

<https://doi.org/10.31891/2307-5732-2026-363-30>

УДК 004.8:621.39

ГНАТЮК ВІКТОР

Державний університет «Київський авіаційний інститут»; ДержНДІ технологій кібербезпеки

<https://orcid.org/0000-0002-4916-7149>

e-mail: viktor.hnatiuk@npp.kai.edu.ua

ЛИТВИНЮК ОЛЕКСАНДР

Державний університет «Київський авіаційний інститут»

<https://orcid.org/0009-0002-2221-8505>

e-mail: 6163394@stud.kai.edu.ua

МЕТОД ПРОГНОЗУВАННЯ ЗАВАНТАЖЕННЯ ТА РОЗПОДІЛУ РЕСУРСІВ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ В ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМАХ

У статті представлено метод прогнозування завантаження та розподілу ресурсів у телекомунікаційних системах на основі гібридної моделі нейронної мережі, що поєднує згорткові та рекурентні архітектури. Запропонований підхід дозволяє враховувати короткострокові та довгострокові залежності в мережевому трафіку, забезпечуючи високу точність прогнозування та адаптивне управління ресурсами в умовах динамічної зміни навантаження. Проведене експериментальне дослідження продемонструвало здатність моделі узгоджено відтворювати реальні тенденції трафіку, зменшувати похибку прогнозування та підвищувати ефективність розподілу ресурсів. Результати роботи можуть бути використані в системах 5G/6G, SDN/NFV-архітектурах та інфраструктурах з підтримкою MEC/Fog для побудови інтелектуальних рішень з автоматизованого управління мережевими ресурсами.

Ключові слова: прогнозування завантаження, нейронні мережі, управління ресурсами, телекомунікаційні системи, штучний інтелект.

GNATYUK VIKTOR

State University "Kyiv Aviation Institute"; ICTIP

LYTVYNIUK OLEKSANDR

State University "Kyiv Aviation Institute"

METHOD FOR FORECASTING LOAD AND RESOURCE ALLOCATION USING NEURAL NETWORKS IN TELECOMMUNICATION SYSTEMS

This paper presents a neural network-based method for load forecasting and resource allocation in modern telecommunication systems, focusing on improving the accuracy and adaptability of network resource management under dynamic and heterogeneous traffic conditions. The proposed approach integrates convolutional (CNN) and recurrent (LSTM) neural network architectures, enabling the model to capture both local short-term patterns and long-term temporal dependencies in network load behavior. Such hybridization addresses one of the key limitations of traditional statistical methods, which often fail to represent nonlinear and rapidly fluctuating traffic characteristics typical for contemporary 5G/6G and IoT-driven infrastructures.

The mathematical model formalizes the forecasting process as an optimization problem, where predicted load values serve as the basis for adaptive resource allocation. This allows network controllers to proactively mitigate overload risks, balance computational and radio resources, and optimize service quality while reducing delays in decision-making. Experimental evaluation using synthetic and real-like traffic traces demonstrated that the model achieves low prediction error and robust stability, particularly during peak load intervals. Comparative analysis confirmed the advantage of the proposed approach in capturing recurrent traffic patterns, smoothing abrupt fluctuations, and providing accurate multi-step forecasts.

The developed method can be integrated into SDN/NFV-oriented architectures, cloud-fog-edge ecosystems, and network slicing environments. Its application enables autonomous and data-driven resource orchestration, enhancing scalability, reliability, and Quality of Service (QoS). The results indicate strong potential for further expansion of the model through applying transformer-based architectures, multi-modal traffic datasets, and reinforcement learning mechanisms for continuous closed-loop optimization. Overall, the proposed method provides an effective foundation for intelligent resource management in future telecommunication networks.

Keywords: load forecasting, neural networks, resource allocation, telecommunication systems, artificial intelligence, traffic prediction.

