести в классификатор SVM. Если результат указывает на то, что веб-сайт поддельный, он будет обработан защитными контрмерами для предотвращения доступа пользователей к этому веб-сайту. Например, доменное имя сайта будет добавлено в черный список на шлюзах сетей операторов.

Дополнение II

Примеры технических мер

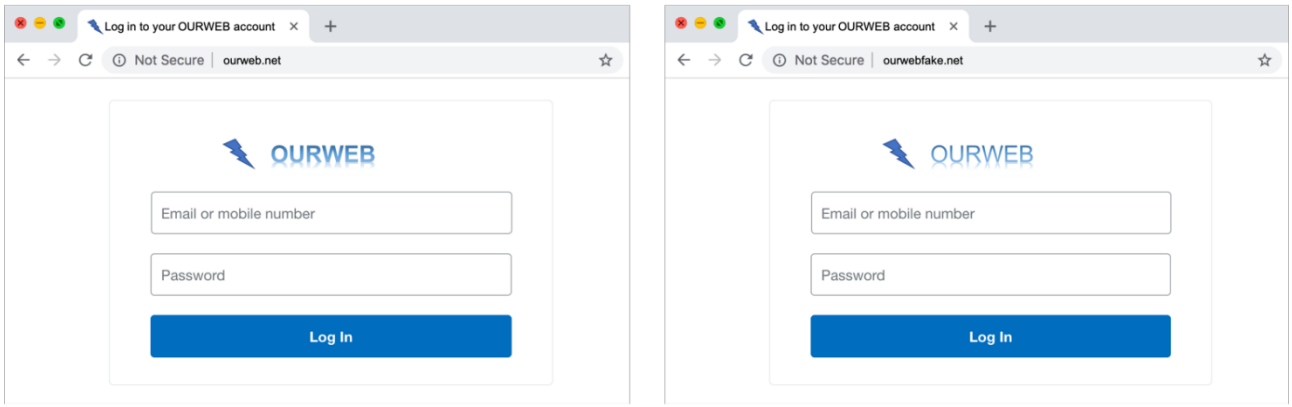
(Данное Дополнение не является неотъемлемой частью настоящей Рекомендации.)

Пример II.1 – Сравнение схожих доменных имен

Предположим, что официальное доменное имя – abc123.cn. Поддельный веб-сайт использует в качестве доменного имени abc123cn.org. Вычислив оценку сходства по Жаккару (размер пересечения наборов символов двух доменных имен делится на размер объединения) для двух доменных имен, получаем оценку сходства 8/11 = 0,73.

Пример II.2 – Обнаружение официального логотипа

На рисунке II.1 слева показан официальный веб-сайт OURWEB, а справа – соответствующий поддельный веб-сайт. Предположим, что официальный логотип OURWEB присутствует в базе логотипов. Логотип поддельного веб-сайта справа очень похож на официальный логотип OURWEB. Используя алгоритм сравнения сходства, можно получить оценку сходства (например, 0,9).



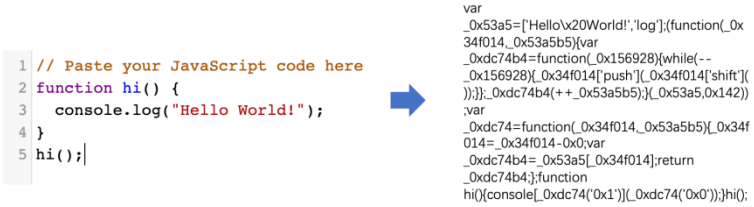
Официальный веб-сайт Поддельный веб-сайт

Рисунок II.1 – Поддельный веб-сайт с официальным логотипом

Пример II.3 – Обнаружение обфускации кода

Некоторые веб-сайты используют JavaScript для динамической генерации веб-контента и онлайн-сервис обфускации (например, https://obfuscator.io/) для шифрования кода JavaScript.

Как показано на рисунке II.2, обфусцированный код значительно отличается от обычного. Чтобы определить, является ли код обфусцированным, можно использовать модель классификации текста на основе ИИ. Алгоритм дает значение точности предсказания для исходного кода (например, 0,7).



Исходный код **Обфусцированный исходный код**

Рисунок II.2 – Примеры обфускации кода JavaScript

Пример II.4 – Обнаружение ввода учетных данных

Как показано на рисунке II.1, поддельный веб-сайт обманом побуждает пользователей ввести свои учетные данные. Сайт можно оценить, проверив, имеется ли на странице форма для ввода учетных данных. Если форма для ввода учетных данных присутствует, сайту дается оценка 1, в противном случае – 0.

