ИНТЕГРАЦИЯ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ В МЕЖСЕТЕВОЙ ЭКРАН CHUWALL

Янославский К.О., Мартинкевич Д.М.

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

Белоусова Е.С. – канд. техн. наук, доцент

Аннотация. В данной работе рассматривается подход обнаружения и блокировки кибератак на уровне межсетевого экрана CHUWALL с использованием алгоритмов машинного обучения. Обосновывается выбор библиотеки CatBoost и фреймворка NFStream для анализа сетевого трафика в реальном времени. Приводятся результаты сравнительного анализа моделей и итогового тестирования системы.

Ключевые слова. Межсетевой экран, анализ трафика, модель машинного обучения, библиотека CatBoost, фреймворк NFStream, методы машинного обучения случайный лес (Random Forest) и градиентный бустинг (Gradient Boosting).

Традиционные средства сетевой защиты ограничиваются использованием классических методов анализа трафика, однако на фоне стремительной эволюции киберугроз и появления новых векторов кибератак подобные методы защиты утрачивают свою эффективность, что требует внедрения инетеллектуальных систем анализа трафика. В данной работе рассматривается разработка и проверка эффективности применения модуля на базе модели машинного обучения, интегрированного в межсетевой экран CHUWALL, для активного противодействия киберугрозам.

Для обеспечения высокой производительности межсетевого экрана CHUWALL [1] была выбрана бинарная классификация сетевых потоков: легитимный трафик (Benign) и вредоносный трафик (Attack). Был выбран подход машинного обучения с учителем (Supervised Learning), который обусловлен возможностью использования заранее размеченных наборов данных, что позволяет обучить модель выявлять скрытые закономерности, характерные для различных типов кибератак.

Для выбора программного обеспечения, которое позволит модели эффективно обрабатывать большые потоки сетевых данных в реальном времени, было проведено сравнение различных библиотек и методов машинного обучения (таблица 1).

Таблица 1 – Сравнение библиотек для обучения модели

Критерий	Scikit-learn (Random Forest, Gradient Boosting)	Pytorch/TF (Deep Learning)	XGB/CatBoost (Gradient Boosting)
Скорость обучения	Средняя	Низкая	Высокая
Скорость работы	Очень высокая	Низкая	Средняя
Точность	Средняя	Высокая	Очень высокая
Сложность внедрения	Низкая	Высокая	Средняя
Основные задачи в ИБ	Быстрый поиск аномалий по готовым признакам, обнаружение известных кибератак	Анализ сырых пакетов и зашифрованного трафика, обнаружение zero-day атак	Максимальная точность в определении кибератак
Устойчивость к выбросам	Средняя	Низкая	Высокая

Сравнительный анализ показал, что классические методы случайного леса и градиентного бустинга обеспечивают высокую скорость работы, в то время как метод глубокого обучения хотя и эффективен для выявления zero-day кибератак и имеет высокую точность, однако для анализа большого потока трафика на межсетевом экране он не подходит из-за своей низкой скорости работы, а также требования значительных вычислительных ресурсов. На основании полученных данных исследование было сужено до сравнения двух наиболее эффективных для поставленных задач методов машинного обучения: Random Forest и Gradient Boosting (таблица 2).

Результаты сравнения показывают преимущество библиотеки CatBoost при реализации градиентного бустинга. Тем не менее, выбор между методами Random Forest и Gradient Boosting неоднозначен. В связи с этим было принято решение об обучении двух моделей на идентичном датасете с последующим их сравнением в равных условиях при выполнении реальных задач. Такой подход позволил оценить и сравнить эффективность выбранных методов не только по теоретическим критериям, но и по результатам тестирования в условиях имитации кибератак.

Таблица 2 – Сравнение методов машинного обучения

Критерий	Scikit-Learn [2] (Random Forest)	Scikit-Learn (Gradient Boosting)	CatBoost [3] (Gradient Boosting)
Скорость обучения	Средняя (только CPU)	Низкая (только CPU)	Высокая (поддержка GPU)
Скорость работы	Высокая	Высокая	Самая высокая
Потребление ресурсов	Среднее	Среднее	Низкое
Склонность к переобучению	Низкое	Среднее	Низкая (использование Order ed Boosting)

На основе сравнения, представленного в таблице 3, в качестве основы для обучения модели был выбран датасет NF-CSE-CIC-IDS2018-v2 [4], который содержит актуальные признаки сетевых потоков в формате Netflow (фреймворка, который в дальнейшем был использован для передачи информации о сессии модели). Первоначально обучения проводились на полном наборе данных, однако значительный дисбаланс классов (соотношение легитимного трафика к атакам составляло примерно 5:1) приводил к неверным предсказаниям обеих моделей. Поэтому был применен метод случайного понижающего сэмплирования, который позволил обеспечить оптимизацию процесса обучения и устранить значительный дисбаланс классов. Исходный объем легитимных сессий был сокращен до 3 млн записей, выбранных случайным образом. Это позволило достичь соотношения классов, близкого к 1:1, что положительно сказалось на результатах моделей.

