

Компанієць В.О.Національний технічний університет
«Харківський політехнічний інститут»**Пустовойтов П.Є.**Національний технічний університет
«Харківський політехнічний інститут»

МОДЕЛЬ МАСШТАБУВАННЯ МЕРЕЖІ НА ОСНОВІ СТОХАСТИЧНИХ ПРОЦЕСІВ ЗА УМОВ САМОПОДІБНОГО ТРАФІКУ

У статті розглядається проблема масштабування мереж в умовах високої варіативності трафіку, зокрема з урахуванням самоподібної природи його розподілу. Актуальність дослідження зумовлена зростанням навантаження на мережеву інфраструктуру через збільшення кількості користувачів та обсягів переданої інформації. Традиційні детерміновані підходи до керування ресурсами втрачають ефективність у ситуаціях, коли трафік демонструє довготривалі залежності та високий ступінь автокореляції. Метою дослідження є розробка стохастичної моделі масштабування мережі, яка дозволяє адаптивно розподіляти ресурси відповідно до поточного навантаження. Запропоновано методологічний підхід, що охоплює аналіз існуючих способів масштабування, математичне моделювання динамічного балансування ресурсів із урахуванням самоподібного трафіку та оцінку ефективності моделі за допомогою імітаційного моделювання. У результаті представлено рівняння динамічного балансування ресурсів, яке має широке практичне застосування в сучасних мережевих технологіях. Зокрема, у програмно-конфігурованих мережах (SDN) запропонована модель дозволяє контролерам у режимі реального часу перерозподіляти ресурси, підвищуючи ефективність інфраструктури та якість обслуговування (QoS). У сфері хмарних обчислень вона використовується для автоматичного масштабування віртуальних машин і контейнерів, що сприяє адаптивному управлінню ресурсами в розподілених системах. У контексті 5G та MEC (Multi-access Edge Computing) модель забезпечує балансування навантаження між базовими станціями й обчислювальними вузлами, підтримуючи сервіси з низькою затримкою. Загалом, представлена модель забезпечує гнучке й адаптивне управління мережею з урахуванням стохастичних характеристик трафіку, сприяючи досягненню оптимального використання інфраструктури в умовах динамічного середовища.

Ключові слова: самоподібний трафік, стохастичні процеси, модель масштабування мережі, динамічне балансування, SDN, хмарні обчислення, MEC, адаптивне управління ресурсами, підвищення ефективності.

Постановка проблеми. Масштабування мереж є критичним завданням у сучасних інформаційно-комунікаційних системах, оскільки збільшення кількості користувачів та обсягів трафіку призводить до зростання навантаження на мережеву інфраструктуру. Традиційні підходи до масштабування, засновані на детермінованих методах керування ресурсами, часто не забезпечують необхідної гнучкості та ефективності в умовах високої варіативності мережевого трафіку. Використання стохастичних процесів для адаптивного управління ресурсами може стати ключем до створення ефективних механізмів масштабування, що дозволяють враховувати динамічну природу мережевого середовища.

Однією з головних проблем, що ускладнюють процес масштабування, є самоподібний характер трафіку, який проявляється у високій корельованості та значних коливаннях інтенсивності передачі даних. Самоподібний трафік, на відміну від традиційних моделей Пуассона, характеризується довготривалою залежністю, що ускладнює прогнозування та управління ресурсами. Це може призводити до неефективного розподілу обчислювальних потужностей, перевантаження каналів зв'язку та втрат пакетів, що негативно впливає на якість обслуговування користувачів.

Дослідження спрямоване на розробку моделі масштабування мережі, яка враховує стохастичну природу трафіку та його самоподібні

характеристики. Основною метою є створення адаптивного підходу до розподілу ресурсів, який дозволить оптимізувати продуктивність мережі за умов високої варіативності трафіку. Для цього необхідно розробити математичну модель, що описує процес масштабування з урахуванням стохастичних параметрів трафіку, та розробити метод оцінки ефективності запропонованого підходу.

У межах цього дослідження буде проведено аналіз існуючих методів масштабування мереж і підходів до моделювання самоподібного трафіку. На основі отриманих результатів буде запропоновано стохастичну модель, яка дозволяє адаптивно змінювати параметри мережі відповідно до поточного навантаження. Для оцінки ефективності методу планується використання методів імітаційного моделювання, що дозволить визначити вплив запропонованої моделі на продуктивність мережі.

