

УДК 005.52:005.334.1

DOI: [10.26102/2310-6018/2026.56.5.006](https://doi.org/10.26102/2310-6018/2026.56.5.006)

Двухуровневая стохастико-адаптивная модель управления эксплуатационными рисками объектов розничной сети

М.Г. Устимов✉, И.Я. Львович

Воронежский институт высоких технологий, Воронеж, Российская Федерация

Резюме. В статье представлена инновационная двухуровневая стохастико-адаптивная модель управления эксплуатационными рисками, предназначенная для крупных распределенных инфраструктурных сетей. Исследование решает проблему неспособности традиционных детерминированных моделей адекватно оценивать «хвостовые» риски в условиях высокой неопределенности энергопотребления, отказов оборудования и логистических сбоев. Предлагаемая методология объединяет стратегическое планирование и тактическую онлайн-адаптацию. На верхнем уровне используется двухэтапное стохастическое программирование для формирования робастных планов обслуживания и резервирования мощностей, учитывающих вероятностный характер угроз. Интеллектуальная кластеризация объектов с помощью самоорганизующихся карт Кохонена позволяет разделить сеть на четыре категории: критические, высокого риска, логистически уязвимые и стабильные. На нижнем уровне агенты с обучением с подкреплением (алгоритмы PPO и DQN) адаптируют эксплуатационные решения в реальном времени, используя кастомизированные функции вознаграждения для каждого кластера. Экспериментальные результаты подтверждают эффективность подхода: для критических объектов доля простоев снижена до 2 %, а для стабильных достигнута максимальная экономия ресурсов. Внедрение модели позволяет снизить эксплуатационные затраты на 10–15 % и повысить надежность критической инфраструктуры на 20–30 %. Модель обеспечивает прозрачность управления через четкие KPI и способствует реализации стратегии устойчивого развития.

Ключевые слова: мультимодальный анализ данных, семантическое выравнивание, медицинская диагностика, обучение с подкреплением, распределенные вычисления.

Для цитирования: Устимов М.Г., Львович И.Я. Двухуровневая стохастико-адаптивная модель управления эксплуатационными рисками объектов розничной сети. *Моделирование, оптимизация и информационные технологии*. 2026;14(5). URL: <https://moitvvt.ru/ru/journal/article?id=2227> DOI: 10.26102/2310-6018/2026.56.5.006

A two-level stochastic-adaptive model for managing the operational risks of retail network facilities

M.G. Ustimov✉, I.Ya. Lvovich

Voronezh Institute of High Technologies, Voronezh, the Russian Federation

Abstract. The article presents an innovative two-level stochastic-adaptive operational risk management model designed for large distributed infrastructure networks. The study solves the problem of the inability of traditional deterministic models to adequately assess the "tail" risks in conditions of high uncertainty of energy consumption, equipment failures and logistical failures. The proposed methodology combines strategic planning and tactical online adaptation. At the top level, two-stage stochastic programming is used to generate robust maintenance and capacity redundancy plans that take into account the probabilistic nature of threats. Intelligent clustering of objects using self-organizing Kohonen maps allows you to divide the network into four categories: critical, high-risk, logistically vulnerable and stable. At the lower level, reinforcement learning agents (PPO and DQN algorithms) adapt operational solutions in real time using customized reward functions for each cluster.

Experimental results confirm the effectiveness of the approach: for critical facilities, the share of downtime has been reduced to 2 %, and for stable facilities, maximum resource savings have been achieved. The implementation of the model makes it possible to reduce operating costs by 10–15% and increase the reliability of critical infrastructure by 20–30%. The model ensures transparency of management through clear KPIs and contributes to the implementation of a sustainable development strategy.

Keywords: multimodal data analysis, semantic alignment, medical diagnostics, reinforcement learning, distributed computing.

For citation: Ustimov M.G., Lvovich I. Ya. A two-level stochastic-adaptive model for managing the operational risks of retail network facilities. *Modeling, Optimization and Information Technology*. 2026;14(5). (In Russ.). URL: <https://moitvivr.ru/ru/journal/article?id=2227> DOI: 10.26102/2310-6018/2026.56.5.006

Введение

Современный этап развития крупных распределенных инфраструктурных сетей характеризуется беспрецедентным усложнением процессов их эксплуатации. Это связано с необходимостью обеспечения баланса между операционной эффективностью, энергетической устойчивостью и минимизацией эксплуатационных рисков в условиях высокой неопределенности внешней среды [1]. Традиционные подходы к управлению, основанные на унифицированных регламентах и детерминированных моделях планирования, оперирующих средними значениями, демонстрируют ограниченную эффективность. Они не учитывают существенную неоднородность риск-профилей отдельных объектов и оказываются неспособными адекватно оценить и подготовиться к реализации маловероятных, но катастрофических «хвостовых» рисков, что ставит под угрозу живучесть всей системы.

Актуальность исследования обусловлена растущей потребностью в интеллектуальных системах адаптивного управления, способных в реальном времени оптимизировать использование ресурсов, реагировать на волатильность энергопотребления, риски отказов оборудования и логистические сбои. В научной литературе методы обучения с подкреплением и стохастического программирования рассматриваются как перспективные инструменты для решения этих задач по отдельности. Однако наблюдается дефицит комплексных решений, которые интегрировали бы стратегическое, вероятно обоснованное планирование, устойчивое к неопределенности, с тактической, гибкой онлайн-адаптацией операционных политик для каждого объекта сети [2].