Таблица 3 – Сравнение датасетов для обучения моделей

Название датасета	Количество легитимных сессий	Количество кибератак	Количество классов кибератак	Год создания
NF-CSE-CIC-IDS2018-v2	16 249 186	2 644 522	15	2018
NF-UNSW-NB15-v2	2 295 222	95 053	9	2015
NF-ToN-IoT-v2	609 391	769 883	9	2019
NF-BoT-IoT-v2	13 859	586 196	4	2018

Классификация сессий осуществляется на основе 27 статистических характеристик потока: длительность соединения, общее количество и длина переданных пакетов, временные интервалы между пакетами и др.

В ходе экспериментов сравнивались модели Random Forest (Scikit-Learn) и Gradient Boosting (CatBoost), обученные на идентичном сокращенном датасете NF-CSE-CIC-IDS2018-v2. Модели тестировались в одинаковых условиях и анализировали одинаковые сохраненные дампы трафика с различными кибератаками. Результаты тестирования и сравнения моделей представлены в таблицах 4–5.

Таблица 4 – Классификация модели, обученной методом Random Forest

	Precision	Recall	F1-score
Метод Random Forest			
Без кибератак	0,97	0,98	0,98
С кибератаками	0,98	0,97	0,97
Метод Gradient Boosting			
Без кибератак	0,98	1,00	0,99
С кибератаками	1,00	0,97	0,98

Таблица 5 – Сравнение уровней уверенности моделей в кибератаках

Кибератака	Scikit-Learn	CatBoost
DoS UDP	35,33%	94,68%
DoS TCP	26,0%	96,79%
SSH BruteForce	27,33%	89,37%
Port Scan	28%	93,85%

Проведенный сравнительный анализ показал значительное превосходство второй модели Gradient Boosting (CatBoost), из-за чего она и была выбрана для интеграции в межсетевой экран и анализа трафика в реальном времени. Программная реализация выполнена на языке Python с использованием фреймворка NFStream.

Алгоритм работы модуля состоит из следующих этапов:

1. Захват трафика на сетевом интерфейсе в реальном времени. Модуль настроен на анализ исключительно входящих сетевых соединений, что позволяет минимизировать влияние ложноположительных срабатываний на исходящий трафик внутренней сети и работу доверенных сервисов.

2. Извлечение признаков потока и передача их в модель CatBoost.

3. Расчет вероятности, что сессия является кибератакой. Если уверенность модели больше 50%, то инициируется блокировка.

4. IP-адрес атакующего заносится в список nftables и все последующего пакеты блокируются.

5. Автоматическая разблокировка по истечении установленного интервала времени.

В качестве сценария проверки работы разработанного модуля анализа трафика моделью машинного обучения и его эффективности в реальных условиях была реализованна DoS-атака на внешний интерфейс межсетевого экрана, при отключенных остальных модулях защиты. Кибератака включала в себя генерацию интенсивного потока TCP-пакетов с использованием сетевого инструмента hping3. В ходе тестирования, представленного на рисунке 1, модуль, работающий в реальном времени, успешно проанализировал сетевые потоки, классифицировал их как аномальные и идентифицировал атакующий узел. В соответствии с заложенной логикой, при превышении порога уверенности модели (Threshold 0,50), система мгновенно занесла IP-адрес нарушителя в правила фильтрации nftables, что подтвердило эффективность разработанного механизма защиты.

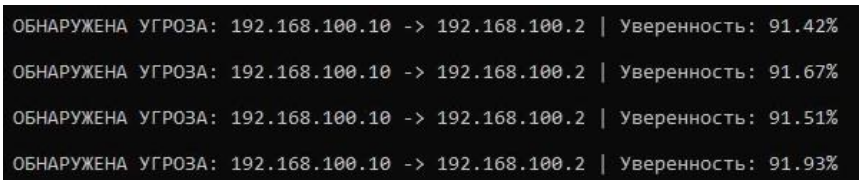

Рисунок 1 – Обнаружение кибератаки моделью

Таким образом, было проведено сравнение различных библиотек и методов для обучения моделей как теоретически, так и практически. Интеграция модели машинного обучения в процесс фильтрации трафика позволяет межсетевому экрану выявлять необычные угрозы, которые могут преодолеть все предыдущие этапы проверки. Применение CatBoost совместно с NFStream обеспечивает необходимую скорость работы, достаточную для анализа сетевых потоков в режиме реального времени.

Список использованных источников:

1. Маршрутизатор с функциями анализа сетевого трафика / Е. С. Белоусова, В. Л. Мальцев, Д. М. Мартинкевич, К. О. Янославский // Будущее через исследования : сборник материалов I Республиканской конференции молодых ученых, г. Минск, 29 января 2026 г. – Минск : БГТУ, 2026. – С. 52–57.

2. Scikit-learn : [сайт]. – 2026. – URL: https://scikit-learn.org (дата обращения: 28.03.2026).

3. CatBoost: unbiased boosting with categorical features / L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, A. Gulin // Journal of Machine Learning Research. – 2011. – P. 2825-2830.

4. Датасет NF-CSE-CIC-IDS-2018-v2. База данных The University of Queensland. [сайт] – URL: https://espace.library.uq.edu.au/view/UQ:e9636b7 (дата обращения: 02.03.2026).