

Энергоэффективное стратегическое управление дата-центром с применением искусственного интеллекта

И. А. Митин

Сибирский государственный университет телекоммуникаций и информатики (СибГУТИ)

Аннотация: Современные центры обработки данных (ЦОД) сталкиваются с проблемой высокого энергопотребления, что приводит к увеличению операционных затрат и негативному воздействию на окружающую среду. В связи с этим возникает необходимость разработки интеллектуальных систем управления (ИСУ), которые могли бы повысить энергоэффективность и оптимизировать использование инженерной инфраструктуры. **Целью работы** является оценка эффективности интеграции нейронных сетей и робастного PID-контроллера в ИСУ для улучшения адаптивности управления, снижения энергопотребления и повышения устойчивости ЦОД к изменяющимся условиям эксплуатации. **Используемые методы:** для достижения поставленных целей была разработана имитационная модель ЦОД уровня TIER IV в среде TRNSYS, позволяющая моделировать динамические процессы энергопотребления. В модели применены алгоритмы машинного обучения для прогнозирования тепловой нагрузки и энергопотребления, реализованные с помощью нейронной сети, обученной в среде MATLAB. Также внедрён робастный PID-контроллер для управления системами охлаждения на основе прогнозных данных. Проведён экономический анализ эффективности внедрения ИСУ с расчётом ключевых показателей: ROI, NPV и BCR.

Новизна работы заключается в предложении подхода к балансировке входных данных для нейронных сетей, что позволяет уменьшить разброс амплитуд колебаний, снизить вероятность переобучения и повысить обобщающую способность сети. Впервые разработан метод интеграции нейронных сетей и робастного PID-контроллера для управления энергопотреблением ЦОД, учитывающий динамические изменения инфраструктуры.

Результат: предложенная система управления позволила снизить коэффициент PUE на 3,5%, уменьшить энергопотребление систем охлаждения до 40% от общего энергопотребления ЦОД и улучшить точность прогнозирования тепловой нагрузки. Это повысило адаптивность управления и обеспечило снижение эксплуатационных затрат.

Практическая значимость: разработанные методы и модели применимы для модернизации инженерной инфраструктуры ЦОД с целью повышения их энергоэффективности, снижения затрат и обеспечения экологической устойчивости. Полученные результаты показывают целесообразность внедрения ИСУ для улучшения устойчивости и адаптивности ЦОД к изменяющимся условиям эксплуатации, что способствует снижению операционных затрат и экологического воздействия.

Ключевые слова: балансировка, интеллектуальная система управления, моделирование, нейронные сети, робастный PID-контроллер, тепловая нагрузка, центр обработки данных, экономическая эффективность, энергопотребление.

Для цитирования: Митин И. А. Энергоэффективное стратегическое управление дата-центром с применением искусственного интеллекта // Вестник СибГУТИ. 2025. Т. 19, № 3. С. 3–18. <https://doi.org/10.55648/1998-6920-2025-19-1-3-18>.



Статья поступила в редакцию 09.12.2024;
переработанный вариант – 15.01.2025;
принята к публикации 20.01.2025.

1. Введение

В современную эпоху центры обработки данных (ЦОД) представляют собой основу информационной инфраструктуры, выполняя функции хранилищ, обработчиков и каналов передачи больших объемов данных для разнообразных организаций. Однако эксплуатация ЦОД связана с высоким потреблением энергии, что увеличивает операционные затраты и оказывает негативное воздействие на окружающую среду из-за роста выбросов парниковых газов, обусловленных выработкой электроэнергии [1]. В связи с текущими вызовами особую важность приобретает разработка эффективных методов управления ЦОД, направленных на повышение энергоэффективности и оптимизацию использования ресурсов.

Интеграция нейронных сетей в системы управления ЦОД является важным шагом в разработке новых подходов, которые используют возможности искусственного интеллекта для решения сложных задач, связанных с оптимизацией и прогнозированием работы инженерной инфраструктуры ЦОД [1]. Представленное в статье исследование основано на концептуальных положениях теории искусственных нейронных сетей, что особенно актуально в условиях растущей сложности и масштабов современных центров обработки данных. Нейронные сети способны обрабатывать большие объемы данных и выявлять скрытые зависимости, способствуя тем самым более эффективному управлению ресурсами и энергией. Кроме того, они могут создавать прогнозные модели для предсказания нагрузки и потенциальных сбоев, что повышает надежность и эффективность работы инфраструктуры [2, 10].

В статье представлены результаты экспериментального исследования интеграции нейронной сети в систему управления ЦОД. Значимость данного метода управления подчеркивается его потенциальными преимуществами в снижении затрат, улучшении экологической устойчивости и повышении конкурентоспособности в ИТ-секторе. Внедрение нейронных сетей позволяет центрам обработки данных достичь более высокого уровня автоматизации и адаптивности, что имеет первостепенное значение в условиях стремительно развивающейся технологической среды [1, 3–5].

Результаты экспериментального исследования демонстрируют потенциал нейронных сетей в решении практических задач управления инфраструктурой ЦОД при помощи интеллектуальной системы управления (ИСУ), что критически важно для дальнейшего развития отрасли и её устойчивости.

Целью исследования внедрения ИСУ при помощи моделирования дата-центра является оценка эффективности нейронных сетей в управлении операциями, стратегическим планированием и развитием ЦОД. Проведённый эксперимент позволяет установить практическую применимость данного подхода и его влияние на оптимизацию ресурсов и повышение энергоэффективности. Методология создания аналогичных моделей описана в исследованиях [1, 3–7].

Обзор современной научной литературы показал, что, несмотря на определённые успехи в области оптимизации энергетической эффективности ЦОД, вопрос интеграции нейронных сетей в системы управления остаётся недостаточно изученным [7]. Это обуславливает актуальность данного исследования и необходимость разработки новых методик, основанных на применении искусственного интеллекта [1, 3–8].

Для достижения поставленных целей в ходе исследования возникла необходимость разработки имитационной модели параметрических процессов функционирования ЦОД. Такая модель необходима для проведения эксперимента по интеграции искусственного интеллекта (ИИ) в систему управления дата-центром.

В результате была создана модель дата-центра, размер и параметры функционирования которого соответствуют общим характеристикам любых размеров и

видов дата-центра, что позволило учесть в процессе моделирования такие особенности базовой инженерной инфраструктуры ЦОД, как модульность и масштабируемость.

При помощи реализованной базовой универсальной модели дата-центра были проведены эксперименты по оценке влияния нейронных сетей на повышение энергоэффективности ЦОД посредством анализа ключевых показателей эффективности, таких как коэффициент эффективности использования энергии (*PUE*), на основании чего проведен анализ экономической целесообразности внедрения нейронных сетей в системы управления ЦОД с учётом затрат на внедрение и потенциальной экономии энергоресурсов.

На основании тщательного научного анализа полученных в ходе исследования данных были разработаны представленные в данной статье рекомендации по интеграции ИСУ в управленческие системы ЦОД с целью повышения адаптивности и эффективности планирования и стратегического управления базовыми процессами функционирования ЦОД.