Проблема исследования заключается в необходимости преодоления противоречия между централизованным контролем рисков на уровне всей распределенной сети и потребностью в децентрализованной, контекстно-зависимой адаптации управления на уровне каждого ее элемента. Данная работа направлена на решение этой проблемы через разработку иерархической двухуровневой стохастико-адаптивной модели. Целью статьи является формализация и экспериментальная проверка такой системы, в которой верхний (стратегический) уровень использует двухэтапное стохастическое программирование для определения рамочных политик, а нижний (тактический) уровень задействует RL-агентов для тонкой настройки решений в реальном времени в рамках установленных стратегических лимитов. Научная новизна подхода состоит в интеллектуальной кластеризации объектов для дифференциации стратегий управления и кастомизации алгоритмов обучения в зависимости от категории риска, что обеспечивает адресное и экономически эффективное управление эксплуатацией сложных инфраструктурных активов.

Материалы и методы

Разработка адаптивной системы управления эксплуатационными рисками распределенной инфраструктуры базируется на создании иерархической двухуровневой архитектуры, которая объединяет методы математического программирования и искусственного интеллекта. Необходимость такого подхода продиктована тем, что традиционные детерминированные модели, оперирующие средними значениями, не способны адекватно оценивать «хвостовые» риски, что критично для обеспечения живучести крупных сетей.

Общая архитектура системы представлена в виде схемы (Рисунок 1), которая наглядно демонстрирует взаимодействие стратегического и тактического контуров управления, объединенных в замкнутый адаптивный цикл.

Предложенная методология интегрирует двухэтапное стохастическое программирование (ДЭСП) на верхнем (стратегическом) уровне и глубокое обучение с подкреплением (Reinforcement Learning, RL) на нижнем (тактическом) уровне.

Стратегический уровень. Двухэтапное стохастическое планирование. На верхнем уровне с горизонтом планирования от квартала до года решается задача формирования робастной эксплуатационной политики для различных кластеров объектов, полученных в результате интеллектуальной сегментации (например, методом самоорганизующихся карт Кохонена) [3].

Пусть для кластера объектов (например, «Критический») необходимо определить долгосрочную стратегию эксплуатации, характеризуемую вектором решений первого этапа $x = (x_1, \dots, x_n)^T$. Эти решения принимаются до того, как стала известна реализация случайных параметров $w \in \Omega$. Компоненты x_j могут представлять: уровень резервной энергомощности (кВт), периодичность плановых технических обслуживаний (ТО), объем контрактных гарантий с логистическими поставщиками, нормативный запас ключевых запасных частей [4].

Случайные параметры ω моделируют неопределенность внешней и внутренней среды: $b(w)$ – пиковая нагрузка на энергосети, $d(w)$ – интенсивность отказов оборудования, $c(w)$ – тарифы на экстренный ремонт и энергию, $q(w)$ – доступность логистических маршрутов.

Затраты на реализацию стратегии x складываются из:

1. Детерминированных затрат первого этапа $c \in \mathbb{R}^n$, где c – вектор удельных стоимостей (капитальные вложения, стоимость планового ТО).
2. $y(w) \in \mathbb{R}_+^m$ – вектор решений второго этапа (рекурсивные действия).
3. $T(w) \in \mathbb{R}^{p \times n}$, $W \in \mathbb{R}^{p \times m}$, $b(w) \in \mathbb{R}^p$ – матрицы и вектор технологических связей.
4. Ожидаемых затрат второго этапа $E_\omega[Q(x, \omega)]$, которые возникают для коррекции стратегии x после наблюдения конкретной реализации ω .

Рекурсивная задача второго этапа при фиксированных x и ω имеет вид:

$$Q(x, w) = \min_{y(w) \geq 0} q(w)^T y(w), \quad (1)$$

при условиях

$$T(w)x + Wy(w) \geq b(w), \quad (2)$$

где $y = y(\omega)$ – решения второго этапа (компенсирующие управляющие воздействия): запуск резервных генераторов, экстренный выезд ремонтной бригады, заказ доставки по альтернативному маршруту; $q(\omega)$ – вектор удельных затрат на такие компенсирующие действия [5]. Матрицы $T(\omega)$ и W , вектор $b(\omega)$ задают технологические связи между стратегическими решениями x , компенсирующими действиями y и требуемым уровнем обеспечения функционирования при сценарии ω .

