ОБНАРУЖЕНИЕ ПРОГРАММ-ВЫМОГАТЕЛЕЙ НА ОСНОВЕ КОМБИНИРОВАННОЙ НЕЙРОННОЙ СЕТИ И АНАЛИЗА ЭНТРОПИИ ДАННЫХ

Мальцев В. Л., Возмитель В. В., Матяс Е. В.
Кафедра защиты информации,
Белорусский государственный университет информатики и радиоэлектороники
Национальный детский технопарк
Минск, Республика Беларусь
E-mail: {viktor.maltsevlul, racfor4}@gmail.com

В условиях растущего числа изощренных кибератак программы-вымогатели представляют собой одну из наиболее серьезных угроз информационной безопасности. Традиционные методы защиты и методы, основанные на сигнатурах и анализе энтропии данных, нередко оказываются неэффективными из-за высокого уровня ложных срабатываний и неспособности надежно отличить зашифрованные данные от легитимных файлов со сжатием. В качестве решения данной проблемы предлагается двухуровневый подход анализа данных, совмещающий в себе разные методы. На первом этапе выполняется быстрая фильтрация файлов по порогу энтропии для отбора «подозрительных» объектов. На втором этапе эти файлы анализируются гибридной нейронной сетью, объединяющей сверточную сеть для выявления локальных паттернов и архитектуру Transformer для анализа глобальных контекстных зависимостей в данных.

Введение

В современном цифровом мире, где информация является одним из самых ценных активов, киберугрозы и кибератаки постоянно эволюционируют, становясь все более изощренными и опасными. Одной из наиболее серьезных и распространенных угроз являются кибератаки с использованием программ-вымогателей. В настоящее время количество таких атак неумолимо растет, нанося колоссальный ущерб как крупным компаниям, так и обычным пользователям. С развитием искусственного интеллекта стали возможными и кибератаки с использованием искусственных нейронных сетей. Одним из перспективных подходов к выявлению несанкционированного шифрования является анализ энтропии данных. Этот метод основывается на алгоритме энтропии Шеннона, который является мерой неопределенности или случайности данных [1][2].

Основной раздел

Идея заключается в том, что зашифрованные данные обладают очень высокой степенью случайности, и их показатель энтропии приближается к максимальному значению, равному примерно 8. Однако этот метод имеет существенный недостаток. Высокую энтропию также имеют и многие популярные и абсолютно легитимные типы файлов. К ним относятся, например, архивы (ZIP, RAR), многие форматы документов (например, .docx), а также видео и графические файлы, которые используют сжатие. Этот факт делает невозможным точное выявление работы программ-вымогателей на основе одного лишь порога энтропии, так как это привело бы к огромному количеству ложных срабатываний. Чтобы наглядно продемонстрировать эту проблему энтропии, был проведен анализ множества файлов в двух состояниях: до и после шифрования. Результаты, представленные на графике распределения энтропии для зашифрованных и незашифрованных файлов, показывают степень их энтропии.

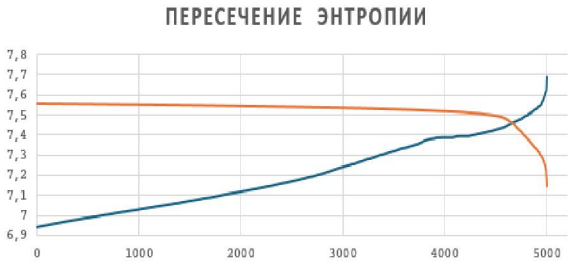

Рис. 1 – График распределения энтропии для зашифрованных и незашифрованных файлов

На графике (рисунок 1) представлено пересечение двух кривых, где синяя линия соответствует обычным, незашифрованным файлам, а оранжевая – тем же файлам, но уже в зашифрованном виде. Эта иллюстрация наглядно показывает две ключевые проблемы:

1. Значительное количество обычных, незашифрованных файлов имеет энтропию, равную или даже превышающую показатели многих зашифрованных файлов.

2. Некоторые зашифрованные файлы, в свою очередь, могут иметь низкий или относительно низкий показатель энтропии. Это может быть связано с использованием алгоритмов шифрования с низким уровнем энтропии на начальных этапах процесса или с частичным шифрованием файла.

Если провести эксперимент по выявлению зашифрованных файлов, установив оптимальное

пороговое значение энтропии на уровне 7.45, можно получить следующие результаты:

- ложно положительные срабатывания: 5.42 процента;
- ложно отрицательные срабатывания: 0.34 процента;
- истинно положительные срабатывания: 94.58 процентов;
- истинно отрицательные срабатывания: 99.66 процентов.