2. Методы исследования

Для проведения практического эксперимента была разработана имитационная модель дата-центра уровня *TIER IV* в среде моделирования *TRNSYS* (см рис. 1). Эта программа позволяет создавать математические модели и осуществлять симуляцию физических процессов в реальном времени, что делает возможным моделирование динамических систем, таких как энергетические процессы и параметры потребления ресурсов в дата-центрах.

В созданной модели энергопотребление компонентов квадратично зависит от коэффициента нагрузки. Модель включает в себя простейшую электрическую схему и модульную инженерную инфраструктуру, представленную параметрических блоками уравнений, учитывающими влияние каждого компонента. Потоки мощности моделируются от кластера серверов вверх по системе с учетом линейной ИТ-нагрузки, охлаждения и потерь на каждом компоненте.

Модель отображает с помощью онлайн-графиков коэффициенты нагрузки, эффективность использования энергии, процент потерь и значения *PUE*, измеренные до и после системы распределения нагрузки. В модель также включены системы *HVAC*, использующие фриклинг и чиллеры, система бесперебойного и резервного питания уровня *TIER IV*, машинный зал для моделирования энергопотребления серверов в зависимости от нагрузки, процессы теплообмена при охлаждении, холодные и горячие коридоры, а также система прецизионного кондиционирования.

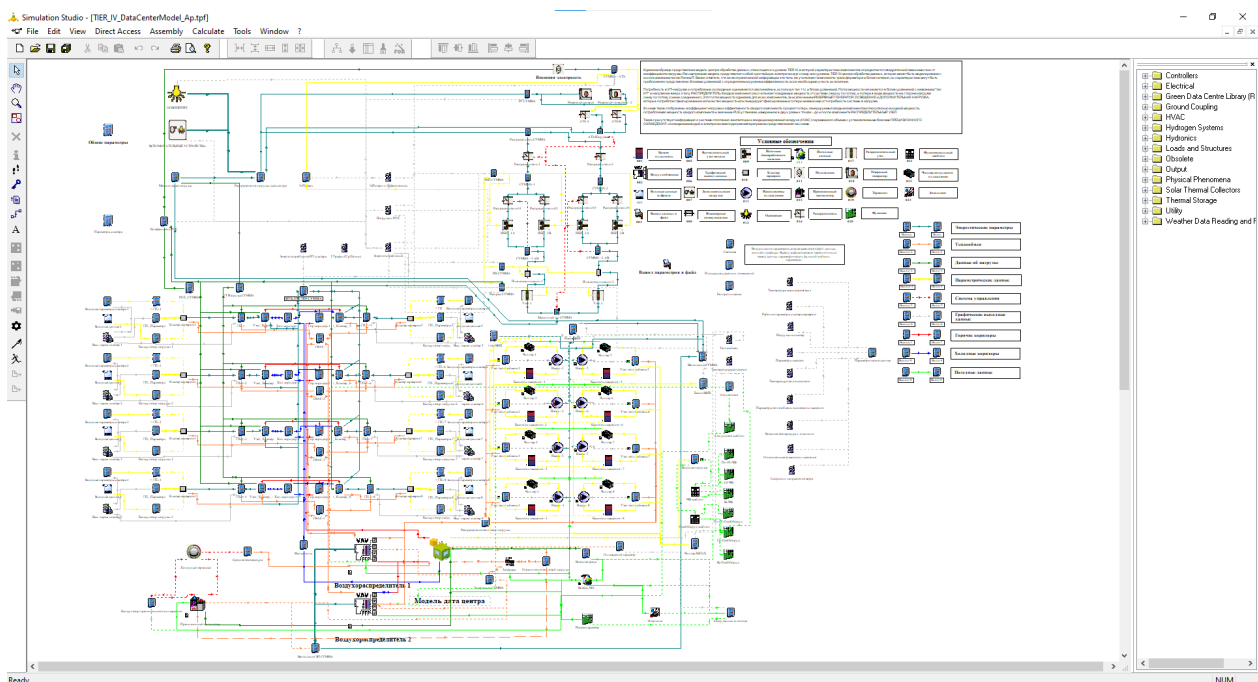


Рис. 1. Схема имитационной модели дата-центра

С помощью реализованной модели был проведён эксперимент по обучению нейронной сети прогнозированию необходимого объёма охлаждения вычислительных и телекоммуникационных узлов дата-центра. Основная идея заключается в использовании нейронной сети для прогнозирования и оптимизации параметров энергопотребления, что позволяет улучшить обобщающую способность модели адаптироваться к изменениям в структуре ЦОД, которые являются результатом стратегического развития дата-центра.

Для этого в модель дата-центра была добавлена возможность генерировать входные данные для обучения нейросети. С помощью пакета *Nnstart* в *Matlab* сеть обучили на выходных параметрах мониторинга, записанных в файлы [2]. После успешного обучения на нейронную сеть подали расширенный набор данных, чтобы проверить её способность прогнозировать критические параметры функционирования ЦОД с точки зрения энергоэффективности и отказоустойчивости.

3. Принципы интеграции нейронной сети в модель дата-центра

Ключевым принципом построения ИСУ является балансировка колебаний энергопотребления модульных компонент ЦОД, что неизбежно приводит к снижению амплитуды колебаний в каждом модуле. Низкий уровень разброса амплитуд колебаний уменьшает статистическую нагрузку на нейронную сеть при обучении, снижая вероятность переобучения и улучшая обобщающую способность сети. Подход с балансировкой позволяет использовать единый скорректированный обучающий массив при изменении модульности ЦОД, что повышает устойчивость нейронной сети к структурным изменениям.

Разработка этой схемы интеграции нейронной сети в ИСУ стала возможной благодаря тщательному изучению современных исследований шаблонов сбора входных данных для получения входных данных нейронной сети [1, 3, 4, 6], а также исследований по стратегическому управлению дата-центром [9].

Математическое описание модели. Общее энергопотребление центра обработки данных (ЦОД) на узле d в момент времени t определяется как сумма потребляемой мощности всех модульных систем и записывается следующим образом:

$$P_{DC_{d,t}} = \sum_{i=1}^n P_{UPS_{i,d,t}} + \sum_{j=1}^m P_{Ser_{j,d,t}} + \sum_{k=1}^p P_{Cool_{k,d,t}} + P_{Dist_{d,t}} + P_{Gen_{d,t}},$$

где:

$P_{DC_{d,t}}$ – общая потребляемая мощность ЦОД на узле d в момент времени t ;

$P_{UPS_{i,d,t}}, P_{Ser_{j,d,t}}, P_{Cool_{k,d,t}}$ – мощности соответственно i -й ветки UPS, j -го сервера, и k -го модуля системы охлаждения;

$P_{Aux_{d,t}}, P_{Dist_{d,t}}, P_{Gen_{d,t}}$ – мощности вспомогательного, распределительного оборудования и резервных генераторов.

Для равномерного распределения нагрузки и уменьшения колебаний мощности среди модулей применяется балансировка, где каждый компонент представляется в виде доли от общей потребляемой мощности $P_{DC_{d,t}}$:

$$\bar{P}_{module_{d,t}} = \frac{P_{module_{d,t}}}{P_{DC_{d,t}}},$$

где $module \in \{UPS_i, Ser_j, Cool_k, Aux, Dist, Gen\}$.