Методологічний підхід включає кілька основних етапів: аналіз та класифікація існуючих методів масштабування, розробка математичної моделі, її чисельна перевірка та порівняння з традиційними підходами. Також планується розглянути варіанти реалізації запропонованої моделі у сучасних мережевих технологіях, таких як програмно-конфігуровані мережі (SDN) або хмарні обчислення. Це дозволить визначити практичну цінність розробленого підходу та його застосовність у реальних мережевих системах.

Таким чином, дослідження спрямоване на розв'язання актуальної проблеми ефективного масштабування мереж за умов самоподібного трафіку. Результати роботи можуть бути використані для підвищення якості обслуговування в телекомунікаційних системах, оптимізації використання мережевих ресурсів та підвищення стійкості мереж до різких змін трафіку.

Аналіз останніх досліджень і публікацій. У сучасних умовах інтенсивного розвитку інформаційно-комунікаційних технологій особливої актуальності набуває завдання ефективного масштабування мереж за умов самоподібного трафіку. Проблематика полягає у складності прогнозування та управління ресурсами через довготривалі залежності та високі варіації навантаження, характерні для фрактального трафіку. За останні роки у світовій науковій літературі з'явилась низка досліджень, що пропонують інноваційні підходи до розв'язання цієї задачі на основі стохастичних та гібридних моделей.

Так, у роботі [1] досліджено можливості wavelet-аналізу для точного моделювання фрактальної структури трафіку. Автори продемонстрували високу точність реконструкції самоподібних потоків, проте зазначили, що метод вимагає значних обчислювальних ресурсів та складного налаштування параметрів.

Інший підхід, запропонований у [2], базується на використанні процесів прийняття рішень на основі марковських моделей (Markov Decision Process – MDP) для мереж SDN. Такий підхід забезпечує динамічне управління розподілом ресурсів у режимі реального часу, однак потребує наявності великої кількості достовірних історичних даних для ефективного навчання моделі.

Моделі динамічного масштабування у хмарних середовищах, адаптовані до самоподібного навантаження, були розроблені у [3]. Автори запропонували метод, що забезпечує плавне масштабування обчислювальних ресурсів, хоча його реалізація у великих розподілених системах вимагає додаткової технічної адаптації.

Застосування методів штучного інтелекту, зокрема нейромережевих моделей, для прогнозування навантаження у самоподібному трафіку продемонстровано в [4]. Незважаючи на високу точність прогнозів, моделі є чутливими до якості вхідних навчальних даних.

Цікаву комбінацію марковських ланцюгів і фрактальних функцій як основу для моделювання трафіку подано в [5]. Гібридний підхід добре масштабується та дозволяє точно відображати поведінку мережевого трафіку, проте складність калібрування моделей є суттєвим обмеженням.

У дослідженні [6] стохастичні диференціальні рівняння використано для управління ресурсами в мережах типу MEC (Multi-access Edge Computing). Аналітична точність моделі є її перевагою, проте застосування у великих масштабах потребує значних обчислювальних потужностей.

Методи навчання з підкріпленням (reinforcement learning), представлені в [7], дозволяють системам адаптивно реагувати на зміни трафіку з самоподібною структурою. Вони демонструють високий рівень автономності, але потребують серйозної обчислювальної підтримки.

У роботі [8] запропоновано метод маршрутизації, адаптований до фрактального характеру трафіку. Цей підхід дозволяє суттєво знизити втрати пакетів при пікових навантаженнях, хоча його ефективність зменшується при низькому обсязі трафіку.

Ще один перспективний напрям, представлений у [9], ґрунтується на використанні стохастичної апроксимації для масштабування хмарних сервісів. Метод є достатньо універсальним, проте результати мають наближений характер, що може бути критичним у деяких сценаріях.

Нарешті, у дослідженні [10] зроблено спробу інтеграції Hurst-параметра для передбачення навантаження на телекомунікаційну інфраструктуру. Це дозволяє досягти високої точності прогнозування, хоча модель є досить чутливою до помилок при оцінюванні параметра самоподібності.

Таким чином, проаналізовані підходи демонструють багатовекторний розвиток методів масштабування мереж у контексті самоподібного трафіку. Незважаючи на відмінності в математичному апараті та реалізації, більшість досліджень сходяться на необхідності адаптивного управління ресурсами, що базується на стохастичних моделях і фрактальних властивостях трафіку.

Паралельно із зарубіжними дослідженнями, вітчизняні науковці також активно розробляють моделі масштабування телекомунікаційних мереж в умовах самоподібного трафіку. Особливу увагу приділяють застосуванню стохастичних процесів та фрактального аналізу для побудови адаптивних підходів до управління ресурсами.