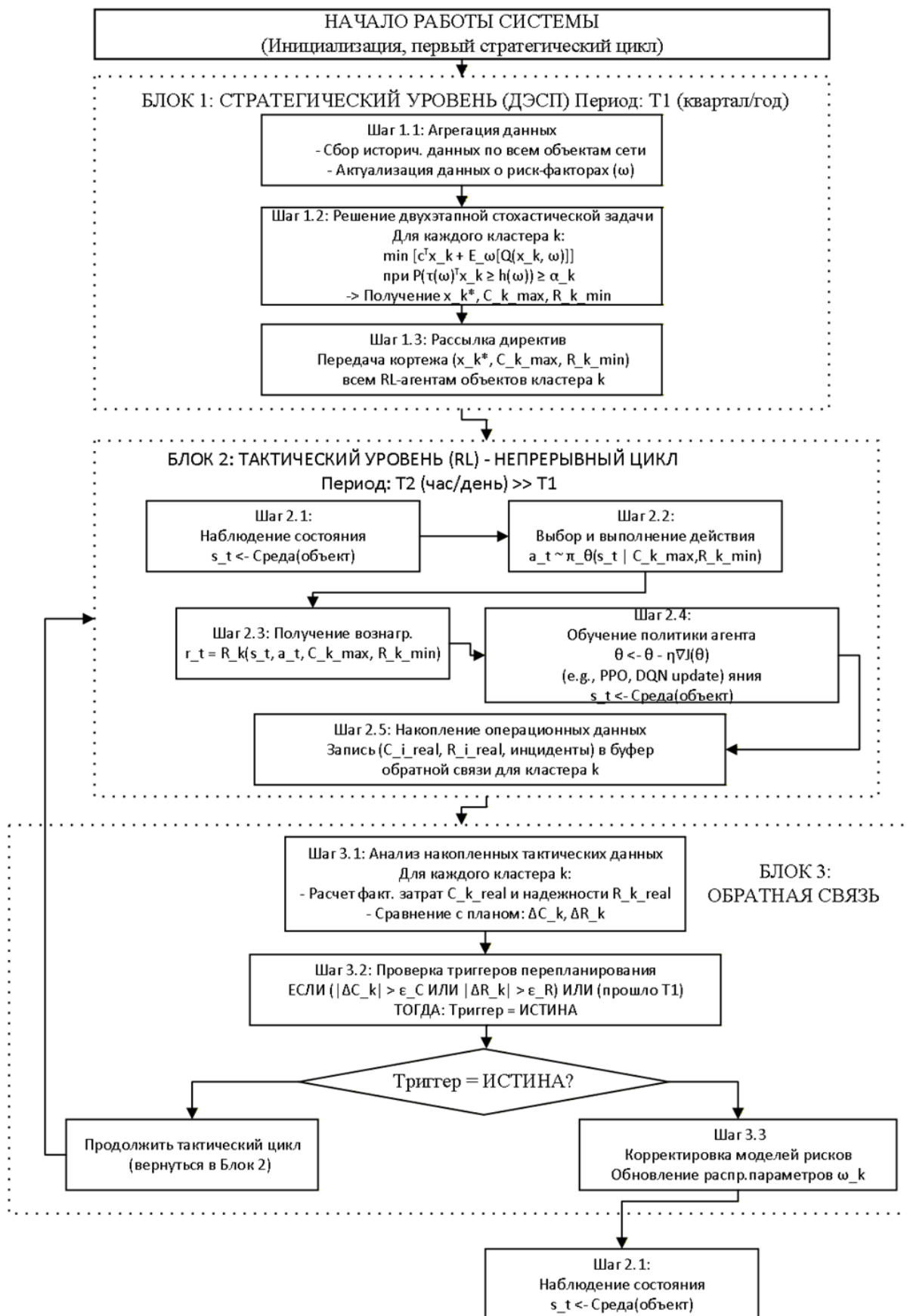


Рисунок 1 – Двухуровневая стохастико-адаптивная система управления
Figure 1 – Two-level stochastic adaptive control system

Общая постановка двухэтапной стохастической задачи для кластера объектов имеет вид:

$$\min_x Z(x) = c^T x + E_w[Q(x, w)], \quad (3)$$

при условиях: $Ax \leq q, x \geq 0$, где $A \in \mathbb{R}^{k \times n}, q \in \mathbb{R}^k$ задают детерминированные бюджетные и ресурсные ограничения.

Для учета приемлемого уровня риска вводятся вероятностные (шансовые) ограничения:

$$P(\tau(w)^T x \geq h(w)) \geq \alpha, \quad (4)$$

где $\tau(w)$ и $h(w)$ описывают целевой показатель (например, уровень надежности), а $\alpha \in (0,1)$ – заданный допустимый уровень риска.

Ограничение (3) означает: «Вероятность того, что выбранная стратегия x обеспечит выполнение целевого показателя $h(w)$ (например, бесперебойность энергоснабжения), должна быть не менее α ». Параметр α (например, 0,95 или 0,99) задается для каждого кластера в соответствии с его риск-профилем, определенным на этапе кластеризации.

Тактический уровень. Адаптивное обучение с подкреплением. Нижний уровень с горизонтом от часа до дня, как показано на Рисунке 1, отвечает за тонкую настройку эксплуатационных политик в реальном времени. Здесь функционируют RL-агенты, задача которых – оптимизировать ежедневные операционные решения в рамках стратегических директив, переданных сверху.

Стохастическая модель задает рамочные условия для тактического RL-агента. Интеграция осуществляется через модифицированную функцию вознаграждения и пространство допустимых действий агента [6].

Пусть для объекта, принадлежащего кластеру k , стохастической моделью определены: целевой уровень надежности R_k^{min} (выведенный из вероятностного ограничения). Лимит ожидаемых эксплуатационных затрат C_k^{max} (результат решения задачи (1)–(3)). Базовые стратегические директивы x_k^* (оптимальный вектор первого этапа).

Тогда функция вознаграждения RL-агента $R_k(s, a)$ на такте t дополняется штрафом за отклонение от стратегических ориентиров:

$$R_k(s_t, a_t) = R_k^{orig}(s_t, a_t) - \lambda_1 \max\{0, C_k^{acc}(t) - C_k^{max}\} - \lambda_2 \max\{0, R_k^{min} - R_k^{curr}(t)\} \quad (5)$$

где $R_k^{orig}(s_t, a_t)$ – базовая функция вознаграждения из раздела, $C_k^{acc}(t)$ – накопленные фактические затраты объекта к такту t . $R_k^{curr}(t)$ – текущий расчётный уровень надёжности объекта (например, 1 – (время простоя / общее время)). λ_1, λ_2 – коэффициенты штрафа, калибруемые в зависимости от приоритетов кластера.

Пространство допустимых действий A_k агента может быть сужено действиями, противоречащими стратегической директиве x_k^* . Например, для кластера «Критический» агенту может быть запрещено отключать резервные системы питания в целях экономии.

Для обучения используются специализированные алгоритмы: PPO (Proximal Policy Optimization) для объектов, где критична стабильность, и DQN (Deep Q-Network) для остальных.

Цикл адаптации, обратной связи и работа с данными. Согласно Рисунку 1, система функционирует как замкнутый контур. По итогам тактических операций собирается статистика реальных затрат и инцидентов. Ввиду чувствительности реальных операционных данных в рамках исследования была применена методика синтеза данных.