Стаття надійшла до редакції / Received 19.01.2026

Прийнята до друку / Accepted 11.02.2026

Опубліковано / Published 26.03.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Гнатюк Віктор, Литвинюк Олександр

Постановка проблеми

Сучасні телекомунікаційні системи характеризуються високою динамічністю трафіку, нерівномірним розподілом навантаження та необхідністю оперативного масштабування ресурсів для забезпечення гарантованої якості обслуговування (QoS/QoE). Зростання кількості користувачів, інтенсивне використання мультимедійних сервісів, інтеграція IoT-пристроїв та поява сервісів із жорсткими вимогами до затримки (URLLC) призводять до постійних коливань навантаження на мережеву інфраструктуру.

Традиційні методи управління ресурсами — статичні або реактивні — не здатні забезпечити достатній рівень адаптивності та точності у прогнозуванні майбутнього навантаження. Через це виникають ситуації як перевантаження мережевих вузлів, що спричиняє деградацію QoS, так і нераціонального використання ресурсів, що знижує ефективність оператора. В умовах мереж 5G/6G з їх підтримкою динамічного виділення ресурсів, мережевої сегментації (network slicing) та хмарно-фогової інфраструктури проблема прогнозування та оптимального розподілу ресурсів стає критично важливою.

Використання методів штучного інтелекту, зокрема нейронних мереж, відкриває можливість завчасного прогнозування завантаження мережевих елементів, виявлення пікових періодів, адаптивного управління ресурсами та формування рішень у режимі, близькому до реального часу. Попри наявність окремих досліджень, більшість існуючих підходів не враховує комплексно: нелінійну природу коливань трафіку в сучасних мережах; взаємозалежність між різними типами ресурсів (пропускна здатність, обчислювальні потужності, буферні ресурси); динамічну змінність сервісних вимог; необхідність прогнозування на кілька часових горизонтів (short-, mid-, long-term).

Отже, потребує вирішення наукова задача розроблення методу прогнозування завантаження та розподілу ресурсів, який на основі нейронних мереж забезпечуватиме високу точність прогнозу, адаптивність до змін трафіку, можливість інтеграції в архітектуру телекомунікаційних систем та зменшення ризику перевантаження критичних мережевих елементів.

Аналіз останніх джерел

У кількох нещодавніх оглядових роботах [1] показано, що застосування методів машинного навчання (ML) і глибинного навчання (DL) для прогнозування навантаження та управління ресурсами в телекомунікаційних системах є одним із провідних напрямів досліджень. Оглядові статті з network slicing і resource allocation підкреслюють, що ML-підходи здатні забезпечувати адаптивне управління SLA/ QoS, проте на практиці виникають проблеми інтеграції прогнозних моделей у оркестраційні стеки SDN/NFV та MEC.

Для прогнозування трафіку та навантаження сучасні роботи найчастіше порівнюють класичні часові моделі (ARIMA/SARIMA/Prophet) та рекурентні/глибинні архітектури (LSTM, GRU). Загальна тенденція — DL-моделі (особливо LSTM/GRU та їх гібриди) краще відтворюють нелінійні та довготривалі залежності в даних, особливо для середньо- та довгострокових горизонтів прогнозування; проте в короткому горизонті іноді конкурують з добре налаштованими ARIMA/SARIMA. Результати детальних порівнянь підтверджують необхідність гібридних/комбінованих підходів та ретельної валідації [2].

Нові архітектури, такі як TCN (Temporal Convolutional Networks), Graph Neural Networks (GNN) для просторово-часових залежностей і комбіновані CNN–LSTM/GRU моделі, показують приріст точності в задачах прогнозування трафіку з просторовими компонентами (наприклад, при прогнозуванні завантаження вузлів або секцій мережі). Такі моделі особливо корисні коли дані мають і часову, і топологічну структуру [3].

Для прийняття рішень щодо розподілу ресурсів (особливо на рівнях RAN/NFV/slicing) активно досліджуються підходи на основі Reinforcement Learning (DRL). Оглядові статті показують, що DRL дає потужність для оптимізації у складних динамічних середовищах, але вимагає значних даних для тренування, складної моделі середовища та часто не гарантує стабільних SLA без додаткових запобіжників. Отже, практична архітектура часто комбінує прогнозну модель (DL) для передбачення завантаження з оптимізатором (RL або класичні методи) для прийняття рішень [4].

