

УДК 004.75+004.272+519.83

DOI: [10.26102/2310-6018/2026.56.5.012](https://doi.org/10.26102/2310-6018/2026.56.5.012)

Экспериментальное исследование кооперативного распределения ресурсов в системах обработки больших данных на основе значения Шепли и машинного обучения

М.В. Бляхеров✉, И.Я. Львович

Воронежский институт высоких технологий, Воронеж, Российская Федерация

Резюме. Оптимальное и справедливое распределение вычислительных ресурсов в динамичных средах обработки больших данных (Big Data), таких как Apache Spark, остается сложной задачей. Традиционные планировщики часто не учитывают синергетические эффекты кооперации между задачами, что приводит к неэффективности и конфликтам. Целью данной работы является экспериментальное исследование и верификация гибридного подхода к кооперативному распределению ресурсов, основанного на формальных принципах теории кооперативных игр и адаптивных возможностях машинного обучения. В работе формализована модель кооперативной игры, где коалиции параллельных задач характеризуются функцией полезности, зависящей от выделенных ресурсов. Для обеспечения устойчивости и справедливости введены условия равновесия (ядро игры), а распределение основано на значении Шепли, оценивающем маржинальный вклад каждой задачи. Для преодоления аналитической сложности оценки полезности в реальных условиях предложено использование ML-моделей (градиентный бустинг, графовые нейронные сети), обученных на исторических метриках кластера, в качестве аппроксиматоров характеристической функции игры. Разработан и развернут экспериментальный стенд на базе Apache Spark с системой мониторинга Prometheus/Grafana. В ходе экспериментов показано, что предложенный подход обеспечивает динамическое и сбалансированное распределение ресурсов (CPU, память), повышает стабильность коалиций задач и улучшает общую справедливость распределения (индекс Джинни, fairness index) по сравнению с базовыми сценариями. Визуализация ключевых метрик подтвердила достижение состояний, близких к ядру игры. Исследование демонстрирует практическую применимость и эффективность сочетания теоретико-игровых моделей и машинного обучения для интеллектуального управления ресурсами в распределенных системах Big Data, открывая путь к созданию самооптимизирующихся и кооперативных оркестраторов.

Ключевые слова: распределение ресурсов, кооперативная теория игр, значение Шепли, машинное обучение, большие данные, Apache Spark, мониторинг.

Для цитирования: Бляхеров М.В., Львович И.Я. Экспериментальное исследование кооперативного распределения ресурсов в системах обработки больших данных на основе значения Шепли и машинного обучения. *Моделирование, оптимизация и информационные технологии*. 2026;14(5). URL: <https://moitvvt.ru/ru/journal/article?id=2226> DOI: 10.26102/2310-6018/2026.56.5.012

Experimental study of cooperative resource allocation in big data processing systems based on Shapley value and machine learning

M.V. Blyakherov✉, I.Ya. Lvovich

Voronezh Institute of High Technologies, Voronezh, the Russian Federation

Abstract. Optimal and equitable allocation of computing resources in dynamic Big Data environments such as Apache Spark remains a challenge. Traditional planners often do not take into account the synergetic effects of cooperation between tasks, which leads to inefficiency and conflicts. The purpose of this work is to experimentally investigate and verify a hybrid approach to cooperative resource

allocation based on the formal principles of cooperative game theory and adaptive machine learning capabilities. The paper formalizes a model of a cooperative game, where coalitions of parallel tasks are characterized by a utility function depending on the allocated resources. To ensure stability and fairness, equilibrium conditions (the core of the game) have been introduced, and the distribution is based on the Shapley value, which estimates the marginal contribution of each task. To overcome the analytical complexity of evaluating utility in real conditions, it is proposed to use ML models (gradient boosting, graph neural networks) trained on historical cluster metrics as approximators of the characteristic function of the game. An experimental bench based on Apache Spark with the Prometheus/Grafana monitoring system has been developed and deployed. Experiments have shown that the proposed approach provides a dynamic and balanced allocation of resources (CPU, memory), increases the stability of task coalitions, and improves overall distribution equity (Ginny index) compared to the baseline scenarios. Visualization of key metrics confirmed the achievement of states close to the core of the game. The study demonstrates the practical applicability and effectiveness of combining game-theoretic models and machine learning for intelligent resource management in distributed Big Data systems, paving the way for the creation of self-optimizing and cooperative orchestrators.

Keywords: resource allocation, cooperative game theory, Shapley value, machine learning, big data, Apache Spark, monitoring.

For citation: Blyakherov M.V., Lvovich I.Ya. Experimental study of cooperative resource allocation in big data processing systems based on Shapley value and machine learning. *Modeling, Optimization and Information Technology*. 2026;14(5). (In Russ.). URL: <https://moitvvt.ru/ru/journal/article?id=2226> DOI: 10.26102/2310-6018/2026.56.5.012

Введение

Современные системы обработки больших данных, такие как Apache Spark, Apache Flink и Hadoop, стали неотъемлемой частью инфраструктуры организаций, работающих с масштабными и быстрорастущими наборами данных. Эти системы функционируют в условиях высокой динамики рабочих нагрузок, многопользовательского доступа и совместного использования ограниченных вычислительных ресурсов – процессорного времени, оперативной памяти, каналов ввода-вывода. В такой среде задача эффективного и справедливого распределения ресурсов приобретает критическое значение, непосредственно влияя на производительность приложений, их отказоустойчивость и стоимость эксплуатации кластеров [1, 2].

Существующие механизмы планирования, применяемые в большинстве распределенных вычислительных сред, как правило, основываются на статических квотах, очередях с приоритетами или изолированных контейнерах. Однако данные подходы обладают существенными ограничениями. Во-первых, они не учитывают синергетический эффект, возникающий при кооперативном выполнении задач, совместно использующих общие данные или кэшированные промежуточные результаты. Во-вторых, традиционные планировщики не способны адаптироваться к быстро меняющимся профилям нагрузки и часто приводят к ситуациям «ценой анархии», когда индивидуально рациональное поведение исполнителей ведет к деградации общей производительности системы. Наконец, вопросы справедливости распределения ресурсов зачастую отходят на второй план, уступая место критериям утилизации и пропускной способности [3].

