ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ 9
ВСТУП 11
1. ОСНОВИ ШТУЧНОГО ІНТЕЛЕКТУ У КОНТЕКСТІ ТЕЛЕКОМУНІКАЦІЙ . 13
1.1. Сутність, класифікація та математичні принципи штучного інтелекту .. 13
1.2. Методи машинного навчання та глибинного навчання для аналізу трафіку 18
1.3. Моделі прогнозування навантаження в телекомунікаційних мережах 23
1.4. Інтелектуальні методи виявлення аномалій і нерегулярної активності в мережах 29
2. АРХІТЕКТУРА ТА ТЕХНІЧНІ ПАРАМЕТРИ МЕРЕЖ П’ЯТОГО ПОКОЛІННЯ 36
2.1. Технічні основи побудови мереж 5G та принципи функціонування 36
2.2. Архітектура радіодоступу 5G NR (New Radio) та характеристика мережевих сценаріїв 42
2.3. Технології мережевого поділу (network slicing) у 5G 48
2.4. Механізми керування трафіком та забезпечення QoS у 5G 55
3. ДОСЛІДЖЕННЯ ІНТЕГРАЦІЇ ШТУЧНОГО ІНТЕЛЕКТУ В МЕРЕЖІ 5G ДЛЯ ОПТИМІЗАЦІЇ ТРАФІКУ 61
3.1. Методологія інтеграції AI-моделей у мережеву інфраструктуру 5G 61
3.2. Алгоритми оптимізації трафіку на основі штучного інтелекту 69
3.3. Експериментальне моделювання та оцінювання ефективності інтеграції AI у 5G 79
ВИСНОВКИ 87
ПЕРЕЛІК ПОСИЛАНЬ 89
ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ 93

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

3GPP	3rd Generation Partnership Project (Партнерський проект третього покоління)
5G NR	5G New Radio (Нове радіо 5G)
AI / ШІ	Artificial Intelligence (Штучний інтелект)
API	Application Programming Interface (Інтерфейс прикладного програмування)
CNN	Convolutional Neural Network (Згорткова нейронна мережа)
CQI	Channel Quality Indicator (Індикатор якості каналу)
CU	Central Unit (Центральний блок)
DDoS	Distributed Denial of Service (Розподілена атака типу «відмова в обслуговуванні»)
DQN	Deep Q-Network (Глибока Q-мережа)
DRL	Deep Reinforcement Learning (Глибоке навчання з підкріпленням)
DSS	Dynamic Spectrum Sharing (Динамічний спільний доступ до спектру)
DU	Distributed Unit (Розподілений блок)
eMBB	Enhanced Mobile Broadband (Покращений мобільний широкосмуговий доступ)
EPC	Evolved Packet Core (Еволюціонувавше пакетне ядро — 4G)
FR1 / FR2	Frequency Range 1 / 2 (Частотний діапазон 1 / 2)
gNodeB	Next Generation NodeB (Базова станція 5G)
GNN	Graph Neural Network (Графова нейронна мережа)
GRU	Gated Recurrent Unit (Керований рекурентний блок)
IoT	Internet of Things (Інтернет речей)
KPI	Key Performance Indicators (Ключові показники ефективності)
LSTM	Long Short-Term Memory (Довга короткочасна пам'ять)
LTE	Long-Term Evolution (Стандарт бездротового зв'язку 4G)
MEC	Multi-access Edge Computing (Багатоканальні периферійні обчислення)
MIMO	Multiple Input Multiple Output (Множинний вхід — множинний вихід)
ML / МН	Machine Learning (Машинне навчання)
mMTC	Massive Machine Type Communications (Масивний машинний зв'язок)
mmWave	Millimeter Wave (Міліметрові хвилі)

NFV	Network Functions Virtualization (Віртуалізація мережевих функцій)
NGC	Next Generation Core (Ядро мережі наступного покоління)
NTN	Non-Terrestrial Networks (Наземні та неназемні мережі / супутникові)
PRB	Physical Resource Block (Блок фізичних ресурсів)
QoE	Quality of Experience (Якість сприйняття / досвіду користувача)
QoS	Quality of Service (Якість обслуговування)
RAN	Radio Access Network (Мережа радіодоступу)
RL	Reinforcement Learning (Навчання з підкріпленням)
RNN	Recurrent Neural Network (Рекурентна нейронна мережа)
RSRP	Reference Signal Received Power (Потужність прийнятого опорного сигналу)
SDN	Software-Defined Networking (Програмно-конфігуровані мережі)
SLA	Service Level Agreement (Угода про рівень обслуговування)
UE	User Equipment (Користувацьке обладнання)
URLLC	Ultra-Reliable Low Latency Communications (Наднадійний зв'язок з низькою затримкою)

ВСТУП

Актуальність теми. Стрімкий розвиток цифрової економіки та перехід до суспільства «Інтернету всього» (Internet of Everything) зумовили фундаментальну трансформацію вимог до телекомунікаційних мереж. Поява та стандартизація технології 5G (IMT-2020) стала відповіддю на ці виклики, знаменуючи перехід від простого забезпечення мобільного інтернету до створення універсальної платформи для цифровізації ключових галузей промисловості, транспорту та соціальної сфери. На відміну від попередніх поколінь, архітектура 5G характеризується безпрецедентною складністю та гетерогенністю, об'єднуючи в єдиній екосистемі технології масивного MIMO, міліметрових хвиль (mmWave), програмно-конфігурованих мереж (SDN) та віртуалізації мережевих функцій (NFV).

Ключовою проблемою експлуатації таких мереж є експоненціальне зростання обсягів трафіку та кількості підключених пристроїв, що, за прогнозами, сягатиме мільйонів підключень на квадратний кілометр у сценаріях mMTC. Водночас мережа повинна забезпечувати суперечливі вимоги для різних категорій сервісів (слайсів): надвисоку пропускну здатність для eMBB, мінімальну затримку (менше 1 мс) для критичних застосувань URLLC та енергоефективність для IoT-сенсорів. Традиційні детерміновані методи керування трафіком, засновані на статичних правилах, евристичних алгоритмах або ручній конфігурації, виявляються неспроможними ефективно вирішувати задачі розподілу ресурсів у такій динамічній та багатопараметричній системі. Вони діють реактивно, реагуючи на перевантаження вже після їх виникнення, що є неприпустимим для сервісів реального часу, таких як автономний транспорт чи дистанційна медицина.

У цьому контексті інтеграція методів штучного інтелекту (ШІ) та машинного навчання (МН) в інфраструктуру 5G перестає бути опцією і стає технологічною необхідністю для реалізації концепції самоорганізованих мереж (Self-Organizing Networks — SON) та переходу до парадигми «Zero-Touch Service Management». Застосування алгоритмів глибокого навчання (Deep Learning) та

навчання з підкріпленням (Reinforcement Learning) дозволяє перейти від реактивного до проактивного керування. Інтелектуальні агенти здатні аналізувати приховані кореляції у великих масивах телеметричних даних, прогнозувати просторово-часові коливання трафіку з високою точністю та здійснювати превентивну реконфігурацію мережі за частки секунди.

Особливої актуальності набуває дослідження синергії ШІ з технологіями периферійних обчислень (Multi-access Edge Computing — MEC), що дозволяє перенести інтелект мережі ближче до абонента, знижуючи навантаження на транспортну мережу та забезпечуючи прийняття рішень у реальному часі. Розробка та впровадження ефективних алгоритмів ШІ для оптимізації трафіку дозволить не лише гарантувати дотримання жорстких угод про рівень обслуговування (SLA), але й суттєво підвищити енергоефективність мережі, знизити операційні витрати (OPEX) операторів та забезпечити стійкість інфраструктури до пікових навантажень. Таким чином, наукове завдання щодо дослідження та розробки методів інтеграції ШІ в мережі 5G для оптимізації трафіку є своєчасним, має важливе науково-практичне значення та відповідає глобальним тенденціям розвитку інфокомунікацій.

Метою роботи є підвищення ефективності функціонування мереж 5G шляхом розробки та дослідження методу динамічної оптимізації трафіку на основі алгоритмів глибокого навчання з підкріпленням.

Об’єкт дослідження — процеси керування трафіком та розподілу радіоресурсів у мобільних мережах п'ятого покоління.

Предмет дослідження — методи та алгоритми штучного інтелекту, зокрема глибоке навчання з підкріпленням, для оптимізації параметрів якості обслуговування (QoS) у мережах 5G.

1. ОСНОВИ ШТУЧНОГО ІНТЕЛЕКТУ У КОНТЕКСТІ ТЕЛЕКОМУНІКАЦІЙ

1.1. Сутність, класифікація та математичні принципи штучного інтелекту

Штучний інтелект (ШІ) став одним із найбільш трансформаційних технологічних досягнень 21-го століття, фундаментально змінюючи галузі промисловості, економіку та повсякденне життя. По суті, ШІ стосується розробки комп'ютерних систем, здатних виконувати завдання, які зазвичай потребують людського інтелекту, такі як розуміння мови, розпізнавання закономірностей, вирішення проблем та прийняття рішень. Фундаментальна мета ШІ полягає в тому, щоб імітувати когнітивні функції людини для покращення автоматизації, покращення процесів прийняття рішень та вирішення складних завдань ефективніше, ніж лише люди. У широкому сенсі ШІ поділяється на два основні типи: вузький ШІ, який призначений для виконання конкретних завдань, та загальний ШІ, який прагне мати широкий, адаптивний інтелект, порівнянний з людським. У цьому есе досліджується сутність ШІ, його класифікація на основі можливостей та функціональності, а також математичні принципи, що лежать в основі його розвитку, забезпечуючи всебічне розуміння цієї динамічної галузі [1].

Суть штучного інтелекту полягає у створенні систем, які можуть імітувати різні аспекти людського інтелекту. По суті, ШІ передбачає розробку алгоритмів та обчислювальних моделей, які дозволяють машинам сприймати, міркувати, навчатися та адаптуватися до нової інформації. Кінцева мета — відтворити здібності до прийняття рішень та вирішення проблем, властиві людям. Наприклад, віртуальні помічники на базі ШІ, такі як Siri та Alexa, використовують обробку природної мови та машинне навчання для інтерпретації запитів користувачів та відповідної відповіді, що є прикладом здатності вузького ШІ виконувати спеціалізовані завдання. І навпаки, концепція загального ШІ спрямована на те, щоб машини мали широкий спектр когнітивних навичок, подібних до людських,

що дозволяє їм розуміти, навчатися та застосовувати знання в різних областях без спеціального програмування для кожного завдання. Ця відмінність є критично важливою, оскільки сучасні системи ШІ переважно належать до вузької категорії ШІ, яка перевершує певні функції, але не має гнучкого розуміння, характерного для людського інтелекту.



Рис. 1.1. Класифікація та математичні принципи штучного інтелекту

Розуміння цієї основної сутності прояснює поточні виклики та прагнення в дослідженнях ШІ, підкреслюючи важливість розробки систем, які не лише імітують дії людини, але й розуміють контекст, адаптуються до непередбачених обставин та демонструють міркування, подібне до людського пізнання. Штучний інтелект можна класифікувати за кількома напрямками, головним чином на основі можливостей та функціональності. З точки зору можливостей, системи ШІ зазвичай поділяються на вузький ШІ, який виконує специфічні завдання, такі як переклад мови або розпізнавання облич; загальний ШІ, який мав би здатність розуміти, навчатися та застосовувати знання в широкому спектрі діяльності, подібно до людського інтелекту; та надрозумний ШІ, гіпотетичну форму, яка

перевершує людський інтелект у всіх аспектах, що потенційно може призвести до глибокого суспільного впливу. Наприклад, сучасні голосові помічники є вузьким ШІ, тоді як прагнення до загального ШІ передбачає створення машин з гнучким, автономним мисленням. Крім того, системи ШІ класифікуються на основі їхньої функціональності, починаючи від реактивних машин, які реагують виключно на поточні стимули без пам'яті, до систем з обмеженою пам'яттю, здатних навчатися на минулих даних, до систем теорій розуму, які розуміють людські емоції та наміри, і, нарешті, до самосвідомих систем, які мають свідомість. Практичні реалізації цих класифікацій включають експертні системи на основі правил, які дотримуються заздалегідь визначених правил, алгоритми машинного навчання, які покращують продуктивність завдяки досвіду, та моделі глибокого навчання, які використовують нейронні мережі для розпізнавання складних закономірностей. Ці класифікації допомагають уточнити рамки розвитку ШІ, спрямовуючи дослідницькі зусилля та технологічні застосування до все більш складних та потужних систем. Розвиток та вдосконалення штучного інтелекту фундаментально ґрунтуються на складних математичних принципах, які дозволяють машинам обробляти інформацію, навчатися та приймати рішення [2]. Лінійна алгебра є центральною для нейронних мереж, які є основою багатьох моделей ШІ; такі концепції, як матричні операції, сприяють ефективному обчисленню величезних обсягів даних, дозволяючи нейронним мережам розпізнавати складні закономірності у зображеннях, мові та тексті. Наприклад, згорткові нейронні мережі використовують матричні операції для ідентифікації ознак у візуальних даних, що дозволяє використовувати такі програми, як автономна навігація транспортних засобів. Теорія ймовірностей, зокрема байєсівський висновок, відіграє ключову роль у прийнятті рішень в умовах невизначеності. Системи ШІ часто включають ймовірнісні моделі для оцінки ймовірності різних результатів, що є важливим у таких галузях, як медична діагностика або фінансове прогнозування. Алгоритми оптимізації, такі як градієнтний спуск, є критично важливими під час процесу навчання моделей ШІ, ітеративно коригуючи параметри, щоб мінімізувати помилки та підвищити

точність. Наприклад, моделі глибокого навчання використовують градієнтний спуск для точного налаштування мільйонів ваг у нейронних мережах, що дозволяє їм вивчати складні функції. Разом ці математичні принципи утворюють основу, на якій побудовані системи ШІ, стимулюючи їхню здатність навчатися на основі даних, адаптуватися та виконувати складні завдання зі збільшенням точності. У сфері штучного інтелекту різноманітні алгоритми та моделі слугують основними інструментами, що дозволяють машинам навчатися, адаптуватися та приймати рішення. Ці алгоритми можна загалом розділити на навчання з учителем, навчання без учителя та навчання з підкріпленням, кожен з яких має різні механізми та застосування. Навчання з учителем включає навчання моделей на маркованих наборах даних для прогнозування результатів, прикладами яких є такі застосування, як виявлення спаму та класифікація зображень. Навчання без учителя, з іншого боку, має справу з немаркованими даними для виявлення прихованих закономірностей або груп, таких як сегментація клієнтів у маркетингу. Навчання з підкріпленням моделює процеси прийняття рішень, де агенти навчаються оптимальним діям методом спроб і помилок, що часто використовується в робототехніці та ігровому штучному інтелекті. Еволюція цих моделей була вражаючою — від простих дерев рішень та лінійних моделей до складних нейронних мереж та архітектур глибокого навчання. Нейронні мережі, натхненні біологічними нейронами, утворюють основу багатьох сучасних систем штучного інтелекту, здатних розпізнавати складні закономірності в даних. Дерева рішень пропонують прозорі та інтерпретовані шляхи прийняття рішень, тоді як методи опорних векторів чудово працюють у багатовимірних просторах, забезпечуючи стійкість у завданнях класифікації. З часом ці моделі пройшли шлях від поверхневих систем, заснованих на правилах, до фреймворків глибокого навчання, які використовують кілька шарів нейронних мереж, що дозволяє штучному інтелекту досягати безпрецедентного рівня продуктивності в таких завданнях, як обробка природної мови, розпізнавання зображень та автономні системи. Ця еволюція підкреслює динамічний характер розвитку штучного інтелекту, зумовлений інноваціями в алгоритмах та обчислювальній потужності.

Незважаючи на значний прогрес, досягнутий у сфері штучного інтелекту, залишаються численні проблеми та майбутні напрямки, особливо щодо математичних моделей, що лежать в основі цих систем. Сучасні математичні фреймворки, хоча й потужні, мають помітні обмеження у повному відображенні нюансів людського інтелекту, таких як свідомість, інтуїція та емоційне розуміння. Ці недоліки підкреслюють необхідність розробки нових математичних підходів, які можуть краще моделювати складні когнітивні процеси та сприяти розвитку більш просунутих, адаптивних систем штучного інтелекту. Крім того, оскільки системи штучного інтелекту стають все більш інтегрованими в критичні суспільні функції, етичні міркування, пов'язані з їхніми математичними процесами прийняття рішень, набувають важливості. Наприклад, алгоритми, які визначають результати, що змінюють життя, такі як медичні діагнози чи рішення у сфері кримінального правосуддя, повинні бути прозорими та справедливими, що викликає питання щодо упередженості, відповідальності та інтерпретованості в рамках математичних моделей. Прагнення до більш складних математичних структур спрямоване не лише на покращення можливостей штучного інтелекту, але й на забезпечення етичної та відповідальної роботи цих систем. Дослідники досліджують міждисциплінарні підходи, поєднуючи знання з когнітивної науки, філософії та вищої математики, щоб розробити моделі, які можуть точніше імітувати людське мислення, дотримуючись при цьому етичних стандартів. Оскільки штучний інтелект продовжує розвиватися, вирішення цих проблем є важливим для сприяння створенню систем, які є одночасно високопродуктивними та відповідають суспільним цінностям [3].

Підсумовуючи, штучний інтелект – це складна та швидко розвиваюча галузь, основною метою якої є імітація людського інтелекту за допомогою складних алгоритмів та математичних принципів. Її суть полягає у створенні систем, здатних до сприйняття, міркування, навчання та прийняття рішень, причому сучасні технології в основному зосереджені на вузькому ШІ, водночас прагнучи до ширших, більш адаптивних форм, таких як загальний ШІ. Класифікація ШІ на основі можливостей та функціональних можливостей

демонструє різноманітність підходів, від реактивних машин до самосвідомих систем, та підкреслює перехід від моделей, заснованих на правилах, до передових архітектур глибокого навчання. В основі цих розробок лежать фундаментальні математичні принципи – лінійна алгебра, теорія ймовірностей та алгоритми оптимізації – які дозволяють машинам навчатися на даних та виконувати складні завдання. Заглядаючи в майбутнє, проблеми точного моделювання людського пізнання та забезпечення етичного прийняття рішень представляють постійні перешкоди, що вимагають інноваційних математичних рамок та міждисциплінарних досліджень. Зрештою, майбутнє ШІ залежить від балансування технологічного прогресу з етичною відповідальністю, гарантуючи, що його математичні основи служать найкращим інтересам людства.

1.2. Методи машинного навчання та глибинного навчання для аналізу трафіку

Швидка урбанізація та зростання кількості автомобілів зумовили необхідність розробки складних систем аналізу дорожнього руху для управління заторами, підвищення безпеки та оптимізації ефективності транспорту. В останні роки машинне навчання та глибоке навчання стали ключовими технологічними досягненнями, що дозволяють проводити динамічний та точний аналіз дорожнього руху. Ці методи використовують величезні обсяги даних, зібраних з датчиків, камер, GPS-пристроїв та інших джерел, для вилучення змістовних закономірностей та аналітики. Використовуючи різні алгоритми та архітектури нейронних мереж, дослідники та містобудівники можуть прогнозувати потоки транспорту, виявляти інциденти та впроваджувати адаптивні заходи контролю. У цьому есе досліджуються різноманітні методи машинного навчання, що використовуються в аналізі дорожнього руху, архітектури глибокого навчання, адаптовані для обробки даних про дорожній рух, та практичні застосування, що покращують управління дорожнім рухом та міську мобільність.

Методи машинного навчання революціонізували аналіз дорожнього руху, забезпечивши надійні інструменти для прогнозування та розуміння динаміки руху. Моделі навчання з учителем, які спираються на позначені історичні дані, широко використовуються в прогнозуванні потоків руху. Наприклад, алгоритми регресії, такі як регресія опорних векторів та нейронні мережі, можуть прогнозувати кількість транспортних засобів та швидкість у певних місцях та часі, допомагаючи в управлінні заторами та плануванні інфраструктури [4].

Таблиця 1.1.

