УДК 004.93 https://doi.org/10.26160/2541-8637-2026-17-85-87

НЕЙРОСЕТЕВАЯ СИСТЕМА ОБНАРУЖЕНИЯ ФИШИНГОВЫХ ВЕБ-РЕСУРСОВ

Вульфин А.М., Кириллова А.Д., Минко А.В.
Уфимский университет науки и технологий, Уфа

Ключевые слова: мультимодальный анализ, фишинг, нейросетевая модель, анализ URL, защита информации, обнаружение угроз.
Аннотация. В работе рассматривается подход к обнаружению фишинговых веб-ресурсов на основе мультимодальной нейросетевой модели, объединяющей анализ URL-адресов и визуального представления веб-страниц. Разработаны модели на основе CNN и EfficientNet-B7, а также проведен анализ эффективности предложенного решения. Результаты демонстрируют высокую точность классификации (F1-score 97,86% для URL и 93,22% для изображений) и подтверждают перспективность мультимодального подхода для противодействия эволюционирующим фишинговым атакам.

NEURAL NETWORK SYSTEM FOR DETECTING PHISHING WEBSITES

Vulfin A.M., Kirillova A.D., Minko A.V.
Ufa University of Science and Technology, Ufa

Keywords: multimodal analysis, phishing, neural network model, URL analysis, information security, threat detection.
Abstract. The paper discusses an approach to detecting phishing web resources based on a multimodal neural network model that combines the analysis of URLs and the visual representation of websites. Models based on CNN and EfficientNet-B7 were developed, and an analysis of the effectiveness of the proposed solution was conducted. The results demonstrate high classification accuracy (F1-score of 97.86% for URLs and 93.22% for images) and confirm the promise of the multimodal approach to counter evolving phishing attacks.

Непрерывный рост числа фишинговых атак [1, 2] требует совершенствования методов защиты. Злоумышленники создают ресурсы, визуально и функционально имитирующие легитимные сайты, используя поддельные URL и адаптивный контент. Традиционные методы защиты не способны эффективно выявлять такие ресурсы, так как не учитывают динамические изменения в структуре страниц и их оформлении.

Цель работы – повышение эффективности обнаружения фишинговых веб-ресурсов с применением мультимодальной нейросетевой модели классификации, объединяющей анализ URL-адресов и визуального представления веб-страниц.

Создание системы обнаружения фишинговых веб-ресурсов на основе нейросетевых методов требует разработки модульной архитектуры, обеспечивающей комплексный анализ данных. Система интегрируется с веб-браузером, обрабатывает текстовые и визуальные данные и предоставляет пользователю результат классификации.

На рисунке 1 отображена диаграмма последовательностей, иллюстрирующая взаимодействие ключевых элементов системы.