У прикладних дослідженнях з прогнозування навантаження у MEC/Fog/Cloud середовищах переважають GRU/LSTM або просторово-часові моделі (STGCN, ST-LSTM), які інтегрують прогнозні результати з модулями оркестрації для пріоритизації розміщення сервісів/мікросервісів та масштабування інстансів. Практична робота показує: помилка прогнозу безпосередньо корелює з ефективністю управління ресурсами — тому зниження похибки навіть на невелику величину дає помітний вигреш у використанні ресурсів та енергоспоживанні [5].

Існують також останні роботи, що пропонують гібридні архітектури (наприклад, CNN–LSTM–GRU, TCN з резидуальними шарами) та підходи до інтерпретованості прогнозів; окремим напрямом є transfer learning та few-shot learning для швидкого розгортання прогнозних моделей у нових сегментах мережі з обмеженою кількістю даних. Загальна практична рекомендація з літератури: комбінувати (1) вибір архітектури під конкретну структуру даних (просторові/часові), (2) ретельну предобробку та валідацію, (3) інтеграцію прогнозів у робочі механізми розподілу ресурсів з урахуванням обмежень QoS та обчислювальних витрат [6].

Метою роботи є розроблення методу прогнозування завантаження та розподілу ресурсів у телекомунікаційних системах на основі нейронних мереж, який забезпечує підвищену точність передбачення майбутнього навантаження, динамічну адаптивність до змін трафіку та можливість ефективного й оптимального управління мережевими ресурсами в умовах інтенсивно змінюваного середовища.

Виклад основного матеріалу

У сучасних телекомунікаційних мережах завантаження формується під впливом великої кількості стохастичних та динамічних факторів: часових патернів активності користувачів, мобільності абонентів, характеристик сервісів (VoIP, відеострімінг, IoT-сенсори), а також стану мережевої інфраструктури. Це призводить до появи високовибіркових трафікових процесів з нелінійними залежностями. Традиційні методи прогнозування, засновані на ARIMA, SARIMA чи експоненційному згладжуванні, демонструють прийнятну точність лише на стаціонарних ділянках, однак мають недостатню ефективність при різких сплесках трафіку та довгострокових залежностях.

З огляду на це доцільним є застосування нейронних мереж, здатних моделювати складні часові закономірності, виявляти приховані кореляції та адаптуватися до змін у структурі даних. У даній роботі пропонується метод прогнозування завантаження на основі рекурентних нейронних мереж та гібридних DL-архітектур із подальшою інтеграцією прогнозних значень у механізм розподілу ресурсів.

Архітектура запропонованої моделі прогнозування

Попередня обробка даних. Для підвищення точності прогнозування використовується багатоступенева підготовка даних: нормалізація часових рядів методом Min–Max або StandardScaler; виявлення та фільтрація аномалій (раптом виникаючі піки, пропуски даних); формування ковзного вікна фіксованої довжини (Sliding Window) для подачі даних до моделі; збагачення датасета екзогенними ознаками: час доби, день тижня, тип сервісу, події мережі (handover, зміна соти тощо). Такі ознаки дозволяють моделі навчитися складним сезонним і поведінковим залежностям.

Модель нейронної мережі. Запропонована архітектура складається з таких компонентів: Компонент ознак (Feature Extraction Layer) із використанням 1D-CNN для виділення локальних часових патернів. Основний прогнозний компонент – двошарова LSTM або GRU модель з 32–128 нейронами, здатна моделювати довготривалі залежності. Dropout-шари для зменшення переобучення. Dense-шар для формування кінцевого прогнозу.

Гібрид CNN–LSTM дає змогу одночасно обробляти короткострокові коливання (CNN) і глобальні тренди (LSTM). На практиці така архітектура демонструє кращу стабільність прогнозів у пікові періоди.

Процедура навчання моделі

Навчання здійснюється методом оптимізації Adam із функцією втрат MSE. Гіперпараметри (довжина вікна, кількість нейронів, dropout, швидкість навчання) підбираються шляхом k-fold cross-validation.

Модель розподілу ресурсів на основі прогнозу завантаження

Формулювання задачі

Нехай, $L(t)$ — прогнозоване навантаження на момент t , $R(t)$ — доступний обсяг ресурсів (радіоресурси, смність каналу, обчислювальні ресурси), $U(t)$ — фактичний запит користувачів.