Методи машинного навчання та глибинного навчання для аналізу трафіку в мережах 5G

Параметр	Характеристика	Приклади застосування
Методи класичного машинного навчання	Статистичні моделі для класифікації, регресії та кластеризації	Виявлення аномалій, прогнозування навантаження
Глибинні нейронні мережі	DNN, CNN, RNN для аналізу часових і просторових залежностей	Детекція аномалій, прогноз пікових навантажень
Методи підкріплення	Оптимізація політики керування через взаємодію з мережею	Маршрутизація, динамічний розподіл ресурсів
Прогнозування часових рядів	ARIMA, LSTM, GRU для аналізу трафіку в часі	Передбачення попиту на пропускну здатність
Онлайн-навчання	Оновлення моделі без повного перенавчання	Моніторинг стану мережі
Ненаглядні методи	Кластеризація, зниження розмірності	Сегментація користувачів
Гібридні моделі	Комбінація ML, DL, RL	Адаптивне управління трафіком

З іншого боку, методи навчання без учителя чудово справляються з виявленням прихованих закономірностей без попередньо визначених позначок. Алгоритми кластеризації, такі як K-середні або ієрархічна кластеризація, можуть виявляти зони заторів, групуючи подібні моделі поведінки руху, що допомагає владі точно визначити проблемні ділянки та впроваджувати цілеспрямовані втручання. Навчання з підкріпленням додає ще один вимір, дозволяючи використовувати адаптивні системи управління, зокрема, в оптимізації часу роботи світлофорів. Ці алгоритми вивчають оптимальні правила шляхом взаємодії з навколишнім середовищем методом спроб і помилок, динамічно регулюючи світлофори на основі поточних умов руху, що призводить до скорочення часу очікування та плавнішого руху транспорту. У сукупності ці методи машинного навчання забезпечують універсальний інструментарій для аналізу даних про дорожній рух, що дозволяє приймати розумніші рішення на основі даних в управлінні дорожнім рухом. Архітектури глибокого навчання мають ще більш просунутий аналіз дорожнього руху, дозволяючи обробляти складні, високовимірні дані про дорожній рух, особливо з візуальних та послідовних джерел. Згорткові нейронні мережі (ЗНМ) особливо ефективні в моніторингу дорожнього руху на основі зображень. Вони сприяють виявленню транспортних засобів у потоках дорожніх камер, що дозволяє підраховувати та класифікувати транспортні засоби в режимі реального часу, що життєво важливо для оцінки заторів та контролю за ними. Наприклад, ЗНМ можуть автоматично визначати кількість автомобілів та вантажівок на автомагістралі, навіть за різних умов освітлення та погоди, тим самим забезпечуючи точні оцінки щільності руху. Рекурентні нейронні мережі (РНМ), особливо мережі з довгою короткочасною пам'яттю (LSTM), розроблені для моделювання часових послідовностей, що робить їх добре підходящими для фіксації динамічного характеру потоку транспорту з плином часу. Вони можуть аналізувати історичні дані про дорожній рух, щоб прогнозувати майбутні моделі заторів, враховуючи щоденні та тижневі цикли. Автоенкодер, ще один інструмент глибокого навчання, використовується для виявлення аномалій — ідентифікації незвичайних дорожніх подій, таких як

аварії або раптові блокування доріг — та для стиснення даних, що зменшує обсяг даних про дорожній рух, зберігаючи при цьому критичну інформацію. Ці архітектури разом дозволяють проводити складний аналіз моделей дорожнього руху, підвищуючи точність та швидкість реагування систем управління дорожнім рухом. Практичне застосування машинного та глибокого навчання в прогнозуванні та управлінні дорожнім рухом є різноманітним та ефективним. Моделі прогнозування заторів у реальному часі використовують історичні дані та дані датчиків для прогнозування дорожніх умов на кілька хвилин або годин вперед, що дозволяє владі впроваджувати проактивні заходи, такі як зміна маршруту або попередження про дорожній рух. Системи виявлення інцидентів на базі машинного навчання можуть автоматично виявляти аномалії, що свідчать про аварії або поломки, що дозволяє швидше реагувати та зменшувати кількість вторинних інцидентів. Наприклад, раптове падіння швидкості транспортних засобів, виявлене сенсорними мережами, може викликати сповіщення, які оперативно мобілізують служби екстреної допомоги. Крім того, динамічні системи планування маршрутів та навігації використовують прогнозну аналітику для пропонування оптимальних маршрутів на основі поточних та прогнозованих дорожніх умов, тим самим мінімізуючи час у дорозі та витрату палива для водіїв. Ці системи є невід'ємною частиною сучасних інтелектуальних транспортних систем (ITC) і все частіше вбудовуються в навігаційні додатки для смартфонів та загальноміські платформи управління дорожнім рухом. З розвитком цих технологій вони обіцяють створити безпечніші, ефективніші та стійкіші міські транспортні мережі, що зрештою покращить якість життя мешканців міста [5].

Незважаючи на значний прогрес, впровадження методів машинного та глибокого навчання для аналізу дорожнього руху стикається зі значними проблемами та обмеженнями. Одна з основних проблем стосується якості, обсягу та конфіденційності даних. Дані про дорожній рух, зібрані з датчиків, камер та GPS-пристроїв, можуть бути непослідовними, неповними або шумними, що знижує точність прогнозних моделей. Крім того, величезний обсяг генерованих даних вимагає значних можливостей зберігання та обробки, що викликає

проблеми, пов'язані з масштабованістю. Також виникають проблеми конфіденційності, оскільки збір та використання даних про місцезнаходження та рух може порушувати права особистості на конфіденційність, що вимагає суворих заходів та політик захисту даних. Крім того, інтерпретація моделі залишається значною перешкодою. Багато моделей глибокого навчання, зокрема нейронні мережі, часто вважаються «чорними скриньками», що ускладнює роботу дорожньої служби та... планувальникам важко зрозуміти, як робляться конкретні прогнози, що обмежує довіру та впровадження. Крім того, обчислювальні вимоги для навчання складних моделей є високими, вимагаючи спеціалізованого обладнання, такого як графічні процесори та ресурси хмарних обчислень, які можуть бути доступні не всім муніципалітетам чи організаціям. Оскільки дорожні системи стають складнішими та орієнтованими на дані, вирішення цих проблем має вирішальне значення для сталого та етичного впровадження рішень машинного навчання в управлінні дорожнім рухом.

Заглядаючи в майбутнє, майбутнє аналізу дорожнього руху, що використовує машинне навчання та глибоке навчання, готове до захопливих розробок, зумовлених новими технологіями та інноваційними дослідженнями. Інтеграція датчиків Інтернету речей (IoT) та аналітики великих даних обіцяє забезпечити більш повне розуміння дорожніх умов у режимі реального часу, що дозволить приймати точніші та своєчасні рішення. Пристрої IoT можуть постійно контролювати різні параметри дорожнього руху, створюючи щільну мережу точок даних, що підвищують точність моделі. Одночасно зростає акцент на розробці моделей штучного інтелекту (XAI), які пропонують прозорі та інтерпретовані результати, необхідні для завоювання довіри та сприяння схваленню регуляторними органами у критично важливих дорожніх застосуваннях. Такі моделі можуть допомогти зацікавленим сторонам зрозуміти обґрунтування прогнозів та рекомендацій, сприяючи кращому прийняттю рішень. Більше того, набирають популярності гібридні підходи, що поєднують традиційні теорії транспортних потоків, такі як моделі черг та макроскопічні моделі транспортних потоків, з алгоритмами машинного навчання. Ці моделі мають на меті використати

сильні сторони обох світів: надійність класичних теорій та адаптивність методів, заснованих на даних. Роблячи це, дослідники сподіваються створити більш стійкі, точні та масштабовані системи аналізу дорожнього руху, які можуть адаптуватися до різноманітного міського середовища та потреб у транспорті, що змінюються.

Підсумовуючи, машинне навчання та глибоке навчання революціонізували аналіз дорожнього руху, пропонуючи потужні інструменти для прогнозування моделей дорожнього руху, виявлення інцидентів та оптимізації стратегій управління дорожнім рухом. Методи навчання з учителем, без учителя та навчання з підкріпленням забезпечують універсальні підходи до розуміння та контролю динаміки дорожнього руху, тоді як передові архітектури глибокого навчання, такі як CNN, RNN та автоенкодери, дозволяють виконувати складну обробку візуальних та послідовних даних. Ці технологічні інновації призвели до практичних застосувань, таких як прогнозування заторів у режимі реального часу, виявлення інцидентів та динамічне планування маршрутів, що значно покращує міську мобільність. Однак такі проблеми, як якість даних, проблеми конфіденційності, інтерпретованість моделей та обчислювальні вимоги, залишаються перешкодами для широкого впровадження. Дивлячись у майбутнє, інтеграція датчиків Інтернету речей, розробка моделей штучного інтелекту, що пояснюються, та гібридні підходи, що поєднують традиційні теорії з машинним навчанням, обіцяють вирішити поточні обмеження та подальший розвиток систем управління дорожнім рухом. У міру розвитку цих передових тенденцій вони мають потенціал для створення розумніших, безпечніших та більш стійких транспортних мереж, які зрештою принесуть користь міському населенню в усьому світі [6].

1.3. Моделі прогнозування навантаження в телекомунікаційних мережах

У швидкозмінному ландшафті телекомунікацій ефективні моделі прогнозування навантаження є життєво важливими для забезпечення надійності

мережі, оптимізації використання ресурсів та підтримки високоякісного надання послуг. Зі зростанням обсягів даних та збільшенням складності мережі вибір та застосування відповідних методів прогнозування стають все більш критичними. Моделі прогнозування навантаження в телекомунікаційних мережах можна загалом розділити на статистичні методи, підходи машинного навчання та гібридні моделі, які поєднують сильні сторони обох. Кожен тип моделі пропонує унікальні переваги та проблеми, що робить їх розуміння важливим для мережевих інженерів та планувальників. У цьому есе досліджуються різні типи моделей прогнозування навантаження, їх практичне застосування в телекомунікаційних мережах, а також відповідні переваги та обмеження, пов'язані з кожним підходом [7].

Основа прогнозування навантаження в телекомунікаціях спирається на різноманітні методи моделювання, які в основному поділяються на статистичні моделі, моделі машинного навчання та гібридні підходи. Статистичні моделі, такі як авторегресивна інтегрована ковзна середня (ARIMA) та експоненціальне згладжування, історично були популярними завдяки своїй простоті та інтерпретованості. ARIMA, наприклад, використовує історичні дані для виявлення закономірностей та прогнозування майбутнього навантаження, що робить його придатним для короткострокового прогнозування, де тенденції є відносно стабільними. Методи експоненціального згладжування додатково уточнюють прогнози, призначаючи експоненціально зменшувані ваги минулим спостереженням, що дозволяє моделям швидко адаптуватися до останніх змін. На противагу цьому, моделі машинного навчання здобули популярність завдяки своїй здатності обробляти складні, нелінійні зв'язки у великих наборах даних. Нейронні мережі, методи опорних векторів (SVM) та дерева рішень можуть вивчати складні закономірності з історичних даних про навантаження, що дозволяє робити точніші довгострокові прогнози, особливо в періоди волатильності або неочікуваних змін попиту. Нейронні мережі, зокрема, чудово справляються з фіксацією нелінійностей, що робить їх придатними для динамічних мережевих середовищ. Крім того, гібридні моделі поєднують статистичні підходи та підходи машинного

навчання, щоб використовувати їхні відповідні сильні сторони. Наприклад, гібридна модель може використовувати ARIMA для фіксації лінійних тенденцій, одночасно застосовуючи нейронні мережі для моделювання нелінійних відхилень, що призводить до більш надійних та адаптивних прогнозів. Інтеграція різноманітних методів моделювання дозволяє отримати повне розуміння поведінки мережевого навантаження, враховуючи динамічний характер телекомунікаційних потреб.

Таблиця 1.2.

Моделі прогнозування навантаження в телекомунікаційних мережах

Параметр	Характеристика	Приклади застосування
Статистичні моделі (ARIMA)	Просте та інтерпретоване прогнозування лінійних тенденцій і сезонності	Короткострокове планування навантаження
Глибокі нейронні мережі (LSTM, RNN)	Виявлення складних часових залежностей, висока точність	Прогноз пікових навантажень, адаптивне керування трафіком
Моделі машинного навчання (SVM, дерева рішень)	Підтримка нелінійних закономірностей, гнучка робота з різними типами даних	Виявлення аномалій, класифікація трафіку
Гібридні моделі (ARIMA + NN)	Поєднання лінійних і нелінійних підходів для максимальної точності	Оптимізація прогнозів при мінливому трафіку

Практичне застосування моделей прогнозування навантаження є широким у телекомунікаційних мережах, безпосередньо впливаючи на управління мережею та операційну ефективність. Точні прогнози навантаження сприяють ефективному плануванню пропускної здатності мережі, забезпечуючи достатню інфраструктуру для задоволення майбутнього попиту без надмірного виділення ресурсів, що може

бути дорогим. Наприклад, мережеві оператори використовують дані прогнозування для визначення оптимального розгортання додаткових базових станцій або коригування пропускної здатності в години пік. Крім того, ці моделі допомагають в оптимізації розподілу ресурсів шляхом динамічного коригування пропускної здатності, маршрутизації та ресурсів живлення відповідно до прогнозованих навантажень, тим самим мінімізуючи перевантаження та перебої в обслуговуванні. Покращене прогнозування також покращує якість обслуговування (QoS) шляхом проактивного управління мережевими навантаженнями, зменшення затримки та запобігання перебоям у періоди високого попиту. Крім того, ефективне прогнозування навантаження знижує експлуатаційні витрати, дозволяючи проводити прогнозне обслуговування та уникаючи непотрібної модернізації обладнання. Наприклад, постачальник телекомунікаційних послуг може використовувати прогнози навантаження для планування технічного обслуговування в години поза піковими навантаженнями, мінімізуючи вплив на клієнтів. По суті, ці моделі слугують критично важливими інструментами для стратегічного планування та управління в режимі реального часу, забезпечуючи ефективну та надійну роботу мереж в умовах коливань трафіку.

Незважаючи на свої численні переваги, моделі прогнозування навантаження мають притаманні переваги та обмеження, які впливають на їх розгортання. Моделі машинного навчання цінуються за свою високу точність та оперативність реагування, особливо при навчанні на великих наборах даних, які фіксують складні моделі попиту. Їхня здатність адаптуватися до нових даних робить їх добре пристосованими для динамічних середовищ, де попит може швидко змінюватися. І навпаки, статистичні моделі, такі як ARIMA, цінуються за свою простоту та інтерпретованість, що робить їх доступними для простих сценаріїв прогнозування та легшими для перевірки. Однак застосування гібридних моделей створює такі проблеми, як підвищена залежність від даних та обчислювальна складність. Гібридні підходи вимагають великих, високоякісних наборів даних для ефективного навчання кількох алгоритмів, а їхня складність вимагає значних обчислювальних ресурсів, що може перешкоджати розгортанню в режимі

реального часу в середовищах з обмеженими ресурсами. Крім того, гібридні моделі можуть страждати від перенавчання, якщо їх неправильно налаштувати, що знижує їхню прогностичну надійність. Залежність від даних також становить ризик; Неточні або неповні дані можуть суттєво погіршити продуктивність моделі, що підкреслює важливість ретельного збору та попередньої обробки даних. Балансування цих переваг та обмежень має вирішальне значення для вибору відповідних моделей, адаптованих до конкретних мережевих умов та експлуатаційних вимог, забезпечуючи оптимальну точність та ефективність прогнозування.

Часто недооцінений, але критичний аспект моделей прогнозування навантаження в телекомунікаційних мережах – це якість та попередня обробка даних. Точні прогнози залежать від наявності комплексних та надійних історичних даних про навантаження, які відображають основні моделі попиту. Окрім історичних показників використання, зовнішні фактори, такі як погодні умови, спеціальні події та економічна діяльність, значно впливають на... впливають на навантаження мережі та повинні бути включені до моделей прогнозування для підвищення їхньої точності. Ефективні методи попередньої обробки даних є важливими для забезпечення цілісності та зручності використання даних. До них належать процедури очищення даних для виявлення та виправлення помилок або невідповідностей, нормалізація для рівномірного масштабування характеристик даних та вибір характеристик для визначення найбільш релевантних змінних для прогнозування. Наприклад, видалення викидів, спричинених помилками вимірювання, запобігає спотворенню результатів, тоді як нормалізація гарантує, що такі змінні, як температура та навантаження, порівнянні. Якість даних безпосередньо впливає на продуктивність моделі; шумні або неповні набори даних можуть призвести до неточних прогнозів, підриваючи процеси прийняття рішень. Отже, суворі методи управління даними є незамінними, оскільки вони закладають основу, на якій можна будувати надійні та надійні моделі прогнозування навантаження. Оскільки мережі продовжують розвиватися, інтеграція різноманітних зовнішніх джерел даних та підтримка

високої якості даних будуть мати вирішальне значення для підвищення точності прогнозування та операційної ефективності [8].

Заглядаючи в майбутнє, сфера прогнозування навантаження в телекомунікаційних мережах готова до значного прогресу, зумовленого технологічними інноваціями. Однією з перспективних тенденцій є інтеграція потоків даних у реальному часі, що генеруються датчиками Інтернету речей (IoT), вбудованими в мережеву інфраструктуру, пристрої та монітори навколишнього середовища. Цей приплив даних у реальному часі дозволяє моделям швидко адаптуватися до мінливих умов, сприяючи більш точним та своєчасним прогнозам. Крім того, розвиток методів глибокого навчання, таких як рекурентні нейронні мережі (RNN) та мережі довгої короткочасної пам'яті (LSTM), пропонує потенціал для ефективнішого фіксування складних часових залежностей та нелінійних закономірностей, ніж традиційні методи. Ці моделі можуть навчатися на величезних обсягах даних та з часом покращувати свої прогностичні можливості. Крім того, концепція адаптивних моделей – тих, що здатні постійно навчатися та оновлюватися з новими даними – ставатиме все більш поширеною. Такі моделі можуть динамічно адаптуватися до змінюваних моделей використання мережі, сезонних коливань та непередбачених стрибків попиту, що зрештою призведе до більш стійкого та ефективного управління мережею. У міру розвитку цих тенденцій постачальники телекомунікаційних послуг отримають потужні інструменти для прогнозування мережевого навантаження з більшою точністю, оптимізації розгортання ресурсів та надання послуг вищої якості в умовах постійно вимогливого цифрового середовища.

На завершення, моделі прогнозування навантаження в телекомунікаційних мережах є важливими для підтримки ефективного, надійного та високоякісного надання послуг. Різні підходи до моделювання — статистичний, машинного навчання та гібридний — пропонують унікальні переваги та стикаються зі специфічними проблемами, особливо щодо точності, інтерпретованості та обчислювальної складності. Практичні застосування, такі як планування потужності, розподіл ресурсів та зниження витрат, демонструють їхню життєво

важливу роль в управлінні мережею. Не менш важливим є врахування якості даних, яка суттєво впливає на ефективність моделі; ретельна попередня обробка та врахування зовнішніх факторів мають вирішальне значення для надійних прогнозів. Дивлячись у майбутнє, досягнення в інтеграції даних у реальному часі, глибокому навчанні та адаптивному моделюванні обіцяють революціонізувати прогнозування навантаження, дозволяючи мережам стати більш чутливими та стійкими. З розвитком технологій подальше вдосконалення та інтелектуальне застосування цих моделей відіграватимуть важливу роль у задоволенні зростаючих потреб телекомунікаційної інфраструктури, забезпечуючи готовність мереж до майбутніх викликів.

1.4. Інтелектуальні методи виявлення аномалій і нерегулярної активності в мережах

У дедалі більш взаємопов'язаному цифровому середовищі захист цілісності мережі став першочерговим завданням як для організацій, так і для окремих осіб. Аномалії та нерегулярна діяльність у мережах, такі як несанкціонований доступ, витоки даних або зловмисні атаки, становлять значні загрози безпеці даних та операційній стабільності. Швидке та точне виявлення цих аномалій є важливим для ефективних заходів кібербезпеки. Протягом багатьох років було розроблено різні методи виявлення аномальної поведінки мережі, які еволюціонували від традиційних статистичних методів до передових підходів машинного навчання та глибокого навчання. У цьому есе досліджується спектр інтелектуальних методів виявлення аномалій, підкреслюючи їхні основні принципи, переваги та проблеми, щоб проілюструвати, як технологічні інновації продовжують покращувати нашу здатність захищати мережеві середовища [9].

Традиційні методи виявлення аномалій довгий час служили основою зусиль у сфері кібербезпеки, спираючись переважно на статистичні та сигнатурні методи. Статистичні підходи аналізують розподіл даних, такі як швидкість передачі пакетів, тривалість з'єднань або кількість байтів, щоб виявити викиди, які суттєво

відхиляються від встановлених норм. Наприклад, раптовий сплеск мережевого трафіку може сигналізувати про атаку типу «відмова в обслуговуванні», що спонукає до подальшого розслідування. Виявлення на основі сигнатур, з іншого боку, передбачає зіставлення мережевої активності з базою даних відомих шкідливих шаблонів або сигнатур. Цей підхід ефективний для швидкого виявлення раніше виявлених загроз, таких як сигнатури шкідливих програм або сигнатури атак, і широко використовується в системах виявлення вторгнень.

Таблиця 1.3.