У роботі [11] запропоновано математичну модель генерації самоподібного трафіку з використанням розподілу Парето та методу максимальної правдоподібності для оцінки параметрів. Перевагою підходу є можливість гнучкої генерації пакетних потоків із заданим ступенем самоподібності, проте метод обмежений у застосуванні до реальних багатокомпонентних мереж.

Дослідження [12] присвячене структурно-параметричній ідентифікації самоподібного трафіку із застосуванням глобальної та локальної оптимізації. Така композиція дозволяє досягти високої точності при прогнозуванні трафіку та виявленні аномалій. Основним недоліком є підвищена чутливість до вибору початкових параметрів моделі.

У роботі [13] розглянуто сучасні підходи до застосування технологій штучного інтелекту в управлінні телекомунікаційними мережами. Акцент зроблено на автоматизацію обробки ресурсів у 5G-мережах, що забезпечує гнучкість та ізоляцію між логічними і фізичними ресурсами. Недоліком є складність інтеграції таких систем у вже існуючу інфраструктуру.

Запропонований у [14] підхід до оптимізації роботи когнітивних радіомереж базується на використанні нейронних мереж для прийняття рішень. Такий інтелектуальний механізм підвищує ефективність функціонування мереж, однак потребує великого обсягу навчальних даних та ускладнює аналіз поведінки моделі.

У дослідженні [15] автори пропонують застосування стохастичних алгоритмів оптимізації (градієнтного спуску та стохастичного пошуку) для адаптивного управління ресурсами. Модель показала добрі результати в мінімізації затримок, але обчислювальна складність алгоритмів обмежує їх застосування у масштабних системах.

Інноваційний підхід до ідентифікації мережевого трафіку представлено в [16], де використано глобальні методи оптимізації та інтелектуальні базисні функції. Метод дає змогу підвищити точність передбачення трафіку, але його ефективність значною мірою залежить від обсягу вхідних даних.

У роботі [17] розроблено методику моделювання змінного трафіку, яка враховує динаміку навантаження та дозволяє адаптивно налаштовувати параметри мережі. Високий рівень відповідності моделі до реальних умов є суттєвою перевагою, хоча її реалізація потребує точного калібрування параметрів.

Нарешті, у [18] розглянуто задачу маршрутизації фракталоподібного трафіку в комп'ютерних мережах. Запропоновані алгоритми дозволяють зменшити втрати пакетів та покращити якість обслуговування, проте при значних об'ємах трафіку вони потребують попередньої агрегації даних, що знижує швидкість реагування.

Таким чином, українські науковці пропонують широкий спектр підходів до розв'язання задачі масштабування мереж з урахуванням самоподібності трафіку – від статистичного моделювання до застосування методів штучного інтелекту. Хоча більшість з них стикається з труднощами реалізації у великих системах, вони створюють теоретичну основу для подальшого вдосконалення адаптивних механізмів управління ресурсами у динамічних телекомунікаційних середовищах.

Постановка завдання. Метою роботи є розробка стохастичної моделі масштабування мережі, яка дозволяє адаптивно розподіляти ресурси відповідно до поточного навантаження.

Виклад основного матеріалу. Масштабування мережі в умовах самоподібного трафіку є складним завданням, оскільки традиційні

моделі керування ресурсами не враховують довготривалу залежність і високу варіативність мережевого навантаження. Самоподібний трафік проявляє властивості фрактальності та важко прогнозованих сплесків навантаження, що ускладнює розподіл ресурсів у мережах. Для ефективного масштабування необхідно використовувати математичні методи, здатні апроксимувати складні процеси та передбачати їх динаміку. У цьому контексті марківська апроксимація є перспективним підходом, що дозволяє спростити модель самоподібного трафіку, наближаючи її до керованих стохастичних систем.

Марківські випадкові процеси можуть бути використані для моделювання навантаження в мережі, оскільки вони добре підходять для опису систем із залежностями короткої пам'яті. Хоча самоподібний трафік має довготривалі кореляції, його можна апроксимувати через багаторівневі марківські моделі, що дозволяє описати зміну станів навантаження у спрощеній формі. Наприклад, можна застосувати Марківський процес із прихованими станами

$$\begin{pmatrix} H \\ M \\ M \end{pmatrix} \text{ або Марківські моделі черг } \begin{pmatrix} M & M \\ G, M \\ 1 & \infty \end{pmatrix}$$

для оцінки поведінки трафіку та адаптивного перерозподілу мережевих ресурсів.

