УДК 004.056.55:004.85

АВТОМАТИЗИРОВАННОЕ ОБНАРУЖЕНИЕ СЕТЕВЫХ АТАК В IoT-СИСТЕМАХ НА ОСНОВЕ LIGHTGBM

Якунин А.А., Майлыханова Б.А.

Казахский национальный аграрный исследовательский университет, Алматы, Казахстан,
andryakunin@gmail.com

Аннотация: рассмотрена задача автоматизированного обнаружения сетевых атак в инфраструктуре Интернета вещей с использованием алгоритма LightGBM. Определены особенности IoT-трафика, влияющие на построение системы обнаружения вторжений: дисбаланс классов, разнородность устройств, ограниченность вычислительных ресурсов и изменение профиля атак во времени. Предложена последовательность обработки данных, включающая очистку сетевых потоков, кодирование категориальных признаков, стратифицированное разделение выборки, настройку LightGBM и оценивание по accuracy, precision, recall, F1-score и ROC-AUC. Показано, что листовой способ роста деревьев и методы Gradient-based One-Side Sampling и Exclusive Feature Bundling делают LightGBM перспективным для анализа многомерного IoT-трафика. Сформулированы требования к экспериментальной проверке, предотвращающие завышение качества модели из-за утечки данных и дисбаланса классов.

Ключевые слова: интернет вещей, обнаружение вторжений, сетевые атаки, LightGBM, градиентный бустинг, машинное обучение.

I. ВВЕДЕНИЕ

Интернет вещей объединяет датчики, исполнительные устройства, шлюзы и информационные сервисы, обеспечивая непрерывный обмен данными и автоматическое управление. Расширение числа подключённых устройств одновременно увеличивает поверхность атаки. Компрометация узла IoT может привести к утечке данных, нарушению доступности сервиса, подмене управляющих команд и использованию устройства в составе ботнета. Особую сложность создают ограниченные ресурсы конечных устройств, неоднородность протоколов и длительный жизненный цикл оборудования.

Сигнатурные системы обнаружения вторжений эффективно распознают известные угрозы, но зависят от своевременного обновления правил. Методы машинного обучения позволяют выявлять статистические различия между нормальным и вредоносным трафиком и применяются для бинарной и многоклассовой классификации атак [1–3]. При этом высокая доля нормальных пакетов и редкость отдельных атак делают показатель accuracy недостаточным: модель может демонстрировать высокую общую точность, пропуская критические классы.

Цель работы – разработать обоснованную схему автоматизированного обнаружения атак в IoT-трафике на основе LightGBM и определить корректный протокол её экспериментальной оценки. Вклад работы состоит в согласовании этапов подготовки данных, обучения, контроля утечки и выбора метрик с особенностями IoT-среды. Работа не заявляет новые экспериментальные значения точности: окончательные показатели должны рассчитываться на выбранном открытом наборе данных по описанному протоколу.

II. УГРОЗЫ И ДАННЫЕ ДЛЯ ОБНАРУЖЕНИЯ АТАК

Для сетей IoT характерны атаки отказа в обслуживании DoS/DDoS, сканирование портов, подбор учётных данных, внедрение вредоносного программного обеспечения, атаки «человек посередине» и активность ботнетов. Признаками могут служить длительность потока, количество и размер пакетов, направление соединения, число байтов, частота запросов, используемый протокол, флаги TCP и статистики временных интервалов. При проектировании модели признаки следует получать только из информации, доступной в момент принятия решения.

Для воспроизводимых исследований подходят открытые наборы Edge-IIoTset и ToN_IoT. Edge-IIoTset содержит нормальный трафик и несколько классов атак, сформированных в реалистичной IoT/IIoT-среде [4]. ToN_IoT объединяет сетевые потоки, телеметрию IoT/IIoT и журналы операционных систем [5]. Эти наборы различаются структурой и распределением классов, поэтому результаты нельзя переносить между ними без повторной проверки.

Подготовка данных включает удаление дубликатов и некорректных значений, кодирование протоколов и состояний соединения, а также разделение на train, validation и test. Кодировщики и средства

балансировки обучаются только на train-части. При этом одинаковые потоки и временно связанные записи не должны попадать одновременно в обучение и тестирование. Для деревьев решений масштабирование числовых признаков, как правило, не является обязательным.

III. МЕТОД LIGHTGBM И АРХИТЕКТУРА ОБНАРУЖЕНИЯ

LightGBM представляет собой реализацию градиентного бустинга над деревьями решений (рисунок 1). Каждое последующее дерево корректирует ошибки текущего ансамбля, а итоговая оценка формируется суммой вкладов деревьев. В отличие от симметричного роста по уровням, LightGBM выбирает лист с наибольшим потенциальным уменьшением функции потерь и расширяет его. Такой leaf-wise-подход может повысить точность при ограниченном числе листьев, но при малой выборке увеличивает риск переобучения [6-8].