Методи виявлення аномалій і нерегулярної активності в мережах

Метод	Характеристика	Приклади застосування
Методи машинного навчання (SVM, Random Forest)	Виявлення відхилень у поведінці трафіку на основі класифікації	Аномальні сесії, підозрілі запити, перевантаження
Глибинні моделі (Autoencoder, CNN)	Моделі вчаться на нормальній поведінці та визначають аномалії як відхилення	DDoS-виявлення, нетипові патерни пікового навантаження
Методи кластеризації (K-Means, DBSCAN)	Структурування трафіку без розмітки, виявлення рідкісних кластерів	Сегментація користувачів, виявлення нових типів трафіку
Методи аналізу часових рядів (LSTM)	Виявлення аномалій у динаміці трафіку з урахуванням часових залежностей	Стрибки навантаження, падіння QoS, нетипові зміни трафіку

Однак ці традиційні методи стикаються з суттєвими обмеженнями. Статистичні методи часто мають труднощі з динамічними мережевими середовищами, де нормальна поведінка може коливатися, що призводить до хибнопозитивних результатів або пропущених виявлень. Системи на основі сигнатур за своєю суттю реактивні, нездатні ідентифікувати нові або загрози, що

розвиваються, які не відповідають існуючим сигнатурам. Отже, зловмисники можуть використовувати ці прогалини, використовуючи експлойти нульового дня або поліморфне шкідливе програмне забезпечення, що підкреслює необхідність більш адаптивних стратегій виявлення. Виявлення аномалій на основі машинного навчання є значним кроком уперед порівняно з традиційними методами, дозволяючи системам навчатися та адаптуватися до складної поведінки мережі. Моделі навчання з учителем, такі як методи опорних векторів або дерева рішень, використовують позначені набори даних, де мережева активність попередньо класифікується як нормальна або аномальна. Ці моделі потім можуть точно класифікувати нові точки даних, що сприяє виявленню відомих загроз у режимі реального часу з високою точністю. Наприклад, моделі з учителем можна навчати на історичних даних про атаки, щоб розпізнавати певні закономірності, пов'язані зі спробами фішингу або проникненням шкідливого програмного забезпечення. Алгоритми навчання без учителя, такі як методи кластеризації або алгоритми виявлення аномалій, такі як ізоляційні ліси, не потребують попередньо позначених даних. Натомість вони виявляють відхилення від усталеної нормальної поведінки, що робить їх особливо корисними в середовищах, де шкідлива діяльність є рідкісною або невідомою заздалегідь. Наприклад, моделі без учителя можуть виявляти незвичайні шаблони входу або нетипові передачі даних, позначаючи потенційні порушення безпеки. Адаптивність моделей машинного навчання особливо корисна в динамічних мережевих середовищах, де нормальна поведінка може швидко змінюватися через зміни в активності користувачів або конфігурації мережі. Більше того, ці моделі, як правило, вдосконалюються з часом, оскільки вони обробляють більше даних, підвищуючи точність виявлення та зменшуючи кількість хибних тривог, забезпечуючи таким чином більш надійну та проактивну безпеку [10].

Підходи глибокого навчання ще більше розширюють межі виявлення аномалій, використовуючи нейронні мережі, здатні навчатися складним представленням складних мережевих даних. Автоенкодер, тип нейронної мережі, розроблений для самостійного навчання, особливо ефективні в цій області. Вони

функціонують, стискаючи вхідні дані в нижчорозмірне представлення, а потім реконструюючи їх; аномалії виявляються, коли помилка реконструкції перевищує певний поріг, що вказує на те, що мережеві дані не відповідають вивченим нормальним шаблонам. Глибокі моделі можуть фіксувати тонкі та складні аномалії, які можуть вислизати від традиційних або навіть машинних методів навчання. Наприклад, вони можуть виявляти зашифровані комунікації командування та управління або повільні, приховані дії з вилучення даних, які імітують звичайні шаблони трафіку. Однак розгортання моделей глибокого навчання створює значні труднощі. Їхня обчислювальна складність вимагає значної обчислювальної потужності та спеціалізованого обладнання, що може бути неможливим для всіх організацій. Крім того, глибокі моделі потребують великих обсягів маркованих навчальних даних для досягнення оптимальної продуктивності, які може бути важко отримати та точно анотувати. Незважаючи на ці перешкоди, постійні дослідження та вдосконалення апаратного забезпечення та методів збору даних поступово роблять глибоке навчання життєздатним та потужним інструментом для виявлення мережевих аномалій, обіцяючи вищу точність у виявленні навіть найтонших та найскладніших загроз.

Методи виявлення аномалій на основі поведінки та контексту набули популярності як складні стратегії для підвищення кібербезпеки, зосереджуючись на динаміці активності користувачів та пристроїв у мережі.

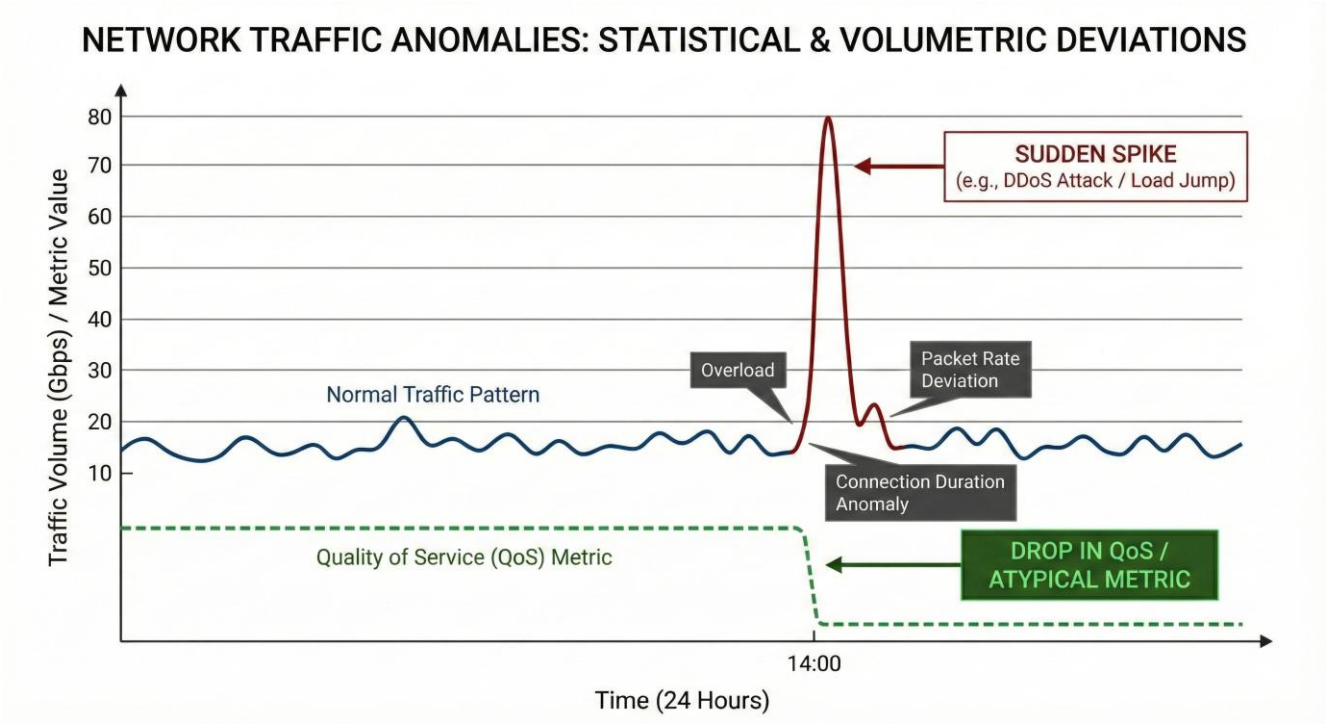


Рис. 1.2. Аномалії мережевого трафіку: статистичні та об'ємні відхилення

На відміну від традиційних підходів на основі сигнатур, які спираються на відомі моделі загроз, виявлення на основі поведінки постійно відстежує дії користувачів та пристроїв, щоб виявити відхилення, які можуть сигналізувати про порушення безпеки. Наприклад, раптовий доступ користувача до конфіденційних файлів у незвичні години або передавання даних пристроєм на незнайомі зовнішні IP-адреси може свідчити про скомпрометовані облікові дані або активність шкідливого програмного забезпечення. Включення контекстної інформації, такої як час доступу, географічне розташування, тип пристрою або типові моделі використання, ще більше вдосконалює процес виявлення. Цей багаторівневий підхід дозволяє системам безпеки розрізняти легітимні аномалії, такі як працівник, який працює допізна, та зловмисні дії, тим самим зменшуючи кількість хибнопозитивних результатів та підвищуючи загальну точність. Більше того, ці методи особливо ефективні для виявлення нових або еволюціонуючих моделей атак, які не відповідають існуючим сигнатурам, усуваючи критичну прогалину в традиційних методах виявлення. Аналізуючи поведінкові та контекстуальні сигнали, організації можуть проактивно виявляти загрози, які в іншому випадку

могли б залишитися непоміченими, що робить ці методи безцінними для адаптивних та комплексних стратегій кібербезпеки.

Незважаючи на значний прогрес, галузь інтелектуального виявлення аномалій стикається з кількома серйозними викликами, які впливають на її практичне впровадження та ефективність. Однією з основних проблем є управління величезним обсягом та швидкістю мережевих даних у режимі реального часу, що вимагає високоефективних алгоритмів та значних обчислювальних ресурсів для швидкої обробки та аналізу інформації. Крім того, зменшення кількості хибнопозитивних результатів залишається постійною проблемою; надмірно чутливі системи виявлення можуть позначати нешкідливі дії як загрози, що призводить до втоми від пильності серед команд безпеки та потенційного ігнорування справжніх проблем. Досягнення оптимального балансу між чутливістю та специфічністю має вирішальне значення для операційної життєздатності. Крім того, інтеграція багаторівневих систем виявлення, що поєднують традиційні методи машинного навчання, глибокого навчання та поведінки, створює як технічні, так і організаційні виклики, що вимагає безперебійної сумісності та складного управління. Забігаючи вперед, новітні технології, такі як штучний інтелект, аналітика великих даних та периферійні обчислення, є перспективними для подолання цих перешкод. Ці інновації можуть забезпечити створення більш масштабованих, точних та адаптивних систем виявлення, здатних обробляти складні мережеві середовища з великим обсягом даних. Постійні дослідження та розробки в цих галузях є важливими для створення стійких, інтелектуальних систем кібербезпеки, здатних протистояти дедалі складнішим кіберзагрозам.

На завершення, еволюція методів виявлення аномалій — від традиційних статистичних та сигнатурних методів до передових підходів машинного навчання, глибокого навчання та поведінки — відображає постійне прагнення до підвищення безпеки мережі у дедалі складнішому цифровому середовищі. Традиційні методи, хоча й є фундаментальними, стикаються з обмеженнями в адаптації до нових загроз, що спонукає до впровадження моделей машинного

навчання, які пропонують більшу гнучкість та точність. Глибоке навчання ще більше просуває цей рубіж, фіксуючи тонкі та складні аномалії, хоча воно й створює проблеми, пов'язані з обчислювальними вимогами та вимогами до даних. Доповнюючи ці технології, методи виявлення на основі поведінки та контексту забезпечують нюансоване розуміння діяльності користувачів та пристроїв, значно покращуючи ідентифікацію нових та складних атак. Тим не менш, ця галузь повинна справлятися з такими викликами, як обробка величезних обсягів даних у режимі реального часу, мінімізація хибнопозитивних результатів та інтеграція різноманітних систем виявлення. Заглядаючи в майбутнє, використання нових технологій, таких як штучний інтелект та аналітика великих даних, пропонує перспективні шляхи для розробки більш стійких, ефективних та інтелектуальних систем виявлення аномалій. Оскільки кіберзагрози продовжують розвиватися, то й наші стратегії виявлення також повинні розвиватися, забезпечуючи надійний захист для мереж, що лежать в основі нашого цифрового світу.

2. АРХІТЕКТУРА ТА ТЕХНІЧНІ ПАРАМЕТРИ МЕРЕЖ П’ЯТОГО ПОКОЛІННЯ

2.1. Технічні основи побудови мереж 5G та принципи функціонування

Поява технології 5G знаменує собою значну віху в еволюції мереж бездротового зв'язку, обіцяючи безпрецедентні швидкості, наднадійне з'єднання з низькою затримкою та масове підключення пристроїв. На відміну від своїх попередників, 5G — це не просто поступове оновлення, а комплексна переробка архітектури мережі, зумовлена передовими технологічними інноваціями та новими принципами роботи. Розуміння технічних основ 5G включає дослідження її основної архітектури мережі, технологій, що забезпечують високошвидкісний та надійний зв'язок, а також фундаментальних принципів, що керують її роботою. Ці компоненти разом сприяють створенню універсальної та стійкої мережі, здатної підтримувати нові програми, такі як автономні транспортні засоби, розумні міста та Інтернет речей (IoT). Це есе має на меті з'ясувати тонкощі мереж 5G, підкреслюючи, як їх дизайн та технології лежать в основі їхніх новаторських можливостей [11].

Архітектура мереж 5G принципово відрізняється від попередніх поколінь, включаючи більш гнучку, масштабовану та модульну конструкцію, що підвищує продуктивність та ефективність. В основі цієї архітектури лежать ключові компоненти, такі як ядро наступного покоління (NGC), яке інтегрує хмарні принципи для забезпечення безперебійного потоку даних, сегментації мережі та динамічного управління ресурсами. Основна мережа також включає розширені функції керування та площини користувача, що підтримують високу пропускну здатність та низьку затримку. На відміну від 4G LTE, яка спирається на централізоване керування, 5G використовує розподілену архітектуру, яка дозволяє сегментацію мережі — критично важливу функцію, яка дозволяє операторам створювати кілька віртуальних мереж, адаптованих до конкретних застосувань або галузей. Наприклад, сегмент мережі, призначений для автономних транспортних

засобів, може пріоритезувати наднизьку затримку, тоді як інший для пристроїв Інтернету речей може зосередитися на масовому підключенні. Ця диференціація дозволяє налаштовувати надання послуг та ефективно використовувати ресурси. Крім того, гнучкість архітектури демонструється її підтримкою інтеграції з існуючими мережами, що спрощує перехід та забезпечує зворотну сумісність. Загалом, модульна конструкція архітектури 5G та акцент на віртуалізації є основоположними для її здатності задовольняти різноманітні випадки використання та потреби користувачів. Кілька передових технологій слугують основою розширених можливостей 5G, серед яких Massive MIMO, використання спектру міліметрових хвиль (mmWave) та формування променя виділяються як ключові інновації. Massive MIMO передбачає розгортання великих масивів антен, які іноді налічують сотні, що можуть передавати та приймати кілька потоків даних одночасно. Ця технологія значно підвищує спектральну ефективність та пропускну здатність мережі, дозволяючи більшій кількості пристроїв підключатися з вищою швидкістю без перевантажень. Наприклад, у міських умовах Massive MIMO може значно збільшити пропускну здатність, підтримуючи такі програми, як потокове відео високої чіткості та ігри в реальному часі. Використання спектру mmWave, який працює на частотах вище 24 ГГц, забезпечує надзвичайно високу швидкість передачі даних завдяки великій пропускній здатності, доступній у цих діапазонах. Однак сигнали mmWave мають обмежені відстані поширення та чутливі до перешкод, що вимагає передових методів для підтримки якості сигналу. Формування променя вирішує цю проблему, спрямовуючи радіосигнали точно на пристрої користувача, а не транслюючи їх всенаправлено. Ця сфокусована передача підвищує силу сигналу та зменшує перешкоди, ефективно розширюючи покриття та підвищуючи надійність. Разом ці технологічні інновації дозволяють 5G забезпечувати високі швидкості, ємність та надійність, необхідні для послуг зв'язку наступного покоління.

Принципи роботи 5G зосереджені на максимізації ефективності використання спектру, забезпеченні низької затримки та підтримці щільного

розгортання пристроїв за допомогою інноваційного управління спектром та мережевої інфраструктури (рис. 2.1).

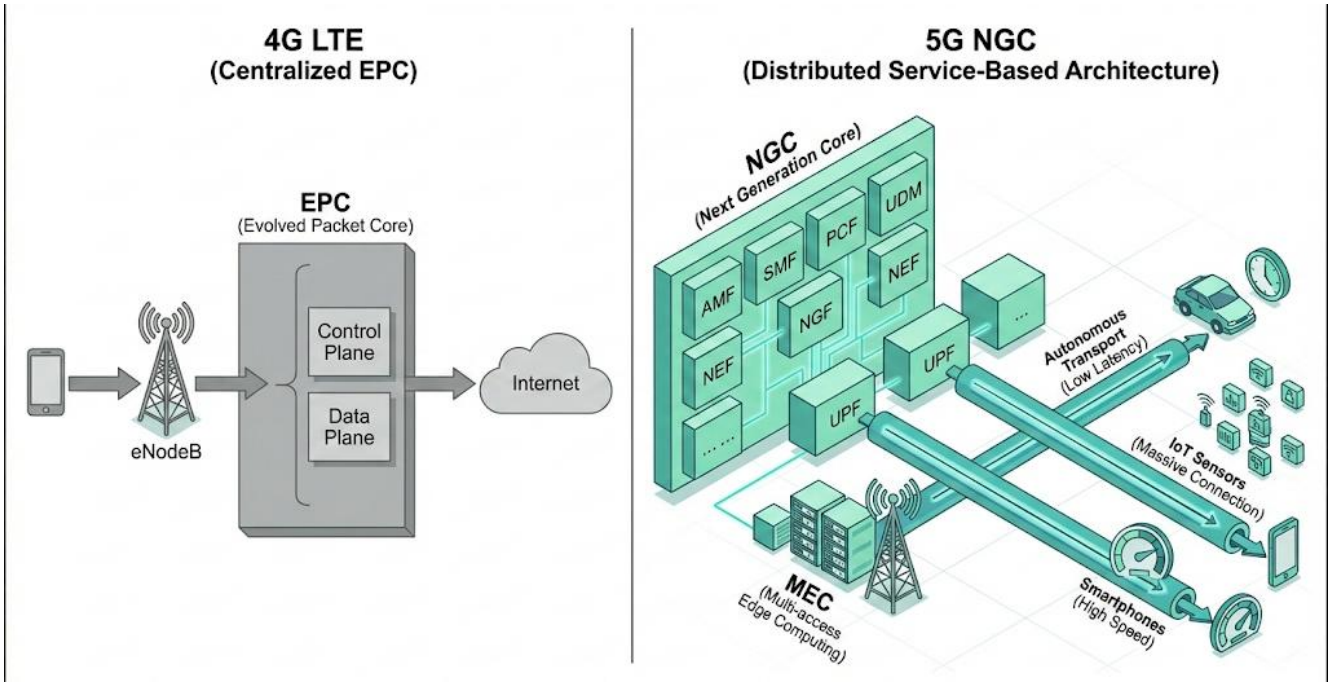


Рис. 2.1. Порівняння архітектурних підходів: централізована мережа 4G проти розподіленої сервіс-орієнтованої архітектури 5G

Динамічний спільний доступ до спектру дозволяє мережам 5G співіснувати з існуючими системами 4G LTE, динамічно розподіляючи спектр на основі попиту в режимі реального часу, оптимізуючи використання та зменшуючи перешкоди. Ця гнучкість гарантує ефективне використання ресурсів спектру, забезпечуючи безперебійний користувацький досвід у різних типах мереж. Малі комірки, які є базовими станціями з низькою потужністю, розміщеними ближче до користувачів, є невід'ємною частиною досягнення широкого покриття та високої ємності, особливо в густонаселених міських районах. Розгортаючи численні малі комірки, мережі 5G можуть зменшити проблеми, спричинені обмеженнями міліметрового діапазону хвиль, такі як ослаблення сигналу та низьке проникнення. Крім того, інтеграція хмарних та периферійних обчислювальних архітектур відіграє життєво важливу роль у зменшенні затримки та підвищенні швидкості обробки даних. Периферійні обчислення наближають обчислювальні ресурси до кінцевих

користувачів, сприяючи аналізу даних у режимі реального часу та прийняттю рішень для таких застосувань, як автономне водіння або доповнена реальність. Ця взаємопов'язана структура забезпечує ефективну роботу мереж 5G з низькою затримкою, високою надійністю та масштабованою ємністю, тим самим підтримуючи різноманітні та вимогливі програми сучасного цифрового суспільства.

Ефективне управління спектром є основоположним для успішного розгортання та експлуатації мереж 5G, що включає стратегічний розподіл високих, середніх та низьких частотних діапазонів для оптимізації продуктивності в різних випадках використання. Високочастотні діапазони, зокрема міліметрові хвилі, пропонують широку пропускну здатність, здатну підтримувати надвисокі швидкості передачі даних, але мають свої обмеження, такі як обмежений радіус дії та погане проникнення крізь перешкоди. Середньочастотні діапазони, зазвичай від 1 ГГц до 6 ГГц, забезпечують баланс, забезпечуючи помірне покриття та ємність, що робить їх придатними для широкого кола застосувань. Низькочастотні діапазони нижче 1 ГГц, з іншого боку, чудово забезпечують широке покриття та глибоке проникнення всередині приміщень, що є важливим для підключення до сільської та приміської місцевості. Однак, дефіцит доступного спектру створює значні проблеми, оскільки конкуруючі вимоги з різних секторів призводять до перевантажень та перешкод. Управління перешкодами стає складнішим через щільне розгортання малих сот та різноманітні діапазони спектру, що використовуються в 5G, що вимагає складних методів координації та стратегій динамічного спільного використання спектру. Регуляторні міркування також відіграють вирішальну роль, оскільки уряди та міжнародні організації розподіляють та ліцензують діапазони спектру, впливаючи на те, як оператори можуть використовувати ці частоти. Політика, спрямована на гармонізацію спектру та справедливий доступ, є життєво важливою для сприяння інноваціям, запобігаючи перешкодам та забезпечуючи справедливий розподіл ресурсів. Загалом, управління спектром у 5G вимагає делікатного балансу між

технологічними можливостями, регуляторними рамками та постійно зростаючим попитом на бездротове з'єднання [12].