Результаты этого теста показывают, что практически каждый 20-й файл из тестового набора был бы ошибочно опознан как зашифрованный. Такая высокая погрешность делает невозможным практическое применение данного метода для корректного выявления программ-вымогателей в реальных условиях.

Для решения проблемы выявления была разработана и обучена нейронная сеть передовой комбинированной архитектуры, состоящая из двух ключевых модулей: CNN (Convolutional Neural Network, сверточная нейронная сеть) и Transformer.

Модуль CNN: модуль специализируется на выявлении локальных зависимостей и паттернов в последовательности данных. В контексте анализа файла он способен выделить из потока байтов характерные микро-признаки, которые могут указывать на процесс шифрования. Поле свертки позволяет эффективно анализировать локальные зависимости между байтами и сократить длину самой последовательности в два раза для последующей обработки.

Модуль Transformer: архитектура, созданная для обработки естественного языка, превосходно справляется с выявлением глобальных зависимостей в данных. После предварительной обработки модулем CNN, Transformer анализирует последовательность в целом, оценивая взаимосвязи между удаленными друг от друга участками файла. Это позволяет нейросети формировать целостное «понимание» структуры файла, а не просто реагировать на отдельные подозрительные фрагменты [2].

Комбинация этих двух модулей позволяет создать высокопроизводительную и точную нейронную сеть, способную улавливать как локальные, так и глобальные признаки вредоносной активности. Для обучения нейронной сети был использован обширный набор данных из открытого источника – NapierOne, суммарное количество файлов в котором составило порядка 300 тысяч. Этот датасет включает в себя как вредоносные, так и безопасные файлы, что позволяет модели «научиться» их различать.

Тестирование нейросети проводилось на специально подобранной выборке зашифрованных и незашифрованных файлов, ключевой особенностью которой было то, что все файлы имели энтропию выше значения 7.4. Оценивать работу на данных с более низкой энтропией не имеет смысла, так как именно в «спорном» диапазоне традиционные методы дают сбой, а количество зашифрованных файлов с низкой энтропией стремится к нулю.

По итогам тестирования на сложных случаях нейросеть продемонстрировала следующие показатели точности:

- ложно положительные срабатывания: 3.75 процентов (ошибочное определение безопасного файла как вредоносного);
- ложно отрицательные срабатывания: 1.63 процента (пропуск реальной угрозы);
- истинно положительные срабатывания: 96.25 процентов (корректное обнаружение угрозы);
- истинно отрицательные срабатывания: 98.37 процентов (корректное определение безопасного файла).

Если же применить двухуровневый подход (сначала быстрый тест на энтропию, а затем анализ нейросетью только «подозрительных» файлов), то итоговые показатели эффективности в абсолютных числах становятся еще более убедительными:

- ложно положительные срабатывания: 0.2 процента;
- ложно отрицательные срабатывания: 0.006 процента;
- истинно положительные срабатывания: 99.8 процентов;
- истинно отрицательные срабатывания: 99.994 процента.

Данные показатели являются абсолютно приемлемыми для использования в защитных решениях на среднестатистических персональных компьютерах.

Заключение

Таким образом, предложенный подход, сочетающий в себе сильные стороны сверточных сетей и архитектуры Transformer, способен практически безошибочно определять факт вредоносного шифрования и, следовательно, эффективно выявлять и блокировать атаки программ-вымогателей.

Список литературы

1. ESET discovers PromptLock, the first AI-powered ransomware[Electronic resource] / ESET North America.: 2025. – Mode of access: `https://www.eset.com/us/about/newsroom/research/eset-discovers-promptlock-the-first-ai-powered-ransomware/`. – Date of access: 02.10.2025.
2. Abásolo, D., Hornero, R., Espino P. Entropy analysis of the EEG background activity in Alzheimer's disease patients // Physiological Measure-ment. 2006. Vol. 27(3). P.241-253. /University of Surrey. – Mode of access: `epubs.surrey.ac.uk/39603/6-006.pdf`. – Date of access: 02.10.2025.
3. Козлов, С. Transformer – новая архитектура нейросетей для работы с последовательностями. – Никосия, Кипр, 1998–2024. – Режим доступа: `https://habr.com/ru/articles/341240/`. – Дата доступа: 02.10.2025.