Высокая вычислительная эффективность обеспечивается двумя механизмами. Gradient-based One-Side Sampling сохраняет наблюдения с большими градиентами и выбирает часть объектов с малыми градиентами. Exclusive Feature Bundling объединяет разреженные взаимно исключающие признаки. Это особенно полезно при кодировании категориальных параметров сетевого трафика, когда возникает большое число разреженных бинарных столбцов.

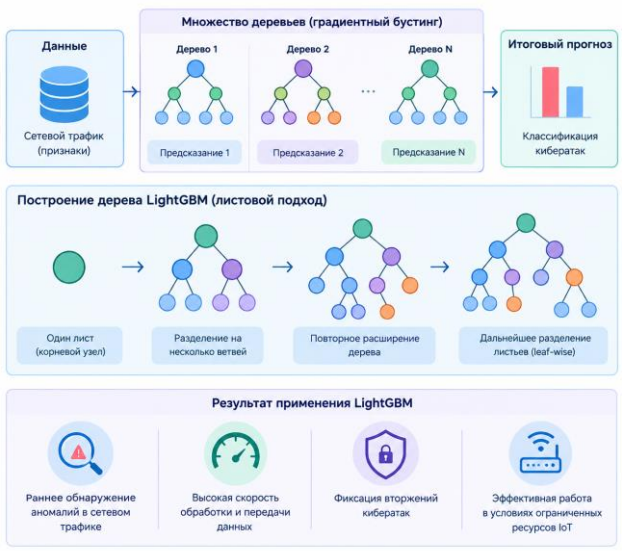


Рисунок 1. Формирование прогноза ансамблем LightGBM и листовой рост дерева

Предлагаемая архитектура состоит из пяти последовательных модулей. Модуль сбора преобразует пакеты в сетевые потоки. Модуль предобработки очищает данные и формирует признаки. Обученный классификатор LightGBM рассчитывает вероятность принадлежности потока к нормальному или атакующему классу. Модуль принятия решения сравнивает вероятность с порогом, выбранным на validation-выборке. Модуль реагирования регистрирует событие, формирует уведомление и передаёт команду системе управления безопасностью.

Таблица 1. Параметры LightGBM и их роль в контроле модели

Параметр	Влияние	Практическая рекомендация
num_leaves	Определяет сложность отдельного дерева	Уменьшать при признаках переобучения
max_depth	Ограничивает глубину дерева	Согласовывать с num_leaves
learning_rate	Задаёт вклад каждого дерева	Снижать вместе с увеличением n_estimators
min_child_samples	Минимум объектов в листе	Увеличивать для редких и шумных признаков
class_weight	Компенсирует дисбаланс классов	Рассчитывать только по train-части
feature_fraction	Доля признаков для дерева	Использовать для регуляризации

Для бинарной классификации параметры модели оптимизируются по логистической функции потерь. Для объекта i с истинной меткой $y_i \in \{0,1\}$ и прогнозом вероятности p_i функция имеет вид $L=-(1/N)\Sigma[y_i \ln(p_i)+(1-y_i)\ln(1-p_i)]$. При многоклассовой постановке используется соответствующая кросс-энтропия. Основные гиперпараметры – num_leaves, max_depth, learning_rate, n_estimators, min_child_samples, feature_fraction и bagging_fraction. Их необходимо подбирать только по validation-части или внутри кросс-валидации (таблица 1).

IV. ПРОТОКОЛ ОЦЕНКИ И ОБСУЖДЕНИЕ

Эксперимент должен начинаться с фиксации схемы разделения данных. Случайное разделение отдельных строк допустимо только при отсутствии временно связанных и повторяющихся потоков. Для потокового IoT-трафика предпочтительно хронологическое разделение: ранние наблюдения используются для обучения, последующие – для тестирования. Если одинаковые сценарии атак генерируются сериями, их следует группировать, чтобы близкие записи не оказались одновременно в обучающей и тестовой частях.

Качество бинарного детектора следует оценивать по матрице ошибок. Precision=TP/(TP+FP) характеризует долю истинных атак среди срабатываний; recall=TP/(TP+FN) показывает долю обнаруженных атак; F1=2·precision·recall/(precision+recall) объединяет обе характеристики. Дополнительно используются ROC-AUC и PR-AUC. При сильном дисбалансе PR-AUC и recall атакующего класса информативнее общей accuracy. Для многоклассовой задачи необходимо публиковать значения precision, recall и F1 для каждого класса, а также macro- и weighted-усреднение.

Сравнение LightGBM должно включать простую базовую модель, например логистическую регрессию, а также альтернативные ансамблевые методы Random Forest и XGBoost. Все модели обучаются на одном разделении данных и одном наборе признаков. Помимо качества распознавания необходимо измерять время обучения, задержку обработки одного потока, объём памяти и размер модели. Такие показатели позволяют определить, следует ли выполнять анализ на шлюзе, периферийном сервере или в облаке (таблица 2).

