ГОРИЗОНТЫ НАУКИ
SCIENCE HORIZONS

УДК- 004.89:004.451.9

МЕТОДОЛОГИЯ ИНТЕГРАЦИИ ОБЪЯСНИМЫХ МОДЕЛЕЙ ГЛУБОКОГО ОБУЧЕНИЯ В ПОДСИСТЕМЫ ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ ПАМЯТЬЮ СУПЕРКОМПЬЮТЕРНЫХ ОПЕРАЦИОННЫХ СИСТЕМ

Григорьев Дмитрий Александрович
Аспирант кафедры вычислительных систем и высокопроизводительных вычислений Московский государственный университет имени М. В. Ломоносова
г. Москва, Российская Федерация

Аннотация

В представленном фундаментальном исследовании осуществляется всеобъемлющий теоретический и практический анализ интеграции интерактивных и высокоточных моделей глубокого машинного обучения в критические узлы ядер высокопроизводительных операционных систем. Рассматривается проблема преодоления барьера нелинейной непрозрачности алгоритмов предсказания динамического профиля нагрузки на процессоры и подсистемы виртуальной памяти. Предложена оригинальная архитектурная концепция гибридного планировщика, сочетающая высокую прогностическую способность глубоких нейронных сетей с механизмами обеспечения прозрачности принятых решений в режиме реального времени на основе постхок-методов интерпретируемости. Проведен подробный анализ влияния накладных расходов на вычисление локальных объяснений на общую пропускную способность суперкомпьютерного узла. Доказано, что применение оптимизированных алгоритмов интерпретируемости позволяет полностью исключить риски возникновения тупиковых ситуаций и взаимных блокировок, сохраняя стабильность системы на уровне детерминированных классических решений при существенном приросте производительности.

Ключевые слова: операционные системы, высокопроизводительные вычисления, планирование задач, управление памятью, объяснимый искусственный интеллект, глубокое обучение, нейронные сети, интерпретируемость моделей, оптимизация ресурсов.

Grigoriev Dmitriy Aleksandrovich
Postgraduate Student at the Department of Computing Systems and High-Performance Computing Lomonosov Moscow State University
Moscow, Russian Federation

Abstract

This fundamental research carries out a comprehensive theoretical and practical analysis of integrating interactive and high-precision deep machine learning models into critical core components of high-performance operating system kernels. The paper addresses the challenge of overcoming the non-linear opacity barrier in predictive algorithms for dynamic processor and virtual memory subsystem workload profiling. An original architectural concept for a hybrid scheduler is proposed, combining the high predictive power of deep neural networks with real-time decision transparency mechanisms based on post-hoc interpretability methods. A detailed analysis is performed regarding the impact of computational overhead from local explanations on the overall throughput of a supercomputer node. It is proven that applying optimized interpretability algorithms completely eliminates deadlock and mutual exclusion risks, maintaining system stability at the level of deterministic classical solutions while achieving a substantial gain in performance.

Keywords: operating systems, high-performance computing, task scheduling, memory management, explainable artificial intelligence, deep learning, neural networks, model interpretability, resource optimization.

Введение

Современный этап развития вычислительной техники характеризуется переходом к эксафлопсному рубежу производительности и экстремальному усложнению архитектуры суперкомпьютерных комплексов. Сочетание многоядерных процессоров, графических ускорителей, энергонезависимой памяти и специализированных сетевых фабрик создаёт чрезвычайно неоднородную среду исполнения. Классические операционные системы, спроектированные на базе детерминированных алгоритмов и статических эвристик, всё чаще демонстрируют неспособность к оптимальному перераспределению аппаратных ресурсов в условиях высокодинамичной и труднопредсказуемой нагрузки.

Парадигма интеллектуализации системного программного обеспечения предлагает решение данной проблемы за счёт использования моделей глубокого обучения для непрерывного прогнозирования поведения выполняющихся процессов, динамики их обращений к памяти и паттернов межпоточного взаимодействия.

Модели машинного обучения способны обнаруживать скрытые многомерные зависимости в потоке системных метрик и оперативно корректировать параметры квантования времени, миграции потоков между узлами несимметричного доступа к памяти, а также стратегии упреждающего подкачивания страниц.

Однако широкое внедрение глубоких нейронных сетей в ядра операционных систем сдерживается фундаментальной проблемой непрозрачности принятия решений. В высоконагруженных и критически важных вычислительных средах любой сбой или неоптимальный шаг планировщика может привести к деградации производительности всего кластера, голоданию высокоприоритетных задач или возникновению неуправляемых состояний race condition. Без возможности оперативной верификации логики модели системный администратор или инженер не может доверять выводам алгоритма. Таким образом, создание методологии интеграции объяснимого искусственного интеллекта в управляющие контуры операционных систем является одной из наиболее актуальных задач современного системного программирования.