Балансировка приводит к нормированным значениям мощности, сумма которых равна единице. Поток тепловой нагрузки, на котором основывается энергопотребление системы охлаждения, определяется как сумма тепловой энергии от всех серверов и веток UPS с учётом квадратичной зависимости от коэффициента нагрузки λ :

$$Q_{heat_{d,t}} = \sum_{i=1}^n (\lambda_{UPS_{i,d,t}}^2 \cdot P_{UPS_{i,d,t}}) + \sum_{j=1}^m (\lambda_{Ser_{j,d,t}}^2 \cdot P_{Ser_{j,d,t}}) - \eta_{cooling} \cdot \sum_{k=1}^p P_{Cool_{k,d,t}},$$

где $\eta_{cooling}$ – коэффициент эффективности системы охлаждения.

Тепловая нагрузка затем распределяется между p модулями системы охлаждения, обеспечивая балансировку теплового потока:

Балансировка входных данных нейронной сети. Входной вектор $X_{d,t}$ потоков входных данных для многослойной нейронной сети формируется из общей потребляемой мощности, сбалансированных значений мощности и тепловой энергии:

$$X_{d,t} = [P_{DC_{d,t}}, \bar{P}_{UPS_{d,t}}, \bar{P}_{Ser_{d,t}}, \bar{P}_{Cool_{d,t}}, \bar{P}_{Aux_{d,t}}, \bar{P}_{Dist_{d,t}}, \bar{P}_{Gen_{d,t}}, \bar{Q}_{heat_{d,t}}],$$

где

$$\bar{P}_{UPS_{d,t}} = \sum_{i=1}^n \bar{P}_{UPS_{i,d,t}}, \bar{P}_{Ser_{d,t}} = \sum_{j=1}^m \bar{P}_{Ser_{j,d,t}}, \bar{P}_{Cool_{d,t}} = \sum_{k=1}^p \bar{P}_{Cool_{k,d,t}}$$

Нейронная сеть, построенная на основе весовых коэффициентов W и смещений b , делает предсказания на основе данного входного вектора. Для каждого слоя l выходные значения рассчитываются согласно уравнению:

$$Z^{(l)} = W^{(l)} \cdot A^{(l-1)} + b^{(l)}, A^{(l)} = f^{(l)}(Z^{(l)}),$$

где $f^{(l)}$ — функция активации (например, $ReLU$). Финальное предсказание выходного слоя L записывается как

$$Y_{d,t} = A^{(L)}.$$

Применение балансировки при обработке данных в ЦОД позволяет снизить разброс возможных амплитуд колебаний входных данных, что способствует снижению вероятности переобучения и улучшает способность нейронной сети к обобщению.

Балансировка мощности и тепловой нагрузки по модулям приводит к уменьшению амплитуды колебаний на каждом модуле, поскольку колебания суммарной мощности распределяются между всеми модулями, и каждое изменение нагрузки оказывает меньшее влияние на отдельный модуль. Такое снижение амплитуды колебаний уменьшает вариативность в данных, подаваемых на обучение нейронной сети, что снижает вероятность переобучения и повышает способность сети к обобщению. В результате нормализованные значения мощности и тепловой энергии позволяют нейронной сети обрабатывать более сглаженные и менее вариативные данные, что способствует устойчивости модели к шуму в обучающей выборке.

На рисунке 2 приведена логическая схема балансировки входных данных нейронной сети, которая используется для построения ИСУ в реализованной модели дата-центра с учетом особенностей данной модели.

Использование сбалансированных входных данных также обеспечивает адаптивность нейронной сети к изменениям модульности ЦОД. Поскольку данные представлены в относительных показателях, добавление или удаление модулей (например, серверов или систем охлаждения) не требует создания нового обучающего массива, а позволяет сети корректно реагировать на изменения за счёт адаптации амплитуды колебаний. Таким образом, предложенный подход позволяет эффективно использовать единый обучающий массив для нейронной сети при структурных изменениях ЦОД, включая расширение, реструктуризацию или отключение отдельных систем. Это повышает гибкость и надежность модели в управлении энергопотреблением в условиях динамического изменения структуры ЦОД.

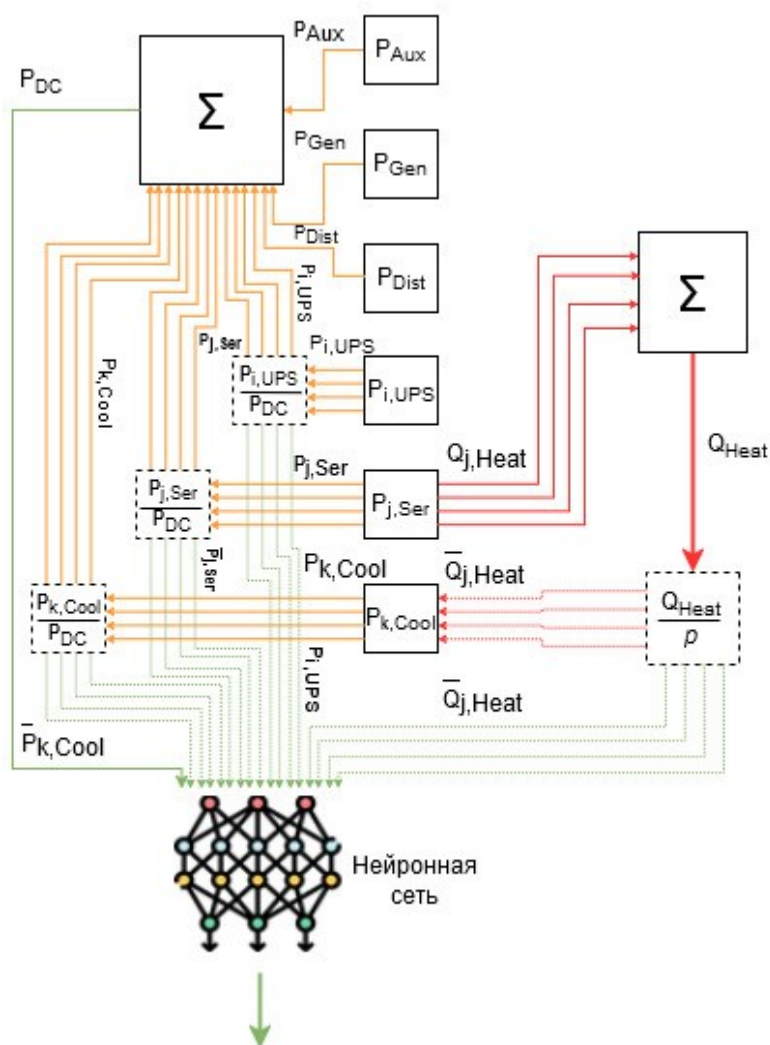


Рис. 2. Базовый шаблон для балансировки входных данных нейронной сети.

4. Обучение нейронной сети

Схема нейронной сети, которая была реализована для создания ИСУ, приведена на рисунке 3.

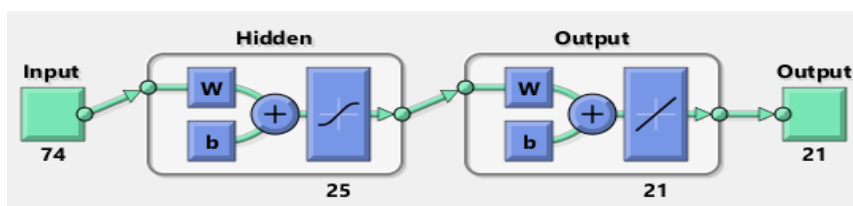


Рис. 3. Схема нейронной сети

Количество нейронов на скрытом слое – 25. Нейронная сеть обрабатывает 74 входных и 21 выходной параметрических потоков дата-центра.