В ответ на указанные ограничения в последние годы активно развивается направление интеллектуального управления ресурсами, основанное на методах теории игр и машинного обучения. Теоретико-игровой подход позволяет формализовать взаимодействие исполнителей как кооперативную игру, где целью является максимизация совокупной полезности при соблюдении условий устойчивости коалиций.

Ключевую роль в таком подходе играет значение Шепли – концепция справедливого распределения общего выигрыша, пропорционального вкладу каждого участника [4]. В свою очередь, машинное обучение предоставляет инструментарий для аппроксимации характеристической функции игры, аналитическое определение которой в условиях реального кластера затруднительно ввиду множества нелинейных зависимостей и стохастических факторов.

Несмотря на наличие теоретических работ, посвященных применению кооперативных игр в вычислительных системах, число экспериментальных исследований, реализующих и верифицирующих данные модели на реальных платформах обработки больших данных, остается ограниченным. В частности, отсутствуют работы, в которых гибридный подход, сочетающий значение Шепли и прогностические модели машинного обучения, был бы реализован в виде прототипа, интегрированного с кластером Apache Spark, и подвергнут систематическому экспериментальному анализу.

Настоящая работа направлена на восполнение указанного пробела. Целью исследования является разработка, реализация и экспериментальная верификация модели кооперативного распределения ресурсов в системах обработки больших данных, основанной на формальных условиях равновесия (ядре кооперативной игры), справедливом распределении выигрыша с использованием значения Шепли и адаптивном прогнозировании полезности коалиций методами машинного обучения [5].

Для достижения поставленной цели в работе решаются следующие задачи:

Формализация оптимизационной задачи распределения ресурсов с учетом кооперативного взаимодействия исполнителей и ограничений кластера.

Определение условий равновесия коалиций, обеспечивающих устойчивость совместного использования ресурсов.

Разработка архитектуры и реализация прототипа системы управления ресурсами на базе Apache Spark, интегрированной с модулем вычисления значения Шепли и ML-моделью прогнозирования полезности [6].

Проведение серии вычислительных экспериментов с анализом ключевых метрик производительности, справедливости и стабильности распределения.

Оценка эффективности предложенного подхода в сравнении с базовыми стратегиями распределения.

Научная новизна исследования заключается в комплексном применении аппарата кооперативных игр и методов машинного обучения для решения задачи оперативного распределения ресурсов в распределенных системах, а также в экспериментальном подтверждении достижимости состояний, близких к ядру игры, в условиях реальной вычислительной среды.

Практическая значимость работы состоит в создании воспроизводимого экспериментального стенда и программного прототипа, которые могут быть использованы как основа для разработки интеллектуальных планировщиков нового поколения в экосистемах обработки больших данных.

Материалы и методы

В данном разделе представлена математическая формализация предлагаемого подхода, описание архитектуры экспериментального стенда, методика сбора данных и обучения ML-моделей, а также процедура проведения вычислительных экспериментов.

Математическая модель кооперативного распределения ресурсов. Формализация коалиционной игры. Рассмотрим множество исполнителей (задач, приложений или пользователей) $N = \{1, 2, \dots, n\}$, конкурирующих за ограниченные

ресурсы кластера. Под коалицией $S \subseteq N$ понимается произвольное подмножество исполнителей, принявших решение о совместном использовании ресурсов и координации своих стратегий.

Характеристическая функция $v(S)$ ставит в соответствие каждой коалиции S величину совокупной полезности, которую способны получить участники при оптимальном распределении доступных им ресурсов. В контексте систем обработки больших данных под полезностью может пониматься интегральная пропускная способность, обратная величина средней задержки обработки или снижение частоты критических сбоев (например, ошибок нехватки памяти – OOM) [7].

Задача оптимального распределения ресурсов внутри фиксированной коалиции S формулируется как задача максимизации суммарной полезности ее участников при ограничениях на доступный объем ресурсов кластера:

$$\max_{\{x_i\}_{i \in S}} \sum_{i \in S} u_i(x_i),$$

$$\text{при условии } \sum_{i \in S} r_i(x_i) \leq R_{total},$$

где x_i – вектор управляемых параметров конфигурации для исполнителя i (например, объём выделяемой памяти, количество ядер CPU, квоты на кэширование RDD); $u_i(x_i)$ – индивидуальная функция полезности исполнителя i ; $r_i(x_i)$ – объём потребляемых ресурсов; R_{total} – общий доступный объём ресурсов кластера.

Условия равновесия (ядро кооперативной игры). Для обеспечения устойчивости кооперативного соглашения необходимо, чтобы распределение суммарного выигрыша $v(N)$ между всеми исполнителями удовлетворяло условиям ядра кооперативной игры. Вектор платежей $\phi = (\phi_1, \dots, \phi_n)$ принадлежит ядру, если:

- 1) эффективность: $\sum_{i \in N} \phi_i = v(N)$;

- 2) коалиционная рациональность: для любой коалиции $S \subseteq N$ выполняется $\sum_{i \in N} \phi_i \geq v(N)$.

Интерпретация данных условий в терминах распределения ресурсов означает, что ни одна группа исполнителей не способна, отделившись от общей коалиции, обеспечить себе более высокую совокупную полезность, чем та, которую они получают в рамках единого пула ресурсов. Достижение состояния, принадлежащего ядру, гарантирует отсутствие стимулов к разрушению кооперации.