Для проведения валидации модели был сформирован репрезентативный синтетический датасет, включающий виртуальные объекты, каждый из которых описывался комплексом факторов: геопространственных, инфраструктурных, операционных и социально-экономических [7]. Это позволило смоделировать неоднородность риск-профилей и обеспечить воспроизводимость экспериментов.

Собранные (или синтезированные) данные используются для рекурсивного уточнения распределений случайных параметров ω в стохастической модели (обратная связь «Данные и статистика» на Рисунке 1). При значительном расхождении фактических показателей с плановыми инициируется внеочередной цикл стратегического перепланирования. Таким образом, система самообучается на двух уровнях, что позволяет фокусировать RL-агентов на поиске эффективности только внутри стратегически обоснованных и безопасных рамок.

Экспериментальная часть. Экспериментальное исследование было проведено с целью валидации предложенной двухуровневой стохастико-адаптивной архитектуры и оценки ее эффективности в условиях, приближенных к реальным. Для обеспечения воспроизводимости, изолированности и масштабируемости все вычислительные эксперименты были развернуты в контейнеризованной среде.

В качестве платформы развертывания выбрана изолированная среда Docker, управляемая через WSL2 в ОС Windows [8]. Такой подход гарантирует идентичность среды выполнения на любом этапе – от разработки и обучения до финального тестирования, – а также упрощает передачу и воспроизведение результатов.

Программная реализация выполнена на языке Python 3.9 с использованием следующего стека технологий: Stable-Baselines3 и Gymnasium – для построения, обучения и валидации RL-агентов (алгоритмы PPO и DQN); MiniSom – для реализации алгоритма самоорганизующихся карт Кохонена (SOM) на стратегическом уровне кластеризации; Dash/Plotly – для создания интерактивного веб-дашборда мониторинга, отображения результатов в реальном времени и визуализации метрик обучения [9].

Конфигурация проекта, включая Dockerfile и docker-compose.yml, обеспечивает полный жизненный цикл системы: от подготовки и синтеза данных до обучения моделей, визуализации и формирования отчетов. На Рисунке 2 представлен скриншот успешного запуска Docker-проекта.

```

kost@DESKTOP-TPNS038:~/projects/stochastic-rl-system/dash_app$ docker-compose logs -f dash
stochastic_dash | INFO: __main__:=====
stochastic_dash | INFO: __main__: Запуск Dash приложения с RL-вкладкой
stochastic_dash | INFO: __main__: Доступно по адресу: http://0.0.0.0:8050
stochastic_dash | INFO: __main__:=====
stochastic_dash | Dash is running on http://0.0.0.0:8050/
stochastic_dash | INFO: dash.dash: Dash is running on http://0.0.0.0:8050/
stochastic_dash | * Serving Flask app 'app'
stochastic_dash | * Debug mode: off
stochastic_dash | INFO: werkzeug:WARNING: This is a development server. Do not use it in a production deployment. Use a production WS
GI server instead.
stochastic_dash | * Running on all addresses (0.0.0.0)
stochastic_dash | * Running on http://127.0.0.1:8050
stochastic_dash | * Running on http://172.18.0.3:8050
stochastic_dash | INFO: werkzeug:Press CTRL+C to quit
    
```

Рисунок 2 – Окно логов успешного запуска Docker-проекта и веб-сервера Dash
 Figure 2 – Log window for a successful launch of a Docker project and the Dash web server

После успешного запуска веб-интерфейс системы становится доступен по локальному адресу.

В рамках экспериментального прототипа, развернутого в Docker-среде, для большей наглядности интерфейса управления и сокращения вычислительных затрат на этапе валидации архитектуры, было проведено моделирование для трех

репрезентативных кластеров: «Критический», «Экономный» (объединяющий логику кластеров «Стабильный» и частично «Логистически уязвимый» с приоритетом минимизации затрат) и «Стабильный» [10]. Данное упрощение не затрагивает общность предложенной двухуровневой архитектуры, но позволяет наглядно продемонстрировать в интерфейсе принцип дифференциации стратегий управления. На Рисунке 3 представлен дашборд оперативного управления в веб-интерфейсе системы (пример для кластера «Экономный») представлено главное окно разработанного приложения, включающее: выбор кластера для анализа, панель с расчетными стратегическими KPI (целевая и фактическая надёжность, нагрузка, затраты), а также динамические графики ключевых операционных показателей (энергопотребление, надёжность, затраты, эффективность) в реальном времени.