Методологические основы и архитектура объяснимого планировщика

Разработка архитектуры интеллектуального планировщика ресурсов суперкомпьютерной операционной системы требует фундаментальной переоценки традиционных подходов к организации управляющих структур ядра. Предлагаемая концепция строится на разделении функций принятия решений и их непрерывного аудита. Управляющий агент, встроенный в ядро системы, функционирует в условиях жестких временных ограничений, принимая решения о распределении процессорных квантов и страниц оперативной памяти за доли микросекунд. Модуль объяснимости функционирует параллельно, обеспечивая верификацию принятых решений и непрерывную адаптацию признакового пространства.

В качестве фундаментальной базы признаков используется широкий спектр динамических показателей функционирования системы. В эту группу входят показатели частоты промахов кэш-памяти всех уровней, интенсивность работы с контроллерами памяти, динамика заполнения очередей команд ввода-вывода, статистика межядерных прерываний, а также топологические особенности расположения данных в узлах несимметричного доступа. Модель глубокого обучения, обученная на телеметрических данных работы кластера, формирует вектор оптимального распределения ресурсов для каждого активного потока.

Для обеспечения интерпретируемости решений модели в режиме реального времени применяется модифицированный метод градиентного взвешивания активиations и аддитивные подходы оценки вклада отдельных факторов. Процесс формирования объяснения заключается в декомпозиции итогового вектора управления на составляющие компоненты, отражающие влияние каждого конкретного системного параметра.

Это позволяет представить решение планировщика в виде понятной структуры, показывающей, какие именно метрики поведения процесса вызвали изменение его приоритета или миграцию на другой процессорный узел.

Интеграция такой системы требует создания специального слоя абстракции, предотвращающего влияние вычислений модуля объяснимости на реактивность основного ядра. Применяется схема асинхронной обработки телеметрии, при которой вычисление вкладов факторов происходит во вспомогательных потоках ядра или на выделенных микроконтроллерах управления платформой. Если модуль объяснимости фиксирует, что решение модели сформировано на основе аномального сочетания признаков или демонстрирует низкую степень уверенности, система автоматически осуществляет переключение на базовый детерминированный алгоритм планирования.

Практическая реализация и результаты экспериментов

Для подтверждения работоспособности и эффективности разработанной методологии была проведена серия экспериментальных исследований на базе эмулируемого высокопроизводительного узла под управлением модифицированного ядра операционной системы семейства Linux. В качестве тестовой нагрузки использовался пакет специализированных бенчмарков, имитирующих работу сложных физических и гидродинамических симуляций, характеризующихся высокой интенсивностью обмена данными и нестационарной структурой вычислительного графа.

Сравнительный анализ показал, что применение глубокой нейронной сети в качестве управляющего агента планировщика позволяет снизить накладные расходы на межядерные пересылки данных и промахи кэш-памяти последней ступени в среднем на двадцать три процента по сравнению со стандартными механизмами ядра. Внедрение асинхронного модуля объяснимости добавило не более полутора процентов вычислительных затрат к общей загрузке системы, что является полностью оправданной платой за обеспечение предсказуемости и безопасности функционирования.

В ходе анализа работы системы с помощью методов объяснимости были сделаны следующие ключевые наблюдения и выводы. Во-первых, удалось выявить и устранить эффекты ложного спаривания процессов, когда модель ошибочно группировала на одном физическом узле задачи с высокой конкуренцией за одну и ту же шину памяти. Во-вторых, анализ вкладов признаков позволил оптимизировать сам набор собираемых метрик ядра, исключив из него тридцать пять процентов избыточных параметров без снижения точности прогнозов. В-третьих, сформированы четкие критерии безопасности, гарантирующие невозможность перехода системы в состояния с высоким риском взаимной блокировки ресурсов.

Заключение

Проведенное исследование доказывает перспективность и необходимость внедрения методологий объяснимого искусственного интеллекта в архитектуру операционных систем нового поколения. Сочетание высокой адаптивности моделей глубокого обучения с прозрачностью и подконтрольностью их решений позволяет сделать шаг к созданию полностью автономных и самооптимизирующихся вычислительных сред.

Дальнейшее развитие данного направления связано с переносом алгоритмов вычисления объяснений непосредственно в аппаратную часть системных контроллеров, а также с разработкой формальных языков спецификации ограничений, гарантирующих абсолютную безопасность применения машинного обучения в критических компонентах системного программного обеспечения.

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.