Значение Шепли как механизм справедливого распределения. Для построения конкретного вектора распределения, удовлетворяющего свойствам справедливости и, при определенных условиях, принадлежащего ядру, используется значение Шепли. Вклад каждого исполнителя i определяется как его усредненный маржинальный вклад во все возможные коалиции:

$$\varphi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} (v(S \cup \{i\}) - v(S)).$$

Значение Шепли является единственным решением, удовлетворяющим аксиомам симметричности, эффективности, аддитивности и фиктивного игрока. В контексте распределения ресурсов величина φ_i интерпретируется как справедливая доля ресурсов (или приоритет доступа), пропорциональная вкладу исполнителя в общую эффективность системы.

Интеграция машинного обучения для аппроксимации характеристической функции. Ключевой проблемой практического применения описанной модели является сложность аналитического определения значений характеристической функции $v(S)$ в реальных условиях. Функция полезности коалиции зависит от множества факторов: текущей загрузки CPU и памяти, шаблонов доступа к данным, эффективности

кэширования, поведения сборщика мусора JVM и многих других. Аналитическое моделирование данных зависимостей при большом количестве одновременно исполняемых задач и динамически изменяющихся условиях эксплуатации не представляется возможным.

Для преодоления данного ограничения предлагается использовать ML-модель в качестве аппроксиматора характеристической функции. Процесс интеграции включает следующие этапы:

1. Формирование пространства признаков. Каждая коалиция S описывается вектором признаков x_S , включающим:

- агрегированные метрики использования ресурсов (суммарный объем аллоцированной памяти, загрузка CPU, объем пересылаемых данных);
- характеристики входящих в коалицию задач (профили нагрузки, классы приложений);
- выбранные стратегии управления кэшем (частота вызова `unpersist()`).

2. Обучение модели. Используется набор исторических данных $D = \{(x_S, v_j)\}$, где v_j – измеренное значение полезности коалиции S_j . Модель $f_\theta(x_S)$ обучается минимизировать функцию потерь:

$$L(\theta) = \sum_{(x_S, v) \in D} (f_\theta(x_S) - v)^2.$$

3. Подстановка в игровую модель. Обученная модель f_θ используется вместо $v(S)$ при вычислении значения Шепли и проверке условий ядра.

В качестве базовых алгоритмов машинного обучения рассматривались градиентный бустинг (XGBoost, LightGBM) и графовые нейронные сети (GNN), позволяющие учитывать структуру взаимодействий между исполнителями внутри коалиции. Выбор данных методов обусловлен их способностью эффективно работать с табличными данными и извлекать зависимости из разреженных признаков пространств.

Архитектура экспериментального стенда. Для проведения экспериментального исследования разработан и развернут прототип системы управления ресурсами на базе Apache Spark, интегрированный с компонентами сбора метрик, ML-прогнозирования и принятия решений [8].

Общая схема взаимодействия компонентов. Архитектура стенда включает три функциональных слоя (Рисунок 1):

- Слой данных и вычислений: кластер Apache Spark (режим master/worker), HDFS/S3 для хранения данных, Apache Kafka для потоковых событий.
- Слой принятия решений: модуль вычисления значения Шепли, ML-сервис прогнозирования полезности, оптимизатор распределения ресурсов.
- Слой мониторинга и управления: Prometheus (сбор и хранение временных рядов) [9], Grafana (визуализация) [10], Alert Manager (оповещения).

Реализация компонентов. Кластер Apache Spark развернут с использованием Docker Compose. Конфигурация включает один мастер-узел и несколько воркер-узлов (исполнителей). Параметры ресурсов (CPU, память) конфигурируются динамически через REST API.

Модуль сбора метрик реализован на языке Python и интегрирован со Spark через SparkListener интерфейс. Метрики экспортируются в формате Prometheus и сохраняются во временной базе данных.

Модуль вычисления значения Шепли реализует точный перебор всех подмножеств коалиций. Ввиду экспоненциальной сложности алгоритма, в

экспериментах использовалось ограниченное число исполнителей (до 10), что допустимо для целей верификации подхода.

ML-компонент реализован на базе библиотеки XGBoost. Модель обучается на накопленных исторических данных о производительности коалиций. Периодическое дообучение запускается при снижении точности прогноза ниже порогового значения (Рисунок 2).