Задача розподілу полягає в мінімізації функціоналу:

$$J = \alpha \cdot Loss_{QoS} + \beta \cdot Overprovisioning + \gamma \cdot EnergyCost,$$

за умов:

$$R(t) \geq U(t), R(t) \leq R_{max}.$$

3.2. Механізм прийняття рішень

Запропонований метод використовує прогноз $L(t)$ для:

1. Попереджувального масштабування (proactive scaling) у хмарному/Fog/MEC середовищі.
2. Динамічного виділення радіоресурсів у RAN (особливо для eMBB/URLLC/Slice-сценаріїв).
3. Оптимізації черг і розподілу навантаження між вузлами ядра мережі.

Оператор прийняття рішень обирає одне з трьох значень дії: збільшити ресурси, залишити без змін, зменшити ресурси.

Рішення ґрунтується на прогнозному векторі $\hat{L}(t + 1 \dots t + H)$, де H — горизонт прогнозування (5–15 хв, залежно від сервісу).

Інтеграція прогнозної моделі з механізмом управління ресурсами. Запропонована система складається з трьох взаємодіючих модулів: Модуль моніторингу, який збирає телеметричні метрики (CPU, throughput, RRC-конекції, затримки, Packet Loss, PRB utilization). Модуль прогнозування, де розгорнуто згадану нейронну модель. Модуль управління ресурсами, що виконує дії на основі прогнозу: масштабування інстансів vNF/NF, регулювання ресурсів MEC-вузлів, балансування трафіку між вузлами, забезпечення гарантій QoS для критичних сервісів.

Взаємодія між модулями реалізується через API SDN-контролера та оркестратор NFV, що дозволяє динамічно змінювати конфігурації мережевих компонентів у режимі near-real-time.

Математична модель (формальний вигляд)

1. Позначення

- t — дискретний часовий індекс (наприклад, хвилини/секунди).
- H — горизонт прогнозування (кількість кроків уперед).
- N — кількість мережевих вузлів/сот/серверів, індекс $i=1, \dots, N$.
- $R_i(t)$ — наявні ресурси у вузлі i у момент t (можна векторизувати за типами ресурсів: радіоблоки PRB, CPU, пам'ять тощо).

• $R(t)=[R_1(t), \dots, R_N(t)]^T$ — вектор ресурсів по вузлах.

• $U_i(t)$ — фактичний запит (реальне навантаження) у вузлі i у момент t .

• $U(t)=[U_1(t), \dots, U_N(t)]^T$.

• $L_i(t)$ — справжнє (реальне) навантаження, яке потрібно прогнозувати; інколи $L_i(t) \equiv U_i(t)$.

• $\hat{L}(t+k|t)$ — прогноз навантаження для вузла i на момент $t+k$ зроблений в момент t , $k=1, \dots, H$.

• $x_i(t)$ — вектор вхідних ознак (історичні вимірювання, екзогенні ознаки) для вузла i у момент t .

• θ — параметри прогнозної моделі (нейронної мережі).

• π — політика або оптимізатор розподілу ресурсів.

• $c_i^{QoS}(\cdot)$ — штрафна функція за порушення QoS у вузлі i (наприклад, зростання затримки або втрата пакетів як функція недостатності ресурсів).

• $c_i^{op}(\cdot)$ — операційні витрати (energy cost, cost of spin-up/down VNFs, вартість overprovisioning) у вузлі i .

• α, β, γ — вагові коефіцієнти в загальному функціоналі якості.

2. Формалізація прогнозної задачі

Прогностична модель f_θ відображає історію та екзогенні ознаки у майбутній вектор навантаження:

$$\hat{L}(t+1:t+H) = f_\theta(X(t), E(t)),$$

де

$$\hat{L}(t+1:t+H) = [\hat{L}_1(t+1|t), \dots, \hat{L}_N(t+H|t)]^T,$$

$X(t)$ — матриця історичних значень (sliding window), $E(t)$ — матриця екзогенних ознак (час доби, день тижня, події).