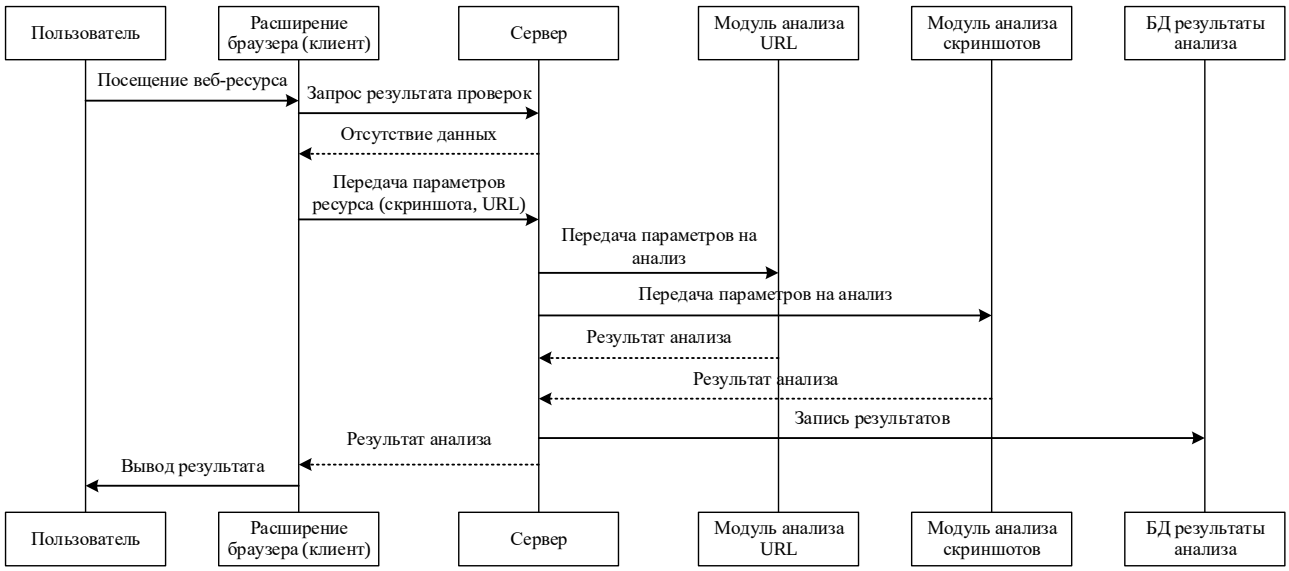

Рис. 1. Диаграмма последовательности

Модель анализа символьного имени (URL) предназначена для автоматизированного выявления фишинговых сайтов на основе анализа структуры и содержания URL. Основные задачи модели: обнаружение аномалий в структуре URL и его бинарная классификация на «фишинговые» и «безопасные» с применением методов глубокого обучения. Для обучения использован сбалансированный датасет MTLP [3]. Модель реализована на основе сверточной нейронной сети (CNN) с ядрами размеров 3, 4 и 5 для эффективного выделения локальных паттернов. Параметры модели: Embedding размерностью 64, три параллельных сверточных слоя с 64 фильтрами, функция потерь Binary Crossentropy, оптимизатор Adam. Разработанная модель демонстрирует высокую эффективность, достигая уровня F1-score = 97,86% на тестовой выборке.

Модель анализа визуального образа главной страницы веб-ресурса предназначена для автоматизированного выявления фишинговых сайтов на основе их графического содержимого. Основные задачи модели: обнаружение визуальных аномалий и бинарная классификация веб-ресурса на «фишинговые» и «безопасные». Для реализации модели выбрана предобученная архитектура EfficientNet-B7 [4, 5], адаптированная для бинарной классификации заменой последнего слоя на линейный классификатор с сигмоидной активацией. Обучение проводилось на скриншотах из MTLP-Dataset с применением аугментаций (отражение, цветовая вариация, поворот). Использовалась функция потерь BCEWithLogitsLoss, оптимизатор AdamW и планировщик OneCycleLR. Модель достигла уровня F1-score = 93,22%, подтверждая эффективность применения глубоких нейросетевых архитектур.

Разработан прототип программного обеспечения системы обнаружения фишинговых веб-ресурсов, реализующий модульную архитектуру для комплексного анализа URL и визуального контента страницы. Диаграмма компонентов включает браузерное расширение, сервер, базу данных, CNN-модель, EfficientNet-B7, Selenium и файловое хранилище.

В ходе тестирования выявлены ключевые ограничения:

1. Временные задержки. Среднее время анализа составляет 20 секунд в конфиденциальном режиме и 10 секунд в неконфиденциальном. Причина –

неоптимизированная интеграция моделей, требующая повторной загрузки весов при каждом запросе.

2. Ложноположительные срабатывания. Модель анализа URL ошибочно классифицирует легитимные ресурсы (vk.com, rutube.ru) как фишинговые из-за совпадения их структуры с паттернами из обучающего датасета.

3. Нестандартизированные входные данные. В неконфиденциальном режиме скриншоты поступают в произвольном разрешении, что не соответствует параметрам обучения модели (480x270 пикселей) и снижает качество анализа.

Несмотря на ограничения, прототип демонстрирует работоспособность мультимодального подхода, объединяющего анализ URL и визуального контента. CNN и EfficientNet-B7 показали высокую точность (97,86% и 93,22% F1-score) в анализе URL и визуальных данных, подтверждая эффективность интеграции разнородных данных. Модульная структура системы позволяет независимо обновлять компоненты.

Ключевым преимуществом является способность системы к объективной оценке даже при ложноположительных срабатываниях модели анализа URL. Например, при анализе сайта vk.com CNN-модель ошибочно классифицирует его как фишинговый, однако модель EfficientNet-B7, анализируя визуальный контент, выдает корректное решение. Это демонстрирует устойчивость мультимодального подхода к систематическим ошибкам отдельных моделей м снижает риск ложных блокировок.

Список литературы

1. Актуальные киберугрозы: IV квартал 2024 года – I квартал 2025 года [Электронный ресурс]. – Режим доступа: https://www.ptsecurity.com/.
2. Спам и фишинг в 2024 году [Электронный ресурс]. – Режим доступа: https://securelist.ru/spam-and-phishing-report-2024/111743/.
3. Çolhak F., Ecevit M.I., Uçar B.E., Creutzburg R., Dağ H. Phishing website detection through multi-model analysis of html content // International Conference on Theoretical and Applied Computing. – Singapore: Springer Nature Singapore. 2024, pp. 171-184.
4. Tan M., Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks // International conference on machine learning. – PMLR, 2019, pp. 6105-6114.
5. Tekkali C.G., Natarajan K. Transfer learning of pre-trained CNNs on digital transaction fraud detection // International Journal of Knowledge-Based and Intelligent Engineering Systems. 2024, vol. 28, no. 3, pp. 571-580.

Сведения об авторах:

Вульфин Алексей Михайлович – д.т.н., профессор кафедры вычислительной техники и защиты информации (ВТиЗИ);

Кириллова Анастасия Дмитриевна – к.т.н., доцент кафедры ВТиЗИ;

Минко Александр Васильевич – аспирант кафедры ВТиЗИ.