2.3.6.

Е.И. Белова, Р.М. Хамитов

АО «Сетевая компания»,
ФГБОУ ВО "Казанский государственный энергетический университет",
институт цифровых технологий и экономики,
кафедра информационные технологии и интеллектуальные системы,
Казань, belovaei.gc@gmail.com, hamitov@gmail.com,

СОВЕРШЕНСТВОВАНИЕ МЕТОДОВ ОБНАРУЖЕНИЯ ФИШИНГОВЫХ АТАК НА БАЗЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

В работе рассматривается исследование совершенствования методов обнаружения фишинговых атак на базе алгоритмов машинного обучения, подробно исследуется модель искусственного интеллекта, предназначенная для классификации веб-сайтов. В рамках методологии и серии экспериментов особое внимание уделяется использованию нейронных сетей в качестве ключевого компонента этой модели.

Ключевые слова: *информационная безопасность, фишинг, модель искусственного интеллекта, классификация веб-сайтов, нейронные сети, методология исследования, эксперименты.*

В современном мире веб-приложения активно собирают и обрабатывают огромные объемы личных данных пользователей. Среди этой информации могут быть упомянуты такие данные, как имена, адреса электронной почты, пароли, а также информация о финансах, включая данные с кредитных карт. С ростом популярности электронных сервисов, таких как онлайн-покупки, социальные сети или облачные хранилища, объем собираемой конфиденциальной информации постоянно увеличивается. Злоумышленники проявляют интерес к этим данным и стремятся получить доступ к ним.

В настоящее время более 60% жителей планеты используют всемирную сеть, а за последнее десятилетие число пользователей удвоилось, в начале 2022 года численность интернет-аудитории достигла почти 5 млрд пользователей. E-mail является самой большой ахиллесовой пятой кибербезопасности любой организации и точкой входа для 91% кибератак [1].

Почтовый фишинг (Phishing) – разновидность атаки, когда злоумышленники отправляют электронные письма, выглядящие как легитимные сообщения от банков, компаний или сервисов, с целью получить конфиденциальную информацию, такую как пароли и данные банковских карт.

В общей стратегии безопасности для выявления и предотвращения атак используются традиционные методы обнаружения фишинга, которые включают в себя различные техники и инструменты, например, фильтрация спама и почтового трафика, поддержание актуальных списков известных фишинговых доменов и адресов, которые используются для блокировки доступа к вредоносным ресурсам, обучение пользователей, использование антивирусных и антифишинговых программ и др.

С развитием технологий и появлением новых видов угроз современные системы кибербезопасности дополняют традиционные методы более сложными и инновационными подходами. Использование искусственного интеллекта и машинного обучения позволяет более точно выявлять угрозы и адаптироваться к постоянно меняющимся характеристикам кибератак, особенно это обуславливается возросшим числом фишинговых атак в 2020 году во всем мире [2].

Глубокое обучение (deep learning) — это тип машинного обучения, который использует искусственные нейронные сети с несколькими скрытыми слоями для автоматического извлечения признаков из большого объема данных [3].

Прежде всего, акцентируется возможность применения глубокого обучения для выявления фишинговых писем. Обученная нейронная сеть способна распознавать специфические характеристики таких сообщений, включая неожиданных отправителей, наличие ссылок на небезопасные сайты и запросы на предоставление личной информации. Этот обученный процесс позволяет системе обнаруживать и блокировать подобные мошеннические письма.

В дополнение подчеркивается применение глубокого обучения для мониторинга активности пользователей на фишинговых сайтах. Например, нейронная сеть обучается распознавать характеристики фишинговых веб-ресурсов, такие как схожесть дизайна с оригинальными сайтами и присутствие определенных элементов, таких как формы для ввода паролей и личных данных. Если такой сайт выявлен, система принимает меры по его блокировке и предупреждению пользователей о возможной угрозе.

Таким образом, сочетание методов машинного и глубокого обучения представляет собой эффективный инструмент в борьбе с фишингом, автоматически обнаруживая и блокируя мошеннические письма и веб-сайты.

В данной статье подробно исследуется модель искусственного интеллекта, предназначенная для классификации веб-сайтов. В рамках методологии и серии экспериментов особое внимание уделяется использованию нейронных сетей в качестве ключевого компонента этой модели. Опишем этапы проведённого исследования (рис.1).

1. Подготовка данных. Были использованы данные, представляющие собой набор признаков и соответствующих меток классов. Из набора признаков были выбраны два ключевых параметра, а именно **qty_dot_url** (количество точек в URL) и **qty_ip_resolved** (количество разрешенных IP-адресов), которые использовались в качестве входных данных для модели классификации. Эти признаки предоставляют модели информацию о структуре URL-адреса и характеристиках, связанных с IP-адресами, которые могут быть использованы для определения характеристик веб-сайтов и классификации их как фишинговые или нефишинговые.