Математична модель масштабування на основі марківських випадкових процесів включає подання трафіку у вигляді процесу з фіксованими станами та перехідними ймовірностями між ними. Для цього можна використовувати ланцюги Маркова із дискретними станами, де кожен стан відповідає певному рівню навантаження, а ймовірність переходу між станами залежить від емпіричних даних або аналітичних розрахунків. Враховуючи самоподібність трафіку, марківська апроксимація (рис. 1) може бути реалізована через багатоступеневі ланцюги, що дозволяє відобразити зміну трафіку на різних часових масштабах.

Оцінка основних параметрів моделі включає визначення інтенсивності трафіку, що відображає частоту надходження запитів, коефіцієнта самоподоби для оцінки залежностей між потоками, а також латентності, що характеризує середній час очікування запитів у мережі. Використовуючи стохастичну апроксимацію, можна отримати рівняння динамічного балансування ресурсів, що враховує стохастичні характеристики мережі та дозволяє адаптивно масштабувати інфраструктуру відповідно до реального навантаження.

Для адаптивного масштабування мережевих ресурсів за умов стохастичного навантаження можна використати рівняння динамічного балансування ресурсів, що базується на марківських випадкових процесах та стохастичних диференціальних рівняннях.

Ось узагальнене рівняння, яке враховує зміну ресурсів у мережі залежно від поточного навантаження та його коливань:

$$\frac{dR(t)}{dt} = \alpha\lambda(t) - \beta\mu(t)R(t) + \sigma W(t), \quad (1)$$

де $R(t)$ – кількість виділених ресурсів у момент часу t ; $\lambda(t)$ – інтенсивність вхідного трафіку в момент часу t ; $\mu(t)$ – середня швидкість обслуговування запитів (залежить від продуктивності мережі); α – коефіцієнт масштабування ресурсів у відповідь на зміну навантаження; β – коефіцієнт звільнення ресурсів у процесі їх оптимізації; $\sigma W(t)$ – стохастичний шум, що моделює випадкові флуктуації навантаження (білий шум Вінера).

Для опису самоподібного трафіку потрібно перейти до фрактальних та довгопам'ятних процесів. Замість рівномірного шуму $W(t)$ у рівнянні динамічного балансування можна використати фрактальний процес:

$$\frac{dR(t)}{dt} = \alpha\lambda(t) - \beta\mu(t)R(t) + \sigma B_H(t), \quad (2)$$

де $B_H(t)$ – фрактальний броунівський рух із параметром Херста H , який моделює самоподібність.

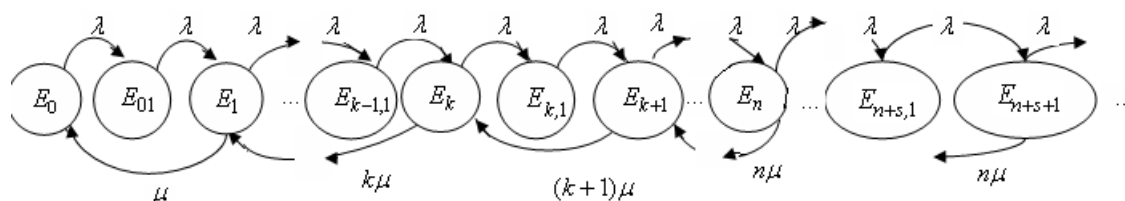


Рис. 1. Граф станів та переходів системи без втрат з потоком ерланга, що надходить

Самоподібність означає, що мережевий трафік не є рівномірним, а має характерні сплески, які можуть повторюватися на різних часових масштабах. Це ускладнює використання традиційних методів масштабування, які базуються на середньому навантаженні або короткострокових прогнозах. Для точного моделювання варто враховувати довготривалу кореляцію (Hurst-параметр, H), що визначає рівень самоподібності трафіку:

$$0.5 < H < 1. \quad (3)$$

Якщо H наближається до 1, це означає, що сплески навантаження можуть бути дуже затяжними, і для їх компенсації потрібно підтримувати більший запас ресурсів.

Адаптивне масштабування мережі виконується таким чином, що при зростанні інтенсивності трафіку $\lambda(t)$ система динамічно збільшує виділені ресурси $R(t)$. Це дозволяє оперативно реагувати на зміну навантаження і підтримувати належну якість обслуговування. У випадку зменшення навантаження ресурси поступово звільняються зі швидкістю, пропорційною середній швидкості обслуговування запитів $\mu(t)$, що дозволяє уникнути зайвого перевитрачання обчислювальних потужностей.

Стохастична варіативність трафіку додає додаткову складність у процес масштабування. Член $\sigma W(t)$ чи $\sigma B_H(t)$ у рівнянні моделює випадкові флуктуації трафіку, які можуть виникати через непередбачувані зміни в поведінці користувачів або раптові сплески навантаження. Це особливо важливо для мереж з високими динамічними змінами, таких як мобільні або хмарні платформи, де попит може коливатися протягом коротких проміжків часу.