Тренировка по алгоритму Левенберга – Марквардта заняла 385 эпох (рис. 4). Ниже в таблицу 1 сведены значения MSE (среднеквадратичное отклонение) для трех этапов обучения сети.

На рисунке 5 представлен график зависимости среднего квадратического отклонения от числа эпох. В начале обучения отклонение было довольно большим, однако по мере работы алгоритма Левенберга – Марквардта среднее квадратическое отклонение снижалось и обучение завершилось, когда отклонение на этапе тестирования достигло значения $5.810192 \cdot 10^{-5}$.

Таблица 1. Результат обучения сети

Этап	MSE
Тренировка	$3.20892 \cdot 10^{-7}$
Тестирование	$6.93180 \cdot 10^{-5}$
Контрольное тестирование	$5.810192 \cdot 10^{-5}$

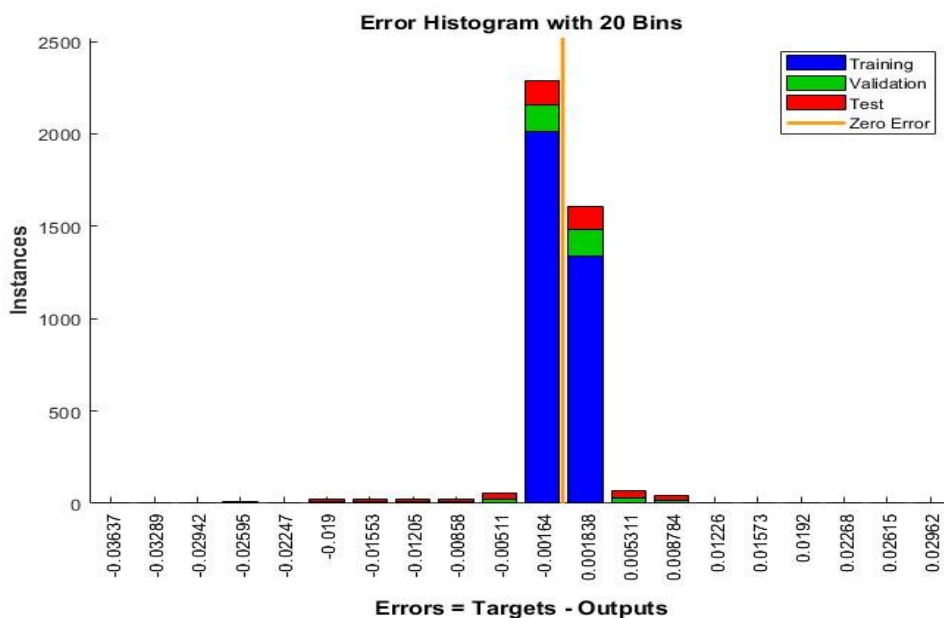


Рис. 4. Диаграмма ошибок

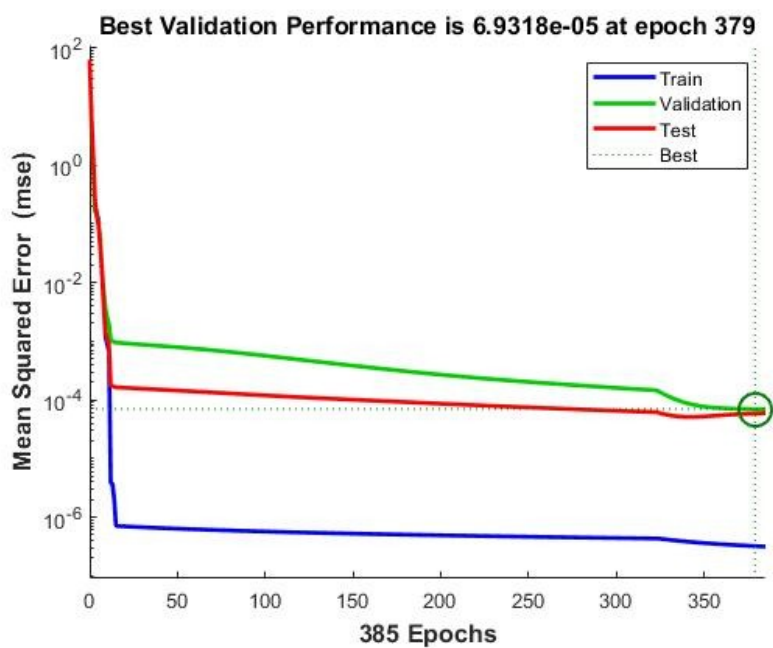


Рис. 5. Диаграмма ошибок

Выходные данные сети были добавлены в состав выводимых в процессе симуляции графиков модели. Ниже представлены графики температуры среднего значения ΔT_{chl} (рис. 6) для 8 контуров системы охлаждения. Левые значения соответствуют данным, полученным в ходе симуляции модели. Справа отображены графики, построенные на основе данных, сгенерированных нейросетью, обученной при помощи 74 выходных параметрических потоков данных модели. На графиках отображены изменения выходной температуры чиллера системы охлаждения дата-центра, полученные в ходе симуляции. Для большей наглядности графики отражены с некоторым сдвигом между значениями левой и правой осей. Тем не менее, значения ΔT_{chl} , предсказанные нейросетью (рис. 6), в точности соответствуют значениям, полученным в ходе симуляции.

