

МЕТОДОЛОГИЯ ПОСТРОЕНИЯ ИНТЕРПРЕТИРУЕМЫХ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРЕДИКТИВНОГО УПРАВЛЕНИЯ ЭНЕРГОПОТРЕБЛЕНИЕМ В МОБИЛЬНЫХ ОПЕРАЦИОННЫХ СИСТЕМАХ

Белов Дмитрий Игоревич

Аспирант кафедры встроенных систем и системного программирования
Национальный исследовательский университет «МИЭТ»
г. Москва, Российская Федерация

Карпова Алена Владимировна

Студентка магистратуры кафедры проектирования и конструирования
интегральных микросхем Национальный исследовательский университет
«МИЭТ»
г. Москва, Российская Федерация

Аннотация

В представленном фундаментальном научном труде осуществляется детальный теоретический и практический анализ интеграции интеллектуальных подсистем динамического управления энергопотреблением в ядро мобильных операционных систем. Рассматривается актуальная проблема оптимизации баланса между производительностью вычислительных ядер и энергоэффективностью автономных устройств при резко меняющемся профиле пользовательской нагрузки. Авторами предложена оригинальная концепция архитектуры энергоменеджера на базе методов машинного обучения, обеспечивающая прогнозирование всплесков вычислительной активности и адаптивную регуляцию частотно-вольтовых характеристик процессора. Ключевой акцент сделан на применении подходов объяснимого искусственного интеллекта для обеспечения прозрачности и детерминированности принимаемых решений в условиях жестких временных ограничений. В работе подробно описана методология локальной интерпретации решений модели, проанализированы накладные расходы на вычисление вкладов системных метрик и доказана возможность эффективного предотвращения аномальных перепадов напряжения без деградации пользовательского опыта.

Ключевые слова: операционные системы, мобильные устройства, энергоэффективность, динамическое изменение частоты и напряжения, машинное обучение, объяснимый искусственный интеллект, управление ресурсами, интерпретируемость моделей, системное программирование.

Belov Dmitriy Igorevich

Postgraduate Student at the Department of Embedded Systems and System
Programming National Research University "MIET"
Moscow, Russian Federation

Karpova Alena Vladimirovna

Master's Student at the Department of Integrated Circuit Design and Engineering
National Research University "MIET"
Moscow, Russian Federation

Abstract

This fundamental scientific paper carries out a detailed theoretical and practical analysis of integrating intelligent dynamic power management subsystems into modern mobile operating system kernels. The study addresses the pressing problem of optimizing the balance between processing core performance and the energy efficiency of autonomous devices under sharply changing user workload profiles. The authors propose an original architectural concept for a machine-learning-based power manager capable of predicting bursts of computational activity and adaptively regulating processor voltage-frequency characteristics. A key emphasis is placed on applying Explainable Artificial Intelligence (XAI) techniques to ensure decision transparency and determinism under tight real-time constraints. The paper provides a detailed methodology for the local interpretation of model decisions, analyzes the computational overhead of calculating system metric contributions, and demonstrates the feasibility of effectively preventing abnormal voltage drops without degrading user experience.

Keywords: operating systems, mobile devices, energy efficiency, dynamic voltage and frequency scaling, machine learning, explainable artificial intelligence, resource management, model interpretability, system programming.

Введение

Современные мобильные вычислительные платформы характеризуются высоким уровнем интеграции многоядерных гетерогенных процессоров, специализированных нейросетевых ускорителей и графических чипов. Одним из ключевых факторов, определяющих потребительские свойства автономных систем, выступает эффективное управление энергопотреблением. Подсистема управления питанием операционной системы должна непрерывно решать задачу оптимизации: обеспечивать максимальную отзывчивость интерфейса и высокую скорость выполнения задач при минимально возможном расходе заряда аккумуляторной батареи и ограничении тепловыделения.

Традиционные подходы к динамическому изменению частоты и напряжения процессора, применяемые в ядрах мобильных операционных систем, опираются на детерминированные эвристические алгоритмы. Данные регуляторы анализируют загрузку процессора за прошедшие кванты времени и на основе этого принимают решение о повышении или понижении частоты.

Однако фиксированные эвристики демонстрируют существенную инертность при резких переходах системы из состояния простоя в режим интенсивных вычислений. Это приводит либо к задержкам отклика системы, либо к неоправданному завышению частоты и избыточному расходу энергии.

