УДК 004.85:004.056.57

ВЛИЯНИЕ АРХИТЕКТУРЫ НЕЙРОННЫХ СЕТЕЙ НА ЭФФЕКТИВНОСТЬ ОБНАРУЖЕНИЯ ПРИЗНАКОВ ШИФРОВАНИЯ

В.Л. Мальцев, В.В. Возмитель, Е.В. Матяс

Учреждение образования «Национальный детский технопарк», г. Минск, Республика Беларусь

Аннотация. В статье исследуется применение моделей глубокого обучения для идентификации зашифрованного содержимого. Рассмотрены полносвязные, сверточные сети, автоэнкодеры и трансформеры, проведен сравнительный анализ их эффективности при классификации байтовых последовательностей с использованием открытых наборов данных. Обоснована необходимость включения сжатых файлов в обучающую выборку для снижения ложных срабатываний в системах защиты информации. Полученные результаты демонстрируют преимущества редукции архитектур для повышения обобщающей способности моделей.

Ключевые слова: информационная безопасность; нейронные сети; шифрование; программы-вымогатели; машинное обучение; глубокое обучение; анализ энтропии; мониторинг файлов.

INFLUENCE OF NEURAL NETWORK ARCHITECTURE ON THE EFFICIENCY OF ENCRYPTION FEATURE DETECTION

V. Maltsau, V. Vozmitel, E. Matyas
Educational Institution "National Children's Technopark", Minsk, Republic of Belarus

Abstract. This article explores the application of deep learning models for identifying encrypted content, examining fully connected networks, convolutional networks, autoencoders, and transformers. A comparative analysis of the efficiency of byte sequence classification is conducted using open datasets. The necessity of including compressed files in the training set is substantiated to reduce the false positive rate in information security systems. The obtained results demonstrate the advantages of architecture reduction for improving model generalizability.
Keywords: information security; neural networks; encryption; ransomware; machine learning; deep learning; entropy analysis; file monitoring.

Введение

В условиях роста числа кибератак вопросы защиты пользовательских данных от программ-шифровальщиков приобретают критическое значение. Традиционные антивирусные решения, основанные на сигнатурном анализе или простых эвристиках (мониторинг энтропии), часто демонстрируют недостаточную точность при появлении новых угроз. Ключевым условием минимизации ущерба является обнаружение вредоносной активности на ранних стадиях, до завершения шифрования основного объема данных.

Основная часть

С целью выявления наиболее эффективных архитектур нейронных сетей для определения факта шифрования данных в процессе файловых операций был использован датасет NapierOne, содержащий более 300 тысяч образцов различных типов файлов. Ключевая проблема классификации заключается в схожести статистических характеристик зашифрованных и сжатых (архивных) данных. Для повышения надежности детектирования в обучающую выборку были включены файлы форматов .zip, .7z, .rar, а также другие документы с высокой плотностью упаковки данных.

В ходе работы были протестированы и проанализированы архитектуры:

1. *Полносвязные нейронные сети (FCNN)*: использовались как базовая модель. Входной вектор формировался на основе гистограммы частот байтов (256 значений). Несмотря на высокую скорость работы, FCNN ограничены в способности учитывать пространственные зависимости между байтами.

2. *Сверточные нейронные сети (CNN)*: входные данные представлялись
в виде одномерных векторов байтов. Применение фильтров позволило извлекать локальные признаки, характерные для заголовков различных файловых структур. Эксперименты показали, что CNN эффективнее справляются с отделением медиафайлов от зашифрованных блоков.

3. *Автоэнкодеры (AE)*: модель обучалась восстановлению «нормальных» (незашифрованных) последовательностей. Резкое увеличение функции потерь при реконструкции данных служило индикатором аномалии, указывающим на процесс шифрования.

4. *Трансформеры (Transformer)*: использование механизмов самовнимания позволило модели анализировать контекст всей последовательности данных. Это обеспечило наиболее точную идентификацию алгоритмов AES-256 в режиме реального времени.

