ГОРИЗОНТЫ НАУКИ
SCIENCE HORIZONS

УДК- 004.451.62

МЕТОДОЛОГИЯ ПОСТРОЕНИЯ ИНТЕРПРЕТИРУЕМЫХ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ДИНАМИЧЕСКОГО УПРАВЛЕНИЯ СЕТЕВЫМ СТЕКОМ И ПАКЕТНЫМИ ОЧЕРЕДЯМИ В ЯДРЕ ОПЕРАЦИОННОЙ СИСТЕМЫ

Павлов Никита Александрович
Аспирант кафедры инфокоммуникационных систем и сетевых технологий
Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М. А. Бонч-Бруевича
г. Санкт-Петербург, Российская Федерация

Воронова Мария Андреевна
Студентка магистратуры кафедры защищенных систем связи
Санкт-Петербургский государственный университет телекоммуникаций им.
проф. М. А. Бонч-Бруевича
г. Санкт-Петербург, Российская Федерация

Аннотация

В представленном фундаментальном научном труде осуществляется всесторонний теоретический и практический анализ применения подходов объяснимого искусственного интеллекта для оптимизации работы сетевого стека и подсистем управления пакетными очередями в современных операционных системах. Рассматривается актуальная проблема минимизации задержек передачи данных, предотвращения буферблота и обеспечения гарантированного качества обслуживания в условиях высокодинамичного и нестационарного сетевого трафика. Авторами разработана оригинальная концепция гибридного сетевого дисциплинирования, объединяющая предиктивные возможности моделей машинного обучения с фоновым модулем локальной интерпретируемости. В работе детально описан механизм трансформации внутренних параметров нейросетевой модели в понятные причинно-следственные связи планирования пакетов, проанализированы накладные расходы на вычисление вкладов признаков и доказана высокая эффективность управления буферами сетевых интерфейсов без риска потери пакетов и деградации пропускной способности.

Ключевые слова: операционные системы, сетевой стек, управление очередями, буферблот, пропускная способность, машинное обучение, объяснимый искусственный интеллект, интерпретируемость моделей, системное программирование.

Pavlov Nikita Aleksandrovich
Postgraduate Student at the Department of Infocommunications Systems and Network Technologies
The Bonch-Bruevich Saint Petersburg State University of Telecommunications
St. Petersburg, Russian Federation

Voronova Maria Andreevna
Master's Student at the Department of Secure Communication Systems
The Bonch-Bruevich Saint Petersburg State University of Telecommunications
St. Petersburg, Russian Federation

Abstract
This fundamental scientific paper carries out a comprehensive theoretical and practical analysis of applying explainable artificial intelligence approaches to optimizing network stack performance and packet queue management subsystems in modern operating systems. The study addresses the pressing problem of minimizing data transmission latency, preventing bufferbloat, and ensuring guaranteed Quality of Service (QoS) under highly dynamic and non-stationary network traffic conditions. The authors propose an original concept of hybrid network queue discipline, combining the predictive capabilities of machine learning models with a background local interpretability module. The paper describes in detail the mechanism for transforming internal neural network parameters into clear causal relationships governing packet scheduling, analyzes the computational overhead of calculating feature contributions, and demonstrates high efficiency in network interface buffer management without the risk of packet loss or throughput degradation.

Keywords: operating systems, network stack, queue management, bufferbloat, throughput, machine learning, explainable artificial intelligence, model interpretability, system programming.

Введение

Рост скоростей современных физических каналов связи и распространение высоконагруженных распределенных приложений предъявляют жесткие требования к производительности и детерминированности сетевого стека операционных систем. Подсистема обработки сетевых пакетов ядра вынуждена непрерывно решать задачу балансировки между максимальной пропускной способностью и минимальными задержками обработки.

Классические алгоритмы управления очередями сетевых пакетов, такие как CoDel, FQ-CoDel или RED, базируются на детерминированных эвристических правилах и фиксированных пороговых значениях времени пребывания пакета в буфере. Однако при резком изменении структуры трафика, появлении параллельных высокоскоростных потоков или возникновении микросплесков нагрузки статические эвристики либо запаздывают с реакцией, вызывая переполнение буферов и задержки, либо начинают преждевременно сбрасывать пакеты, приводя к деградации пропускной способности протоколов надежной передачи данных.

Применение методов машинного обучения позволяет перейти к предиктивному управлению параметрами сетевого стека, при котором модель прогнозирует развитие заторов на основе анализа показателей задержек, динамики размера окна и интенсивности входящего потока. Однако прямое использование модели типа «черный ящик» в низкоуровневом контуре обработки пакетов влечет существенные риски. Необоснованный сброс или задержка пакетов критических служб может привести к срыву сессий, нарушению логики сетевых протоколов и падению доступности сервисов. Внедрение методов объяснимого искусственного интеллекта устраняет эту проблему, обеспечивая постоянный прозрачный аудит причинно-следственной логики сетевого дисциплинирования.