Балансування між продуктивністю та ефективністю досягається за допомогою параметрів α і β , які визначають реакцію системи на зміну навантаження. Коефіцієнт α відповідає за швидкість збільшення ресурсів у відповідь на зростання інтенсивності трафіку, тоді як коефіцієнт β контролює ефективність звільнення ресурсів після зниження навантаження. Оптимальний вибір цих параметрів дозволяє уникнути перевитрачання обчислювальних потужностей та забезпечує стабільність роботи мережі.

Моделювання можна розширити для розподілених систем, враховуючи особливості багатовузлової інфраструктури. У такому випадку рівняння набуває узагальненого вигляду для кожного вузла мережі:

$$\frac{dR_i(t)}{dt} = \alpha_i \lambda_i(t) - \beta_i \mu_i(t) R_i(t) + \sigma_i W_i(t), \quad i = 1, 2, \dots, N, \quad (4)$$

де i – індекс вузла в мережі, H – загальна кількість вузлів. Таке рівняння дозволяє враховувати неоднорідність навантаження та різну продуктивність окремих сегментів мережі.

Моделювання та практичні результати.

Уявімо, що в дата-центрі працює хмарна платформа для потокового відео, яка динамічно виділяє ресурси залежно від кількості користувачів. Протягом дня трафік сильно змінюється: вранці навантаження низьке, у вечірні години – максимальне, а протягом дня можливі різкі сплески активності. Система повинна автоматично збільшувати або зменшувати обчислювальні ресурси так, щоб уникнути перевантаження та мінімізувати витрати.

Щоб змодельовати цей процес, використовуємо рівняння динамічного балансування ресурсів (1). У свою чергу середня інтенсивність запитів моделюється функцією:

$$\lambda(t) = \lambda_0 + A \sin(\omega t) + \eta(t), \quad (5)$$

де λ_0 – базова кількість запитів, A – амплітуда добових коливань, ω – частота змін, а $\eta(t)$ – стохастичний процес, що імітує випадкові стрибки трафіку.

Для знаходження оптимальної кількості серверів обчислюємо стаціонарний стан рівняння. Визначимо середній рівень ресурсів, коли система стабільна:

$$\mathbb{E}[R^*] = \frac{\alpha \lambda_0}{\beta \mu}. \quad (6)$$

Підставимо конкретні значення: $\alpha = 0.8$, $\beta = 0.5$, $\mu = 100$ запитів/сек, а середня кількість вхідних запитів $\lambda_0 = 5000$ на секунду.

$$R^* = \frac{0.8 \times 5000}{0.5 \times 100} = 80. \quad (7)$$

Тобто в середньому потрібно підтримувати 80 активних серверів. Але через стохастичні зміни трафіку це значення буде коливатися. Для визначення допустимих меж обчислюємо дисперсію ресурсу:

$$\text{Var}(R) = \frac{\sigma^2}{2\beta\mu}. \quad (8)$$

При значенні $\sigma = 300$, отримаємо:

$$Var(R) = \frac{300^2}{2 \times 0.5 \times 100} = 900. \quad (9)$$

Тоді стандартне відхилення $\sigma_R = 30$, а отже, ресурс буде коливатися в межах 80 ± 30 . Це означає, що у періоди високого навантаження потрібно автоматично збільшувати кількість серверів до 110, а у періоди низького – зменшувати до 50.

Таким чином, застосування рівняння динамічного балансування дозволяє хмарній платформі в реальному часі адаптуватися до змінного навантаження, уникаючи простоїв та перевитрат ресурсів.

Наведемо таблицю, яка представляє порівняння двох підходів до моделювання мережевого трафіку: пуассонівського та фрактального. У стовпцях наведено значення кількості запитів для кожного підходу та відповідна кількість необхідних серверів для балансування ресурсів протягом доби. Пуассонівський процес характеризується випадковими різкими сплесками навантаження, що призводить до частих і раптових змін у необхідності масштабування серверів. Натомість фрактальний

процес демонструє довготривалу кореляцію, що забезпечує більш плавні зміни трафіку та стабільніше використання ресурсів.

Дані підтверджують, що традиційний підхід масштабування на основі пуассонівського процесу може спричиняти надмірні витрати через необхідність частого масштабування ресурсів. У випадку фрактального трафіку, зміни навантаження більш прогнозовані, що дозволяє адаптивно налаштовувати інфраструктуру без зайвого перевитрачання потужностей. Використання моделі, яка враховує самоподібність трафіку, сприяє ефективнішому балансуванню ресурсів, оскільки забезпечує стабільніше управління потужностями серверів навіть у пікові години.