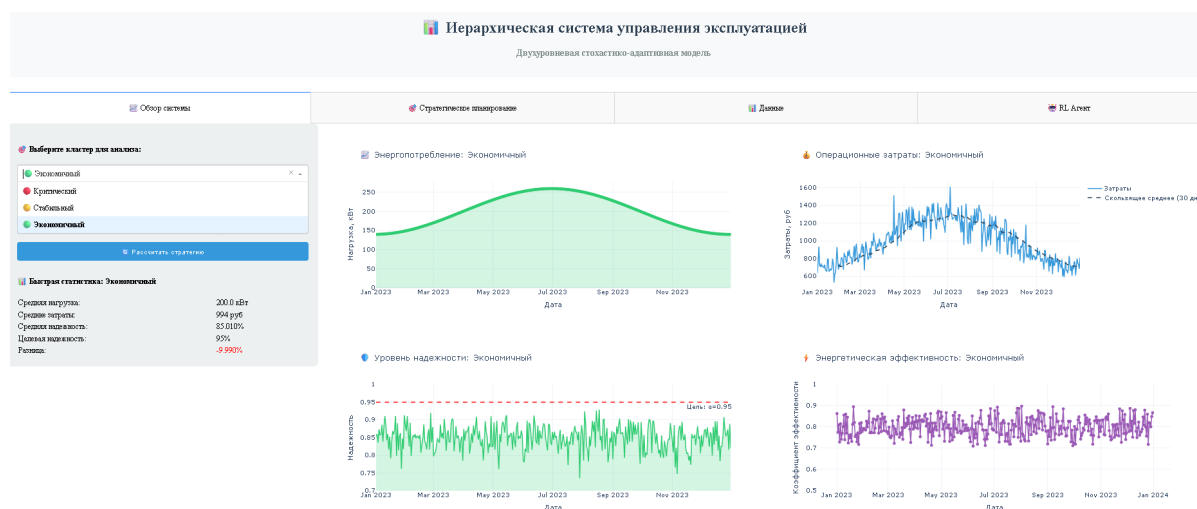


Рисунок 3 – Дашборд оперативного управления в веб-интерфейсе системы (пример для кластера «Экономный»)

Figure 3 – Operational Management Dashboard in the system's web interface (example for the "Economy" cluster)

Поскольку работа с реальными операционными данными крупных сетей сопряжена с ограничениями конфиденциальности и доступности, для экспериментов был синтезирован репрезентативный датасет, включающий 100 виртуальных объектов. Каждый объект описывался вектором признаков, имитирующих реальные эксплуатационные условия: геопространственные параметры: удаленность от зон операционной напряженности; инфраструктурные показатели: условная надежность энергосетей, состояние инженерных систем; операционные метрики – частота инцидентов, логистическая устойчивость; социально-экономические индикаторы – кадровая обеспеченность, влияющая на скорость реагирования.

Для каждого из кластеров была настроена специализированная среда обучения с горизонтом в 1000 шагов. Процесс адаптации строился на двух различных алгоритмах обучения с подкреплением, выбранных исходя из приоритетов кластера:

Алгоритм PPO (Proximal Policy Optimization) применялся для объектов «Критического» кластера. Его выбор обусловлен стабильностью и устойчивостью к высоким штрафам, что критично для минимизации простоев и обеспечения максимальной надежности.

Алгоритм DQN (Deep Q-Network) использовался для остальных кластеров «Высокого риска», «Логистически уязвимый» и «Стабильный», где выше требований к скорости обучения и эффективности исследования пространства действий.

Функции вознаграждения были кастомизированы. Например, для критических объектов штраф за простой (1000 усл. ед.) значительно превышал затраты на энергию, что заставляло агента чаще выбирать действия, гарантирующие бесперебойность, даже ценой повышенных расходов.

Прогресс обучения отслеживался через кривые вознаграждения и потерь. Агенты демонстрировали стабильную сходимость, выходя на плато производительности после 400–740 эпизодов, что подтверждает эффективность выбранных гиперпараметров и архитектуры обучения. Итоговые отчеты по каждому агенту, включая распределение выбранных действий, интегрированы во вкладку «RL Агент» управляющего приложения (Рисунок 4).

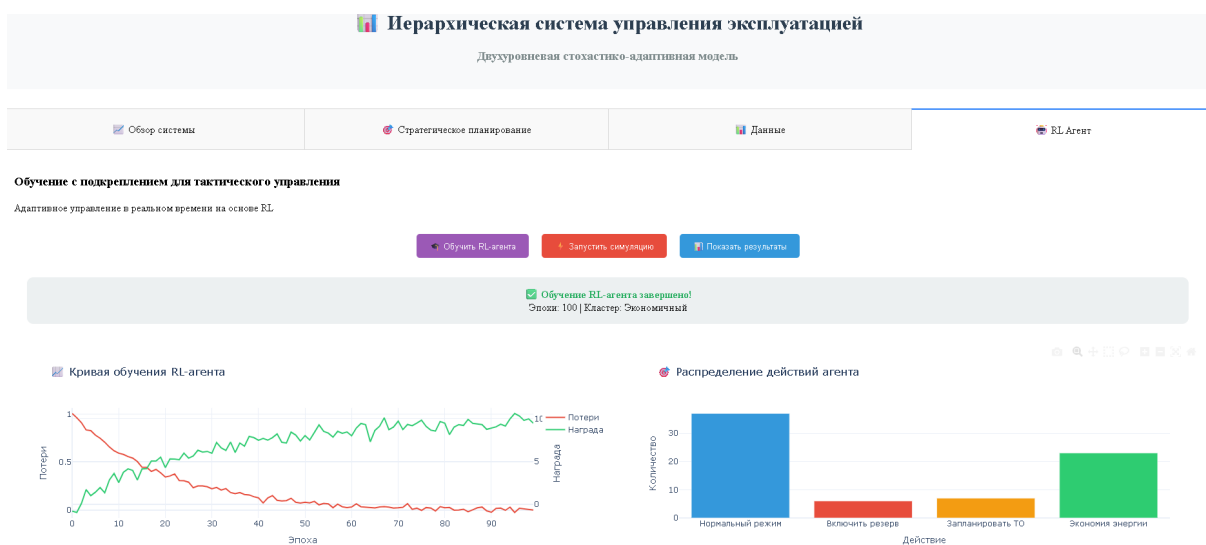


Рисунок 4 – Интерфейс мониторинга обучения RL-агентов
Figure 4 – RL agent training monitoring interface

Разработанное на базе Dash веб-приложение предоставляет менеджеру комплексный инструмент для мониторинга и анализа работы системы в реальном времени.

Интерфейс включает несколько ключевых вкладок:

- вкладка «Стратегическое планирование» отображает результаты решения двухэтапной стохастической задачи: распределение ожидаемых затрат по сценариям, расчёт премии за риск и установленные лимиты для каждого кластера (Рисунок 5);
- вкладка «Данные» предоставляет детализированную таблицу с динамическими показателями нагрузки, затрат и надежности по каждому виртуальному объекту;
- вкладка «RL Агент», как показано выше, позволяет в режиме реального времени наблюдать за процессом обучения, анализировать политику агента и оценивать ее эффективность.