Анализ интегрируется в систему мониторинга на базе Windows Minifilter. Драйвер перехватывает IRP-пакеты, извлекает буфер данных из операций записи и передает нейросетевому модулю. Тестирование показало, что комбинированный подход – анализ энтропии Шеннона совместно с предсказанием нейросети – снижает уровень ложных срабатываний до 0,5% при работе с офисными документами и архивами.

Для оценки влияния архитектурной емкости на обобщающую способность проведена серия экспериментов по редукции моделей. Исходная сверточная сеть (BaseFC) демонстрировала переобучение: разрыв точности между обучающей и тестовой выборками достигал 5%. Последовательное удаление слоев с усилением сжатия в узком месте привело к версии MiniFC, в которой дельта точности сократилась до 0,7%, а метрики F1-score и Precision улучшились на 4–5 % по сравнению с исходной конфигурацией. Дальнейшее сокращение параметров (NanoFC) вызвало недообучение.

Аналогичная закономерность выявлена для трансформеров. Исходная модель (BaseTR) после непродолжительного обучения (до 11 эпох) демонстрировала рост точности до 82%, затем впадала в катастрофическое переобучение. Версия MiniTR, полученная усечением выходных слоев и усилением сжатия, устранила этот дефект и превзошла исходную модель: точность достигла 88,1%, F1-score 0,887, коэффициент корреляции Мэтьюса – 0,771. Дальнейшая редукция (NanoTR) привела к росту полноты (Recall) до 99,3% за счет увеличения ложноположительных срабатываний, что допустимо в сценариях, где критически важно не пропустить ни одного зашифрованного файла. Результаты подтверждают, что избыточная параметрическая емкость вредна, а целенаправленная редукция архитектуры повышает обобщающую способность без дополнительных регуляризационных механизмов.

Заключение

Результаты подтверждают эффективность нейросетевых моделей для обнаружения ransomware-активности. Наилучший баланс точности и вычислительной эффективности достигается сверточными сетями и трансформерами при условии редукции архитектуры: удаление избыточных параметров устраняет переобучение и повышает обобщающую способность (версии MiniFC, MiniTR). Включение сжатых файлов в обучающую выборку снижает уровень ложных срабатываний. Реализация на базе Windows Minifilter позволяет блокировать подозрительные операции записи до необратимого изменения данных, что делает систему перспективным компонентом EDR-решений.

Список использованных источников

1. Гаттс, Дж. Глубокое обучение на Python / Дж. Гаттс. – СПб.: Питер, 2018. – 400 с.

2. Таненбаум, Э. Современные операционные системы / Э. Таненбаум, Х. Бос. – 4-е изд. – СПб.: Питер, 2015. – 1120 с.

3. Шай, Ш. Основы машинного обучения: от теории к алгоритмам / Ш. Шай, Ш. Бен-Давид; пер. с англ. – М.: ДМК Пресс, 2017. – 436 с.

References

1. Chollet, F. (2018) Deep Learning with Python. Saint Petersburg: Piter. (In Russian)

2. Tanenbaum, A., Bos, H. (2015) Modern Operating Systems. 4th ed. Saint Petersburg: Piter. (In Russian)

3. Shalev-Shwartz, S., Ben-David, S. (2017) Understanding Machine Learning: From Theory to Algorithms. Moscow: DMK Press. (In Russian)

Сведения об авторах

Мальцев В.Л. заведующий лабораторией «Информационная безопасность» учреждения образования «Национальный детский технопарк», viktor.maltsevlul@gmail.com.

Возмитель В.В. учащийся учреждения образования «Национальный детский технопарк», racfor4@gmail.com.

Матяс Е.В. учащийся учреждения образования «Национальный детский технопарк».

Information about the authors

Maltsau V. Head of the Information Security Laboratory, Educational Institution "National Children's Technopark", viktor.maltsevlul@gmail.com.

Vozmitel V. student, Educational Institution "National Children's Technopark", Minsk, Republic of Belarus, racfor4@gmail.com.

Matyas E. student, Educational Institution "National Children's Technopark"