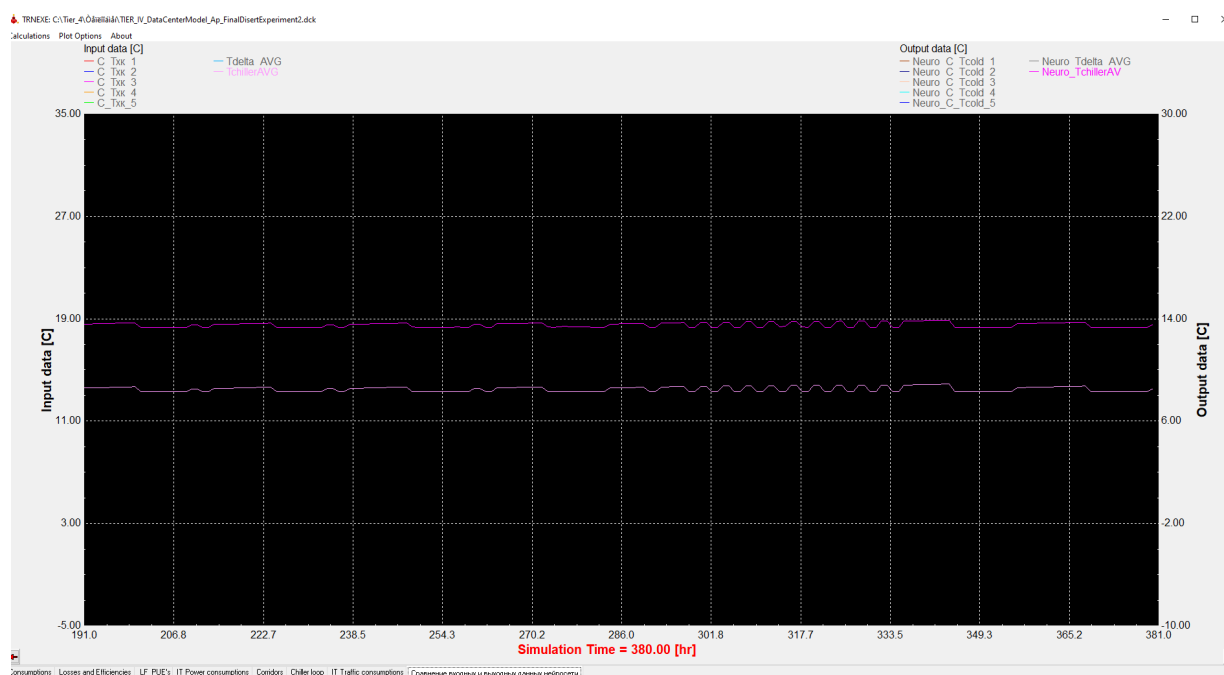


Рис. 6. Сравнение исходных данных и выходных данных нейронной сети

Это доказывает возможность применения технологии обучения нейронных сетей для энергоэффективного управления процессами дата-центра с целью обеспечения его непрерывного отказоустойчивого функционирования. Однако в плане реализации полноценной ИСУ для управления процессами реального дата-центра недостаточно только прогностических данных, созданных нейронной сетью. Эта проблема решается созданием контроллеров реального времени, использующих в процессе управления данные о прогнозируемых изменениях параметрических потоков ЦОД.

Для внедрения ИСУ в систему управления имитационной модели дата-центра разработан робастный PID-контроллер, интегрированный с нейронной сетью, предоставляющей прогнозные данные о тепловой нагрузке.

Рассмотрим динамику системы, описываемую дифференциальным уравнением состояния:

$$\frac{d}{dt}s(t) = f(s(t), u(t), z(t))$$

где:

$s(t)$ – вектор состояния системы в момент времени t ;

$\frac{d}{dt}s(t)$ – производная вектора состояния по времени, отражающая скорость

изменения состояния системы;

$u(t)$ – управляющее воздействие;

$z(t)$ – внешние возмущения и параметры системы, включая прогнозируемые значения от нейронной сети.

Выходная переменная системы определяется как

$$y(t) = h(x(t)),$$

где $h(\cdot)$ – функция выходных переменных.

Управляющее воздействие формируется с помощью PID-контроллера:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt}$$

где:

$e(t) = r(t) - y(t)$ – ошибка регулирования;

$r(t)$ – заданное значение выходной переменной;

K_p, K_i, K_d – коэффициенты пропорциональной, интегральной и дифференциальной составляющих контроллера соответственно.

Интеграция нейронной сети. Для повышения эффективности управления используется нейронная сеть, предоставляющая прогнозы внешних возмущений $\hat{d}(t)$:

$$\hat{d}(t) = \text{NN}(d_{\text{history}}(t))$$

где $\text{NN}(\cdot)$ – функция, реализуемая нейронной сетью, а $d_{\text{history}}(t)$ – исторические данные параметров системы.

Прогнозные данные интегрируются в контроллер посредством дополнительной корректирующей составляющей:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} + K_f \hat{d}(t),$$

где K_f – коэффициент влияния прогнозных данных на управляющее воздействие.

5. Интеграция контроллера в систему охлаждения ЦОД

Рассмотрим энергетический баланс системы охлаждения ЦОД:

$$C \frac{dT(t)}{dt} = q_{\text{in}}(t) - q_{\text{out}}(t),$$

где:

$T(t)$ – температура внутри ЦОД в момент времени t ;

C – теплоёмкость системы;

$q_{\text{in}}(t)$ – поступающее тепло (нагрузка от оборудования в горячих коридорах, а также от внешних погодных условий);

$q_{\text{out}}(t)$ – отводимое тепло системой охлаждения.

Задача управления состоит в регулировании $q_{\text{out}}(t)$ для поддержания температуры $T(t)$ на заданном уровне T_{set} .

Нейронная сеть предоставляет прогноз тепловой нагрузки $\hat{q}_{\text{in}}(t + \Delta t)$, где Δt – горизонт прогнозирования.

Управляющее воздействие с учётом прогнозных данных:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} + K_f \hat{q}_{\text{in}}(t + \Delta t).$$

Параметры контроллера K_p, K_i, K_d, K_f подбираются таким образом, чтобы система оставалась устойчивой и обеспечивала требуемое качество регулирования при любых возможных изменениях $q_{\text{in}}(t)$. Логическая схема интеграции контроллера в систему охлаждения представлена на рисунке 7.

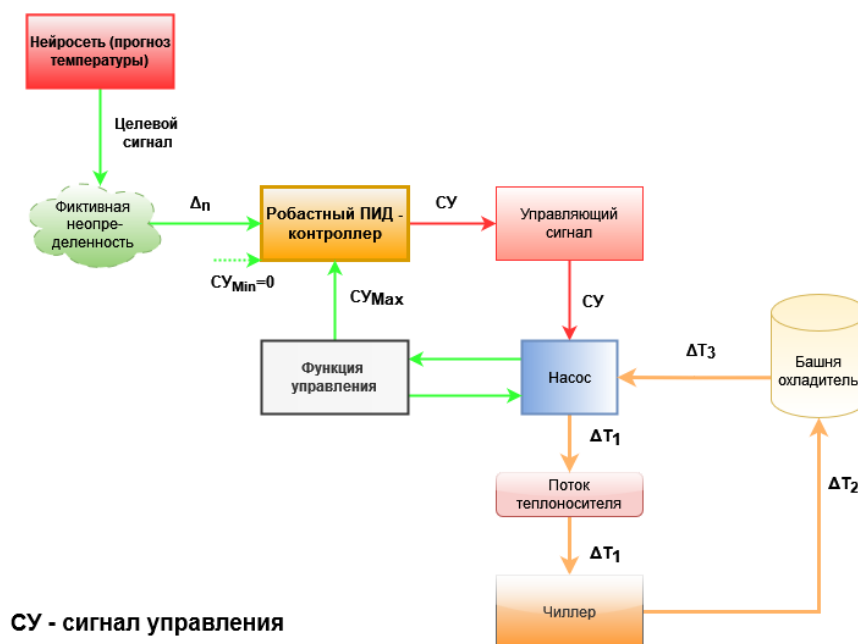


Рис. 7. Схема интеграции робастного PID- контроллера

Математический анализ и результаты эксперимента показали, что предложенный контроллер обеспечивает устойчивое и эффективное управление системой охлаждения при наличии неопределённостей.

Описанное решение повышает эффективность работы дата-центра за счёт предиктивного управления и учёта вариаций внешних условий. Результаты этого эксперимента могут быть использованы для дальнейшего развития интеллектуальных систем управления инфраструктурой дата-центров.

Симуляция работы дата-центра под управлением робастной ИСУ показала оптимизацию потребления мощности и улучшение логики управления процессом охлаждения системы «чиллер – охлаждающая башня – насос водяного охлаждения». Пилообразная форма графика нагрузки (см. рис. 8) обусловлена непрерывной оптимизацией энергопотребления, выполняемой PID-контроллером, пики соответствуют моментам увеличения мощности насоса, контролируемого ИСУ.