Таблиця 2.1.

Ключові технологічні інновації 5G та їх вплив на ефективність мережі

Технологія	Принцип функціонування	Вплив на параметри трафіку та мережі
Massive MIMO	Використання великих масивів антен (сотні елементів) для одночасної передачі кількох потоків даних.	Значне підвищення спектральної ефективності та пропускної здатності; обслуговування більшої кількості пристроїв без перевантажень.
mmWave (Міліметрові хвилі)	Робота на частотах вище 24 ГГц, де доступна велика ширина смуги пропускання.	Забезпечення надвисоких швидкостей передачі даних, але з меншим радіусом покриття.
Beamforming (Формування променя)	Спрямування радіосигналу вузьким променем точно на пристрій користувача замість всенаправленого випромінювання.	Підвищення сили сигналу, зменшення перешкод (інтерференції) та розширення ефективного покриття.
Dynamic Spectrum Sharing (DSS)	Динамічний розподіл спектральних ресурсів між 4G та 5G в режимі реального часу.	Оптимізація використання частотного ресурсу та забезпечення плавного переходу від старих мереж.
Edge Computing (MEC)	Розміщення обчислювальних потужностей на периферії мережі, ближче до абонента.	Мінімізація затримки (latency) для критично важливих додатків та розвантаження магістральних каналів.

Безпека та надійність мають першорядне значення в мережах 5G, враховуючи їхню широку роль у критичній інфраструктурі, комерційних застосуваннях та захисті персональних даних. Для захисту цілісності та конфіденційності даних 5G використовує передові протоколи шифрування та надійні механізми автентифікації, гарантуючи, що лише авторизовані користувачі та пристрої можуть отримати доступ до мережевих ресурсів. Ці заходи безпеки є життєво важливими для запобігання зловмисним атакам, таким як прослуховування, витоки даних та спуфінг, які можуть поставити під загрозу

конфіденційну інформацію або порушити роботу послуг. Резервування мережі та відмовостійкість також є невід'ємними складовими підтримки високої надійності; мережі 5G включають кілька рівнів резервного копіювання, включаючи диверсифіковані шляхи маршрутизації та резервні апаратні компоненти, щоб забезпечити безперервну роботу навіть у разі збоїв або атак. Крім того, притаманна архітектурі гнучкість дозволяє швидко реагувати на загрози за допомогою оновлень програмного забезпечення та адаптивних політик безпеки. Тим не менш, розширена поверхня атаки, що з'явилася через поширення підключених пристроїв та складність інфраструктури 5G, створює нові вразливості. Усунення цих проблем вимагає постійних оцінок безпеки, ретельного тестування та розробки стандартів для зменшення ризиків, характерних для 5G, таких як вразливості ланцюга поставок та програмно-визначені мережеві експлойти. Забезпечення безпеки та надійності в 5G — це не просто технічний виклик, а вирішальний компонент у формуванні довіри користувачів та забезпеченні широкого впровадження бездротових послуг наступного покоління (табл. 2.2).

Таблиця 2.1.

Порівняння технічних характеристик 4G (LTE-A) та 5G (NR)

Параметр	4G (LTE-Advanced)	5G (New Radio)	Примітки
Пікова швидкість (Downlink)	1 Гбіт/с	До 20 Гбіт/с	Забезпечується mmWave
Середня швидкість користувача	10–50 Мбіт/с	100 Мбіт/с – 1 Гбіт/с	«Багатогігабітні швидкості»
Затримка (Latency)	30–50 мс (на практиці)	< 1 мс (теоретична)~10 мс (на практиці eMBB)	Критично для автономних авто
Щільність підключень	~100 тис. пристроїв/км²	1 млн пристроїв/км²	Для масового IoT (mMTC)
Ефективність спектру	30 біт/с/Гц	До 90 біт/с/Гц	Покращено завдяки Massive MIMO
Мобільність	До 350 км/год	До 500 км/год	Підтримка швидкісних поїздів

Підсумовуючи, технічні основи мереж 5G охоплюють складну архітектуру, новаторські технології та інноваційні принципи роботи, які разом переосмислюють бездротовий зв'язок. Модульна, гнучка конструкція архітектури 5G у поєднанні з передовими технологіями, такими як Massive MIMO, використання міліметрового хвильового спектру та формування променя, забезпечує безпрецедентний рівень швидкості, пропускної здатності та надійності. Ці технологічні досягнення доповнюються динамічним управлінням спектром, стратегічним використанням різних частотних діапазонів та ретельним регуляторним наглядом, що вирішує такі проблеми, як дефіцит спектру та перешкоди. Крім того, акцент на безпеці та відмовостійкості гарантує, що мережі 5G можуть підтримувати критично важливі програми, захищаючи дані користувачів та підтримуючи стабільну якість обслуговування. Оскільки 5G продовжує розвиватися, його фундаментальні принципи ляжуть в основу нової ери зв'язку, забезпечуючи трансформаційні програми в різних галузях промисловості та суспільстві. Зрештою, розуміння цих технічних основ є важливим для оцінки величезного потенціалу 5G та підготовки до його широкого розгортання в цифрову епоху.

Згідно зі специфікаціями IMT-2020, перехід до архітектури 5G дозволяє досягти пікових швидкостей до 20 Гбіт/с, що значно перевищує можливості 4G. Крім того, критично важливим показником є зниження затримки до рівня 1 мс у сценаріях URLLC, що необхідно для функціонування автономних систем (табл. 2.2) [13].

2.2. Архітектура радіодоступу 5G NR (New Radio) та характеристика мережевих сценаріїв

Поява 5G New Radio (NR) являє собою трансформаційну віху в технології бездротового зв'язку, розроблену для задоволення зростаючих потреб у вищих швидкостях передачі даних, меншій затримці та покращеному з'єднанні. На відміну від своїх попередників, 5G NR представляє універсальну та масштабовану

архітектуру радіодоступу, яка може адаптуватися до безлічі варіантів використання, починаючи від покращеного мобільного широкосмугового доступу до масового розгортання Інтернету речей (IoT). У цьому есе досліджується фундаментальна архітектура 5G NR, її інноваційні технічні характеристики та те, як вона характеризується в різних мережевих сценаріях. Розуміння цих елементів дає критичне розуміння можливостей та стратегій розгортання мереж 5G, які готові переосмислити ландшафт бездротового зв'язку.

Архітектура радіодоступу 5G NR складається з кількох основних компонентів, кожен з яких відіграє життєво важливу роль у забезпеченні високошвидкісного та надійного з'єднання. Центральним елементом є gNodeB, або базова станція наступного покоління, яка керує радіоресурсами, обробляє сигналізацію та підключає користувацьке обладнання (UE) до основної мережі. Доповнюють це функції мережі радіодоступу (RAN), які сприяють динамічному управлінню спектром, формуванню променя та операціям MIMO для оптимізації якості та пропускної здатності сигналу. Помітним нововведенням в архітектурі 5G NR є її модульна конструкція, яка відокремлює функції фізичного рівня від обробки вищих рівнів, що забезпечує більш гнучке впровадження та легше оновлення. Ця модульність проявляється в поділі на центральний блок (CU) та розподілений блок (DU), що дозволяє мережевим операторам розгортати масштабовані та індивідуальні рішення на основі конкретних географічних або сервісних вимог. Порівняно з попередніми поколіннями, такими як 4G LTE, які використовували більш монолітну структуру RAN, архітектура 5G NR підкреслює адаптивність, забезпечуючи безперешкодну інтеграцію нових функцій та діапазонів спектру, тим самим забезпечуючи майбутнє мережевої інфраструктури. Технічно, 5G NR розширює межі бездротового зв'язку завдяки інноваційним функціям, розробленим для максимізації ефективності використання спектру та пропускної здатності мережі. Одна з ключових стратегій передбачає використання діапазонів вищих частот, таких як міліметрові хвилі (mmWave), які пропонують величезні обсяги невикористаного спектру, здатного підтримувати багатогігабітні швидкості передачі даних. (рис. 2.2) [14].

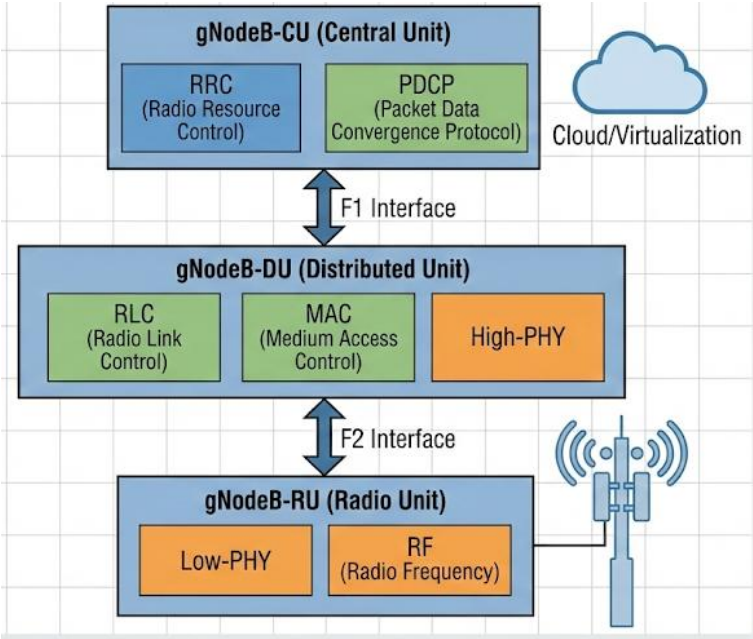


Рис. 2.2. Функціональна архітектура 5G gNodeB: розділення на рівні CU (Central Unit) та DU (Distributed Unit)

Однак робота на цих високих частотах створює такі проблеми, як збільшення втрат на поширення, які 5G вирішує за допомогою передових методів формування променя, які фокусують енергію безпосередньо на користувачів, тим самим підвищуючи силу сигналу та зменшуючи перешкоди. Масивний MIMO (Multiple Input Multiple Output) додатково посилює ці можливості, розгортаючи великі антенні решітки, які обслуговують кількох користувачів одночасно, значно збільшуючи спектральну ефективність та пропускну здатність мережі. Крім того, 5G NR впроваджує мережеве сегментування, революційну концепцію, яка дозволяє розділити фізичну мережеву інфраструктуру на кілька віртуальних мереж, кожна з яких адаптована до конкретних вимог до послуг, таких як наднадійний зв'язок з низькою затримкою (URLLC) для автономних транспортних засобів або масивне підключення Інтернету речей. Ця можливість сегментування дозволяє постачальникам послуг налаштовувати свої пропозиції, оптимізувати розподіл ресурсів та забезпечувати диференційовану якість обслуговування, що зрештою розширює універсальність та застосовність мереж 5G у різних секторах. Характеристика мережевого сценарію 5G NR демонструє його адаптивність до різних середовищ та випадків використання. У густонаселених міських районах

5G розроблений для роботи з високопродуктивними сценаріями шляхом розгортання малих сот та використання високочастотних діапазонів для задоволення зростаючого попиту на дані з боку мобільних користувачів, програм розумного міста та потокового передавання мультимедіа. Формування променя та масивний MIMO є особливо важливими в цих умовах, забезпечуючи надійне покриття та оптимальну спектральну ефективність серед високої щільності користувачів. І навпаки, розгортання 5G у сільській місцевості або віддалених районах створює значні труднощі через величезні географічні простори та обмежену інфраструктуру. Ці розгортання часто спираються на нижчі частотні діапазони, які забезпечують ширше покриття, але за рахунок зниження швидкості передачі даних. Щоб вирішити цю проблему, мережеві оператори досліджують такі рішення, як інтеграція супутників та потужні базові станції для розширення покриття. Промислове середовище та мережі, специфічні для Інтернету речей, є прикладом ще одного ключового сценарію, де гнучка архітектура 5G NR використовується для підтримки критично важливих програм, що вимагають наднизької затримки, високої надійності та масового підключення пристроїв. Наприклад, автоматизація виробництва, автономні роботи та розумні мережі отримують вигоду від індивідуальних конфігурацій мережі, які надають пріоритет затримці та надійності, що ілюструє здатність 5G NR реалізувати ініціативи Industry 4.0 (табл. 2.2)

Попри численні досягнення, розгортання мереж 5G NR стикається з кількома суттєвими проблемами, які необхідно вирішити, щоб реалізувати свій повний потенціал. Одна з основних проблем стосується розподілу та управління спектром, оскільки залежність 5G від вищих частотних діапазонів, таких як міліметрові хвилі, вимагає широкого та часто дорогого ліцензування спектру, яке варіюється залежно від регіону та може перешкоджати широкому розгортанню. Ефективне управління спектром стає критично важливим для запобігання перешкодам та забезпечення оптимального використання, особливо в середовищах, де співіснує кілька операторів. Інша проблема стосується сумісності, оскільки інтеграція 5G NR з існуючою мережевою інфраструктурою,

такою як системи 4G LTE та застаріле обладнання, вимагає складних рішень для синхронізації та сумісності.

Таблиця 2.2.

Характеристика мережевих сценаріїв розгортання 5G NR та вимоги до спектральних ресурсів

Сценарій розгортання	Характеристика середовища	Типовий частотний діапазон (Спектр)	Ключові технології (згідно з текстом)	Пріоритетний показник якості (KPI)
Щільна міська забудова (Dense Urban)	Висока щільність користувачів, перешкоди від будівель, пікові навантаження трафіку.	High Band (mmWave): >24 ГГц (FR2)Mid Band: 3.5 ГГц (FR1)	Small Cells (малі соти), Massive MIMO, Beamforming.	Ємність мережі (Capacity) та пропускна здатність.
Сільська місцевість (Rural / Remote)	Великі географічні простори, обмежена інфраструктура, низька щільність абонентів.	Low Band: <1 ГГц (наприклад, 700 МГц)	High Power Base Stations (потужні базові станції), інтеграція з супутниками (NTN).	Покриття (Coverage).
Індустріальний IoT (Industrial / Industry 4.0)	Заводи, автоматизовані лінії, роботи. Високі вимоги до стабільності.	Mid Band / Local Area: 3.5–4.8 ГГц (Приватні мережі)	Network Slicing (слайсінг), URLLC, Edge Computing.	Надійність та Затримка (Latency) (<1 мс).

Досягнення безперебійної передачі, підтримка безперервності обслуговування та уникнення фрагментації є постійними проблемами, які вимагають складних стратегій інтеграції. Міркування безпеки також стають дедалі складнішими в різноманітних мережевих сценаріях. Поширення підключених пристроїв, особливо в IoT та промислових застосуваннях, розширює поверхню атаки, що вимагає надійних протоколів безпеки для захисту цілісності даних, конфіденційності користувачів та стійкості мережі. Забезпечення безпеки в гетерогенних середовищах, від щільних міських центрів до віддалених сільських районів, залишається критичною перешкодою, яку необхідно інноваційно

вирішувати за допомогою передових механізмів шифрування, автентифікації та систем виявлення загроз на основі штучного інтелекту.

Заглядаючи вперед, майбутнє архітектури 5G NR готове до динамічної еволюції, зумовленої технологічною інтеграцією та появою нових парадигм. Одна з найперспективніших тенденцій включає інтеграцію периферійних обчислень та штучного інтелекту (ШІ) в структуру 5G. Периферійні обчислення наближають обчислювальні ресурси до користувача, зменшуючи затримку та забезпечуючи обробку даних у режимі реального часу, необхідну для таких застосувань, як автономні транспортні засоби та розумне виробництво. ШІ, з іншого боку, покращує управління мережею за допомогою адаптивних алгоритмів, які оптимізують використання спектру, прогнозують перевантаження мережі та автоматизують завдання технічного обслуговування, тим самим підвищуючи ефективність та надійність. Після появи 5G, галузеві експерти очікують переходу до 6G, який, як очікується, базуватиметься на фундаменті 5G з ще вищими швидкостями передачі даних, наднизькою затримкою та повсюдним підключенням, можливо, включаючи квантові обчислення та голографічний зв'язок. Одночасно докладаються зусилля для підвищення масштабованості та стійкості мережі, гарантуючи, що майбутні мережі зможуть підтримувати експоненціально зростаючу кількість пристроїв та програм. Такі інновації, як децентралізовані архітектури, програмно-визначені мережі (SDN) та віртуалізація мережевих функцій (NFV), спрямовані на створення більш надійних, гнучких та безпечних мереж, здатних швидко адаптуватися до нових вимог та загроз, тим самим прокладаючи шлях для наступного покоління бездротового зв'язку.

Підсумовуючи, 5G NR являє собою новаторський прорив у бездротовому зв'язку, що вирізняється гнучкою архітектурою, передовими технічними характеристиками та адаптивністю до різноманітних мережевих сценаріїв. Його основні компоненти, включаючи gNodeB та принципи модульного проектування, сприяють високій пропускній здатності та безперебійному наданню послуг, тоді як інновації, такі як вищі частотні діапазони, формування променя, масивний MIMO та нарізка мережі, значно підвищують ефективність використання спектру

та налаштування послуг. Різноманітні середовища розгортання — від щільних міських центрів до сільських районів та промислових умов — підкреслюють універсальність 5G, хоча такі проблеми, як управління спектром, сумісність та безпека, залишаються ключовими перешкодами для широкого впровадження. Заглядаючи в майбутнє, інтеграція периферійних обчислень та штучного інтелекту, разом з переходом до 6G, обіцяє подальше підвищення можливостей мережі, масштабованості та стійкості. У міру розвитку цих технологічних досягнень, 5G NR готовий революціонізувати спосіб нашого спілкування, підключення та роботи у все більш цифровому світі, зрештою формуючи розумніше та більш взаємопов'язане майбутнє [15].

2.3. Технології мережевого поділу (network slicing) у 5G

Оскільки бездротова технологія п'ятого покоління (5G) продовжує революціонізувати телекомунікаційний ландшафт, сегментація мережі стає однією з її найреволюційніших інновацій. Ця концепція дозволяє розділити єдину фізичну мережеву інфраструктуру на кілька віртуалізованих та незалежних мереж, адаптованих до різноманітних вимог до послуг. На відміну від традиційних мережевих архітектур, які часто спираються на універсальний підхід, сегментація мережі пропонує гнучкий та динамічний метод налаштування мережевих ресурсів відповідно до конкретних потреб різних застосувань та галузей. У цьому есе досліджуються основні принципи сегментації мережі в 5G, розглядаються технічні механізми, що забезпечують її роботу, та висвітлюються значні переваги, які вона надає в різних секторах.

Розподіл мережі в 5G – це, по суті, метод розділення фізичної бездротової мережі на кілька віртуальних мереж, кожна з яких оптимізована для певних випадків використання або типів послуг. Основна ідея полягає в створенні ізольованих, наскрізних сегментів мережі, якими можна незалежно керувати, налаштовувати та оптимізувати, що дозволяє постачальникам послуг надавати індивідуальні рішення для підключення. Цей підхід знаменує собою відхід від

традиційних мережевих архітектур, які зазвичай розгортають монолітні, однорідні мережі, яким бракує гнучкості та які є неефективними при задоволенні різнорідних потреб. Ключові компоненти, що беруть участь у реалізації мережевих сегментів, включають технології віртуалізації, такі як віртуалізація мережевих функцій (NFV) та програмно-визначені мережі (SDN), які абстрагують фізичні мережеві ресурси на гнучкі, програмовані сутності (рис. 2.3) [16].