Пример II.5 – Сторонняя служба безопасности

Онлайн-службы URL-запросов поставщиков средств безопасности (например, https://www.urlvoid.com/) помогают определить, является ли веб-сайт поддельным. Если обнаруженный веб-сайт классифицируется как поддельный, ему дается оценка 1, в противном случае – 0.

Пример II.6 – Анализ ссылок

На рисунке I.1 логотип может относиться к ресурсам официального сайта OURWEB. Анализ ссылок позволяет определить, что графический ресурс взят с веб-сайта ourweb.net (а не ourwebfake.net). Если на сайте есть ссылки на официальные графические ресурсы, ему может быть присвоена оценка 1, в противном случае – 0.

Пример II.7 – Объединение результатов

Результаты выявления поддельных веб-сайтов различными методами объединяются в вектор. Пример оценки веб-сайта приведен в таблице II.1. Второй столбец таблицы образует оценочный вектор. Этот вектор можно ввести в классификатор, чтобы определить, является ли веб-сайт поддельным.

Таблица II.1 – Пример оценки веб-сайта

Методы идентификации	Оценка
Сравнение схожих доменных имен	0,8
Обнаружение официального логотипа	0,9
Обнаружение обфускации кода	0,7
Обнаружение ввода учетных данных	1
Сторонняя служба безопасности	0
Анализ ссылок	1

Таблица II.2 – Сопоставление методов идентификации и характеристик поддельных веб-сайтов

Методы идентификации	Характеристики
Сравнение схожих доменных имен	Похожее доменное имя
Обнаружение официального логотипа	Визуально похож Заимствует визуальные элементы
Обнаружение обфускации кода	Обфусцированный исходный код
Обнаружение ввода учетных данных	Вводятся учетные данные
Сторонняя служба безопасности	
Анализ ссылок	Ссылки на известные веб-сайты

Библиография

[b-ITU-T X.1126] Рекомендация МСЭ-Т X.1126 (2017 г.), *Руководящие указания по смягчению негативных последствий от зараженных терминалов в сетях подвижной связи*.

[b-ITU-T X.1244] Рекомендация МСЭ-Т X.1244 (2008 г.), *Общие аспекты противодействия спаму в мультимедийных IP-приложениях*.

[b-SVM] Nature biotechnology, 2006, 24(12) 1565-1567*: What is a support vector machine?*

СЕРИИ РЕКОМЕНДАЦИЙ МСЭ-Т

Серия A	Организация работы МСЭ-Т
Серия D	Принципы тарификации и учета и экономические и стратегические вопросы международной электросвязи/ИКТ
Серия E	Общая эксплуатация сети, телефонная служба, функционирование служб и человеческие факторы
Серия F	Нетелефонные службы электросвязи
Серия G	Системы и среда передачи, цифровые системы и сети
Серия H	Аудиовизуальные и мультимедийные системы
Серия I	Цифровая сеть с интеграцией служб
Серия J	Кабельные сети и передача сигналов телевизионных и звуковых программ и других мультимедийных сигналов
Серия K	Защита от помех
Серия L	Окружающая среда и ИКТ, изменение климата, электронные отходы, энергоэффективность; конструкция, прокладка и защита кабелей и других элементов линейно-кабельных сооружений
Серия M	Управление электросвязью, включая СУЭ и техническое обслуживание сетей
Серия N	Техническое обслуживание: международные каналы передачи звуковых и телевизионных программ
Серия O	Требования к измерительной аппаратуре
Серия P	Качество телефонной передачи, телефонные установки, сети местных линий
Серия Q	Коммутация и сигнализация, а также соответствующие измерения и испытания
Серия R	Телеграфная передача
Серия S	Оконечное оборудование для телеграфных служб
Серия T	Оконечное оборудование для телематических служб
Серия U	Телеграфная коммутация
Серия V	Передача данных по телефонной сети
Серия X	**Сети передачи данных, взаимосвязь открытых систем и безопасность**
Серия Y	Глобальная информационная инфраструктура, аспекты протокола Интернет, сети последующих поколений, интернет вещей и умные города
Серия Z	Языки и общие аспекты программного обеспечения для систем электросвязи

Отпечатано в Швейцарии
Женева, 2022 год