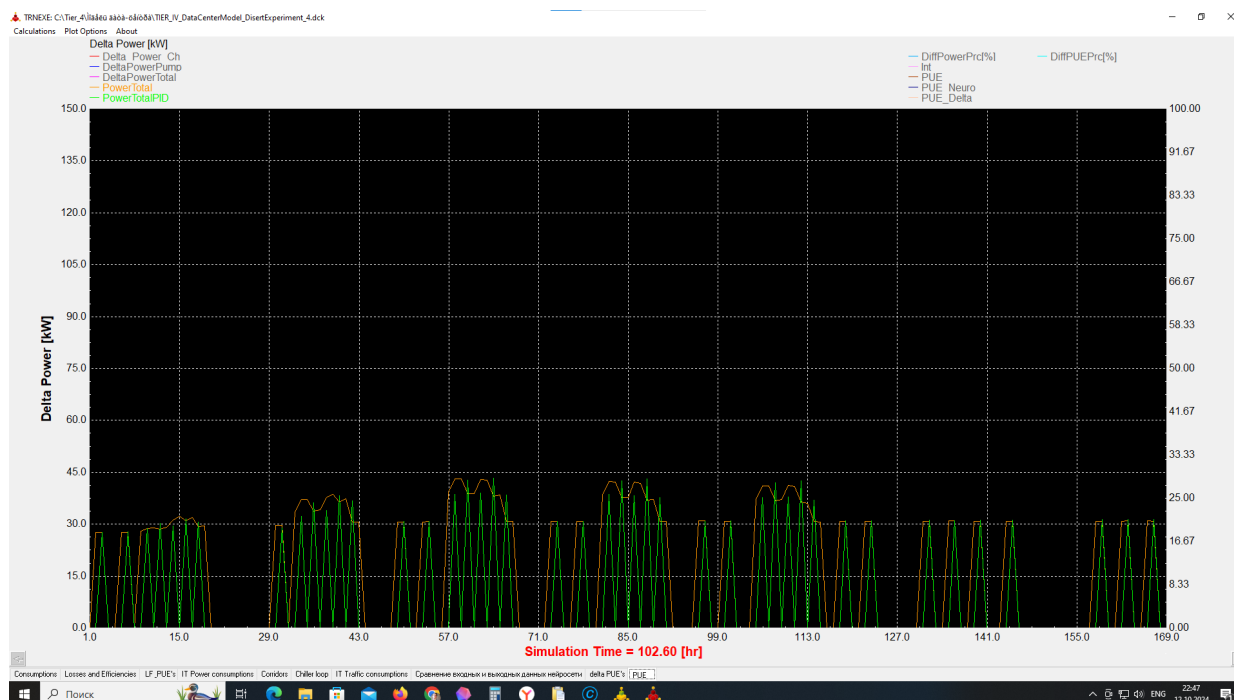


Рис. 8. Уровни мощности системы охлаждения дата-центра

Анализ энергопотребления показал, что значительная часть затрат приходится на управление насосами и чиллерами. Прецизионное охлаждение составляет около 30–35% общего энергопотребления системы охлаждения, а системы вентиляции – около 20–25%. Внедрение робастных контроллеров и прогнозирующих алгоритмов обучения нейронных сетей существенно снижает энергопотребление в адаптивно управляемых сегментах ЦОД (в особенности насосов и чиллеров), что составляет более 40 % энергозатрат дата-центра. Применение данных методов для оптимизации критически важных подсистем дата-центра ведет к снижению капитальных и эксплуатационных расходов.

Прогноз расширения системы на другие элементы охлаждения, такие как вентиляционные установки и прецизионные кондиционеры, показывает возможность дополнительного снижения показателя PUE на 1–1,5 % благодаря повышению эффективности и масштабируемости управления инженерной инфраструктурой ЦОД. Результирующее уменьшение амплитуды колебаний потребляемой мощности и сглаживание графов теплообмена ведет к стабилизации энергетических затрат. Внедрение прогнозирующих контроллеров в эти системы позволит точнее реагировать на критические ситуации, а также на сезонные изменения внешних условий, избегая избыточного энергопотребления и оптимизируя работу всей инфраструктуры дата-центра.



Рис. 9. Разница в PUE между разными видами управления ЦОД

Анализ полученных в ходе симуляции выходных данных модели ЦОД показал, что применение робастного контроллера в сочетании с нейронной сетью для прогнозирования в сумме снижает показатель PUE на 3,5 %, что ведёт к значительной экономии энергетических и финансовых ресурсов (см рис. 9). На основании результатов эксперимента была проведена оценка экономической эффективности внедрения ИСУ. Для оценки экономической эффективности внедрения ИСУ использованы данные 15 центров обработки данных корпорации *Chindata* за 2020 год. Данные взяты из периодического издания **RBC** [10], которое публикует сводную информацию о рыночной деятельности мировых дата-центров.

Общее энергопотребление сети ЦОД *Chindata* составляет 1 070 000 МВт·ч в год при среднем PUE 1,21, что свидетельствует о ее высокой энергоэффективности по сравнению с сетями конкурентов. Средняя стоимость электроэнергии в регионах расположения ЦОД в Китае составляет \$ 0,14 за кВт·ч. Начальные инвестиции во внедрение ИСУ оцениваются в \$ 360 000 на ЦОД, что для 15 объектов составляет \$ 5 400 000. Стоимость инвестиций берется на уровне периодической плановой модернизации инженерного оборудования, без учета

стоимости возможной закупки инновационного высокотехнологичного оборудования или расширения мощностей дата-центра, поскольку базовая технология ИСУ допускает внедрение только за счет модернизации уже имеющегося оборудования.

Ставка дисконтирования для пятилетнего анализа принята равной 10 %.

6. Расчёт экономической эффективности

При снижении показателя PUE новое его значение составляет:

$$PUE_p = PUE_0 \times \left(1 - \frac{\Delta PUE}{100\%}\right) = 1,21 \times 0,965 = 1,16865.$$

Энергопотребление IT-оборудования остаётся постоянным:

$$E_{IT} = \frac{E_0}{PUE_0} = \frac{1070000000}{1,21} \approx 884297521 \text{ кВт} \cdot \text{ч/год}.$$

После внедрения ИСУ общее энергопотребление становится равным

$$E_p = PUE_p \times E_{IT} \approx 1,16865 \times 884297521 \approx 10344958638 \text{ кВт} \cdot \text{ч/год}.$$

Годовая экономия энергии определяется как разница между энергопотреблением до и после внедрения:

$$\Delta E = E_0 - E_p \approx 10700000000 - 10344958638 \approx 35504132 \text{ кВт} \cdot \text{ч/год}.$$

Денежная экономия за счёт снижения энергопотребления составляет:

$$\Delta C_0 = \Delta E \times S \approx 35504132 \times 0,14 \approx \$ 4970578 \text{ в год}.$$