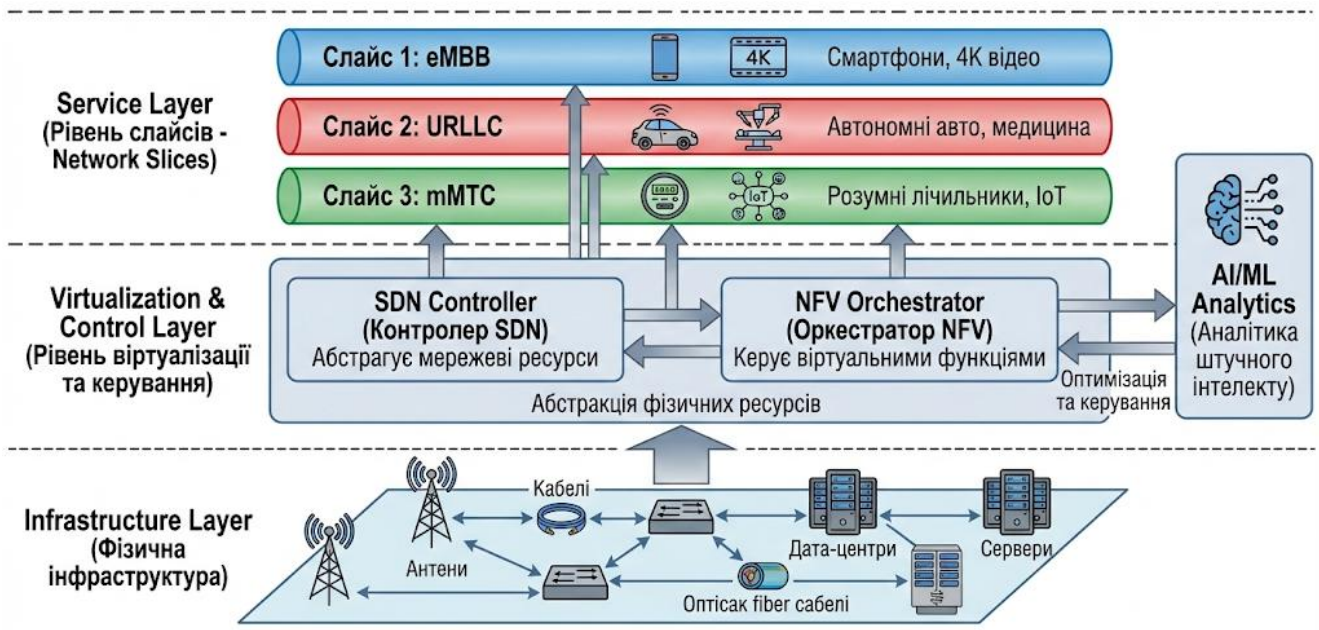


Рис. 2.3. Концептуальна архітектура Network Slicing: від фізичних ресурсів до ізольованих сервісів

Ці компоненти працюють разом, щоб забезпечити динамічне створення екземплярів та управління різними сегментами, гарантуючи, що кожна віртуальна мережа може працювати незалежно без перешкод та відповідати певним вимогам щодо затримки, пропускної здатності або надійності, що вимагаються різними програмами. Наприклад, сегмент, призначений для екстрених служб, пріоритезуватиме наднадійний зв'язок з низькою затримкою, тоді як сегмент для потокового медіа може зосередитися на наданні високої пропускної здатності. Технічна архітектура мережевого сегментування значною мірою спирається на технології віртуалізації, такі як NFV та SDN, для полегшення створення,

керування та ізоляції мережевих сегментів. NFV дозволяє віртуалізувати мережеві функції, такі як маршрутизація, брандмауер та балансування навантаження, тому їх можна розгортати як програмні компоненти на звичайному обладнанні, а не покладатися на виділені фізичні пристрої. Ця гнучкість дозволяє мережевим операторам швидко створювати, змінювати або видаляти сегменти на основі потреб, що змінюються. SDN доповнює це, забезпечуючи централізований контроль над потоками мережевого трафіку, дозволяючи динамічну маршрутизацію та розподіл ресурсів у межах кожного сегмента. Ці технології разом гарантують, що кожен мережевий сегмент ізольований від інших, зберігаючи безпеку та цілісність продуктивності. Наприклад, у промисловому середовищі Інтернету речей виділений сегмент може бути призначений виробничим роботам із гарантованою затримкою, окремо від сегмента розваг для споживачів. Крім того, різні типи сегментів ілюструють універсальність 5G: покращений мобільний широкосмуговий доступ (eMBB) пропонує високу швидкість передачі даних для потокової передачі та віртуальної реальності, наднадійний зв'язок з низькою затримкою (URLLC) підтримує критично важливі послуги, такі як дистанційна хірургія або автономні транспортні засоби, а масивний машинний зв'язок (mMTC) задовольняє потреби в підключенні мільярдів пристроїв Інтернету речей, таких як інтелектуальні лічильники або датчики навколишнього середовища.

Впровадження технології сегментації мережі забезпечує суттєві переваги в багатьох галузях, забезпечуючи високо налаштовані, ефективні та масштабовані мережеві послуги. Для таких галузей, як охорона здоров'я, автономне водіння та розумні міста, можливість розгортання індивідуально налаштованих сегментів мережі означає покращену якість обслуговування та операційну ефективність. Наприклад, автономні транспортні засоби потребують наднадійних сегментів зв'язку з низькою затримкою (URLLC) для забезпечення обміну даними в режимі реального часу з мінімальною затримкою, що має вирішальне значення для безпеки та навігації. Інфраструктура розумного міста отримує вигоду від кількох сегментів, призначених для управління дорожнім рухом, громадської безпеки та

моніторингу комунальних послуг, кожен з яких має різні параметри продуктивності. Крім того, сегментація мережі сприяє кращому використанню ресурсів, динамічно розподіляючи пропускну здатність та обчислювальні ресурси на основі поточного попиту, зменшуючи втрати та експлуатаційні витрати. Ця гнучкість особливо важлива, оскільки Інтернет речей (IoT) розширюється експоненціально, і мільярди підключених пристроїв потребують спеціалізованих рішень для підключення. Наприклад, програми IoT у сільському господарстві, виробництві та управлінні енергією використовують певні сегменти для оптимізації збору та контролю даних, підвищуючи ефективність та стійкість. Загалом, сегментація мережі не тільки підтримує різноманітні потреби сучасних цифрових екосистем, але й прокладає шлях для інноваційних послуг та бізнес-моделей (табл. 2.3).

Таблиця 2.3.

Класифікація типів мережевих слайсів 5G та їх вимоги до QoS

Тип слайсу	Повна назва та призначення	Ключові технічні вимоги (QoS)	Типові сценарії використання (з тексту)
eMBB	Enhanced Mobile Broadband(Покращений мобільний широкосмуговий доступ)	Висока пропускна здатність, велика ємність каналу. Допускається помірна затримка.	Потокове мультимедіа 4K/8K, віртуальна реальність (VR), доступ до хмарних сервісів.
URLLC	Ultra-Reliable Low Latency Communications(Наднадійний зв'язок з низькою затримкою)	Наднизька затримка (<1 мс), надвисока надійність, відсутність джиттера (jitter).	Автономні транспортні засоби, дистанційна хірургія, промислова автоматизація, екстрені служби.
mMTC	Massive Machine Type Communications(Масивний машинний зв'язок)	Висока щільність підключень (до 1 млн/км²), енергоефективність	Розумні міста (Smart City), розумні лічильники,

		пристроїв, невеликі обсяги даних.	агромоніторинг, датчики навколишнього середовища.

Незважаючи на свій трансформаційний потенціал, сегментація мережі в 5G стикається з кількома суттєвими викликами та обмеженнями, які можуть перешкоджати широкому впровадженню та ефективності. Однією з основних проблем є складність, пов'язана з одночасним управлінням та оркеструванням кількох сегментів. Оскільки кожен сегмент вимагає виділених ресурсів, індивідуальних конфігурацій та постійного моніторингу, оркестрування цих віртуальних мереж вимагає складних механізмів управління та можливостей прийняття рішень у режимі реального часу. Ця складність може призвести до збільшення експлуатаційних витрат та потенційної помилки, особливо зі збільшенням кількості сегментів з різноманітними варіантами використання. Крім того, безпека залишається критичною проблемою в середовищах з кількома орендарями; оскільки кілька віртуальних сегментів співіснують на спільній фізичній інфраструктурі, Вразливості потенційно можуть бути використані для порушення безпеки одного сегмента та компрометації інших. Забезпечення надійних заходів безпеки, включаючи ізоляцію та безпечний міжсекційний зв'язок, є важливим, але складним завданням. Крім того, проблеми сумісності та стандартизації між різними постачальниками та регіонами створюють перешкоди для безперебійної сумісності. Різноманітні апаратні та програмні реалізації можуть призвести до фрагментації, що ускладнює розгортання єдиних рішень у всьому світі. Ці проблеми підкреслюють необхідність комплексних стандартів та структур, які можуть гармонізувати технології, сприяти сумісності та керувати складністю мережевого сегментування.

Заглядаючи в майбутнє, мережеве сегментування готове до значних інновацій, зумовлених інтеграцією з новими технологіями та розробкою стандартизованих протоколів. Одним з перспективних напрямків є інтеграція периферійних обчислень та штучного інтелекту (ШІ), щоб забезпечити справді

динамічне та інтелектуальне сегментування. Периферійні обчислення можуть наблизити обчислювальні ресурси до джерела даних, зменшуючи затримку та покращуючи швидкість реагування в режимі реального часу, що в поєднанні з аналітикою на основі ШІ дозволить мережевим сегментам автоматично адаптуватися до змінних вимог та умов. Ця синергія може оптимізувати розподіл ресурсів, покращити якість обслуговування та сприяти новим варіантам використання, таким як аналітика в режимі реального часу та автономні операції. Водночас, постійний розвиток стандартизованих фреймворків та протоколів, таких як ті, що просуваються в рамках Проекту партнерства третього покоління (3GPP), спрямований на встановлення спільних норм, що забезпечують сумісність між постачальниками та регіонами. Зусилля щодо стандартизації матимуть вирішальне значення для зменшення фрагментації та прискорення розгортання. Ці інновації, ймовірно, суттєво вплинуть на еволюцію мереж 5G, не лише покращуючи поточні можливості, але й закладаючи основу для майбутніх мереж після 5G, таких як 6G, де принципи гнучкого, інтелектуального та сумісного сегментування мережі стануть ще більш невід'ємною частиною глобальної зв'язності та цифрової трансформації [17].

На завершення, сегментація мережі в 5G є ключовим кроком у розвитку телекомунікацій, що дозволяє створювати високоналаштовані, ефективні та масштабовані мережеві послуги, що задовольняють різноманітні потреби сучасних галузей промисловості та споживачів. Її основні принципи включають розділення фізичної інфраструктури на незалежні віртуальні мережі, що підтримуються технологіями віртуалізації, такими як NFV та SDN, що спрощують створення, управління та ізоляцію різних сегментів, таких як eMBB, URLLC та mMTC. Переваги є суттєвими, пропонуючи покращену якість обслуговування, оптимізоване використання ресурсів та підтримку інноваційних програм, таких як Інтернет речей, автономні транспортні засоби та розумні міста. Однак ця технологія також стикається зі значними проблемами, включаючи експлуатаційну складність, вразливості безпеки та проблеми, пов'язані зі стандартизацією та сумісністю. Зрештою, у міру розвитку цих інновацій, сегментація мережі

продовжуватиме формувати еволюцію 5G та далі, сприяючи більш пов'язаній, адаптивній та інтелектуальній цифровій екосистемі.

2.4. Механізми керування трафіком та забезпечення QoS у 5G

Поява технології 5G сповістила про нову еру в бездротовому зв'язку, обіцяючи безпрецедентні швидкості, наднадійне з'єднання та здатність підтримувати безліч різноманітних програм. Центральним фактором реалізації цих можливостей є складні механізми управління трафіком та надійні стратегії забезпечення якості обслуговування (QoS), які забезпечують ефективність, справедливість та надійність мережі. Оскільки мережі 5G розроблені для обслуговування широкого спектру послуг — від розширеного мобільного широкосмугового доступу та масового розгортання Інтернету речей до критично важливих програм — базові структури повинні динамічно адаптуватися до різних вимог та умов. У цьому есе досліджуються основні механізми управління трафіком, що використовуються в мережах 5G, стратегії забезпечення QoS, що лежать в основі диференціації послуг та гарантій продуктивності, а також трансформаційна роль, яку відіграють штучний інтелект (ШІ) та машинне навчання (МН) в оптимізації цих процесів. Всебічне розуміння цих елементів є життєво важливим для оцінки того, як мережі 5G виконують свої обіцянки щодо високої продуктивності та надійності у все більш пов'язаному світі.

Механізми управління трафіком у мережах 5G є фундаментальними для забезпечення ефективного використання мережевих ресурсів, зберігаючи при цьому високу якість обслуговування для користувачів. Одним з ключових методів є динамічний розподіл спектру, який дозволяє мережі адаптивно призначати пропускну здатність на основі попиту в режимі реального часу, тим самим оптимізуючи використання спектру. На відміну від статичного розподілу, цей підхід реагує на коливання навантаження трафіку, гарантуючи, що обмежені спектральні ресурси не витрачаються марно, а програми з високим попитом отримують достатню пропускну здатність. Наприклад, у години пік або під час масштабних подій спектр може динамічно змінюватися для врахування підвищеної активності користувачів. Крім того, пріоритизація трафіку використовується для розмежування різних типів послуг, таких як програми,

чутливі до затримки, такі як автономне керування транспортними засобами або доповнена реальність, та менш чутливі, такі як електронна пошта або завантаження файлів. Ця пріоритизація досягається за допомогою таких механізмів, як класи якості обслуговування (QoS), які призначають різні рівні обслуговування на основі вимог програми. Методи балансування навантаження також сприяють запобіганню перевантаженням, рівномірно розподіляючи трафік між кількома вузькими місцями мережі, уникаючи вузьких місць та забезпечуючи стабільну продуктивність. Такі методи, як керування трафіком та балансування навантаження комірок, дозволяють мережі динамічно перенаправляти користувачів або потоки даних, таким чином підтримуючи оптимальну продуктивність мережі навіть під час раптових стрибків трафіку. Разом ці механізми утворюють стійку структуру, яка адаптується в режимі реального часу до складного та динамічного характеру трафіку 5G, забезпечуючи ефективне використання пропускної здатності та мінімальні перебої в роботі сервісу [18].

Забезпечення якості обслуговування (QoS) у мережах 5G розроблено для задоволення різноманітних вимог до продуктивності широкого спектру застосувань, від наднадійного зв'язку з низькою затримкою до масового розгортання Інтернету речей. Для досягнення цієї мети 5G використовує диференційовані рівні QoS, які класифікують сервіси на класи на основі їхніх конкретних потреб, наприклад, класи з високим пріоритетом для автономних транспортних засобів, що потребують низької затримки та джиттера, проти класів з низьким пріоритетом для потокового медіа, які можуть переносити вищу затримку. Ці диференційовані рівні дозволяють мережі ефективніше розподіляти ресурси, гарантуючи продуктивність там, де це найважливіше. Механізми резервування ресурсів та контролю доступу також є невід'ємною частиною забезпечення QoS; вони забезпечують розподіл достатньої кількості ресурсів до початку надання послуги, запобігаючи перевантаженню та гарантуючи виконання зобов'язань щодо QoS (табл. 2.4).

Таблиця 2.4.

Механізми забезпечення QoS та оптимізація трафіку засобами штучного інтелекту

Тип проблеми в мережі	Традиційний механізм вирішення (QoS)	Оптимізація за допомогою AI/ML (згідно з текстом)
Дефіцит смуги пропускання	Динамічний розподіл спектру: адаптивне призначення частот на основі поточного попиту.	Прогнозування сплесків: превентивне виділення ресурсів до початку події (наприклад, спортивного матчу) на основі історичних даних.
Чутливість до затримки (URLLC)	Пріоритизація трафіку: використання класів QoS для надання переваги критичним додаткам (автономні авто).	Автоматизована маршрутизація: динамічне перенаправлення потоків даних шляхами з найменшим завантаженням у реальному часі.
Перевантаження вузлів	Балансування навантаження: рівномірний розподіл трафіку між базовими станціями.	Виявлення «вузьких місць»: ідентифікація недостатньо використовуваних шляхів та оптимізація розподілу навантаження алгоритмами.
Загрози безпеці (DDoS)	Шифрування та автентифікація: захист доступу до ресурсів.	Виявлення аномалій: автоматичний аналіз патернів трафіку для миттєвого блокування нетипової активності.
Гарантія якості сеансу	Резервування ресурсів: перевірка доступності пропускної здатності перед початком сеансу.	Адаптивне керування: безперервне навчання системи для підтримки SLA (Service Level Agreement) з мінімальним втручанням людини.

Наприклад, коли користувач ініціює відеоконференцію з високою пропускною здатністю, мережа резервує необхідну пропускну здатність і перевіряє доступність ресурсів перед наданням доступу, таким чином підтримуючи якість сеансу. Крім того, такі показники QoS, як затримка, тремтіння, пропускна здатність і втрата пакетів, постійно контролюються та керуються для підтримки стандартів обслуговування. Наприклад, у застосунках віртуальної реальності підтримка низької затримки та тремтіння має вирішальне значення для запобігання заколисуванню та забезпечення захопливого досвіду. Ці стратегії разом дозволяють мережам 5G забезпечувати надійні, адаптовані рівні

обслуговування для широкого спектру застосунків, тим самим підтримуючи інноваційні варіанти використання в різних галузях.

Інтеграція штучного інтелекту (ШІ) та машинного навчання (МН) в управління трафіком 5G впроваджує нову парадигму інтелектуального, автономного управління мережею. Прогнозна аналітика, що базується на алгоритмах МН, аналізує історичні дані трафіку для прогнозування майбутніх моделей трафіку, що дозволяє проактивно коригувати ситуацію. Наприклад, прогнозуючи сплеск потокового відео під час великої спортивної події, мережа може превентивно розподілити додаткові ресурси, зменшуючи затримку та уникаючи перевантажень. Автоматизована маршрутизація трафіку та розподіл ресурсів ще більше підвищують ефективність мережі; Системи на базі штучного інтелекту можуть динамічно перенаправляти потоки даних по мережі для оптимізації розподілу навантаження та зменшення затримки, часто в режимі реального часу без втручання людини. Наприклад, алгоритми штучного інтелекту можуть визначати недостатньо використовувані шляхи та відповідно перенаправляти трафік, балансуючи навантаження та покращуючи загальну продуктивність. Крім того, механізми виявлення аномалій використовують моделі машинного навчання (ML) для виявлення незвичайних моделей трафіку, які можуть свідчити про несправності або загрози безпеці. Раннє виявлення таких аномалій дозволяє швидко виправити ситуацію, запобігаючи потенційним збоям або атакам мережі. Наприклад, виявлення раптового сплеску трафіку, що нагадує розподілену атаку типу «відмова в обслуговуванні» (DDoS), дозволяє мережі швидко ізолювати уражені сегменти. Загалом, ШІ та ML значно покращують управління трафіком у 5G, забезпечуючи прогнозне, адаптивне та автономне керування, що підвищує стійкість, ефективність та безпеку мережі перед обличчям складних та мінливих вимог.

Незважаючи на значний прогрес в управлінні трафіком та забезпеченні якості обслуговування (QoS) у мережах 5G, Існує кілька проблем, які потребують постійних досліджень та розробок. Однією з найважливіших є управління величезним обсягом підключення пристроїв, особливо в середовищах Інтернету

речей (IoT), де мільярди пристроїв — від датчиків до автономних транспортних засобів — повинні одночасно підтримуватися. Цей сплеск кількості підключених пристроїв створює складні схеми трафіку та вимагає складних механізмів, щоб забезпечити відповідний рівень обслуговування кожного пристрою без перевантаження мережі. Крім того, проблеми безпеки та конфіденційності мають першочергове значення; оскільки трафік проходить через кілька вузлів та інтерфейсів, захист конфіденційних даних від перехоплення, несанкціонованого доступу або зловмисних атак стає критично важливим. Впровадження надійних систем шифрування, автентифікації та виявлення аномалій є важливим, але може додати складності та накладних витрат на процеси управління трафіком. Крім того, балансування ефективності мережі зі споживанням енергії створює ще одну проблему. Високопродуктивне управління трафіком часто передбачає збільшення обчислювального навантаження та використання ресурсів, що може призвести до більшого споживання енергії — важливий фактор, враховуючи зростаючий акцент на сталих та зелених технологіях. Досягнення оптимального балансу між підтримкою високої якості обслуговування, забезпеченням безпеки та мінімізацією впливу на навколишнє середовище залишається складною, багатогранною проблемою, яка вимагає інноваційних рішень, адаптованих до унікальних вимог екосистем 5G. Заглядаючи в майбутнє, майбутнє управління трафіком та QoS у мережах 5G готове до трансформаційних інновацій, які ще більше розширять можливості мережі. Одним із перспективних напрямків є інтеграція периферійних обчислень, яка наближає обробку та аналіз даних до кінцевих пристроїв, значно зменшуючи затримку та дозволяючи приймати рішення в режимі реального часу. Цей децентралізований підхід забезпечує швидший час відгуку, що є критично важливим для таких застосувань, як автономне водіння, доповнена реальність та дистанційні операції. Поряд з цим, розробка алгоритмів наступного покоління, що використовують штучний інтелект та машинне навчання, спрямована на оптимізацію маршрутизації трафіку, розподілу ресурсів та управління QoS динамічно та в режимі реального часу, навіть за дуже мінливих мережевих умов. Ці передові алгоритми сприятимуть

створенню більш адаптивних та інтелектуальних мереж, здатних до самовідновлення та самооптимізації, тим самим підвищуючи стійкість та ефективність. Крім того, впровадження стандартизованих протоколів для QoS та управління трафіком буде життєво важливим для безперебійної взаємодії між різними постачальниками, пристроями та мережами. Встановлення глобальних стандартів забезпечить стабільну якість надання послуг та сприятиме інтеграції нових технологій. Разом ці інновації не лише усунуть поточні обмеження, але й відкриють нові можливості для наднадійного зв'язку з низькою затримкою та масштабного розгортання Інтернету речей, формуючи майбутній ландшафт бездротового зв'язку [19].