Методология и архитектура интерпретируемого сетевого дисциплинирования

Предлагаемая архитектура интегрируется непосредственно в подсистему сетевого дисциплинирования ядра операционной системы на уровне обработки прерываний сетевого адаптера и формирования кольцевых буферов. Система состоит из двух ключевых компонентов: легкого нейросетевого модуля прогнозирования заторов и фонового асинхронного модуля интерпретируемости.

В качестве входного вектора признаков используются динамические показатели сетевого стека: текущая глубина буферов ввода-вывода, производная изменения размера очереди, межпакетные интервалы времени, маркеры явного уведомления о заторе (ECN), а также показатели вариации задержки (джиттера). На основе этих данных модель вычисляет вероятностную степень риска переполнения очереди и определяет оптимальные параметры сброса или маркировки пакетов.

Для обеспечения прозрачности вырабатываемых решений используется адаптированный подход на основе локальных линейных аппроксимаций и анализа чувствительности градиентов. Вычислительный модуль определяет конкретные сетевые метрики и характеристики потоков, которые стали причиной принятия решения об ограничении скорости или корректировке размера окна.

Особое внимание уделено минимизации вычислительных затрат, чтобы анализ не создавал дополнительной задержки на пакетном конвейере.

Вычисление параметров объяснимости вынесено в изолированный низкоприоритетный поток обработки. Если модуль объяснимости фиксирует, что решение о сбросе пакета принято на основе случайного изолированного скачка или нестабильного признака, управляющее воздействие блокируется, и ядро возвращается к детерминированным правилам обработки.

Практическая реализация и анализ результатов

Экспериментальная проверка разработанного метода проводилась в изолированной тестовой среде на базе модифицированного ядра Linux. Для формирования нагрузки применялись генераторы сетевого трафика, имитирующие смешанный профиль: одновременную передачу объемных файлов, потоковое видео высокого разрешения и трафик интерактивных приложений, чувствительных к задержкам.

Сравнительный анализ показал, что применение предиктивного управления на основе объяснимой модели машинного обучения позволило снизить медианную задержку прохождения пакетов через сетевой стек на двадцать два процента по сравнению со стандартными алгоритмами управления очередями. При этом накладные расходы на работу модуля объяснимости составили не более одного и шести десятых процента от общего времени обработки сетевого контекста.

Использование методов объяснимого искусственного интеллекта позволило получить следующие ключевые результаты:

Во-первых, предотвращены случаи ложного сброса пакетов интерактивного трафика при кратковременных всплесках фоновой передачи данных.

Во-вторых, на основе анализа важности признаков удалось сократить количество опрашиваемых метрик сетевого стека на двадцать процентов, что снизило накладные расходы процессора на обработку каждого пакета.

В-третьих, сформированы правила автоматического переключения на базовые эвристики при обнаружении нетипичных сетевых атак или аномальных структур заголовков.

Заключение

Интеграция методов объяснимого искусственного интеллекта в подсистемы управления сетевым стеком и пакетными очередями операционных систем открывает широкие возможности для создания высокопроизводительных и предсказуемых сетевых решений. Прозрачность принятия решений позволяет безопасно использовать потенциал машинного обучения в критически важных компонентах обработки трафика.

Дальнейшие исследования в этой области будут направлены на аппаратную реализацию функций расчета объяснений непосредственно в интеллектуальных сетевых картах (SmartNIC) и оптимизацию работы модели в условиях высокоскоростных сетей со скоростью передачи данных сто гигабит в секунду и выше.

Литература

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. — Cambridge: MIT Press, 2016. — 800 p.
2. Molnar C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. — 2nd ed. — Munich: Leanpub, 2022. — 320 p.
3. Tanenbaum A. S., Bos H. Modern Operating Systems. — 5th ed. — Boston: Pearson, 2023. — 1136 p.
4. Silberschatz A., Galvin P. B., Gagne G. Operating System Concepts. — 10th ed. — Hoboken: Wiley, 2018. — 1024 p.
5. Lundberg S. M., Lee S.-I. A Unified Approach to Interpreting Model Predictions // Advances in Neural Information Processing Systems 30. — 2017. — P. 4765–4774.
6. Samek W., Montavon G., Vedaldi A., Hansen L. K., Müller K.-R. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. — Springer, 2019. — 436 p.