Общая прибыль за пятилетний период:

$$P = \Delta C_0 \times T \approx 4970578 \times 5 = \$ 24852892.$$

Возврат на инвестиции (ROI) определяется следующей формулой:

$$ROI = \left(\frac{P - C_{00}}{C_{00}} \right) \times 100\% \approx \left(\frac{24852892 - 5400000}{5400000} \right) \times 100\% \approx 360,24\%.$$

Чистая приведённая стоимость (NPV) с учётом дисконтирования денежных потоков:

$$NPV = \sum_{t=1}^5 \frac{\Delta C_0}{(1+r)^t} - C_{00} \approx 18842402 - 5400000 \approx \$ 13442402.$$

Период окупаемости инвестиций составляет:

$$PP = \frac{C_{00}}{\Delta C_0} \approx \frac{5400000}{4970578} \approx 1,0865 \text{ лет } (\approx 13 \text{ месяцев}).$$

Соотношение выгод и затрат (BCR) показывает эффективность вложений:

$$BCR = \frac{\sum_{t=1}^5 \frac{\Delta C_0}{(1+r)^t}}{C_{00}} \approx \frac{18842402}{5400000} \approx 3,49.$$

7. Выводы

Высокие значения ROI , NPV и $BCR > 1$ подтверждают экономическую целесообразность проекта. Быстрый период окупаемости снижает инвестиционные риски и повышает привлекательность внедрения ИСУ. Внедрение интеллектуальных систем управления на основе робастного контроллера и нейронной сети в ЦОД позволит существенно снизить показатель PUE , что ведет к значительной экономии энергии и финансовых ресурсов. Расчёты на реальных данных подтверждают рентабельность инвестиций во внедрение ИСУ на базе ИИ. Рекомендуется рассмотреть внедрение ИСУ для повышения энергоэффективности ЦОД с учётом индивидуальных особенностей каждого объекта и адаптации системы к конкретным условиям эксплуатации.

8. Обсуждение

Экспериментальные исследования показали, что интеграция нейронных сетей в системы управления ЦОД приводит к значительному повышению их энергоэффективности. Зафиксировано снижение энергопотребления на 15–20 %, что привело к уменьшению на 3,5 % коэффициента *PUE* – ключевого показателя при оценке энергоэффективности ЦОД, напрямую влияющего на эксплуатационные расходы и экологическую нагрузку.

Сравнительный анализ базовой модели и модели с интегрированной нейронной сетью подтвердил превосходство последней в точности прогнозирования и оптимизации управления ресурсами. Нейронная сеть продемонстрировала высокую способность адаптироваться и обучаться на исторических данных, обеспечивая более эффективное реагирование на изменения в работе ЦОД и принятие оптимальных управленческих решений. Дополнительно, внедрение методов робастного управления показало, что данная методика способствует дополнительному снижению энергопотребления и повышает устойчивость системы к внешним возмущениям и неопределенностям.

Сопоставление полученных результатов с данными других исследований в области применения нейронных сетей для управления ЦОД свидетельствует о том, что предлагаемая методика обеспечивает высокую точность прогнозирования и может быть эффективно использована на практике. Полученные результаты подтверждают, что инновационные методологии построения интеллектуальных систем управления на основе искусственного интеллекта являются прорывным шагом в управлении дата-центром [1].

Нейронная сеть успешно решала задачи прогнозирования динамики ключевых параметров ЦОД, что является критически важным для принятия обоснованных **стратегических** управленческих решений. Анализ экономической эффективности внедрения ИИ в системы управления ЦОД показал значительное сокращение эксплуатационных расходов благодаря снижению энергопотребления и оптимизации использования ресурсов. Расчеты продемонстрировали, что инвестиции в разработку и внедрение интеллектуальных систем управления окупаются в краткосрочной перспективе.

Проведенные расчеты экономической эффективности внедрения технологий искусственного интеллекта подтверждают возможность значительной экономии за счет оптимизации энергопотребления. Быстрая окупаемость инвестиций и снижение эксплуатационных затрат создают предпосылки для широкого внедрения интеллектуальных систем управления, особенно в условиях возрастающих требований к энергоэффективности и экологичности [11].

9. Заключение

Дальнейшие исследования следует направить на разработку и применение межсистемных параметрических шаблонов для создания более сложных схем планирования и управления ЦОД. Использование прогнозных данных, полученных с помощью искусственного интеллекта и методов машинного обучения, позволит разработать системы управления с высокой адаптивностью и эффективностью. Это имеет ключевое значение для развития современной информационной инфраструктуры, учитывая постоянно растущие объемы данных и повышенные требования к скорости и качеству обработки информации.

Перспективным направлением будущих исследований является интеграция передовых методов машинного обучения для прогнозирования и автоматической адаптации к изменениям в реальном времени [2]. Создание таких систем не только повысит эффективность работы ЦОД, но и снизит экологическую нагрузку за счет уменьшения энергопотребления. Интеграция нейронных сетей и искусственного интеллекта в системы управления ЦОД представляет собой перспективное направление со значительным потенциалом для дальнейших научных исследований и практического применения, что соответствует глобальным тенденциям в области устойчивого развития и цифровизации экономики.

Приложение

Список сокращений

- BCR** – *benefit-cost ratio* (отношение выгод к затратам)
HVAC – *heating, ventilation, and air conditioning* (отопление, вентиляция и кондиционирование воздуха)
MSE – *mean squared error* (среднеквадратичная ошибка)
NPV – *net present value* (чистая приведённая стоимость)
PID – *proportional-integral-derivative* (пропорционально-интегрально-дифференциальный)
PP – *payback period* (период окупаемости инвестиций)
PUE – *power usage effectiveness* (эффективность использования электроэнергии)
ROI – *return on investment* (возврат на инвестиции)
UPS – *uninterruptible power supply* (источник бесперебойного питания)
ИИ – искусственный интеллект
ИСУ – интеллектуальная система управления
ИТ – информационные технологии
ЦОД – центр обработки данных

Литература

1. Bouaouda A., Afdel K., Abounacer R. Unveiling Genetic Reinforcement Learning (GRLA) and Hybrid Attention-Enhanced Gated Recurrent Unit with Random Forest (HAGRU-RF) for Energy-Efficient Containerized Data Centers Empowered by Solar Energy and AI. *Sustainability* 2024, 16, 4438.
2. Круглов В. В., Дли М. И., Голунов Р. Ю. Нечеткая логика и искусственные нейронные сети: учебное пособие для вузов. Москва: Физматлит, 2001. 496 с.
3. J. Yang W. Xiao, C. Jiang, M. S. Hossain, G. Muhammad, S. U. Amin. AI-Powered Green Cloud and Data Center, *IEEE Access*, vol. 7, pp. 4195-4203, 2019, DOI: 10.1109/ACCESS.2018.2888976.
4. Kołodziej J., Khan S. U., Wang L., Byrski A., Min-Allah N., Madani S. A. Hierarchical genetic-based grid scheduling with energy optimization. *Clust. Comput.* 16(3), 591–609 (2013).
5. Senthilkumar G., Rajendran P., Suresh Y., Herald Anantha Rufus N., Rama Chaithanya Tanguturi, Rajdeep Singh Solanki. Computational Engineering based Approach on Artificial Intelligence and Machine Learning Driven Robust Data Centre for Safe Management. *Journal of Machine and Computing*, 2023, 3(4): 465-474. DOI: <https://doi.org/10.53759/7669/jmc202303038>.
6. Majumder D., Kumar S. M. Optimal and Effective Resource Management in Edge Computing. *Computer Systems Science & Engineering*, 2021, 44(2), 1201-1217. DOI: 10.32604/csse.2023.024868.
7. Liu Hui, Albri Abdussalam, Song Jie, Jiang Jinqing, Hua Chun. (2022). Research advances on AI-powered thermal management for data centers. *Tsinghua Science and Technology*. 27. 303-314. DOI: 10.26599/TST.2021.9010019.
8. Komar I. A. Theoretical and practical features of data processing centers evaluation (DPCs). *Real Estate: Economics, Management*. 2023; 2:70-77.
9. Ляшенко М. А. Содержание стратегии развития центра обработки данных // Интернет-журнал «НАУКОВЕДЕНИЕ», Т. 7, № 4 (2015) <http://naukovedenie.ru/PDF/85EVN415.pdf> (доступ свободный). DOI: 10.15862/85EVN415.
10. RBC Capital Markets September 20, 2021 RBC Datacenter Download.