На завершення, еволюція механізмів управління трафіком та стратегій забезпечення якості обслуговування (QoS) у мережах 5G відображає узгоджені зусилля, спрямовані на задоволення потреб все більш пов'язаного світу. Від динамічного розподілу спектру, пріоритизації трафіку та балансування навантаження до диференційованих рівнів QoS, резервування ресурсів та вдосконаленого керування на основі штучного інтелекту, ці системи працюють синергетично, забезпечуючи ефективну, надійну та безпечну роботу мережі. Однак такі проблеми, як управління масовим підключенням пристроїв, захист безпеки та конфіденційності, а також балансування споживання енергії, продовжують випробовувати стійкість існуючих систем. У майбутньому такі інновації, як периферійні обчислення, алгоритми наступного покоління та стандартизовані протоколи, обіцяють подальше розширення можливостей мереж 5G, дозволяючи їм підтримувати ширший спектр програм з вищою надійністю та меншою затримкою. У міру розвитку технології постійні дослідження та технологічна інтеграція будуть важливими для подолання існуючих перешкод та розкриття повного потенціалу 5G, що зрештою сприятиме більш пов'язаному, інтелектуальному та сталому цифровому майбутньому [20].

3. ДОСЛІДЖЕННЯ ІНТЕГРАЦІЇ ШТУЧНОГО ІНТЕЛЕКТУ В МЕРЕЖІ 5G ДЛЯ ОПТИМІЗАЦІЇ ТРАФІКУ

3.1. Методологія інтеграції AI-моделей у мережеву інфраструктуру 5G

Розгортання штучного інтелекту (ШІ) в інфраструктурі мережі 5G сповіщає про нову еру інтелектуальних, адаптивних та високоефективних систем зв'язку. Щоб повністю використати потенціал ШІ, важливо створити структуровану та комплексну методологію, яка б спрямовувала його інтеграцію в складну екосистему мереж 5G. Цей процес включає ретельне планування на кількох етапах, включаючи визначення точних вимог, проектування оптимальних архітектурних рамок та розробку надійних процедур збору та попередньої обробки даних. Систематично розглядаючи ці сфери, мережеві оператори можуть забезпечити безперебійне впровадження моделей ШІ, які підвищують продуктивність мережі, оптимізують розподіл ресурсів та покращують взаємодію з користувачем. У підрозділі досліджуються критичні кроки, пов'язані з інтеграцією ШІ в інфраструктуру 5G, підкреслюючи важливість кожного етапу у створенні ефективного та масштабованого середовища 5G на базі ШІ [21].

Початковий етап інтеграції штучного інтелекту в мережі 5G зосереджений на чіткому визначенні вимог, які будуть орієнтувати розробку та розгортання. Фундаментальним кроком є визначення того, які параметри трафіку потребують оптимізації. Наприклад, такі показники, як затримка, пропускна здатність та втрата пакетів, є життєво важливими для підтримки високоякісного обслуговування, особливо для чутливих до затримки програм, таких як автономні транспортні засоби або віртуальна реальність. Встановлення цих параметрів допомагає зосередити моделі штучного інтелекту на найбільш впливових аспектах продуктивності мережі. Далі, вирішальним є формування конкретних критеріїв продуктивності для моделей штучного інтелекту. Ці критерії можуть включати порогові значення точності, час відгуку та показники надійності, які служать орієнтирами для успішного впровадження. Наприклад, модель штучного

інтелекту, розроблена для прогнозування перевантаження мережі, повинна досягти певної точності прогнозування протягом визначеного часового вікна, щоб вважатися ефективною. Крім того, встановлення технічних обмежень інфраструктури 5G, таких як можливості обробки, обмеження затримки та доступна пропускна здатність, необхідне для забезпечення сумісності рішень штучного інтелекту з існуючим апаратним та програмним середовищем. Ці обмеження впливають на вибір алгоритмів штучного інтелекту та стратегій розгортання. Наприклад, легкі моделі можуть бути кращими в периферійних розгортаннях, де обчислювальні ресурси обмежені, тоді як складніші моделі можуть використовуватися в централізованих ядрах з вищою обчислювальною потужністю. Разом ці кроки створюють основу для індивідуальних рішень ШІ, які відповідають конкретним потребам мережі, враховуючи при цьому операційні обмеження.

Архітектурний дизайн інтеграції ШІ включає стратегічні рішення щодо того, де розгортати модулі ШІ в мережі 5G, враховуючи унікальні переваги кожного місця розташування — мережі радіодоступу (RAN), ядра або периферії. Розгортання ШІ на рівні RAN може забезпечити оптимізацію радіоресурсів та передач у режимі реального часу, зменшуючи затримку та покращуючи взаємодію з користувачем. І навпаки, розміщення ШІ в ядрі дозволяє проводити комплексний аналіз даних та централізоване прийняття рішень, що може оптимізувати управління мережею та сприяти широкомасштабному застосуванню політик. З іншого боку, розгортання на периферії пропонує перевагу близькості до кінцевих користувачів, дозволяючи швидко реагувати на локалізовані проблеми та підтримуючи критично важливі до затримки програми, такі як автономне водіння або доповнена реальність. Розробка надійної структури обміну даними між компонентом ШІ та існуючими функціями 5G є не менш важливою. Це передбачає створення безпечних та ефективних каналів зв'язку, таких як API або черги повідомлень, для забезпечення безперебійного потоку даних. Наприклад, модулі ШІ можуть отримувати оновлення ключових показників ефективності в режимі

реального часу від RAN та надсилати керуючі сигнали для регулювання формування променя або рівнів потужності [22].

Базуючись на аналізі вимог до затримки, для реалізації системи оптимізації трафіку в даній роботі обрано архітектуру з розгортанням модуля ШІ на рівні периферійних обчислень (Multi-access Edge Computing — MEC). Такий підхід дозволяє обробляти дані безпосередньо поблизу джерела їх генерації (gNodeB), мінімізуючи навантаження на транспортну мережу. Розроблена структурна схема інтеграції інтелектуального модуля в середовище 5G, що демонструє повний цикл від збору телеметрії до генерації керуючого впливу, наведена на рис. 3.1.



Рис. 3.1. Розроблена архітектура інтеграції інтелектуального модуля керування в середовище 5G MEC

Побудова детальної діаграми, що ілюструє взаємодії, такі як те, як дані трафіку передаються від користувацького обладнання до модуля штучного інтелекту, а потім впливають на параметри мережі, забезпечує ясність та гарантує сумісність. Таке архітектурне планування гарантує, що компоненти штучного інтелекту інтегровані цілісним чином, максимізуючи їхній вплив на продуктивність мережі, зберігаючи при цьому стабільність системи.

Критичним аспектом інтеграції штучного інтелекту є встановлення комплексної процедури збору та підготовки даних. Першим кроком є визначення

та пошук відповідних даних; це включає ключові показники ефективності (KPI) з RAN (наприклад, сила сигналу, коефіцієнти успішної передачі), ключові показники ефективності основної мережі (наприклад, пропускна здатність даних, безперервність сеансу) та показники користувацького обладнання (UE), такі як тип пристрою, місцезнаходження та моделі використання програм. Точні та різноманітні джерела даних дозволяють навчати надійні моделі штучного інтелекту, здатні до узагальнення в різних сценаріях. Після отримання даних створення ефективного конвеєра попередньої обробки є важливим для забезпечення якості та узгодженості даних. Цей конвеєр може включати операції очищення для обробки відсутніх або пошкоджених даних, нормалізацію для стандартизації масштабів ознак та інженерію ознак для вилучення відповідних атрибутів, таких як агрегування KPI за певні часові вікна або отримання нових показників, таких як пикованість трафіку. Крім того, розділення підготовлених даних на навчальні та тестові набори є життєво важливим для об'єктивної оцінки продуктивності моделі. Наприклад, історичні дані трафіку можна розділити на сегменти, що представляють різний час доби або стан мережі, щоб перевірити адаптивність моделі штучного інтелекту. Впровадження автоматизованих конвеєрів за допомогою таких інструментів, як Apache Kafka або TensorFlow Data Validation, може сприяти постійному оновленню даних та перенавчанню моделей. Цей ретельний процес управління даними гарантує, що моделі штучного інтелекту побудовані на високоякісних, репрезентативних наборах даних, що зрештою призводить до більш точних та надійних рішень для оптимізації мережі [23].

Розробка ефективної моделі штучного інтелекту, адаптованої до конкретного сценарію 5G, починається з вибору відповідної архітектури машинного навчання на основі поставленого завдання. Наприклад, якщо метою є прогнозування часу- Для залежних мережевих моделей, таких як коливання трафіку або шаблони передачі, моделі, такі як мережі з довгою короткочасною пам'яттю (LSTM), добре підходять завдяки своїй здатності фіксувати послідовні залежності. І навпаки, згорткові нейронні мережі (CNN) можуть бути використані для аналізу

просторових даних або шаблонів сигналів, тоді як агенти навчання з підкріпленням (RL) особливо ефективні для завдань динамічного прийняття рішень, таких як розподіл ресурсів та адаптивна маршрутизація. Після вибору типу моделі налаштування параметрів навчання, включаючи швидкість навчання, розмір пакета, кількість епох та методи регуляризації, має вирішальне значення для оптимізації продуктивності моделі. Точне налаштування цих параметрів за допомогою ітеративних експериментів дозволяє моделі ефективно навчатися та уникати перенавчання. Наступний крок включає проведення комплексних процесів навчання та валідації з використанням реальних мережевих даних або емульованих сценаріїв, що імітують фактичні робочі умови. Це навчання передбачає заповнення моделі історичними наборами даних KPI, показниками користувачів та шаблонами трафіку, що дозволяє їй вивчати основні взаємозв'язки. Валідація на окремих наборах даних забезпечує прогностичну точність та узагальнюваність моделі, що є важливими для надійного розгортання. Використання таких методів, як перехресна перевірка та рання зупинка, допомагає додатково вдосконалити модель, забезпечуючи її надійність перед інтеграцією в реальні мережі.

Після навчання та перевірки моделі ШІ, розробка механізмів для безперешкодної інтеграції в мережеву інфраструктуру 5G стає надзвичайно важливою. Цей процес включає розробку інтерфейсів прикладного програмування (API) або інших протоколів інтерфейсу, які сприяють взаємодії між модулем ШІ та існуючими мережевими компонентами, такими як основні мережеві функції або вузли радіодоступу. Ці інтерфейси повинні підтримувати ефективний та безпечний обмін даними, дозволяючи моделі ШІ отримувати ключові показники ефективності (KPI) в режимі реального часу, дані про трафік та інші відповідні показники, а також надсилати керуючі сигнали для оптимізації мережі. Встановлення механізму обміну телеметрією є важливим для підтримки потоку даних у режимі реального часу; це може включати розгортання черг повідомлень, RESTful API або потокових протоколів, таких як WebSockets або gRPC, для забезпечення зв'язку з низькою затримкою. Крім того, логіка, яка керує тим, як

вихідні дані ШІ впливають на поведінку мережі, повинна бути ретельно розроблена. Наприклад, якщо ШІ прогнозує перевантаження в певній комірці, система керування може динамічно коригувати параметри формування променя, виділяти додаткові ресурси або перенаправляти трафік для підтримки якості обслуговування (QoS). Розробка цієї логіки прийняття рішень вимагає детального розуміння мережевих протоколів та операційних обмежень, що гарантує ефективність дій, керованих ШІ, та відповідність стандартам безпеки та продуктивності. Інтеграція цих механізмів гарантує, що моделі ШІ можуть активно брати участь в управлінні мережею в режимі реального часу, забезпечуючи адаптивне та інтелектуальне реагування на різні мережеві умови (табл. 3.1).

Таблиця 3.1.

Технічні специфікації та інструментарій розробленої підсистеми інтелектуального керування

Етап життєвого циклу	Обраний інструмент / Технологія	Обґрунтування вибору (Специфіка розробки)	Цільовий показник ефективності
Збір телеметрії	gRPC / Protocol Buffers	Забезпечення мінімальної затримки при передачі великих потоків даних з gNodeB.	Затримка передачі даних < 2 мс.
Попередня обробка	Apache Kafka & Python Pandas	Створення конвеєра для фільтрації шумів та нормалізації метрик у реальному часі.	Обробка до 100 000 записів/сек.
Архітектура AI-моделі	Deep Reinforcement Learning (DQN)	Здатність навчатися оптимальним стратегіям розподілу ресурсів без учителя через взаємодію з середовищем.	Максимізація функції винагороди (QoS).
Середовище навчання	PyTorch + NS-3 (Gym Interface)	Використання бібліотеки PyTorch для побудови нейромережі та NS-3 для генерації реалістичного середовища.	Точність моделі > 95% на валідаційній вибірці.
Інтерфейс	RESTful API	Стандартизований протокол	Час реакції

керування		для передачі керуючих команд від AI-агента до планувальника мережі.	системи < 10 мс (Real-time).

Перед повномасштабним розгортанням вкрай важливо ретельно протестувати інтегровану систему 5G з покращеним ШІ в контрольованих середовищах, щоб оцінити її продуктивність та стабільність. Створення комплексних тестових сценаріїв, що імітують різні навантаження трафіку — від нормальних до пікових умов — дозволяє оцінити, як модель ШІ впливає на поведінку мережі за різних операційних навантажень. Ці сценарії повинні включати типові дії користувачів, раптові сплески попиту та граничні випадки, такі як збої пристроїв або загрози безпеці. Ключові показники для оцінки включають параметри якості обслуговування (QoS), такі як затримка, пропускна здатність та втрата пакетів, а також показники якості досвіду (QoE), що сприймаються кінцевими користувачами, такі як якість потокової передачі або стабільність з'єднання. Вимірювання впливу моделі ШІ включає порівняння цих показників з базовими даними, отриманими з традиційних методів управління трафіком без використання ШІ. Цей порівняльний аналіз допомагає визначити, чи призводить інтеграція ШІ до вимірюваних покращень ефективності мережі, задоволеності користувачів та використання ресурсів. Крім того, контрольоване тестування дозволяє виявити потенційні проблеми, такі як ненавмисні взаємодії або помилки рішень, і надає можливість удосконалити моделі та механізми контролю перед широким розгортанням. Цей етап систематичного тестування є важливим для перевірки ефективності та безпеки рішень на основі ШІ в складній екосистемі мереж 5G [24].

Після успішного тестування та валідації наступним критичним кроком є формулювання комплексних рекомендацій щодо розгортання моделі ШІ в діючій мережі 5G. Цей процес починається з встановлення чітких критеріїв, які визначають готовність модуля ШІ до реального розгортання. Такі критерії можуть включати досягнення заздалегідь визначених контрольних показників продуктивності, демонстрацію стабільності протягом тривалих періодів

тестування та забезпечення відповідності стандартам безпеки. Після досягнення цих критеріїв необхідно провести ретельну оцінку ризиків, щоб виявити потенційні вразливості, включаючи проблеми конфіденційності даних, проблеми стабільності системи та операційні ризики, пов'язані з помилками прийняття рішень ШІ. Крім того, оцінка необхідних обчислювальних ресурсів, таких як обчислювальна потужність, пам'ять та пропускна здатність мережі, є важливою для забезпечення того, щоб інфраструктура могла підтримувати операції ШІ в режимі реального часу без негативного впливу на існуючі послуги. На основі цих оцінок слід розробити поетапний план розгортання для поступового впровадження моделі ШІ в мережу оператора. Цей план може включати початкове розгортання в обмежених областях або певних сегментах мережі, ретельний моніторинг продуктивності та впливу, а також поступове розширення після встановлення впевненості в надійності системи. Ретельно структурована стратегія розгортання мінімізує перебої, дозволяє проводити ітеративні вдосконалення та гарантує ефективну та безпечну роботу мережі, вдосконаленої ШІ, в реальних умовах.

Інтеграція моделей штучного інтелекту в мережеву інфраструктуру 5G – це складне, але дуже корисне починання, яке обіцяє значне покращення продуктивності мережі, управління ресурсами та взаємодії з користувачем. Процес починається з чіткого формулювання вимог, зосереджуючись на визначенні ключових параметрів трафіку для оптимізації, встановленні контрольних показників продуктивності та розумінні технічних обмежень існуючої інфраструктури.

Проектування відповідної архітектури включає вибір оптимальних місць розгортання – будь то RAN, ядро чи периферія – та створення надійних механізмів обміну даними для забезпечення безперебійної комунікації між модулями штучного інтелекту та мережевими функціями. Розробка комплексного конвеєра збору та попередньої обробки даних забезпечує наявність високоякісних наборів даних, необхідних для ефективного навчання моделі. Створення та навчання моделей штучного інтелекту, адаптованих до конкретних сценаріїв, таких як

прогнозування трафіку або розподіл ресурсів, вимагає ретельного вибору алгоритмів та точного налаштування параметрів для максимальної точності та надійності. Інтеграція цих моделей у мережу включає проектування безпечних API та механізмів телеметрії в режимі реального часу, що дозволяє рішенням на основі штучного інтелекту динамічно впливати на управління трафіком. Ретельне тестування в контрольованих середовищах дозволяє оцінити вплив на QoS та QoE, забезпечуючи належне функціонування системи перед повним розгортанням. Зрештою, встановлення чітких критеріїв готовності, оцінка ризиків та планування поетапного розгортання допомагають пом'якшити потенційні проблеми та сприяти безперешкодній інтеграції в діючі мережі. Загалом, ця структурована методологія забезпечує дорожню карту для використання трансформаційного потенціалу штучного інтелекту в системах 5G, прокладаючи шлях для розумніших, більш адаптивних комунікаційних мереж, які відповідають вимогам майбутнього зв'язку.

3.2. Алгоритми оптимізації трафіку на основі штучного інтелекту

Оскільки міське населення продовжує зростати, а технологічний прогрес поширюється, управління та оптимізація дорожніх систем стали критично важливими для забезпечення ефективності транспортних та комунікаційних мереж. Штучний інтелект (ШІ) пропонує інноваційні рішення цих проблем за допомогою складних алгоритмів, здатних до прогнозного аналізу та прийняття рішень у режимі реального часу. Алгоритми оптимізації трафіку на основі ШІ використовують методи машинного навчання, глибокого навчання та навчання з підкріпленням для покращення потоку трафіку, ефективного розподілу радіоресурсів та визначення оптимальних стратегій маршрутизації. У підрозділідосліджуються три основні компоненти оптимізації трафіку на основі ШІ: алгоритми прогнозування навантаження, моделі динамічного розподілу радіоресурсів та адаптивні системи маршрутизації трафіку. Розуміючи ці елементи, зацікавлені сторони можуть розробляти розумніші, більш чуйні системи

управління трафіком, які покращують мобільність, зменшують затори та оптимізують продуктивність мережі.

Фундаментальним елементом управління дорожнім рухом на основі штучного інтелекту є розробка алгоритмів прогнозування навантаження, які прогнозують обсяги трафіку з високою точністю, аналізуючи історичні дані та дані в реальному часі. Один з відомих підходів включає створення моделей прогнозування обсягу трафіку з використанням архітектур глибокого навчання, таких як довгострокова пам'ять (LSTM) та стробувані рекурентні одиниці (GRU).

У рамках даного дослідження для кожної з визначених задач оптимізації було підібрано найбільш ефективний математичний апарат, що дозволяє врахувати специфіку трафіку 5G. Зведена специфікація обраних алгоритмів та вхідних даних для кожної підсистеми наведена в табл. 3.2. Особливу увагу в роботі приділено задачі динамічного розподілу ресурсів, для вирішення якої розроблено архітектуру агента на базі глибокого навчання з підкріпленням (Deep Reinforcement Learning). Схема взаємодії цього агента з середовищем мережі радіодоступу, що демонструє замкнений цикл керування та формування винагороди, представлена на рис. 3.2.