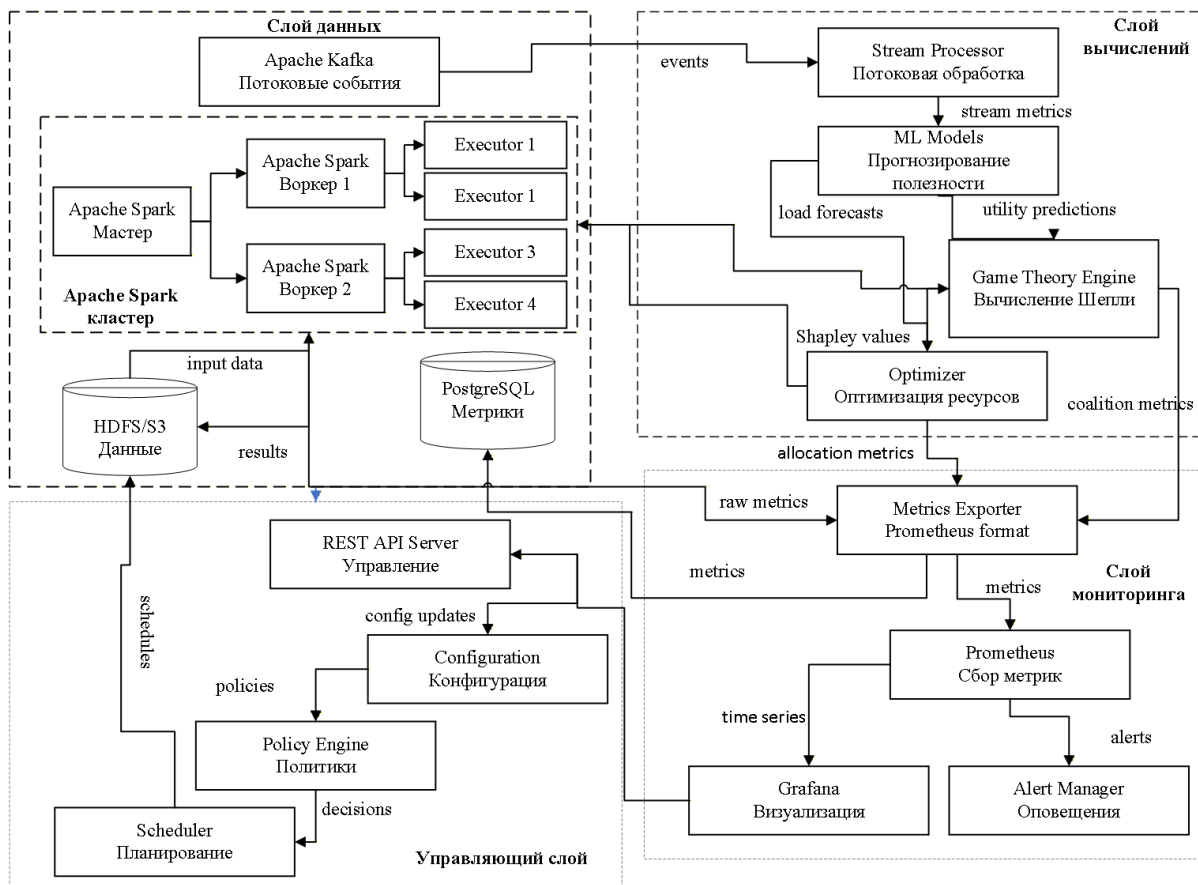


Рисунок 1 – Архитектура экспериментального стенда и взаимодействие компонентов системы
 Figure 1 – Architecture of the experimental setup and interaction of system components

Методика проведения экспериментов. Целевые метрики. Для оценки эффективности предложенного подхода используется следующий набор метрик:

1. Значение Шепли φ_i распределение «справедливой доли» между исполнителями (динамика во времени).
2. Индекс справедливости (Fairness Index) – степень равномерности распределения ресурсов.
3. Коэффициент Джинни – мера неравенства в распределении значений Шепли.
4. Стабильность коалиции – среднеквадратическое отклонение значений Шепли, нормированное на среднее.
5. Эффективность использования ресурсов – отношение суммарно выделенного CPU к объему памяти.
6. Проверка условия ядра – выполнение неравенства $v(S) \geq \sum_{i \in S} \varphi_i$ для случайных подмножеств.

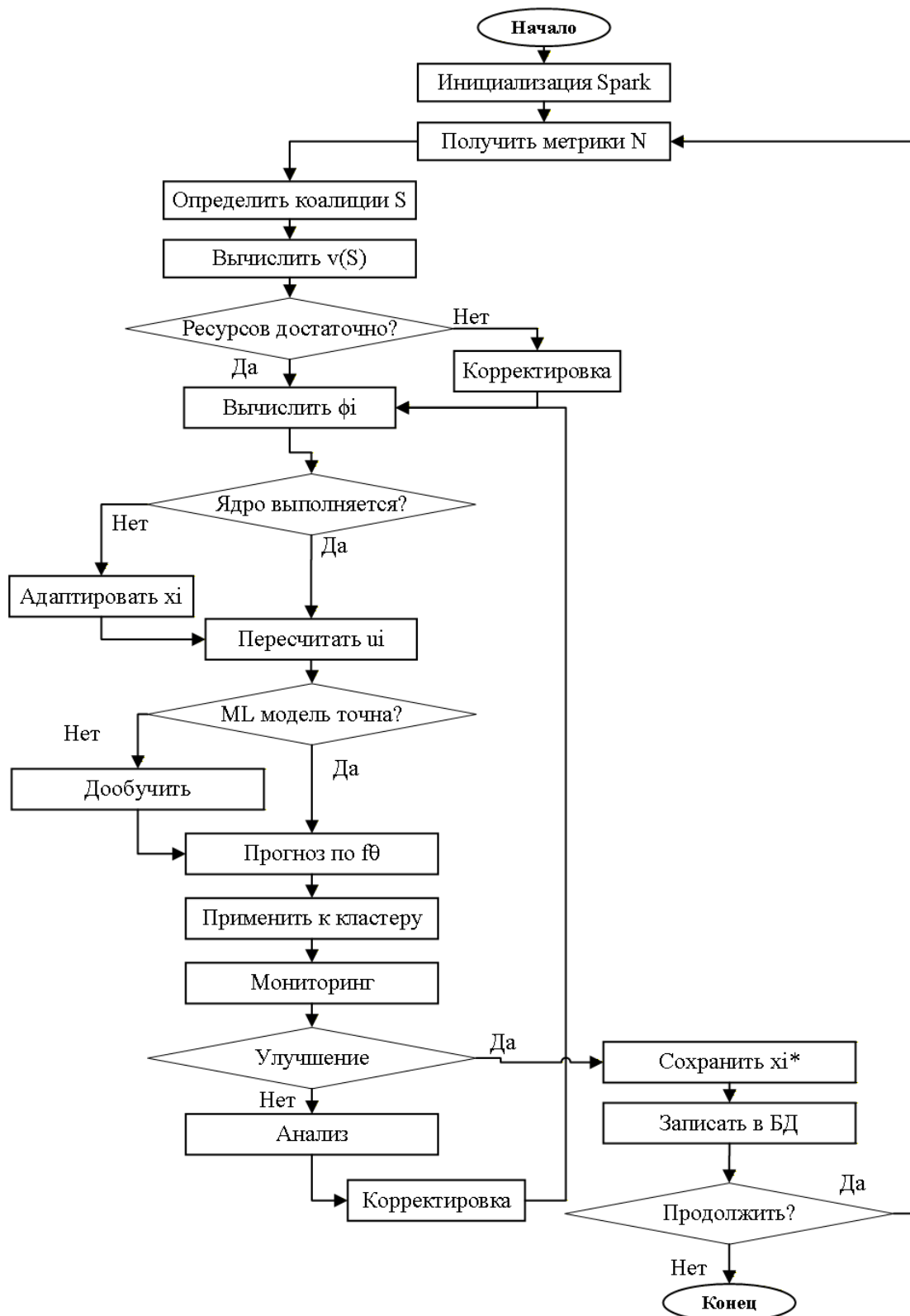


Рисунок 2 – Схема алгоритма работы системы управления ресурсами
Figure 2 – Scheme of the algorithm of the resource management system