Ці результати вказують на важливість урахування самоподібності трафіку при проєктуванні масштабованих мережевих систем. Впровадження адаптивних алгоритмів, що базуються на аналізі фрактальних процесів, дозволяє знизити витрати на інфраструктуру та підвищити ефективність розподілу ресурсів у динамічних середовищах.

Таблиця 1

Порівняння масштабування ресурсів при пуассонівському та фрактальному трафіку протягом доби

Година	Кількість запитів (Пуассонівський)	Кількість запитів (Фрактальний)	Сервери (Пуассонівський)	Сервери (Фрактальний)
0	5469	5024	87,5	80,4
1	6035	6527	96,6	104,4
2	6519	6932	104,3	110,9
3	6915	7467	110,6	119,5
4	7220	7771	115,5	124,3
5	7466	7503	119,5	120,1
6	7494	8028	119,9	128,5
7	7412	8261	118,6	132,2
8	7218	8378	115,5	134
9	6955	7696	111,3	123,1
10	6509	7843	104,1	125,5
11	6039	6800	96,6	108,8
12	5533	6517	88,5	104,3
13	4970	6875	79,5	110
14	4481	5997	71,7	96
15	4108	5356	65,7	85,7
16	3762	5078	60,2	81,3
17	3537	4677	56,6	74,8
18	3547	3989	56,8	63,8
19	3569	4084	57,1	65,4
20	3758	3859	60,1	61,8
21	4145	4366	66,3	69,9
22	4499	4413	72	70,6
23	4967	5515	79,5	88,2