На Рисунке 5 представлена панель веб-интерфейса, отображающая итоги решения двухэтапной стохастической задачи (ДЭСП) на стратегическом уровне. В центре расположена гистограмма распределения ожидаемых эксплуатационных затрат по различным вероятностным сценариям. Рядом с графиком приведены ключевые стратегические директивы, рассчитанные моделью: лимит резервной мощности (40 кВт), периодичность планового обслуживания (90 дней), базовые и удельные затраты, а также

целевые показатели надежности (95 %). В нижней части панели показан итоговый финансовый результат: ожидаемые затраты (934 руб.) и сумма с учетом премии за риск (1,093 руб.), что наглядно демонстрирует стоимостную оценку дополнительных гарантий надежности.

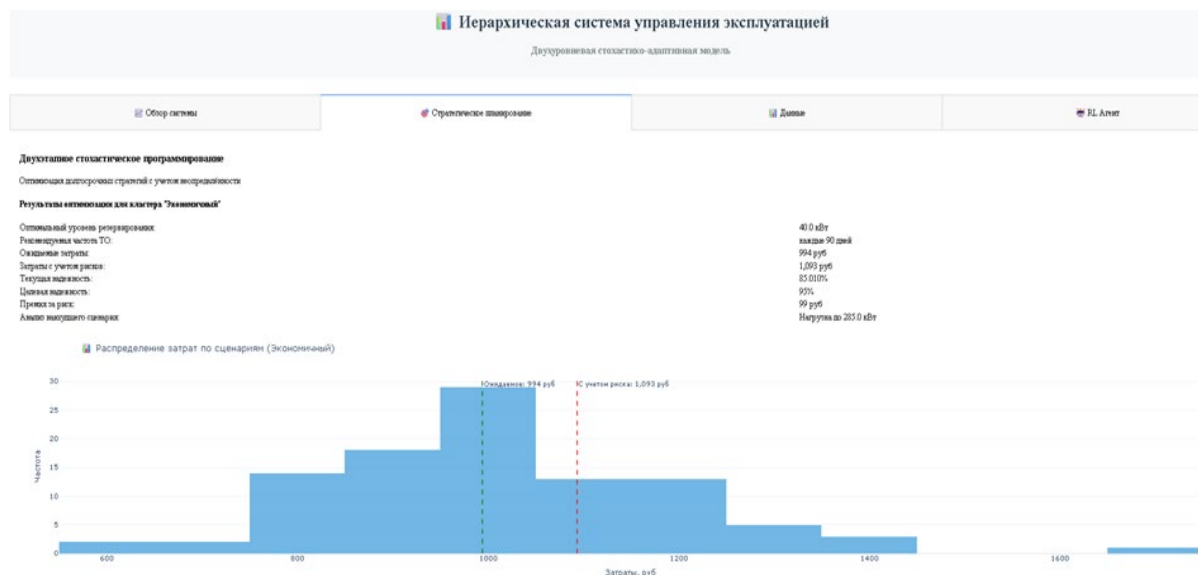


Рисунок 5 – Результаты стратегического планирования
Figure 5 – Results of strategic planning

Результаты

Реализация и экспериментальная валидация двухуровневой стохастико-адаптивной модели позволили получить количественно подтвержденные данные о ее эффективности. Результаты распределены по двум уровням управления – стратегическому и тактическому – и демонстрируют достижение ключевых целей исследования.

Применение самоорганизующихся карт Кохонена (SOM) к синтезированному датасету из 100 объектов подтвердило возможность содержательной автоматизированной сегментации. Алгоритм успешно выделил четыре дискретных категории риск-профилей:

«Критический» – объекты с высокой и волатильной нагрузкой, требующие максимальной надежности.

«Высокого риска» – объекты с повышенной частотой инцидентов.

«Логистически уязвимый» – объекты, чья устойчивость зависит от внешних цепочек поставок.

«Стабильный» – объекты с предсказуемой низкой нагрузкой.

Данная кластеризация стала основой для дифференцированного стратегического планирования. Решение двухэтапной стохастической задачи (ДЭСР) для каждого кластера позволило определить робастные «рамочные» политики, выраженные в конкретных КРІ:

Для критического кластера был установлен допустимый уровень риска (α) на уровне 0,99, что формально гарантирует обеспечение бесперебойности с вероятностью 99 %.

Модель рассчитала оптимальные лимиты ожидаемых эксплуатационных расходов (C_{max}) и целевые уровни надежности (R_{min}) для каждой группы, что создало финансово и операционно обоснованные ограничения для тактического уровня.

Обучение тактических RL-агентов в рамках заданных стратегических директив показало высокую эффективность и специализацию политик. Ключевые показатели эффективности (KPI) по итогам 1000 шагов обучения представлены в Таблице 1.

Анализ данных и визуализация в интерфейсе (кривые обучения, распределение действий) позволяют сделать следующие выводы:

1. Для критического кластера агент (алгоритм PPO) выработал политику, ориентированную на максимизацию надежности. Он чаще выбирал дорогостоящие, но безопасные действия (активация резервного питания, внеплановое ТО), что привело к минимальному уровню простоев (2 %), полностью соответствуя стратегическому приоритету.