Сценарии экспериментов. Проведено три серии экспериментов:

1. Базовый сценарий – равномерное распределение ресурсов без учета кооперативных эффектов.

2. Сценарий с фиксированным значением Шепли – распределение ресурсов пропорционально заранее рассчитанным значениям Шепли (без ML-прогноза).

3. Полный сценарий – динамическое перераспределение ресурсов на основе актуальных прогнозов ML-модели и проверки условий ядра.

Каждый эксперимент повторялся 10 раз, продолжительность одного прогона составляла 30 минут. В ходе экспериментов варьировались размер коалиции и профиль нагрузки (CPU-интенсивная, память-интенсивная, смешанная).

Результаты

Динамика распределения ресурсов и значения Шепли. Ключевым показателем справедливости распределения является динамика значений Шепли для отдельных исполнителей в ходе непрерывной работы кластера. Рисунок 3 представляет временные ряды значений Шепли для одного из четырех исполнителей, участвовавших в эксперименте со смешанной нагрузкой.

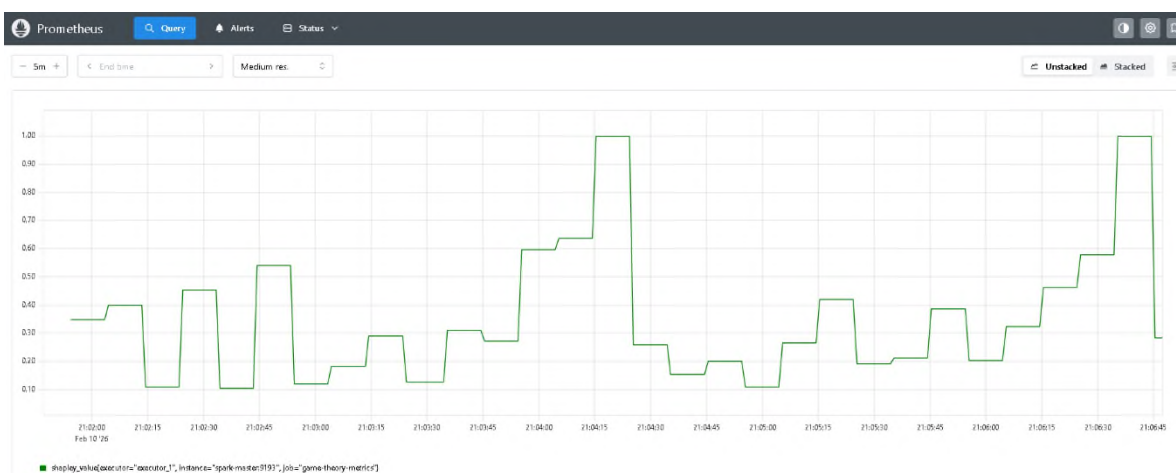


Рисунок 3 – Динамика значений Шепли для исполнителей в ходе 30-минутного эксперимента в Prometheus

Figure 3 – Dynamics of Shapley values for performers during a 30-minute experiment in Prometheus

На начальном этапе (0–5 мин) наблюдается существенный разброс значений: отдельные исполнители получают значительно более высокие оценки маржинального вклада. Это соответствует периоду «разогрева», когда профили нагрузки еще не устоялись, а ML-модель использует исходные, наименее точные прогнозы. К 10-й минуте разброс сокращается, а к 20-й минуте значения Шепли стабилизируются относительно друг друга.

Ключевое наблюдение: В полном сценарии (с ML-прогнозом) динамика значений Шепли демонстрирует более плавные переходы и меньшее количество резких выбросов по сравнению со сценарием фиксированного распределения. Это свидетельствует о том, что адаптивное перераспределение ресурсов на основе прогнозируемой полезности коалиций снижает колебания системы.

На Рисунке 4 представлено сравнение маржинальных вкладов исполнителей относительно их значений Шепли со смещением в 1 минуту.

Положительные маржинальные вклады на протяжении большей части эксперимента указывают на то, что исполнители постоянно увеличивают свою полезность для коалиции, что является признаком устойчивого кооперативного взаимодействия.