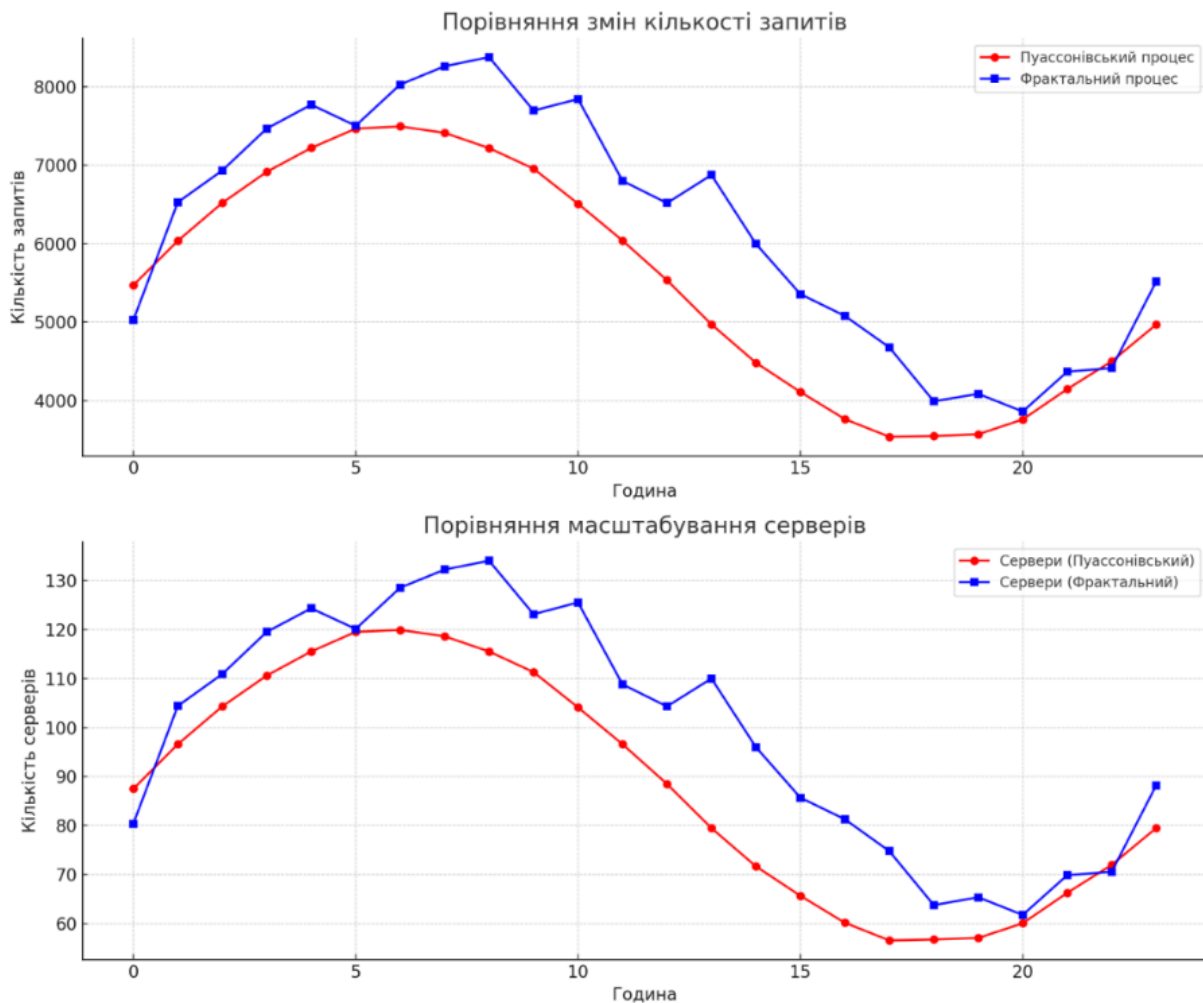


Рис. 1. Масштабування ресурсів мережі при динамічній зміні навантаження

Останній графік ілюструє динаміку зміни кількості запитів та необхідної кількості серверів у двох підходах до моделювання трафіку: пуассонівському та фрактальному. Графік кількості запитів показує, що пуассонівський процес має різкіші коливання, що обумовлено випадковою природою прибуття трафіку. Фрактальний трафік, навпаки, демонструє плавніші зміни з періодичними довготривалими сплесками, що відповідає реальним спостереженням у мережевих системах.

На другому графіку, який відображає необхідну кількість серверів, помітна значна відмінність між підходами. У традиційному масштабуванні, заснованому на пуассонівському процесі, часті та раптові зміни навантаження призводять до різких стрибків у кількості необхідних серверів. Це може створювати проблеми в управлінні ресурсами, оскільки система змушена постійно додавати або зменшувати потужності, що може бути неекономічним. У випадку самоподібного трафіку процес масштабування

більш стабільний: кількість серверів змінюється плавно, що дозволяє уникнути зайвих витрат та забезпечити ефективніше використання ресурсів.

Ці результати свідчать про те, що використання моделей, які враховують самоподібність трафіку, дозволяє створювати більш ефективні системи управління мережевими ресурсами. Вони дають змогу не тільки передбачати зміну навантаження, а й адаптувати інфраструктуру до довготривалих трендів у трафіку, що покращує якість обслуговування та зменшує перевитрати.

Висновки. Рівняння динамічного балансування ресурсів має широке практичне застосування в сучасних мережевих технологіях. Однією з ключових сфер його використання є програмно-конфігуровані мережі (SDN). У таких мережах контролери SDN можуть застосовувати це рівняння для динамічного перерозподілу ресурсів, реагуючи на зміни навантаження в режимі реального часу. Це дозволяє покращити ефективність використання мережевої інфраструктури та підвищити рівень якості обслуговування (QoS).

У сфері хмарних обчислень модель використовується для автоматичного масштабування віртуальних машин та контейнерів відповідно до поточного трафіку. Це дає можливість гнучко керувати ресурсами у великих розподілених системах, забезпечуючи їхню адаптацію до змінних умов роботи. Використання стохастичних процесів у прогнозуванні навантаження дозволяє мінімізувати затримки та уникнути надмірного споживання ресурсів.

Ще однією важливою областю застосування є 5G та MEC (Multi-access Edge Computing). У цих мережах рівняння допомагає балансувати

навантаження між базовими станціями та обчислювальними вузлами, що особливо актуально для підтримки сервісів з низькою затримкою, таких як автономний транспорт, розширена реальність та критично важливі IoT-додатки.

Загалом, ця модель забезпечує адаптивне керування мережевими ресурсами, створюючи баланс між продуктивністю та ефективністю, а також враховуючи стохастичні характеристики трафіку. Це дозволяє досягти оптимального використання інфраструктури в умовах динамічних змін навантаження.