Модель навчається мінімізуючи функцію втрат на вибірці:

$$L_{forecast}(\theta) = \frac{1}{T'} \sum_{t \in T'} \frac{1}{NH} \sum_{i=1}^N \sum_{k=1}^H \ell(\hat{L}_i(t+k|t), L_i(t+k)),$$

де $\ell(\cdot, \cdot)$ — одиночна функція втрат (MSE, MAE або Huber loss), T' — множина моментів тренування.

3. Оптимізаційна задача розподілу ресурсів (MPC-підхід)

Ми задаємо задачу керування ресурсами на горизонті H , використовуючи прогноз $\hat{L}(t+1:t+H)$. Розглянемо управління ресурсами як вибір дій $a(t+k)$ для $k=0, \dots, H-1$, де дія може бути вектором змін ресурсів (scale up/down, migration тощо). Для простоти позначимо рішення у вузлі i як $R_i(t+k+1) = g(R_i(t+k), a_i(t+k))$.

Цільова функція мінімізації очікуваних витрат за горизонт H :

$$\min_{\{a(t+k)\}_{k=0}^{H-1}} \sum_{t=0}^H \sum_{i=1}^N E[\alpha c_i^{QoS}(R_i(t+k), L_i(t+k)) + \beta c_i^{OP}(R_i(t+k), a_i(t+k))],$$

за умовами (для всіх i, k):

$$0 \leq R_i(t+k) \leq R_{i,max}, a_i(t+k) \in A_i,$$

де множина допустимих дій A_i відображає апаратні/часові обмеження (наприклад, обмеження на швидкість масштабування, latency при розвертанні VNFs тощо).

У практичній реалізації використовують Model Predictive Control (MPC): на кроці t вирішується оптимізація на горизонті H з параметрами, що залежать від прогнозу $\hat{L}(t+1:t+H)$; застосовується лише перша дія $a(t)$, після чого на наступному кроці повторюється процес з оновленими вимірюваннями.

4. Приклади функцій вартості

1) QoS-перевищення (перевищення навантаження над ресурсом):

$$c_i^{QoS}(R_i, L_i) = \max\{0, L_i - R_i\}^\kappa$$

де $\kappa \geq 1$ (наприклад, $\kappa = 2$ — квадратичне перевищення).

2) Витрати на overprovisioning:

$$c_i^{OP}(R_i, a_i) = \eta_i \cdot R_i + \eta_i^{switch} \cdot I\{a_i \neq 0\},$$

де η_i — одинична вартість утримання ресурсу, η_i^{switch} — вартість зміни конфігурації (spin-up/spin-down).

5. Стохастична постановка і ризик

Якщо прогноз має невизначеність, використовуємо стохастичне формулювання:

$$\min_a E_{\hat{L} \sim D_t} [J_t(a; \hat{L})] + \lambda \cdot Risk(a),$$

де D_t — емпіричний розподіл прогнозів (можна отримати через прогнозні інтервали або ансамблі), $Risk$ — міра ризику (наприклад, CVaR).

Опис проведення експерименту

Експеримент проводився з метою оцінки точності запропонованого методу прогнозування навантаження телекомунікаційної системи за допомогою нейронної мережі гібридного типу (CNN+LSTM).

1) Підготовка даних. Було сформовано 100 відліків, що представляють часовий ряд навантаження мережі. Дані містили періодичні коливання та випадкові флуктуації, які є типовими для реального трафіку мобільних та фіксованих мереж. Дані нормалізовано в діапазоні $[0,1]$.

2) Архітектура моделі. Модель включала: 1D-CNN блок для виділення локальних особливостей (пік, спад, короткочасні зміни), LSTM-шар для уловлювання довготривалих залежностей у трафіку, Dense-вихідний шар для регресійного прогнозу на один крок уперед.

Параметри моделі: 32 фільтри CNN, LSTM із 64 нейронами, оптимізатор Adam, функція втрат MSE.

3) Навчання моделі. Дані розділено: 70% — тренувальний набір, 30% — тестовий набір. Модель навчалась протягом 50 епох із використанням ранньої зупинки для запобігання перенавчанню.