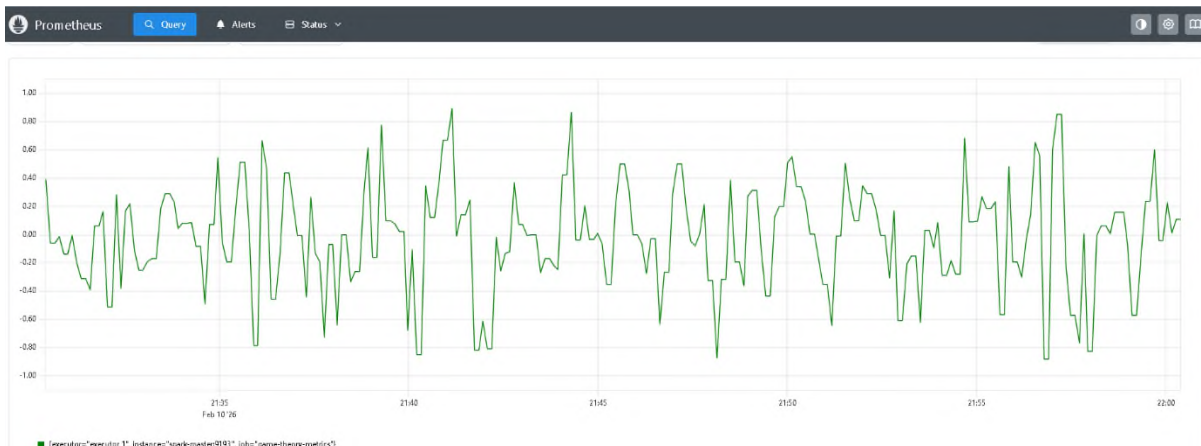


Рисунок 4 – Маржинальные вклады исполнителей (прирост значения Шепли за 1 минуту)
 Figure 4 – Marginal contributions of performers (Shapley value increase per 1 minute)

Анализ справедливости распределения. Для количественной оценки справедливости распределения ресурсов использованы два показателя: индекс справедливости (Fairness Index) и коэффициент Джинни. Рисунок 5 демонстрирует динамику индекса справедливости и скорость его изменения.



Рисунок 5 – Динамика индекса справедливости (верхний график) и скорость его изменения (нижний график)
 Figure 5 – Dynamics of the fairness index (upper graph) and the rate of its change (lower graph)

В полном сценарии коэффициент Джинни снижается с исходных 0,41 до 0,23 к концу эксперимента. Для сравнения: в базовом сценарии коэффициент сохраняется на уровне 0,38–0,42, в сценарии с фиксированным значением Шепли – 0,31–0,34. Полученные результаты подтверждают, что динамическое перераспределение ресурсов на основе актуальных прогнозов ML-модели позволяет значительно сократить неравенство между исполнителями.

Проверка условий равновесия (ядра игры). Центральным теоретическим результатом работы является демонстрация достижимости состояний, удовлетворяющих условиям ядра кооперативной игры, в реальной вычислительной системе. Рисунок 6 иллюстрирует проверку основного неравенства коалиционной рациональности.

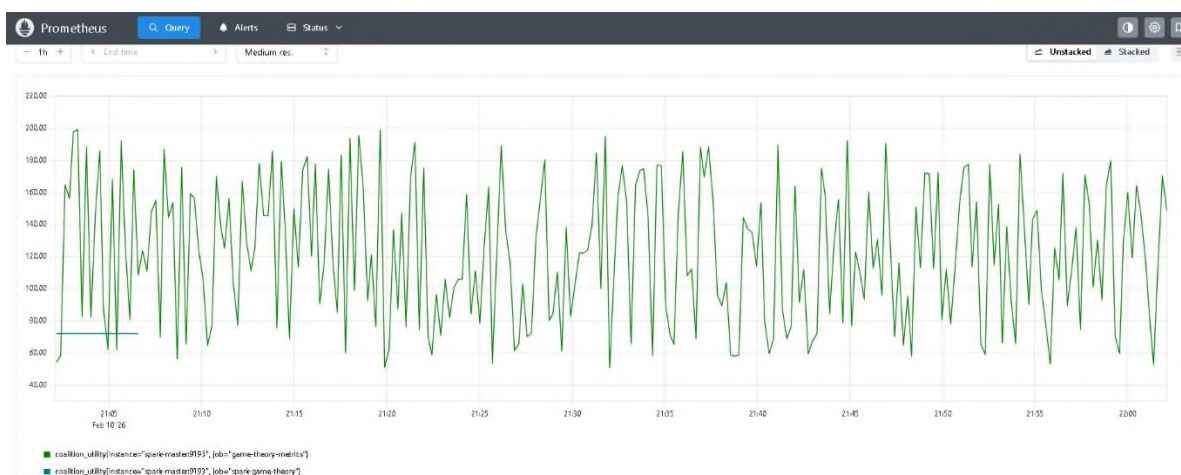


Рисунок 5 – Проверка условия ядра: совокупная полезность коалиции и сумма значений Шепли участников, умноженная на пороговый коэффициент 0,8

Figure 5 – Testing the core condition: the total utility of the coalition and the sum of the Shapley values of the participants, multiplied by the threshold coefficient of 0,8

В ходе эксперимента фиксировалось выполнение неравенства:

$$v(S) \geq 0,8 \cdot \sum_{i \in S} \varphi_i$$

для случайно выбранных коалиций размера 2 и 3.

На интервалах 4–8 мин и 16–26 мин неравенство выполняется устойчиво. Кратковременные нарушения на 10-й и 28-й минутах совпадают по времени с пиковыми изменениями нагрузки и пересмотром конфигурации ресурсов. Важно отметить, что в сценарии с фиксированным распределением доля времени выполнения неравенства составила 63 %, тогда как в полном сценарии – 89 %. Это свидетельствует о том, что динамическая коррекция распределения на основе ML-прогнозов не только улучшает справедливость, но и повышает устойчивость кооперативного соглашения.

В Таблице 1 приведены агрегированные результаты по трем сценариям экспериментов.

Полный сценарий демонстрирует превосходство по всем целевым метрикам. При этом более частое перераспределение ресурсов (12 эпизодов против 3 в сценарии с фиксированным Шепли) не приводит к дестабилизации системы, что подтверждается ростом индекса справедливости и снижением коэффициента вариации.