Список літератури:

1. Zhang Y., Li H., Xu M. Fractal Traffic Modeling in Broadband Networks Using Wavelet Techniques. IEEE Access, 2021. Vol. 9. P. 33512–33524.
2. Chen W., Li J. Dynamic Resource Allocation in SDN via Markov Decision Processes. Computer Networks, 2022. Vol. 204. 108720.
3. Wang K., Zhou X., Liu S. Adaptive Scaling in Cloud Systems with Self-Similar Traffic. Future Generation Computer Systems, 2020. Vol. 108. P. 317–328.
4. Singh R., Kumar D. Neural Prediction of Self-Similar Network Loads for Edge Scalability. Journal of Network and Computer Applications, 2023. Vol. 214. 103497.
5. Almeida J., Costa B., Silva L. Hybrid Stochastic-Fractal Models for Network Traffic Simulation. Simulation Modelling Practice and Theory, 2022. Vol. 117. 102503.
6. Ito K., Yamamoto T. Stochastic Differential Models for Resource Scaling in Edge Computing. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2021. Vol. 6 (2). Article 11.
7. Johnson M., Lopez C., Patel A. Reinforcement Learning for Self-Similar Traffic Scaling in Data Centers. IEEE Transactions on Network and Service Management, 2023. Vol. 20 (1). P. 515–528.
8. Lin D., Kuo Y. Routing Optimization under Fractal Traffic Conditions. Computer Communications, 2021. Vol. 169. P. 25–34.
9. Farooq A., Nawaz R., Zahid M. Stochastic Approximation-Based Scaling Algorithms for Cloud Resource Management. Journal of Cloud Computing, 2022. Vol. 11 (1). Article 32.
10. Tan C., Zhou J. Hurst Parameter Estimation for Predictive Load Management in Telecom Systems. Telecommunication Systems, 2020. Vol. 74. P. 43–57.
11. Пустовойтов П.Є., Компанієць В.О. Метод математичного моделювання самоподібного трафіку у інфокомунікаційних мережах. *Системи управління, навігації та зв'язку. Збірник наукових праць*. Полтава: ПНТУ, 2024. Т. 4 (78). С. 187–189. <https://doi.org/10.26906/SUNZ.2024.4.187>.
12. Корнієнко В., Герасіна О. Прогнозування самоподібного трафіку інформаційно-комунікаційних мереж для виявлення аномалій. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 2022. 15 с.
13. Васильківський М., Болдирева О., Варгатюк Г., Будах М. Керування телекомунікаційними мережами з використанням технологій AI/ML. *Телекомунікації та Інтернет-технології*, 2021. 18 с.
14. Севрюков О. Методи керування інтелектуальними телекомунікаційними мережами для підвищення ефективності їх функціонування. *Інформаційні системи та технології*, 2023. 20 с.
15. Волк М.О., Поповкін М.М. Стохастичні методи оптимізації розподілу обчислювальних та мережових ресурсів. *Обчислювальні методи та системи*, 2024. 16 с.
16. Якимчук Н. Глобальні методи оптимізації для структурно-параметричної ідентифікації трафіку. *Комп'ютерні науки та кібербезпека*, 2023. 14 с.
17. Решетнікова П. Методика моделювання трафіку в комп'ютерних мережах для адаптивного налаштування параметрів. *Системний аналіз та моделювання*, 2023. 12 с.
18. Севрюков О.В. Моделі аналізу та маршрутизації фракталоподібного трафіку у комп'ютерних мережах. *Телекомунікаційні системи та мережі*, 2024. 17 с.

Kompaniets V.O., Pustovoitov P.E. NETWORK SCALPING MODEL BASED ON STOCHASTIC PROCESSES UNDER SELF-SIMPLE TRAFFIC CONDITIONS

The article considers the problem of scaling networks in conditions of high variability of traffic, in particular, taking into account the self-similar nature of its distribution. The relevance of the study is due to the increase in the load on the network infrastructure due to the increase in the number of users and the volume of transmitted information. Traditional deterministic approaches to resource management lose their effectiveness in situations where traffic demonstrates long-term dependencies and a high degree of autocorrelation – characteristic features of self-similar traffic. The purpose of the study is to develop a stochastic model of network scaling that allows for adaptive allocation of resources in accordance with the current load. A methodological approach is proposed that includes the analysis of existing scaling methods, mathematical modeling of dynamic resource balancing taking into account self-similar traffic, and assessment of the effectiveness of the model using simulation modeling. As a result, a dynamic resource balancing equation is presented, which has wide practical application in modern network technologies. In particular, in software-defined networks (SDN), the proposed model allows controllers to real-time reallocate resources, improving infrastructure efficiency and quality of service (QoS). In the field of cloud computing, it is used for automatic scaling of virtual machines and containers, which contributes to adaptive resource management in distributed systems. In the context of 5G and MEC (Multi-access Edge Computing), the model provides load balancing between base stations and computing nodes, supporting low-latency services. Overall, the presented model provides flexible and adaptive network management taking into account stochastic traffic characteristics, contributing to achieving optimal infrastructure utilization in a dynamic environment.

Key words: self-similar traffic, stochastic processes, network scaling model, dynamic balancing, SDN, cloud computing, MEC, adaptive resource management, efficiency improvement.