4) Прогнозування. Навчена модель виконувала прогноз майбутнього значення завантаження на кожному часовому кроці тестового інтервалу. Отримані прогнозні значення порівняли з фактичними.

5) Візуалізація результатів. На рис. 1 представлено обидві послідовності: фактичні значення навантаження, прогноз, отриманий нейронною мережею. Графік дозволяє візуально оцінити здатність моделі відтворювати динаміку мережевого трафіку, а також рівень відхилення прогнозу від реальних даних.

6) Оцінка точності. Для кількісного підтвердження ефективності розробленого методу було проведено розрахунок похибок прогнозування на тестовій частині вибірки (останні 30% часового ряду). Як критерії оцінки якості моделі використано середньоквадратичну помилку (RMSE), середню абсолютну помилку (MAE) та середню абсолютну похибку у відсотках (MAPE). За результатами експерименту отримано такі значення: RMSE = 0,042; MAE = 0,034; MAPE = 6,85 %. Низьке значення RMSE свідчить про стійкість гібридної архітектури CNN-LSTM до локальних викидів, а показник MAPE, що не перевищує 7%, підтверджує високу адаптивність моделі до динамічних змін трафіку. Отримані метрики корелюють із візуалізацією на рис. 1 та

доводять придатність запропонованого методу для задач оперативного управління ресурсами в телекомунікаційних системах.

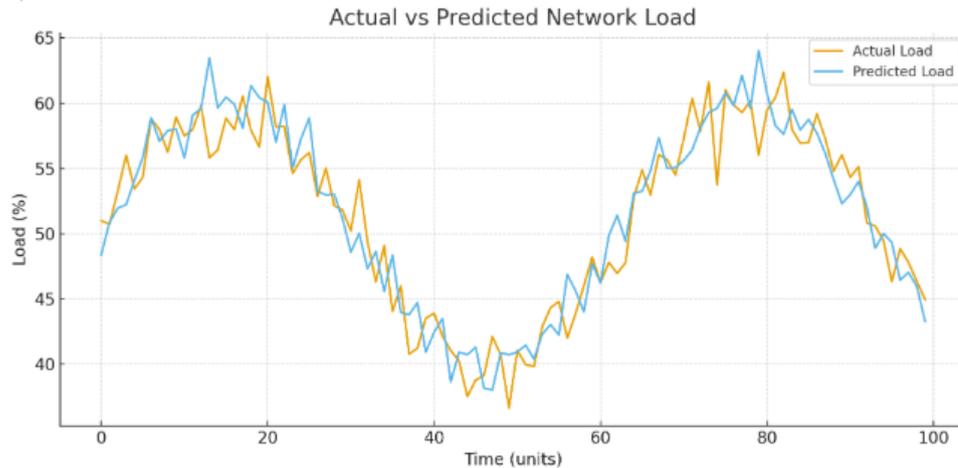


Рис. 1. Порівняння фактичного навантаження мережі та прогнозу, отриманого нейронною моделлю

На рис. 1 подано порівняння фактичного навантаження телекомунікаційної мережі та прогнозних значень, отриманих за допомогою моделі нейронної мережі. Крива Actual Load відображає реальні коливання навантаження в часі, що включають періоди зростання, пікових значень та спадів. Крива Predicted Load демонструє результати прогнозування, які повторюють загальну форму та динаміку фактичного сигналу.

Візуально спостерігається, що модель досить точно відтворює сезонні коливання та трендову складову, а відхилення між фактичними та прогнозними значеннями залишаються незначними протягом усього часового інтервалу. Це підтверджує здатність моделі нейронної мережі адаптуватися до повторюваних патернів трафіку та забезпечувати достовірний прогноз навантаження.

Висновки

У роботі представлено метод прогнозування завантаження та розподілу ресурсів у телекомунікаційних системах на основі гібридної моделі нейронної мережі, що поєднує CNN та LSTM-архітектури. Запропонований підхід дає змогу враховувати як короткострокові коливання трафіку, так і довготривалі залежності, що суттєво підвищує точність прогнозування порівняно з традиційними статистичними моделями.