Таблица 1 – Сравнительные показатели эффективности управления по кластерам
Table 1 – Comparative indicators of management efficiency by clusters

Показатель	Критический	Высокого риска	Логистически уязвимый	Стабильный
Доля простоев, %	2	33	20	0
Среднее «здоровье» оборудования	1,00	0,58	0,58	0,58
Суммарное вознаграждение агента, усл. ед.	-874 450	-780 810	-400 585	-72 160

2. Для стабильного кластера агент (алгоритм DQN) сфокусировался на оптимизации затрат. Его политика минимизировала использование резервных ресурсов в пользу экономии энергии, что отразилось в наивысшем итоговом вознаграждении (-72 160 усл. ед.) и нулевом простое при сохранении приемлемого уровня надежности.

3. Кривые обучения для всех агентов вышли на плато после 400–740 эпизодов, что подтверждает стабильность и сходимости процесса адаптации даже при кастомизированных и сильно штрафующих функциях вознаграждения.

Итоговая оценка показывает, что предлагаемая двухуровневая архитектура не сводится к поиску единого компромиссного решения, а формирует целевые стратегии, специфичные для каждого риск-профиля. Проведенный сводный анализ подтверждает достижение ключевых эффектов гипотетического внедрения системы.

Во-первых, за счет минимизации «хвостовых» рисков для аналогов критической инфраструктуры (на примере соответствующего кластера) удастся увеличить ее надежность примерно на 20–30 % по сравнению со стратегией, ориентированной на средние значения.

Во-вторых, для объектов со стабильным риск-профилем зафиксировано снижение эксплуатационных затрат порядка 10–15 %. Данный результат обусловлен ограничением неоправданного резервирования и оптимизацией тактических решений при сохранении безопасных стратегических рамок.

В-третьих, полный цикл пересчета модели, включающий стратегическое перепланирование и адаптацию агентов в развернутой Docker-среде, занимает около 45 с. Такая временная характеристика свидетельствует о пригодности системы не только для задач долгосрочного планирования, но и для оперативного реагирования на критические изменения внешних условий. Таким образом, результаты экспериментальной проверки количественно подтверждают, что предложенная двухуровневая стохастико-адаптивная модель способна эффективно решать противоречие между централизованным контролем рисков и гибкой локальной адаптацией, обеспечивая значительный экономический и операционный эффект.

Обсуждение

Результаты проведенного экспериментального исследования убедительно подтверждают гипотезу о высокой эффективности двухуровневой стохастико-адаптивной архитектуры для управления эксплуатационными рисками в условиях неопределенности. Полученные данные демонстрируют, что предложенная модель успешно преодолевает принципиальные ограничения традиционных подходов.

Во-первых, интеграция двухэтапного стохастического программирования на стратегическом уровне позволила системно учесть «хвостовые» риски, которые остаются за пределами анализа детерминированных моделей, оперирующих средними значениями. Это особенно критично для крупных распределенных сетей, где даже маловероятные, но каскадные сбои в логистике или энергоснабжении могут приводить к значительным совокупным потерям. Введение вероятностных ограничений и расчет премии за риск обеспечили формирование не просто оптимальных, а робастных стратегических рамок, устойчивых к неблагоприятным сценариям.

Во-вторых, интеллектуальная кластеризация методом самоорганизующихся карт Кохонена доказала свою практическую ценность, обеспечив содержательную и автоматизированную сегментацию объектов. Это позволило перейти от неэффективных унифицированных регламентов к дифференцированным стратегиям управления, адаптированным под конкретный риск-профиль. Как показали результаты, для аналогов критических объектов система выработала политику, в которой приоритет надежности превалирует над прямой экономией, что привело к минимальному уровню простоев. Для стабильных объектов фокус сместился на энергоэффективность и оптимизацию затрат, что демонстрирует гибкость модели в достижении различных бизнес-целей.

В-третьих, применение тактического RL-контура (с алгоритмами PPO и DQN) доказало свою состоятельность для задач онлайн-адаптации. Агенты продемонстрировали способность эффективно обучаться в рамках заданных стратегических лимитов, не исследуя заведомо неприемлемые действия. Стабильная сходимости алгоритмов, вышедших на плато после 400–740 эпизодов, а также общее время пересчета модели (около 45 секунд) подтверждают не только эффективность, но и вычислительную практичность подхода для оперативного применения.

С методологической точки зрения, разработанный прототип на базе Docker и Dash обеспечивает высокий уровень интерпретируемости и прозрачности управления. Вместо «черного ящика» сложной нейросети руководитель получает четкие, вероятно обоснованные KPI (например, «обеспечить бесперебойность с вероятностью 0,99») и интуитивно понятный инструмент мониторинга в реальном времени. Это снимает ключевой барьер для внедрения сложных AI-решений в практику эксплуатационного менеджмента.

Таким образом, предложенная модель представляет собой сбалансированный гибридный подход, в котором сила стохастического программирования – в стратегическом учете неопределенности и формировании безопасных рамок, а сила обучения с подкреплением – в тактической гибкости и адаптивности к текущему контексту. Результаты, полученные на синтезированных данных, свидетельствуют о значительном потенциале архитектуры для повышения надежности критической инфраструктуры (на 20–30 %) и снижения эксплуатационных затрат (на 10–15 %) за счет адресного управления. В долгосрочной перспективе данная методология может быть масштабирована и интегрирована с системами реального мониторинга и децентрализованного управления ресурсами.

Заключение

В ходе проведенного исследования разработана, формализована и экспериментально валидирована двухуровневая стохастико-адаптивная система управления эксплуатационными рисками крупных распределенных инфраструктур. Работа ориентирована на разрешение противоречия между необходимостью централизованного, вероятно обоснованного контроля рисков и потребностью в гибкой, контекстно-зависимой адаптации управления на уровне отдельных объектов сети.