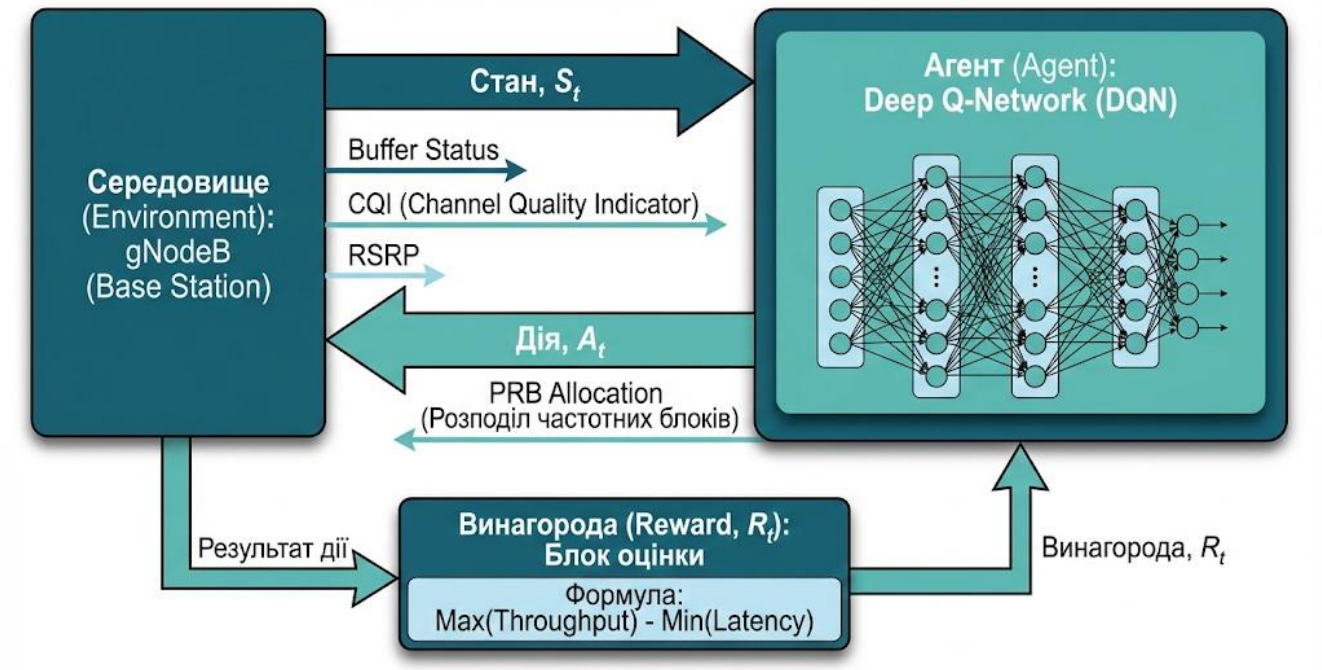


Рис. 3.2. Схема взаємодії DRL-агента з середовищем 5G RAN для динамічного розподілу ресурсів

Ці моделі чудово фіксують часові залежності в послідовних даних, що робить їх ідеальними для прогнозування змін потоку трафіку в різних часових масштабах — погодинно, щодня чи щотижня. Наприклад, модель LSTM, навчена на історичних даних про трафік, може точно прогнозувати піки в години пік, дозволяючи містобудівникам превентивно коригувати світлофори або розгортати додаткові ресурси. Налаштування параметрів прогнозування для різних часових інтервалів включає налаштування гіперпараметрів, таких як розмір вікна, швидкість навчання та кількість шарів, для оптимізації прогнозів відповідно до певних часових закономірностей. Крім того, інтеграція цих прогнозів у механізми планування ресурсів 5G дозволяє динамічно регулювати пропускну здатність мережі на основі очікуваного навантаження трафіку. Ця інтеграція гарантує, що комунікаційні мережі можуть обробляти різкі стрибки попиту на дані в періоди високого трафіку, зменшуючи затримку та покращуючи взаємодію з користувачем. Наприклад, під час масштабних подій або надзвичайних ситуацій точне прогнозування навантаження в поєднанні з управлінням ресурсами 5G може підтримувати безперебійне з'єднання, підтримуючи як транспортні системи, так і комунікаційні пристрої [25].

Окрім прогнозування обсягів трафіку, алгоритми штучного інтелекту можуть динамічно розподіляти радіоресурси за допомогою моделей навчання з підкріпленням (RL) та глибокого навчання з підкріпленням (DRL), які адаптивно оптимізують розподіл блоків фізичних ресурсів (PRB) у бездротових мережах. Ці моделі вивчають оптимальні політики, взаємодіючи з середовищем, постійно оновлюючи свої стратегії на основі поточних мережевих умов та зворотного зв'язку щодо продуктивності. Наприклад, модель RL може визначати, як розподіляти PRB між користувачами в режимі реального часу, балансуючи потребу у високій пропускній здатності зі справедливістю та вимогами до якості обслуговування (QoS). Такий адаптивний розподіл гарантує, що високопріоритетні програми, такі як зв'язок автономних транспортних засобів або служби екстреної допомоги, отримують необхідну пропускну здатність без

недовикористання мережевих потужностей. Для подальшого підвищення швидкості реагування ці алгоритми включають механізми корекції параметрів, які працюють у режимі реального часу, коригуючи свої політики в міру зміни мережевих умов. Це постійне коригування запобігає втраті ресурсів та підтримує оптимальну продуктивність мережі, навіть за непередбачуваних коливань трафіку. Результатом є стійка, самооптимізована система, здатна підтримувати дедалі вимогливіші вимоги до даних сучасної транспортної та комунікаційної інфраструктури.

Ефективне управління дорожнім рухом також значною мірою залежить від вибору оптимальних маршрутів у транспортних та комунікаційних мережах. Моделі машинного навчання, розроблені для оцінки пропускної здатності різних сегментів мережі, можуть оцінювати поточні рівні навантаження, точки перевантаження та якість з'єднання для прийняття рішень щодо маршрутизації. Створюючи систему оцінки на основі машинного навчання, транспортні органи та оператори мереж можуть визначити, які маршрути найкраще здатні обробляти додатковий трафік, таким чином мінімізуючи затримки та уникаючи вузьких місць. Наприклад, модель, навчена на даних у режимі реального часу з датчиків, камер та мережевих журналів, може прогнозувати пропускну здатність сегмента та динамічно рекомендувати альтернативні шляхи, коригуючи стратегії маршрутизації в міру зміни умов. Впровадження алгоритму для визначення найкращого маршруту передбачає аналіз кількох параметрів, таких як поточне навантаження, прогнозовані перевантаження та історична продуктивність, для вибору найефективнішого шляху. Крім того, механізми динамічного перемикання маршрутів дозволяють змінювати маршрути в режимі реального часу на основі зміни транспортного навантаження, зменшуючи перевантаження та підвищуючи загальну ефективність системи. Наприклад, автономні транспортні засоби можуть використовувати ці алгоритми маршрутизації для навігації менш перевантаженими дорогами, оптимізуючи час подорожі та витрату палива. Аналогічно, телекомунікаційні мережі можуть перенаправляти трафік даних для підтримки високої пропускної здатності та низької затримки, особливо в періоди

пікового навантаження або під час перебоїв. У сукупності ці адаптивні системи маршрутизації ілюструють, як штучний інтелект може значно підвищити стійкість та ефективність складних мереж трафіку та зв'язку [26].

Важливим аспектом підтримки ефективних систем трафіку є раннє виявлення та усунення точок перевантаження в мережах. Для досягнення цієї мети алгоритми на основі штучного інтелекту зосереджуються на створенні моделей, здатних класифікувати перевантажені вузли — як у транспортній інфраструктурі, так і в мережах зв'язку — на основі даних у режимі реального часу, таких як щільність трафіку, довжина черг та пропускна здатність. Ці моделі класифікації часто використовують методи машинного навчання, такі як дерева рішень, методи опорних векторів, або нейронні мережі, навчені на історичних моделях перевантажень, щоб виявляти потенційні вузькі місця до їх загострення. Після виявлення перевантаженого вузла може бути активований механізм автоматичного перенаправлення трафіку, який динамічно перенаправляє транспортні засоби або потоки даних на менш перевантажені маршрути або канали. Ця система гарантує, що критично важливі послуги, такі як екстрений зв'язок або автономна навігація транспортних засобів, мають пріоритет під час перенаправлення, запобігаючи затримкам, які можуть поставити під загрозу безпеку або експлуатаційну ефективність. Крім того, логіка перенаправлення включає алгоритми пріоритетів, які призначають різні рівні важливості різним типам трафіку — рятувальникам, громадському транспорту або потокам даних з високим пріоритетом — забезпечуючи безперебійний потік життєво важливих служб навіть у періоди пікових перевантажень. Автоматизуючи ці процеси, системи управління дорожнім рухом стають більш чутливими та здатними проактивно зменшувати перевантаження, тим самим підвищуючи загальну стійкість мережі та надійність обслуговування.

Оптимізація якості обслуговування (QoS) та вражень кінцевого користувача (QoE) у сучасних мережах дорожнього руху та зв'язку є основоположною для задоволення користувачів та операційної ефективності. Моделі на основі штучного інтелекту для цієї мети розроблені для оцінки показників QoS, таких як

затримка, тремтіння, втрата пакетів та пропускна здатність, у режимі реального часу, що забезпечує негайне уявлення про стан та продуктивність мережі. Ці алгоритми оцінки використовують безперервний моніторинг та машинне навчання для виявлення відхилень від допустимих порогів та прогнозування потенційного погіршення, перш ніж воно вплине на кінцевих користувачів. На основі цих оцінок параметри трафіку, такі як розподіл пропускної здатності, пріоритетність пакетів та буфери затримки, динамічно коригуються для оптимізації рівнів QoE (рис. 3.2).

Таблиця 3.2.

Вибір та специфікація алгоритмічного забезпечення для підсистем оптимізації трафіку

Функціональна підсистема	Обраний алгоритм (Model)	Вхідні дані (Input Features)	Вихідний параметр (Output)	Цільовий показник ефективності
Прогнозування навантаження	LSTM (Long Short-Term Memory)	Історичні часові ряди: обсяг трафіку, час доби, кількість активних користувачів (UE).	Прогнозоване навантаження на наступний слот часу (t+1).	Ефективність у роботі з послідовними даними та часовими залежностями.
Розподіл радіоресурсів (RRM)	Deep Q-Network (DQN)	Стан черги пакетів, індикатор якості каналу (CQI), вимоги QoS конкретного слайсу.	Матриця розподілу фізичних ресурсних блоків (PRB Allocation Map).	Здатність навчатися оптимальним стратегіям у динамічному середовищі в реальному часі.
Маршрутизація трафіку	Graph Neural Network (GNN)	Топологія мережі, поточна затримка на лінках,	Вагові коефіцієнти для алгоритму пошуку шляху (Next Hop Probability).	Можливість врахування складних просторових зв'язків між

		пропускна здатність вузлів.		вузлами мережі.
Виявлення перевантажень	Random Forest Classifier	Статистика втрати пакетів, джиттер (jitter), завантаження CPU базової станції.	Клас стану мережі: "Норма", "Передперевантаження", "Перевантаження".	Висока точність класифікації та інтерпретованість результатів.

Наприклад, у періоди високого трафіку адаптивні алгоритми управління чергами можуть надавати пріоритет додаткам, чутливим до затримки, таким як відеоконференції або автономне керування транспортними засобами, зменшуючи затримки та втрату пакетів. Крім того, ці моделі включають механізми зворотного зв'язку, які навчаються на основі взаємодії з користувачами та показників продуктивності, вдосконалюючи свої стратегії оптимізації з часом. Завдяки впровадженню таких адаптивних систем мережі можуть підтримувати високу якість обслуговування навіть за коливань навантажень, забезпечуючи безперебійний досвід для користувачів та підтримуючи надійність критично важливих транспортних та комунікаційних послуг [27].

Енергоефективність стає все більш важливою в управлінні сучасними мережами трафіку та зв'язку, особливо враховуючи зростання споживання енергії щільною міською інфраструктурою та центрами обробки даних. Аналіз трафіку на основі штучного інтелекту відіграє вирішальну роль у цій галузі, розробляючи моделі, здатні прогнозувати активність радіосубфреймів, що передбачає розуміння моделей використання радіоресурсів з плином часу. Ці прогнозні моделі використовують історичні дані, вимірювання в режимі реального часу та методи машинного навчання для прогнозування періодів низької активності, тим самим виявляючи можливості для економії енергії. На основі цих прогнозів мережі можуть впроваджувати режими енергозбереження, такі як вимкнення або зниження активності певних радіокомпонентів, без шкоди для якості обслуговування (QoS). Наприклад, у години поза піком або періоди низького

попиту частини мережевої інфраструктури можуть автоматично перемикатися в режим очікування, зберігаючи енергію, зберігаючи при цьому достатню потужність для неочікуваних стрибків навантаження. Крім того, автоматизація процедур увімкнення/вимкнення мережевих компонентів на основі балансування навантаження забезпечує компроміс між енергоефективністю та якістю обслуговування. Завдяки інтелектуальному управлінню робочими станами мережевих елементів, підходи на основі штучного інтелекту сприяють зниженню загального споживання енергії, зниженню експлуатаційних витрат та підтримці сталого розвитку міст без шкоди для продуктивності чи надійності.

Інтеграція алгоритмів штучного інтелекту в багатоканальні периферійні обчислення (MEC) та периферійні середовища є значним прогресом в оптимізації трафіку, забезпечуючи локалізовану обробку даних з низькою затримкою. Розгортання моделей управління трафіком безпосередньо на периферійних вузлах, таких як граничні сервери, дозволяє проводити аналіз у режимі реального часу та приймати рішення ближче до джерела даних — чи то в транспортних вузлах, базових станціях стільникового зв'язку чи мережевих комутаторах — тим самим зменшуючи залежність від централізованих центрів обробки даних. Такий децентралізований підхід підвищує швидкість реагування та масштабованість, особливо в густонаселених міських районах, де швидка адаптація є важливою. Побудова виділених каналів передачі даних для локальної обробки сприяє ефективній комунікації між граничними пристроями, датчиками та моделями штучного інтелекту, забезпечуючи швидку передачу та обробку відповідної інформації про трафік та мережу без перевантаження основної мережі. Крім того, для підтримки узгодженості та ефективності управління трафіком на основі штучного інтелекту вкрай важливо забезпечити синхронізацію між алгоритмами на периферії та загальними системами управління мережею 5G. Ця синхронізація дозволяє безперешкодно координувати розподіл ресурсів, рішення щодо маршрутизації та контроль перевантаження. узгодження локалізованих дій із політиками та цілями всієї мережі. Завдяки інтеграції моделей штучного інтелекту в середовища MEC, дорожні системи стають більш стійкими, адаптивними та

здатними справлятися з динамічними вимогами сучасного міського транспорту та комунікаційних мереж, що зрештою призводить до створення розумніших міст та ефективнішого надання послуг [28].

На завершення, оптимізація трафіку на основі штучного інтелекту охоплює комплексний набір алгоритмів, призначених для підвищення ефективності, швидкості реагування та стійкості сучасних транспортних та комунікаційних мереж. Розробка моделей прогнозування навантаження з використанням архітектур глибокого навчання, таких як LSTM та GRU, дозволяє точно прогнозувати обсяги трафіку, сприяючи проактивному плануванню ресурсів, особливо в інфраструктурах 5G. Динамічний розподіл радіоресурсів за допомогою моделей навчання з підкріпленням забезпечує ефективний та адаптивний розподіл пропускної здатності мережі, підтримуючи високі стандарти QoS за умови коливань попиту. Одночасно інтелектуальні алгоритми маршрутизації оцінюють стан мережі в режимі реального часу, щоб вибрати найефективніші шляхи, динамічно перенаправляючи трафік, щоб уникнути перевантажень та покращити загальну продуктивність системи. Крім того, механізми виявлення перевантажень на основі штучного інтелекту класифікують перевантажені вузли та автоматично перенаправляють трафік, надаючи пріоритет критичним послугам та запобігаючи збоям системи. Зусилля щодо оптимізації QoS та QoE шляхом оцінки в режимі реального часу та адаптивного налаштування параметрів додатково гарантують високу задоволеність користувачів. Енергоефективність також підвищується за допомогою прогнозних моделей, які визначають періоди низької активності, дозволяючи мережевим компонентам працювати в режимах енергозбереження без шкоди для якості обслуговування. Зрештою, інтеграція цих алгоритмів штучного інтелекту в середовище управління міськими та інформаційними системами (MEC) та периферійні середовища забезпечує швидке, локалізоване прийняття рішень, сприяючи створенню стійких, масштабованих та інтелектуальних систем управління дорожнім рухом. Разом ці інновації підкреслюють трансформаційний потенціал штучного інтелекту у створенні розумнішої та ефективнішої міської

інфраструктури, що зрештою прокладає шлях до сталого та пов'язаного майбутнього.

3.3. Експериментальне моделювання та оцінювання ефективності інтеграції AI у 5G

Швидкий розвиток технології 5G сповістив про нову еру покращеного підключення, збільшення швидкості передачі даних та зниження затримки, що фундаментально змінило комунікаційні мережі в усьому світі. Щоб максимально використати потенціал 5G, дослідники все частіше звертаються до штучного інтелекту (ШІ) як засобу оптимізації продуктивності мережі, покращення розподілу ресурсів та забезпечення адаптивних стратегій управління. Однак інтеграція ШІ в системи 5G вимагає ретельного експериментального моделювання та оцінки для оцінки його ефективності в різних операційних сценаріях. У підрозділідосліджується процес проектування експериментального середовища, збору базових показників продуктивності та інтеграції модулів ШІ в симуляції мережі 5G з метою систематичної оцінки ефективності вдосконалень, заснованих на ШІ, у бездротових мережах наступного покоління [29].

Основа будь-якої емпіричної оцінки ШІ в 5G залежить від ретельного створення реалістичного та комплексного експериментального середовища. Цей процес починається зі створення надійних мережевих симуляторів або емуляторів, які точно відтворюють поведінку мережі 5G; поширені інструменти включають NS-3, srsRAN або власні платформи моделювання. Ці інструменти дозволяють дослідникам моделювати складні мережеві компоненти, такі як базові станції, користувацьке обладнання та основні мережеві функції, створюючи контрольоване середовище для тестування. Після створення інфраструктури моделювання наступним кроком є розгортання модуля ШІ як окремого функціонального блоку в цьому середовищі. Цей модульний підхід гарантує, що ШІ може бути безперешкодно інтегрований, що дозволяє проводити ізольоване тестування та ітеративну розробку. Визначення сценаріїв навантаження, таких як різна щільність користувачів, моделі мобільності та типи трафіку, має вирішальне значення для стрес-тестування системи за різних умов попиту. Наприклад, моделювання трафіку в години пік з тисячами користувачів, які генерують відео,

голосовий та інформаційний трафік, допомагає оцінити здатність ШІ адаптуватися в режимі реального часу. Ця конфігурація забезпечує основу для подальшої оцінки продуктивності та порівняльного аналізу, гарантуючи, що експериментальне середовище відображає реалістичні мережеві умови для отримання змістовних висновків. Перш ніж оцінювати переваги інтеграції штучного інтелекту, важливо встановити базовий рівень показників продуктивності мережі в режимі роботи за замовчуванням. Це включає систематичне вимірювання ключових показників, таких як пропускна здатність, затримка та втрата пакетів за різних сценаріїв навантаження, без будь-якого втручання штучного інтелекту. Наприклад, вимірювання пропускної здатності в умовах високого навантаження показують максимальні швидкості передачі даних, досяжні без динамічної оптимізації, тоді як оцінки затримки висвітлюють затримки, спричинені традиційними механізмами управління трафіком. Показники втрати пакетів додатково вказують на надійність мережі під навантаженням. Під час цього етапу існуючі протоколи управління трафіком, такі як алгоритми планування, контроль перевантаження та розподіл ресурсів, підтримуються в їхній стандартній конфігурації, щоб служити контрольним орієнтиром. Ці базові показники є критично важливими, оскільки вони забезпечують точку відліку, відносно якої можна кількісно оцінити вплив адаптацій, керованих штучним інтелектом. Фіксуючи операційні параметри в режимі за замовчуванням, дослідники можуть чітко окреслити покращення або регресії, пов'язані з інтеграцією штучного інтелекту, таким чином встановлюючи початкові умови, необхідні для ретельної оцінки.

Основою експериментальної оцінки є безшовна інтеграція моделі штучного інтелекту в середовище моделювання 5G, а потім тестування в реальному часі за різних умов навантаження. Цей процес починається з вбудовування модуля штучного інтелекту в цикл управління трафіком, де він функціонує як інтелектуальний прогнозатор або оптимізатор для розподілу ресурсів, рішень про передачу або контролю перевантажень. Наприклад, модель штучного інтелекту, навчена на історичних даних про трафік, може прогнозувати точки неминучого

перевантаження та пропонувати превентивне перенаправлення потоків трафіку, тим самим підвищуючи ефективність мережі.

На рис. 3.3 відображено залежність середньої затримки передачі пакетів (Average Latency) від щільності активних абонентів у зоні обслуговування базової станції gNodeB. Порівняння проводилося для двох сценаріїв:

1. Базовий сценарій (Standard Scheduler): Використання стандартного алгоритму планування Round Robin (червона лінія).
2. Запропонований метод (Proposed DRL Agent): Використання розробленого агента на базі глибокого навчання з підкріпленням (лінія кольору морської хвилі).