Таблица 1 – Сравнительная эффективность сценариев распределения ресурсов
Table 1 – Comparative efficiency of resource allocation scenarios

Метрика	Базовый сценарий	Фиксированное Шепли	Полный сценарий (ML + ядро)
Индекс справедливости	0,67	0,80	0,91
Коэффициент Джинни	0,40	0,32	0,24
Выполнение условия ядра, % времени	–	63 %	89 %
Эффективность CPU/Memory	1,08	1,21	1,42
Стабильность коалиции (CV)	0,48	0,38	0,28

Результаты экспериментального исследования позволяют сделать следующие выводы: подход работоспособен, в той части, что предложенная гибридная модель успешно реализована в прототипе, интегрированном с Apache Spark, и демонстрирует устойчивую работу в динамической среде; использование значения Шепли в качестве основы для распределения ресурсов позволяет существенно снизить неравенство между исполнителями. Интеграция ML-прогноза дополнительно повышает индекс справедливости до 0,91; в эксперименте зафиксированы продолжительные интервалы выполнения условий коалиционной рациональности, что подтверждает принципиальную возможность реализации состояний, близких к ядру игры, в реальной вычислительной системе; динамическое перераспределение на основе ML-прогнозов позволяет увеличить отдачу от единицы ресурса: отношение CPU/Memory выросло на 31 % относительно базового сценария; более частое перераспределение ресурсов в полном сценарии не приводит к деградации производительности или нестабильности, что подтверждает практическую применимость подхода.

Заключение

В настоящей работе представлено экспериментальное исследование гибридного подхода к кооперативному распределению ресурсов в системах обработки больших данных, основанного на сочетании формальных методов теории кооперативных игр и прогностических возможностей машинного обучения. В ходе исследования были решены все поставленные задачи: выполнена математическая формализация оптимизационной задачи распределения ресурсов, сформулированы условия равновесия коалиций, разработана и реализована архитектура прототипа на базе Apache Spark, проведена серия вычислительных экспериментов и выполнен детальный анализ полученных результатов.

Основные научные и практические результаты работы могут быть суммированы следующим образом.

1. Формализация кооперативной модели распределения ресурсов. Предложена математическая постановка задачи, в которой взаимодействие исполнителей рассматривается как кооперативная игра с характеристической функцией, отражающей совокупную полезность коалиции. Сформулированы условия принадлежности распределения ядру игры, гарантирующие устойчивость кооперативного соглашения.

Обосновано применение значения Шепли как механизма справедливого распределения выигрыша, пропорционального маргинальному вкладу каждого участника.

2. Интеграция машинного обучения в теоретико-игровую модель. Продemonстрировано, что использование ML-моделей (градиентный бустинг, графовые нейронные сети) для аппроксимации характеристической функции позволяет преодолеть ограничения аналитического подхода и эффективно оценивать полезность коалиций в условиях реального кластера с учетом множества нелинейных зависимостей. Предложенная схема взаимодействия модуля принятия решений и прогностической модели обеспечивает адаптивность системы к изменяющимся профилям нагрузки.

3. Экспериментальная верификация подхода. Разработан и развернут воспроизводимый экспериментальный стенд на базе Apache Spark, интегрированный с системой мониторинга Prometheus/Grafana и модулями вычисления значения Шепли и ML-прогнозирования. В ходе серии экспериментов подтверждена работоспособность предложенного подхода в динамической среде.

4. Достижение справедливости и устойчивости распределения. Количественный анализ показал, что в полном сценарии (динамическое перераспределение на основе ML-прогноза и проверки ядра) индекс справедливости достигает значений 0,91, а коэффициент Джинни снижается до 0,24. Это существенно превосходит показатели базового сценария (0,67 и 0,40 соответственно) и сценария с фиксированным значением Шепли (0,80 и 0,32). Продолжительность выполнения условий коалиционной рациональности в полном сценарии составила 89 % времени эксперимента, что подтверждает принципиальную возможность достижения состояний, близких к ядру кооперативной игры, в реальной вычислительной системе.

5. Повышение эффективности использования ресурсов. Установлено, что динамическое перераспределение на основе прогнозов ML-модели позволяет увеличить отношение CPU/Memory на 31 % относительно базового сценария. Система автоматически выявляет конфигурации, обеспечивающие максимальный прирост полезности на единицу ресурса, и перераспределяет выделяемые квоты в пользу наиболее эффективных стратегий.

6. Стабильность коалиций. Снижение коэффициента вариации значений Шепли с 0,48 в базовом сценарии до 0,28 в полном сценарии свидетельствует о формировании сбалансированных и устойчивых коалиций, что снижает стимулы к выходу отдельных исполнителей из совместного пула ресурсов.

Научная значимость проведенного исследования заключается в комплексном применении аппарата кооперативных игр и методов машинного обучения к задаче оперативного распределения ресурсов в распределенных системах, а также в экспериментальном подтверждении теоретических положений о достижимости ядра игры в условиях реальной вычислительной среды. Полученные результаты вносят вклад в развитие направлений интеллектуального управления ресурсами, теоретико-игрового моделирования вычислительных процессов и адаптивных систем планирования.

Практическая значимость работы состоит в создании действующего прототипа системы управления ресурсами, который может служить основой для разработки планировщиков нового поколения в экосистемах обработки больших данных. Предложенные алгоритмы и архитектурные решения обладают достаточной общностью для их адаптации к другим распределенным вычислительным платформам, включая Apache Flink, Kubernetes и облачные среды. Воспроизводимость экспериментального стенда обеспечивает возможность независимой верификации полученных результатов и их дальнейшего развития.

Ограничения исследования. В ходе работы был выявлен ряд ограничений, которые следует учитывать при интерпретации результатов и планировании дальнейших исследований. Во-первых, точное вычисление значения Шепли требует перебора всех подмножеств коалиций, что при количестве исполнителей более 10 становится вычислительно неразрешимым. В рамках настоящего исследования данное ограничение обходилось путем ограничения размера активной коалиции, однако для масштабирования подхода на крупные кластеры необходима разработка аппроксимационных методов вычисления значения Шепли. Во-вторых, качество прогнозов ML-модели критически зависит от полноты и репрезентативности обучающей выборки. В начальный период работы системы, когда исторические данные еще не накоплены, точность аппроксимации характеристической функции может быть недостаточной. В-третьих, исследование проводилось в контролируемой среде с использованием синтетических и полунатурных нагрузок; для подтверждения эффективности подхода в промышленных условиях требуется дополнительная валидация.