2. Нормализация данных. Для обеспечения стабильной сходимости нейронной сети признаки были нормализованы с использованием **StandardScaler** из библиотеки **scikit-learn.**

3. Создание и обучение модели нейронной сети. Была построена нейронная сеть, состоящая из одного скрытого слоя с 128 нейронами и выходного слоя с функцией активации sigmoid. Модель была скомпилирована с оптимизатором Adam, бинарной функцией потерь и метрикой точности.

4. Для анализа HTML-кода сайтов и извлечения признаков были использованы библиотеки requests и beautifulsoup4. Была разработана функция extract_features_from_url, которая анализирует HTML-код сайта и извлекает признаки, такие как количество ссылок и изображений.

5. Сохранение и загрузка scaler. После обучения был сохранен объект StandardScaler, который затем может быть загружен для нормализации признаков новых данных. Классификация нового сайта. Используя обученную модель и сохраненный StandardScaler, был реализован процесс классификации нового сайта на основе извлеченных признаков.

В данной статье представлена модель искусственного интеллекта для классификации веб-сайтов, используемая нейронные сети. Эксперименты, проведенные на выборке данных, демонстрируют эффективность предложенной модели в задаче различения между фишинговыми и нефишинговыми веб-сайтами. В результате исследования подчеркнута важность использования нейронных сетей в задачах классификации веб-сайтов. Однако, существует несколько направлений для дальнейших исследований, таких как: расширение признакового пространства; улучшение функции извлечения признаков; работа с несбалансированными данными.

```
  #1 Определение признаков (X) и меток классов (y)
  X = data[['qty_dot_url', 'qty_ip_resolved']]
  y = data['phishing']
  #2 Нормализация признаков
  scaler = StandardScaler()
  X_train = scaler.fit_transform(X_train)
  X_test = scaler.transform(X_test)
  #3 Создание модели нейронной сети
  model = tf.keras.models.Sequential([
     tf.keras.layers.Input(shape=(X_train.shape[1],)),  # Входной
слой
     tf.keras.layers.Dense(128, activation='relu'),# Скрытый
слой
     tf.keras.layers.Dense(1, activation='sigmoid')# Выходной слой
  # Компиляция модели
  model.compile(optimizer='adam', loss='binary_crossentropy',
metrics=['accuracy'])
  # Обучение модели
  model.fit(X_train, y_train, epochs=10, batch_size=64,
validation_data=(X_test, y_test))
  # После обучения
  scaler = StandardScaler()
  X_train = scaler.fit_transform(X_train)
  #5 Сохраните scaler
  import joblib
  joblib.dump(scaler, 'scaler.pkl')
  # Загрузите scaler
  scaler = joblib.load('scaler.pkl')
  # Получение признаков для сайта
  url_to_classify = 'https://www.elibrary.ru'
  site_features = extract_features_from_url(url_to_classify)
  if site_features is not None:
  # Нормализация признаков сайта
     site_features = scaler.transform([site_features])
  # Классификация сайта
     prediction = (model.predict(site_features) >
0.5).astype("int32")
     if prediction == 1:
        print(f"Сайт {url_to_classify} - фишинг")
     else:
        print(f"Сайт {url_to_classify} - не фишинг")
  else:
     print("Не удалось извлечь признаки для классификации.")
```

Рис.1 – Процесс классификации сайта

В целом, результаты данного исследования предоставляют основу для разработки более сложных и эффективных систем обнаружения фишинговых веб-сайтов с использованием методов машинного обучения.

Список литературы

1. *Корнюхина С.П., Лапонина О.Р.,* Исследование возможностей алгоритмов глубокого обучения для защиты от фишинговых атак // International Journal of Open Information Technologies ISSN: 2307-8162 vol. 11, no.6, 2023 [Электронный ресурс]. Режим доступа: https://elibrary.ru/download/elibrary_54101991_33467843.pdf
(Дата обращения 18.11.2023).
2. *Афанасьева Н.С., Елизаров Д.А., Мызникова Т.А.,* Классификация фишинговых атак и меры противодействия им // Инженерный вестник Дона, №5, 2022. URL: ivdon.ru/ru/magazine/archive/n5y2022/7641
3. *Натальсон, А.В.* Формирование цифровых компетенций в области кибербезопасности объектов цифровой энергетики / А. В. Натальсон // Вестник НЦБЖД. – 2023. – № 3(57). – С. 54-60. (Дата обращения 27.11.2023).