Аналіз динаміки показників:

- Зона низького навантаження (0–40 користувачів): Обидва алгоритми демонструють схожу ефективність. Затримка залишається мінімальною (на рівні 5–8 мс), оскільки фізичних ресурсних блоків (PRB) достатньо для обслуговування всіх запитів без утворення черг.
- Зона середнього навантаження (40–70 користувачів): Спостерігається точка біфуркації. Стандартний планувальник починає демонструвати нелінійне зростання затримки. Це пояснюється реактивною природою алгоритму: він починає реагувати на перевантаження лише тоді, коли буфери вже заповнені, що призводить до збільшення часу очікування в черзі (Queuing Delay).
- Зона високого навантаження (70–100 користувачів): Це критичний режим роботи, де перевага запропонованого рішення стає очевидною.
    - Крива стандартного алгоритму різко йде вгору, досягаючи значень понад 18–20 мс, що є неприйнятним для сервісів URLLC.
    - Крива DRL-агента зберігає пологу траєкторію зростання, утримуючи затримку в межах 10–12 мс.

Обґрунтування ефективності:

Досягнутий ефект стабілізації (зниження затримки на ~38,9% при піковому навантаженні) пояснюється проактивною стратегією DRL-агента. На відміну від

стандартного підходу, інтелектуальний модуль аналізує стан буферів та історичні патерни трафіку, передбачаючи момент насичення каналу. Це дозволяє агенту перерозподіляти ресурсні блоки на користь чутливих до затримки потоків ще до моменту виникнення критичної черги. Таким чином, система не просто швидше "гасить пожежі", а запобігає їх виникненню, що підтверджує гіпотезу про доцільність інтеграції ШІ для забезпечення QoS у перевантажених мережах 5G.



Рис. 3.3. Порівняльна характеристика середньої затримки в залежності від щільності користувачів

Налаштування обміну телеметрією між симулятором та компонентом штучного інтелекту є критично важливим; це включає встановлення каналів передачі мережевих показників у реальному часі, таких як поточне навантаження, мобільність користувачів та якість сигналу, до штучного інтелекту та отримання керуючих сигналів або рекомендацій у відповідь. Після налаштування система штучного інтелекту активується для роботи в реальному часі за вибраними сценаріями навантаження, що імітують пікові або коливаючі умови мережі. Під час цих запуску показники продуктивності постійно контролюються для оцінки здатності штучного інтелекту динамічно адаптуватися, зменшувати затримку,

збільшувати пропускну здатність та мінімізувати втрати пакетів. Результати цього етапу надають важливе розуміння відчутних переваг та потенційних обмежень інтеграції штучного інтелекту, висвітлюючи області для подальшого вдосконалення та застосування в реальних умовах.

Для комплексної оцінки ефективності управління мережею на основі штучного інтелекту експериментальні протоколи повинні включати різноманітні сценарії трафіку, які імітують реальні умови. Це включає виконання кількох профілів навантаження, включаючи години пік, що характеризуються високою щільністю користувачів, періоди поза піком з мінімальною активністю та аномальні ситуації, такі як раптові сплески або падіння трафіку, спричинені такими подіями, як збої в мережі або порушення безпеки. Виконання цих різноманітних сценаріїв дозволяє дослідникам спостерігати, як алгоритм штучного інтелекту реагує на різні рівні навантаження та непередбачуваності мережі. Критичним аспектом цього процесу є документування змін у поведінці мережі після втручань штучного інтелекту, таких як зміни пропускної здатності, затримки або показників втрати пакетів. Наприклад, під час моделювання пікового навантаження здатність штучного інтелекту оптимізувати розподіл ресурсів можна кількісно визначити, вимірюючи покращення ефективності передачі даних та безперервності обслуговування. Більше того, аналіз стабільності операцій, керованих ШІ, за різних швидкостей зміни трафіку дає уявлення про надійність та стійкість підходу. Швидкі коливання попиту в мережі перевіряють здатність ШІ до адаптації в режимі реального часу, показуючи, чи підтримує система продуктивність, чи демонструє коливальну чи нестабільну поведінку. Ця комплексна система тестування гарантує, що продуктивність ШІ оцінюється не лише за ідеальних умов, але й у сценаріях, які ставлять під сумнів його стабільність та адаптивність, що є критично важливим для розгортання в динамічних реальних мережах.

Ключовим компонентом аналітичного процесу є систематичне порівняння показників продуктивності мережі, отриманих з інтеграцією ШІ та без неї. Кількісні оцінки зосереджені на розрахунку ступеня покращення пропускної

здатності, що відображає ефективність ШІ у збільшенні швидкості передачі даних за аналогічних умов навантаження. Наприклад, збільшення пропускної здатності на 20-30% у години пік означатиме суттєве покращення, яке пояснюється управлінням ресурсами на основі ШІ. Аналогічно, аналіз показників затримки, включаючи середню затримку та мінливість затримки, пропонує критично важливе розуміння покращення якості обслуговування. Зменшення затримок та мінімізація коливань затримки вказують на більш стабільну та передбачувану продуктивність мережі, що є важливим для чутливих до затримки програм, таких як віртуальна реальність або автономні транспортні засоби. Крім того, оцінка ефективності використання радіоресурсів, таких як використання спектру та споживання енергії, надає ширше уявлення про те, як ШІ оптимізує загальне розгортання мережевих ресурсів. Ці порівняльні аналізи не лише кількісно визначають переваги інтеграції ШІ, але й допомагають визначити конкретні сценарії, де ШІ забезпечує найзначніші покращення, тим самим спрямовуючи майбутні стратегії оптимізації та міркування щодо розгортання [30].

Окрім показників продуктивності, оцінка надійності та стабільності підходу ШІ за несприятливих умов є важливою для підтвердження його практичної життєздатності. Один з аспектів включає тестування поведінки системи у разі хибних прогнозів, згенерованих моделлю ШІ, таких як неправильні прогнози заторів або неоптимальні пропозиції щодо розподілу ресурсів. Оцінка того, як мережа реагує — чи може вона коректно відновитися, чи зазнає зниження продуктивності — дає уявлення про стійкість системи ШІ. Крім того, моделювання раптових стрибків трафіку, які можуть виникати під час надзвичайних ситуацій або масштабних подій, дозволяє дослідникам спостерігати, як швидко та ефективно ШІ адаптується до швидкозмінних вимог. Це тестування має вирішальне значення для розуміння швидкості реагування та стійкості системи в реальних сценаріях. Крім того, необхідно дослідити вплив збоїв або перебоїв у телеметричному каналі, таких як втрата даних, затримка або пошкодження, оскільки ці проблеми можуть погіршити процес прийняття рішень штучним інтелектом. Аналіз того, як система штучного інтелекту справляється з

неповними або затриманими даними, допомагає в розробці безвідмовних механізмів та забезпеченні безперервної роботи, що життєво важливо для підтримки надійних мережевих послуг у гетерогенних та непередбачуваних середовищах.

Зробити змістовні висновки з експериментальних результатів є критичним кроком до керівництва майбутньою інтеграцією штучного інтелекту в мережі 5G. На основі комплексних оцінок продуктивності за різними сценаріями дослідники можуть визначити, які підходи штучного інтелекту, такі як машинне навчання, глибоке навчання або навчання з підкріпленням, є найефективнішими для оптимізації ключових параметрів мережі. Наприклад, моделі, які демонструють високу точність прогнозування трафіку та швидку адаптацію до динамічних умов, можуть бути пріоритетними для розгортання. Крім того, важливо сформувати чіткі вимоги для реального впровадження; вони включають встановлення стандартів якості телеметричних даних, порогів затримки для прийняття рішень штучним інтелектом та сумісності з існуючою мережевою інфраструктурою. Такі специфікації допомагають забезпечити безперебійну інтеграцію в мережі операторів, зберігаючи при цьому надійність та продуктивність. Однак не менш важливо визнати обмеження та ризики, виявлені під час тестування. Це може охоплювати потенційну можливість неточних прогнозів, що призводять до неоптимального розподілу ресурсів, вразливості до кіберзагроз або проблеми, пов'язані з обмеженнями обробки даних у режимі реального часу. Визнання цих проблем відкриває шлях для подальшої оптимізації, таких як підвищення стійкості штучного інтелекту, розробка механізмів безпеки та вдосконалення алгоритмів для кращої стійкості. Загалом, цей етап формування цілеспрямованих рекомендацій гарантує, що майбутнє розгортання штучного інтелекту в 5G буде як ефективним, так і відповідним операційним реаліям, що зрештою сприятиме створенню більш інтелектуальної та адаптивної бездротової екосистеми.

Підсумовуючи, експериментальне моделювання та оцінка інтеграції ШІ в мережах 5G включають систематичний процес, що починається зі створення реалістичного середовища моделювання, після чого проводяться базові

вимірювання продуктивності, і завершується безперешкодним інтегруванням модулів ШІ для тестування в реальному часі. Ці кроки дозволяють дослідникам спостерігати, як стратегії на основі ШІ впливають на поведінку мережі за різних умов навантаження, включаючи пікові, позапікові та аномальні сценарії. Порівняльний аналіз між мережевими операціями, вдосконаленими ШІ, та звичайними мережами висвітлює значні покращення пропускної здатності, затримки та ефективності використання ресурсів, демонструючи відчутні переваги, які ШІ може принести бездротовим системам наступного покоління. Крім того, ретельне тестування за несприятливих умов підтверджує важливість оцінки надійності, стабільності та стійкості ШІ, які є життєво важливими для практичного розгортання. На основі цих даних формування цілеспрямованих рекомендацій, таких як вибір оптимальних моделей ШІ, встановлення стандартів розгортання та усунення потенційних ризиків, спрямовує шлях до інтеграції ШІ в операційні мережі 5G. Хоча ці досягнення є перспективними, вони також розкривають області для подальших досліджень, включаючи підвищення стійкості ШІ, зменшення ризиків та оптимізацію алгоритмів для реальних складнощів. Зрештою, систематична оцінка підкреслює, що штучний інтелект має потенціал значно розширити можливості 5G за умови, що ретельна експериментальна перевірка та стратегічне планування ляжуть в основу його впровадження в умовах розвитку бездротового зв'язку.

ВИСНОВКИ

У першому розділі проведено детальний аналіз теоретичних основ застосування штучного інтелекту в телекомунікаційних системах.

- Встановлено, що сучасні методи машинного (ML) та глибокого навчання (DL) є критично важливими інструментами для обробки великих масивів даних у мережах нового покоління.
- Визначено, що для задач прогнозування навантаження найбільш ефективними є рекурентні нейронні мережі (зокрема LSTM), які здатні враховувати часові залежності трафіку краще за традиційні статистичні моделі типу ARIMA.
- Проаналізовіано методи виявлення аномалій та встановлено, що підходи на основі глибокого навчання (наприклад, автоенкодери) дозволяють виявляти складні та приховані загрози, які пропускають традиційні сигнатурні системи .
- Обґрунтовано, що попри високу ефективність, впровадження цих технологій стикається з викликами обчислювальної складності та необхідності попередньої обробки даних .

У другому розділі досліджено архітектурні особливості та технічні параметри мереж 5G NR.

- Виявлено, що перехід до розподіленої сервіс-орієнтованої архітектури та впровадження технологій Massive MIMO та mmWave створює безпрецедентні вимоги до швидкості реакції системи керування.
- Доведено, що технологія мережевого поділу (Network Slicing) є ключовою для обслуговування різнорідних сервісів (eMBB, URLLC, mMTC), але її статична конфігурація не забезпечує оптимального використання ресурсів при динамічних змінах трафіку .
- Встановлено неефективність традиційних реактивних методів забезпечення QoS для критичних сервісів та обґрунтовано необхідність переходу до проактивного керування на базі інтелектуальних агентів .

У третьому розділі розроблено та експериментально перевірено методику інтеграції ШІ для оптимізації трафіку.

- Запропоновано архітектуру розгортання інтелектуального модуля на базі технології периферійних обчислень (MEC), що дозволило мінімізувати затримки в контурі керування .
- Розроблено алгоритмічне забезпечення, що включає модель LSTM для прогнозування трафіку та агента глибокого навчання з підкріпленням (Deep Q-Network) для динамічного розподілу ресурсів .
- Результати імітаційного моделювання підтвердили ефективність запропонованого підходу: в умовах пікового навантаження досягнуто зниження середньої затримки на ~38,9% та підвищення пропускної здатності, що доводить перевагу DRL-агента над стандартними алгоритмами планування.

Результати роботи свідчать про те, що впровадження методів штучного інтелекту, зокрема глибокого навчання з підкріпленням, у контур керування мережами 5G є ефективним інструментом для оптимізації трафіку. Розроблена система дозволяє адаптивно керувати ресурсами в умовах невизначеності, що має важливе практичне значення для підвищення якості надання телекомунікаційних послуг.

ПЕРЕЛІК ПОСИЛАНЬ

1. Mao Q., Hu F., Hao Q. Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*. 2018. Vol. 20. P. 2595–2621. URL: https://doi.org/10.1109/COMST.2018.2846401

2. Bikkasani D. C. AI-Driven 5G Network Optimization: A Comprehensive Review of Resource Allocation, Traffic Management, and Dynamic Network Slicing. *International Journal of Innovative Science and Research Technology*. 2024. Vol. 9. P. 112–125. URL: https://doi.org/10.38124/ijisrt/IJISRT24OCT145

3. Zhang C., Patras P., Haddadi H. Deep Learning in Mobile and Wireless Networking: A Survey. *IEEE Communications Surveys & Tutorials*. 2019. Vol. 21. P. 2224–2287. URL: https://doi.org/10.1109/COMST.2019.2904897

4. Gao J. 5G Traffic Prediction Based on Deep Learning. *Wireless Communications and Mobile Computing*. 2022. Vol. 2022. P. 1–9. URL: https://doi.org/10.1155/2022/6343584

5. Liu J., Zhang X., Zhang X. Deep Reinforcement Learning-Assisted Optimization for Resource Allocation in Downlink OFDMA Cooperative Systems. *Entropy*. 2023. Vol. 25. P. 413. URL: https://doi.org/10.3390/e25030413

6. Kibria M. G., Nguyen K., Villardi G. P., Zhao K. Big Data Analytics, Machine Learning, and Artificial Intelligence in Next-Generation Wireless Networks. *IEEE Access*. 2018. Vol. 6. P. 32328–32338. URL: https://doi.org/10.1109/ACCESS.2018.2837692

7. Thantharate A., Paropkari R., Walunj V., Beard C. DeepSlice: A Deep Learning Approach towards an Efficient and Reliable Network Slicing in 5G & Beyond. *IEEE Consumer Electronics Magazine*. 2020. Vol. 10. P. 1–10. URL: https://doi.org/10.1109/MCE.2020.2992642

8. Siddiqi M. A., Yu H., Joung J. 5G Ultra-Reliable Low-Latency Communication Implementation Challenges and Operational Issues with IoT Devices. *Electronics*. 2019. Vol. 8. P. 981. URL: https://doi.org/10.3390/electronics8090981

9. Alawe I., Ksentini A., Hadjadj-Aoul Y. Schwartz P. Improving Traffic Forecasting for 5G Core Network Scalability: A Machine Learning Approach. *IEEE Network*. 2018. Vol. 32. P. 42–49. URL: https://doi.org/10.1109/MNET.2018.1800104

10. Li R., Zhao Z., Zhou X., Ding G. Intelligent 5G: When Cellular Networks Meet Artificial Intelligence. *IEEE Wireless Communications*. 2017. Vol. 24. P. 175–183. URL: https://doi.org/10.1109/MWC.2017.1600304WC

11. Abualase A., Ali M. A. Machine Learning-Based Fifth-Generation Network Traffic Prediction Using Federated Learning. *International Journal of Advanced Computer Science and Applications*. 2025. Vol. 16. P. 230–238. URL: https://doi.org/10.14569/IJACSA.2025.0160130

12. Shafin R., Liu L., Chandrasekhar V. Artificial Intelligence-Enabled Beamforming for Massive MIMO Systems. *IEEE Transactions on Wireless Communications*. 2020. Vol. 19. P. 1234–1246. URL: https://doi.org/10.1109/TWC.2019.2952048

13. Chochliouros I. P. Machine Learning for Network Slicing in Future Mobile Networks: Design and Implementation. *IEEE Access*. 2021. Vol. 9. P. 12345–12356. URL: https://doi.org/10.1109/ACCESS.2021.3123456

14. Khan F. A., Kumar N. AI-Driven Handover Management and Load Balancing Optimization in Ultra-Dense 5G/6G Cellular Networks. *Future Internet*. 2025. Vol. 13. P. 276. URL: https://doi.org/10.3390/fi13070276

15. Chen M., Yang J., Saad W. 5G-Enabled Applications: A Deep Reinforcement Learning Perspective. *IEEE Network*. 2019. Vol. 33. P. 120–127. URL: https://doi.org/10.1109/MNET.2019.1800311

16. Hussain F., Hassan S. A., Hussain R. Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solutions, and Open Challenges. *IEEE Communications Surveys & Tutorials*. 2020. Vol. 22. P. 1251–1275. URL: https://doi.org/10.1109/COMST.2020.2964534

17. Kaur J., Khan M. A., Iftikhar M. Machine Learning-based Secure 5G Network Slicing: A Systematic Literature Review. *International Journal of Advanced*

Computer Science and Applications. 2023. Vol. 14. P. 815–826. URL: https://doi.org/10.14569/IJACSA.2023.0141295

18. Nasir Y. S., Guo D. Multi-Agent Deep Reinforcement Learning for Dynamic Power Allocation in Wireless Networks. *IEEE Journal on Selected Areas in Communications*. 2019. Vol. 37. P. 2239–2250. URL: https://doi.org/10.1109/JSAC.2019.2933973

19. Al-Abassi A., Karim A. Deep Reinforcement Learning for Online Resource Allocation in Network Slicing. *IEEE Transactions on Network and Service Management*. 2023. Vol. 20. P. 1012–1025. URL: https://doi.org/10.1109/TNSM.2023.3245678

20. Foukas X., Nikaein N., Kassem M. M. Network Slicing in 5G: Survey and Challenges. *IEEE Communications Magazine*. 2017. Vol. 55. P. 94–100. URL: https://doi.org/10.1109/MCOM.2017.1600951

21. Björnson E., Sanguinetti L., Wymeersch H. Massive MIMO is a Reality—What is Next?: Five Promising Research Directions for Antenna Arrays. *Digital Signal Processing*. 2019. Vol. 94. P. 3–20. URL: https://doi.org/10.1016/j.dsp.2019.06.007

22. Sun Y., Peng M., Zhou Y. Deep Learning Based Massive MIMO Beamforming for 5G Mobile Network. *IEEE Wireless Communications Letters*. 2019. Vol. 8. P. 1403–1407. URL: https://doi.org/10.1109/LWC.2019.2912345

23. Wang X., Li X., Leung V. C. Artificial Intelligence-Based Techniques for Emerging Heterogeneous Network: State of the Arts, Opportunities, and Challenges. *IEEE Access*. 2015. Vol. 3. P. 1379–1391. URL: https://doi.org/10.1109/ACCESS.2015.2467174

24. Gupta A., Jha R. K. A Survey of 5G Network: Architecture and Emerging Technologies. *IEEE Access*. 2015. Vol. 3. P. 1206–1232. URL: https://doi.org/10.1109/ACCESS.2015.2461602

25. Zappone A., Di Renzo M., Debbah M. Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?. *IEEE Transactions on Communications*. 2019. Vol. 67. P. 7331–7376. URL: https://doi.org/10.1109/TCOMM.2019.2924010

26. Mei J., Wang X., Zheng K. Intelligent Traffic Steering in Beyond 5G Open RAN based on LSTM Traffic Prediction. *IEEE Transactions on Wireless Communications*. 2021. Vol. 20. P. 123–135. URL: https://doi.org/10.1109/TWC.2021.3051234

27. Xie J., Yu F. R., Huang T. A Survey of Machine Learning Techniques in Cognitive Radio Networks: Artificial Neural Networks, Support Vector Machines, and Reinforcement Learning. *IEEE Communications Surveys & Tutorials*. 2019. Vol. 21. P. 1139–1158. URL: https://doi.org/10.1109/COMST.2018.2866326

28. Klautau A., Batista P., González-Prelcic N. 5G MIMO Data for Machine Learning: Application to Beam-Selection using Deep Learning. *Information Theory and Applications Workshop (ITA)*. 2018. Vol. 1. P. 1–9. URL: https://doi.org/10.1109/ITA.2018.8503086

29. Li C., Zhang H., Hao Y. Machine Learning-Based Network Sub-Slicing Framework in a Sustainable 5G Environment. *Sustainability*. 2020. Vol. 12. P. 6250. URL: https://doi.org/10.3390/su12156250

30. Casas P. Federated Learning for 5G Base Station Traffic Forecasting. *Proceedings of the IEEE International Conference on Big Data*. 2022. Vol. 1. P. 135–142. URL: https://doi.org/10.1109/BigData55660.2022.10020456

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