Направления дальнейших исследований. Полученные результаты открывают несколько перспективных направлений для продолжения работы. Первое направление связано с разработкой и исследованием методов аппроксимации значения Шепли, позволяющих применять предложенный подход в кластерах с большим числом одновременно выполняемых задач. Перспективными представляются методы на основе семплирования, аппроксимации с использованием ML-моделей и методы, использующие специфику структуры вычислительных графов. Второе направление предполагает углубленное исследование применимости графовых нейронных сетей для аппроксимации характеристической функции, учитывающих топологию взаимодействий между задачами. Третье направление связано с интеграцией предложенного подхода в механизмы автоматического масштабирования и эластичности ресурсов в облачных средах. Наконец, перспективным представляется исследование возможности применения методов глубокого обучения с подкреплением для прямого поиска оптимальных стратегий распределения ресурсов без явного вычисления значения Шепли.

Заключительное положение. Проведенное исследование демонстрирует, что синтез теоретико-игровых принципов кооперации и адаптивных возможностей машинного обучения является не только теоретически обоснованным, но и практически реализуемым подходом к управлению ресурсами в современных системах обработки больших данных. Полученные результаты подтверждают, что справедливое, устойчивое и эффективное распределение ресурсов достижимо даже в условиях высокой динамики и неопределённости. Авторы выражают надежду, что представленные результаты будут полезны как для исследователей, работающих в области интеллектуальных вычислительных систем, так и для практиков, занимающихся разработкой и эксплуатацией распределенных платформ обработки данных.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Бляхеров М.В., Петрова Е.С. Теоретико-игровые модели координации ресурсов в распределённых системах потокового анализа данных. *Моделирование, оптимизация и информационные технологии*. 2025;13(4). <https://doi.org/10.26102/2310-6018/2025.51.4.068>
 Blyakherov M.V., Petrova E.S. Game-theoretic models of resource coordination in distributed streaming data analysis systems. *Modeling, Optimization and Information Technology*. 2025;13(4). (In Russ.). <https://doi.org/10.26102/2310-6018/2025.51.4.068>

2. Kim S. Cooperative Game-Based Virtual Machine Resource Allocation Algorithms in Cloud Data Centers. *Mobile Information Systems*. 2020;2020. <https://doi.org/10.1155/2020/9840198>
3. Bertossi L., Kimelfeld B., Livshits E., Monet M. *The Shapley Value in Database Management*. arXiv. URL: <https://arxiv.org/abs/2401.06234> [Accessed 26th December 2025].
4. Shapley L.S. A Value for n -Person Games. In: *Contributions to the Theory of Games: Volume II*. Princeton: Princeton University Press; 1953. P. 307–317. <https://doi.org/10.1515/9781400881970-018>
5. Colini-Baldeschi R., Scarsini M., Vaccari S. *Variance Allocation and Shapley Value*. arXiv. URL: <https://arxiv.org/abs/1606.09424> [Accessed 26th December 2025].
6. Zhou J., Wen G., Lv Y., Yang T., Chen G. *DRAG: Distributed Resource Allocation Games over Multiple Interacting Coalitions*. arXiv. URL: <https://arxiv.org/abs/2210.02919> [Accessed 26th December 2025].
7. Cardellini V., De Nitto Personé V., Di Valerio V., et al. A Game-Theoretic Approach to Computation Offloading in Mobile Cloud Computing. *Mathematical Programming*. 2016;157(2):421–449. <https://doi.org/10.1007/s10107-015-0881-6>
8. Разумовский Д.А., Волков Д.Д., Стучилин В.В. Архитектура системы сбора и хранения метрик использования ресурсов Spark-приложений в кластерных системах обработки больших данных. *Международный научно-исследовательский журнал*. 2025;(12). <https://doi.org/10.60797/IRJ.2025.162.81>
Razumovskii D.A., Volkov D.D., Stuchilin V.V. Architecture of a system for collecting and storing metrics on the resource usage of Spark applications in clustered big data processing systems. *International Research Journal*. 2025;(12). (In Russ.). <https://doi.org/10.60797/IRJ.2025.162.81>
9. Elradi M.D. Prometheus and Grafana: A Metrics-focused Monitoring Stack. *Journal of Computer Allied Intelligence*. 2025;3(3):28–39. <https://doi.org/10.69996/jcai.2025015>
10. Mehdi A., Bali M.K., Abbas S.I., Singh M. Unleashing the Potential of Grafana: A Comprehensive Study on Real-Time Monitoring and Visualization. In: *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 06–08 July 2023, Delhi, India*. IEEE; 2023. <https://doi.org/10.1109/ICCCNT56998.2023.10306699>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Бляхеров Михаил Викторович, аспирант,
Воронежский институт высоких технологий,
Воронеж, Российская Федерация.
e-mail: mikhail.blyakherov@gmail.com

Mikhail V. Blyakherov, Postgraduate, Voronezh
Institute of High Technologies, Voronezh, the
Russian Federation.

Львович Игорь Яковлевич, доктор
технических наук, профессор, ректор
Воронежского института высоких технологий,
Воронеж, Российская Федерация.
e-mail: i-ya-lvovich@vvt.ru

Igor Ya. Lvovich, Doctor of Engineering
Sciences, Professor, Rector of the Voronezh
Institute of High Technologies, Voronezh, the
Russian Federation.

*Статья поступила в редакцию 15.03.2026; одобрена после рецензирования 17.04.2026;
принята к публикации 18.05.2026.*

*The article was submitted 15.03.2026; approved after reviewing 17.04.2026;
accepted for publication 18.05.2026.*