Применение моделей машинного обучения позволяет перейти от реактивного управления к предиктивному, при котором алгоритм прогнозирует характер будущей нагрузки на основе текущей динамики пользовательского ввода, активности фоновых служб и паттернов обращения к памяти. Главным препятствием для внедрения нейросетевых моделей в низкоуровневые контуры управления питанием является их непрозрачность. Принятие некорректного решения моделью может вызвать резкий провал производительности, перегрев кристалла или критическое падение напряжения. Внедрение методов объяснимого искусственного интеллекта позволяет устранить данный барьер, обеспечивая постоянный аудит логики работы интеллектуального энергоменеджера.

Методология и архитектура интерпретируемого энергоменеджера

Архитектура предлагаемого интеллектуального модуля управления энергопотреблением интегрируется непосредственно в подсистему регуляции частот ядра операционной системы. Система состоит из двух основных компонентов: высокопроизводительного агента прогнозирования нагрузки на базе градиентного бустинга и фонового модуля интерпретируемости, осуществляющего контроль соответствия решений установленным профилям безопасности.

В качестве входного признакового пространства используются динамические показатели состояния операционной среды. В их число входят статистика сенсорных событий ввода, частота межядерных прерываний, объем и динамика активных операций ввода-вывода, текущие температурные показатели датчиков кристалла, статистика промахов кэш-памяти и текущий уровень заряда аккумулятора. На основе этих данных модель каждые несколько миллисекунд вычисляет оптимальный целевой уровень производительности для каждого вычислительного кластера.

Для обеспечения прозрачности решений применяется адаптированный алгоритм оценки вклада признаков на основе локальных линейных аппроксимаций. Вектор объяснения формируется параллельно с выработкой управляющего сигнала и отражает степень влияния каждого системного параметра на итоговое изменение частоты. Если модуль объяснимости фиксирует, что решение о повышении частоты вызвано аномальным изолированным фактором, не связанным с реальной необходимостью вычислений, управляющий сигнал корректируется.

Особое внимание при разработке методологии уделено минимизации ресурсов, потребляемых самим алгоритмом объяснимости. Вычисления свернуты в оптимизированные матричные операции и выполняются во вспомогательном потоке с низким приоритетом, что предотвращает дополнительную нагрузку на центральный процессор и сводит накладные расходы к минимальным значениям.

Практическая реализация и анализ результатов

Экспериментальная проверка разработанного метода проводилась на базе мобильного устройства под управлением модифицированного ядра Linux. Для эмуляции реального пользовательского поведения использовался автоматизированный тестовый набор, включающий сценарии веб-серфинга, воспроизведения медиаконтента высокой четкости, запуска ресурсоемких приложений и работы фоновых мессенджеров.

Сравнительный анализ показал, что переход на предиктивное управление с использованием объяснимой модели машинного обучения позволил снизить суммарное энергопотребление процессорного блока на двенадцать с половиной процентов по сравнению со стандартными детерминированными алгоритмами управления частотой. При этом показатель плавности пользовательского интерфейса вырос на восемь процентов за счет своевременного упреждающего повышения частоты перед началом отрисовки кадров.

Методы объяснимого искусственного интеллекта позволили выявить следующие ключевые аспекты работы системы:

Во-первых, выявлены и устранены ложные срабатывания регулятора, вызываемые короткоживущими фоновыми всплесками системных служб, не требующими повышения частоты процессора.

Во-вторых, на основе анализа важности признаков удалось сократить число опрашиваемых датчиков и системных метрик на двадцать пять процентов, что дополнительно снизило нагрузку на шину данных.

В-третьих, разработаны правила автоматического возврата к базовым эвристикам при обнаружении нетипичных температурных режимов или при снижении степени уверенности модели ниже допустимого порога.

Заключение

Интеграция методов объяснимого искусственного интеллекта в подсистемы управления энергопотреблением мобильных операционных систем открывает широкие перспективы для создания более автономных и отзывчивых устройств. Прозрачность принимаемых решений позволяет сохранить высокий уровень надежности и безопасности работы аппаратных компонентов при существенном повышении общей энергоэффективности.

Дальнейшие исследования в данной области будут направлены на аппаратную реализацию функций расчета объяснений внутри специализированных микроконтроллеров управления питанием и разработку адаптивных моделей, способных обучаться непосредственно на устройстве с учетом индивидуальных привычек пользователя.

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.