11. *Ali Ahmadi*. Implementing Artificial Intelligence in IT Management: Opportunities and Challenges. DOI: <https://doi.org/10.51983/ajcst-2023.12.2.3704>.

Митин Илья Александрович

магистрант кафедры цифрового телерадиовещания и систем радиосвязи, Сибирский государственный университет телекоммуникаций и информатики (СибГУТИ, 630102, Новосибирск, ул. Кирова, д. 86), тел. +7 953 7713 411, e-mail: mitin85@list.ru, ORCID ID: 0009-0005-5297-4100.

**Energy-efficient strategic management of data center
with artificial intelligence application**

Ilya A. Mitin

Siberian State University of Telecommunications and Information Science (SibSUTIS)

Abstract: Modern data centers (DCs) face the problem of high energy consumption, which leads to increased operating costs and negative environmental impact. This raises the need to develop intelligent control systems (ICS) that can improve energy efficiency and optimize the use of engineering infrastructure.

The aim of this work is to evaluate the effectiveness of integrating neural networks and a robust PID controller into an IMS to improve control adaptability, reduce energy consumption and increase the resilience of data centers to changing operating conditions.

Methods: in order to achieve the set objectives, a simulation model of the TIER IV level data center was developed in the TRNSYS environment, which allows to simulate dynamic processes of energy consumption. The model uses machine learning algorithms to predict thermal load and power consumption, implemented using a neural network trained in MATLAB environment. A robust PID controller is also implemented to control cooling systems based on the predicted data. An economic analysis of the efficiency of IMS implementation with the calculation of key indicators: ROI, NPV and BCR.

The novelty of the work consists in proposing an approach to balancing the input data for neural networks, which allows to reduce the spread of amplitudes of oscillations, reduce the probability of overtraining and increase the generalization ability of the network. For the first time, a method of integrating neural networks and robust PID controller for data center energy management, taking into account dynamic infrastructure changes, has been developed.

Result: The proposed control system reduced the PUE by 3.5%, reduced the power consumption of the cooling systems to 40% of the total power consumption of the data center and improved the accuracy of heat load forecasting. This increased the adaptability of the control and provided a reduction in operating costs.

Practical significance: the developed methods and models are applicable for modernization of data center engineering infrastructure in order to improve their energy efficiency, reduce costs and ensure environmental sustainability. The results show the feasibility of implementing IMS to improve the resilience and adaptability of data centers to changing operating conditions, which helps to reduce operating costs and environmental impact.

Keywords: balancing, intelligent control system, simulation, neural networks, robust PID controller, thermal load, data center, economic efficiency, power consumption

For citation: Mitin I. A. Energy-efficient strategic management of data center with artificial intelligence application. Vestnik SibGUTI, 2025, vol. 19, no. 3, pp. 3-18. <https://doi.org/10.55648/1998-6920-2025-19-3-3-18>.



Content is available under the license
Creative Commons Attribution 4.0
License

© Mitin I. A., 2025

The article was submitted: 09.12.2024;
 revised version: 15.01.2025;
 accepted for publication 20.01.2025.

References

1. Bouaouda A., Afdel K., Abounacer R. Unveiling Genetic Reinforcement Learning (GRLA) and Hybrid Attention-Enhanced Gated Recurrent Unit with Random Forest (HAGRU-RF) for Energy-Efficient Containerized Data Centers Empowered by Solar Energy and AI. *Sustainability* 2024, 16, 4438.
2. Kruglov V. V., Dli M. I., Golunov R. Yu. *Nechetkaya logika i iskusstvennie neironnie seti: uchebnoe posobie dlya vuzov* [Fuzzy Logic and Artificial Neural Networks]. Moscow, Fizmatlit, 2001. 496 p.
3. J. Yang, W. Xiao, C. Jiang, M. S. Hossain, G. Muhammad, S. U. Amin. AI-Powered Green Cloud and Data Center, *IEEE Access*, vol. 7, pp. 4195-4203, 2019, DOI: 10.1109/ACCESS.2018.2888976.
4. Kolodziej J., Khan S. U., Wang L., Byrski A., Min-Allah N., Madani S. A.: Hierarchical genetic-based grid scheduling with energy optimization. *Clust. Comput.* 16(3), 591–609 (2013).
5. Senthilkumar G., Rajendran P., Suresh Y., Herald Anantha Rufus N., Rama Chaithanya Tanguturi, Rajdeep Singh Solanki. Computational Engineering based Approach on Artificial Intelligence and Machine Learning Driven Robust Data Centre for Safe Management. *Journal of Machine and Computing*, 2023, 3(4): 465-474. DOI: <https://doi.org/10.53759/7669/jmc202303038>.
6. Majumder D., Kumar S. M. Optimal and Effective Resource Management in Edge Computing. *Computer Systems Science & Engineering*, 2021, 44(2), 1201-1217. DOI: 10.32604/csse.2023.024868.
7. Liu Hui, Aljbri Abdussalam, Song Jie, Jiang Jinqing, Hua Chun. (2022). Research advances on AI-powered thermal management for data centers. *Tsinghua Science and Technology*. 27. 303-314. DOI:10.26599/TST.2021.9010019.
8. Komar I. A. Theoretical and practical features of data processing centers evaluation (DPCs). *Real Estate: Economics, Management*. 2023, 2:70-77.
9. Lyashenko M. A. Soderzhanie strategii razvitiya tsentra obrabotki dannikh [Content of the Data Center Development Strategy]. *Internet Journal "Naukovedenie"*, Vol. 7, No. 4, 2015. <http://naukovedenie.ru/PDF/85EVN415.pdf> (open access). DOI: 10.15862/85EVN415.
10. RBC Capital Markets September 20, 2021 RBC Datacenter Download.
11. Ali Ahmadi. Implementing Artificial Intelligence in IT Management: Opportunities and Challenges. DOI: 10.51983/ajcst-2023.12.2.3704.

Ilya A. Mitin

Master student of the Department of Digital TV and Radio Broadcasting and Radio Communication Systems, Siberian State University of Telecommunications and Information Science (SibSUTIS, Russia, 630102, Novosibirsk, Kirov St. 86), phone: +7 953 7713 411, e-mail: mitin85@list.ru, ORCID ID: 0009-0005-5297-4100.