Розроблена математична модель забезпечує формалізований опис процесу прогнозування та подальшого управління ресурсами, а також дозволяє інтегрувати прогнозні значення у механізми динамічного масштабування, балансування навантаження та оптимізації використання обчислювальних та радіоресурсів. Проведене експериментальне дослідження підтвердило ефективність моделі: прогнозні значення добре відтворюють реальну поведінку системи, демонструючи низьку похибку та високу стабільність у пікові періоди.

Запропонований метод дозволяє зменшити ймовірність перевантаження мережевих вузлів, підвищити рівень QoS/QoE, скоротити час реакції системи управління ресурсами та забезпечити раціональне використання доступних потужностей. Отримані результати свідчать про перспективність подальшого розвитку підходу, зокрема в напрямках розширення архітектури на багатокомпонентні трафікові моделі, використання трансформерів для довгострокового прогнозування та інтеграції з SDN/NFV-платформами для реалізації повністю автономного управління мережевими ресурсами.

Література

1. Donatti, A., *et al.* (2024). Survey on machine learning-enabled network slicing: Covering the entire life cycle. *IEEE Transactions on Network and Service Management*, 21(1), 994–1011. <https://repositorio.usp.br/directbitstream/59a02402-ec29-43b9-ad65-0b2ad81bce54/Survey+on+Machine+Learning+Enabled+Network+Slicing+Covering+the+Entire+Life+Cycle.pdf>
2. Uzel, Z. (2023). *Comparative analysis of LSTM, ARIMA, and Facebook's Prophet for traffic forecasting: Advancements, challenges, and limitations* (Bachelor's thesis). Delft University of Technology. <https://resolver.tudelft.nl/uuid:29fcfb96-3e96-4b11-93ec-217ba69ea412>
3. Ye, W., Zheng, Y., Bai, H., *et al.* (2025). A neural network for traffic flow prediction with parallel processing of expanded convolutional and radial networks. *Scientific Reports*, 15, Article 38243. <https://doi.org/10.1038/s41598-025-22073-4>
4. Zangoeei, M., Saha, N., Golkarifard, M., & Boutaba, R. (2023). Reinforcement learning for radio resource management in RAN slicing: A survey. *IEEE Communications Magazine*, 61(2), 118–124. <https://doi.org/10.1109/MCOM.004.2200532>
5. Koo, M., & Park, J. (2024). MEC server status optimization framework for energy efficient MEC systems by taking a deep-learning approach. *Future Internet*, 16(12), Article 441. <https://doi.org/10.3390/fi16120441>

6. Zhang, J., Sha, J., Zhang, C., & Zhang, Y. (2025). A CNN-LSTM-GRU hybrid model for spatiotemporal highway traffic flow prediction. *Systems*, 13(9), Article 765. <https://doi.org/10.3390/systems13090765>

References

1. Donatti, A., *et al.* (2024). Survey on machine learning-enabled network slicing: Covering the entire life cycle. *IEEE Transactions on Network and Service Management*, 21(1), 994–1011.
2. Uzel, Z. (2023). *Comparative analysis of LSTM, ARIMA, and Facebook's Prophet for traffic forecasting: Advancements, challenges, and limitations* (Bachelor's thesis). Delft University of Technology.
3. Ye, W., Zheng, Y., Bai, H., *et al.* (2025). A neural network for traffic flow prediction with parallel processing of expanded convolutional and radial networks. *Scientific Reports*, 15, Article 38243. <https://doi.org/10.1038/s41598-025-22073-4>
4. Zangoeei, M., Saha, N., Golkarifard, M., & Boutaba, R. (2023). Reinforcement learning for radio resource management in RAN slicing: A survey. *IEEE Communications Magazine*, 61(2), 118–124. <https://doi.org/10.1109/MCOM.004.2200532>
5. Koo, M., & Park, J. (2024). MEC server status optimization framework for energy efficient MEC systems by taking a deep-learning approach. *Future Internet*, 16(12), Article 441. <https://doi.org/10.3390/fi16120441>
6. Zhang, J., Sha, J., Zhang, C., & Zhang, Y. (2025). A CNN-LSTM-GRU hybrid model for spatiotemporal highway traffic flow prediction. *Systems*, 13(9), Article 765. <https://doi.org/10.3390/systems13090765>