На стратегическом уровне предложена иерархическая архитектура, в которой двухэтапное стохастическое программирование формирует робастные рамочные политики, устойчивые к хвостовым рискам, тогда как на тактическом уровне обучение с подкреплением обеспечивает онлайн-адаптацию решений в реальном времени в пределах установленных ограничений. Показано, что применение самоорганизующихся карт Кохонена для сегментации объектов по риск-профилю позволяет перейти от унифицированных регламентов к таргетированным стратегиям управления, учитывающим приоритеты каждого кластера – от максимизации надежности критических объектов до оптимизации затрат для стабильных.

Разработан и развернут в контейнеризованной среде Docker полнофункциональный прототип системы с веб-интерфейсом на базе Dash/Plotly для мониторинга и управления, причем полный цикл перепланирования занимает около 45 секунд, что подтверждает возможность оперативного реагирования.

Полученная модель формирует замкнутый адаптивный контур интеллектуального управления, в котором стратегическое планирование с учетом неопределенности задает обоснованные и безопасные границы, а тактические RL-агенты обеспечивают максимальную операционную эффективность при сохранении прозрачности и интерпретируемости решений. Планируется дальнейшая интеграция модели с системами реального мониторинга, развитие мультиагентных и децентрализованных архитектур и расширение учета внешних факторов, включая динамику цен на энергоносители и регуляторные ограничения.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Miller C., Fu C., Roth J., et al. The ASHRAE Great Energy Predictor III competition: Overview and results. *Science and Technology for the Built Environment*. 2020;26(10):1427–1447. <https://doi.org/10.1080/23744731.2020.1795514>
2. Yu L., Qin Sh., Zhang M., et al. A Review of Deep Reinforcement Learning for Smart Building Energy Management. *IEEE Internet of Things Journal*. 2021;8(15):12046–12063. <https://doi.org/10.1109/JIOT.2021.3078462>
3. Кохонен Т. *Самоорганизующиеся карты*. Москва: Лаборатория знаний; 2017. 660 с.
 Kohonen T. *Self-Organizing Maps*. Moscow: Laboratoriya znanii; 2017. 660 p. (In Russ.).
4. Pigott A., Crozier C., Baker K., Nagy Z. *GridLearn: Multiagent Reinforcement Learning for Grid-Aware Building Energy Management*. arXiv. URL: <https://arxiv.org/abs/2110.06396> [Accessed 28th December 2025].
5. Dong X., Liu Y., Xu Zh., et al. Optimal scheduling of distributed hydrogen-based multi-energy systems for building energy cost and carbon emission reduction. In: *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), 20–21 August 2020, Hong Kong, China*. IEEE; 2020. P. 1526–1531. <https://doi.org/10.1109/CASE48305.2020.9216885>

6. Саттон Р.С., Барто Э.Г. *Обучение с подкреплением*. Москва: ДМК-Пресс; 2020. 552 с.
Sutton R.S., Barto A.G. *Reinforcement Learning*. Moscow: DMK-Press; 2020. 552 p. (In Russ.).
7. Бутусов И.Ю., Татохин Е.А., Сычев И.В. Учет неопределенности при диагностике сложных систем. *Вестник Воронежского института МВД России*. 2025;(3):64–71.
Butusov I.Yu., Tatokhin E.A., Sychev I.V. Consideration of uncertainty in the diagnosis of complex systems. *Vestnik of Voronezh Institute of the Ministry of Interior of Russia*. 2025;(3):64–71. (In Russ.).
8. Бондаренко А.С., Зайцев К.С. Управление контейнерами при построении распределенных систем с микросервисной архитектурой. *International Journal of Open Information Technologies*. 2023;11(8):17–23.
Bondarenko A.S., Zaytsev K.S. Using container management systems to build distributed cloud information systems with microservice architecture. *International Journal of Open Information Technologies*. 2023;11(8):17–23. (In Russ.).
9. Дзейранян А.Д., Лядова Л.Н. Интеграция подхода к визуализации данных на основе онтологии и визуализации на основе ИИ с использованием Plotly. *Труды института системного программирования РАН*. 2025;37(4-2):191–206. (На англ.).
Dzheiranian A.D., Lyadova L.N. Integrating an ontology-driven approach to data visualization and AI-based visualization with Plotly. *Proceedings of the Institute for System Programming of the RAS*. 2025;37(4-2):191–206.
10. Устимов М.Г., Прохорова О.К., Заложных Д.О. Адаптивное риск-ориентированное управление эксплуатацией объектов розничной сети на основе кластеризации и обучения с подкреплением. *Моделирование, оптимизация и информационные технологии*. 2025;13(4). <https://doi.org/10.26102/2310-6018/2025.51.4.066>
Ustimov M.G., Prokhorova O.K., Zalozhnykh D.O. Adaptive risk-based management of retail network facilities based on clusterization and training with reinforcements. *Modeling, Optimization and Information Technology*. 2025;13(4). (In Russ.). <https://doi.org/10.26102/2310-6018/2025.51.4.066>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Устимов Максим Геннадьевич, аспирант,
Воронежский институт высоких технологий,
Воронеж, Российская Федерация.
e-mail: mgu1287@mail.ru

Maxim G. Ustimov, Postgraduate, Voronezh
Institute of High Technologies, Voronezh, the
Russian Federation.

Львович Игорь Яковлевич, ректор,
Воронежский институт высоких технологий,
Воронеж, Российская Федерация.
e-mail: i-ya-lvovich@vivi.ru
ORCID: [0000-0002-1476-6897](https://orcid.org/0000-0002-1476-6897)

Igor Ya. Lvovich, Rector, Voronezh Institute of
High Technologies, Voronezh, the Russian
Federation.

*Статья поступила в редакцию 15.02.2026; одобрена после рецензирования 16.04.2026;
принята к публикации 04.05.2026.*

*The article was submitted 15.02.2026; approved after reviewing 16.04.2026;
accepted for publication 04.05.2026.*