Таблица 2. Критерии итогового сравнения моделей

Критерий	Что показывает	Почему важен для IoT
Recall атак	Долю обнаруженных атак	Пропуск атаки может нарушить работу системы
Precision	Достоверность тревожных сообщений	Ложные тревоги перегружают оператора
Macro-F1	Баланс качества по всем классам	Учитывает редкие типы атак
PR-AUC	Качество ранжирования при дисбалансе	Устойчивее accuracy при редких атаках
Задержка	Время классификации потока	Определяет пригодность для реального времени
Память	Ресурсы модели при развёртывании	Критично для шлюзов и edge-узлов

Преимуществами LightGBM являются работа с нелинейными зависимостями, устойчивость к различным масштабам числовых признаков и высокая скорость на больших выборках. Ограничения связаны с переобучением при слишком большом числе листьев, чувствительностью к изменению распределения трафика и необходимостью повторного обучения при появлении новых устройств и атак. Следовательно, развёрнутая система должна поддерживать мониторинг дрейфа данных, журналирование ошибок и периодическую актуализацию модели.

V. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ

Практическое внедрение целесообразно выполнять на уровне IoT-шлюза или периферийного сервера. Конечные устройства передают телеметрию штатным способом, а шлюз формирует агрегированные признаки потока и обращается к локальному сервису классификации. Такое размещение не требует установки сложной модели на каждый датчик, уменьшает задержку по сравнению с облачной обработкой и позволяет сохранять работоспособность при временной потере внешнего соединения. В

облако могут передаваться журналы тревог, агрегированная статистика и обезличенные выборки для последующего дообучения.

Выходом классификатора должна быть не только метка, но и вероятность класса. Порог срабатывания выбирается по validation-выборке с учётом стоимости двух типов ошибок. Слишком низкий порог увеличивает число ложных тревог, а слишком высокий повышает риск пропуска атаки. Для критически важных устройств приоритет следует отдавать recall, тогда как для массового мониторинга необходимо одновременно ограничивать частоту ложных срабатываний. Решение об автоматической блокировке соединения рекомендуется принимать только для событий с высокой уверенностью; остальные события направляются оператору.

VI. ЗАКЛЮЧЕНИЕ

Предложена схема автоматизированного обнаружения сетевых атак в IoT-системах на основе LightGBM. Схема охватывает формирование потоковых признаков, предотвращение утечки данных, учёт дисбаланса классов, настройку модели и многокритериальную оценку. LightGBM является перспективным алгоритмом для этой задачи благодаря листовому росту деревьев и эффективной обработке многомерных разреженных данных. Вместе с тем вывод о практическом превосходстве должен подтверждаться сравнением с базовыми моделями на едином тестовом наборе и оценкой не только точности, но и задержки и потребления памяти. Дальнейшая работа предполагает реализацию описанного протокола на Edge-IIoTset или ToN_IoT и проверку устойчивости модели к новым типам атак и изменению профиля сетевого трафика.

ЛИТЕРАТУРА

[1] Бабичева М. В., Третьяков И. А. Применение методов машинного обучения для автоматизированного обнаружения сетевых вторжений // Вестник Дагестанского государственного технического университета. Технические науки. 2023. Т. 50, № 1. С. 53–61.

[2] Балыбердин А. В., Крылов Г. О. Повышение точности выявления аномалий для систем обнаружения вторжения с помощью ансамблевого обучения // Безопасность информационных технологий. 2025. Т. 32, № 1. С. 153–171.

[3] Татарникова Т. М., Богданов П. Ю. Обнаружение атак в сетях Интернета вещей методами машинного обучения // Информационно-управляющие системы. 2021. № 6 (115). С. 42–52.

[4] Ferrag M. A., Friha O., Hamouda D., Maglaras L., Janicke H. Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning // IEEE Access. 2022. Vol. 10. P. 40281–40306. DOI: 10.1109/ACCESS.2022.3165809.

[5] Alsaedi A., Moustafa N., Tari Z., Mahmood A., Anwar A. TON_IoT Telemetry Dataset: A New Generation Dataset of IoT and IIoT for Data-Driven Intrusion Detection Systems // IEEE Access. 2020. Vol. 8. P. 165130–165150. DOI: 10.1109/ACCESS.2020.3022862.

[6] Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 3146–3154.

[7] Bentéjac C., Csörgő A., Martínez-Muñoz G. A Comparative Analysis of Gradient Boosting Algorithms // Artificial Intelligence Review. 2021. Vol. 54. P. 1937–1967. DOI: 10.1007/s10462-020-09896-5.

[8] Чаругин В. В., Чесалин А. Н. Анализ и формирование наборов данных сетевого трафика для обнаружения компьютерных атак // International Journal of Open Information Technologies. 2023. Т. 11, № 6. С. 100–106.

AUTOMATED NETWORK ATTACK DETECTION IN IOT SYSTEMS USING LIGHTGBM

A. Yakunin, B. Mailykhanova

Kazakh National Agrarian Research University, Almaty, Kazakhstan, andryakunin@gmail.com

Abstract: this paper proposes a LightGBM-based workflow for automated network attack detection in IoT infrastructure. It combines flow preprocessing, stratified splitting, and evaluation by precision, recall, F1-score, ROC-AUC, and PR-AUC. The protocol addresses class imbalance and data leakage and supports deployment on an IoT gateway or edge server.

Keywords: Internet of Things; intrusion detection; LightGBM; machine learning.