

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ**  
**ТЕХНОЛОГІЙ**

Кваліфікаційна наукова  
праця на правах рукопису

**ГЕРАСИМЧУК ВЛАДИСЛАВ СЕРГІЙОВИЧ**

УДК 621.396:004.7(043)

**ДИСЕРТАЦІЯ**

**МЕТОДИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ УПРАВЛІННЯ**  
**В ТЕЛЕКОМУНІКАЦІЙНИХ МЕРЕЖАХ ПРИ НАДАННІ**  
**ПОСЛУГ КІНЦЕВОМУ КОРИСТУВАЧЕВІ**

Спеціальність 172 – «Телекомунікації та радіотехніка»

Галузь знань 17 – «Електроніка та телекомунікації»

Подається на здобуття наукового ступеня доктора філософії.

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

\_\_\_\_\_ Владислав ГЕРАСИМЧУК

Наукові керівники: ЗАЇКА Віктор Федорович,  
д-р техн. наук, професор;  
ПЕПА Юрій Володимирович,  
канд. техн. наук, доцент.

Київ – 2026

## АНОТАЦІЯ

*Герасимчук В.С.* Методи підвищення ефективності управління в телекомунікаційних мережах при наданні послуг кінцевому користувачеві. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 172 – «Телекомунікації та радіотехніка» (галузь знань 17 – «Електроніка та телекомунікації»). – Державний університет інформаційно-комунікаційних технологій, м. Київ, 2026.

Дисертаційна робота присвячена створенню методів, моделей і алгоритмів підвищення ефективності управління телекомунікаційними мережами при наданні послуг кінцевим користувачам шляхом адаптивного розподілу мережевих ресурсів, оптимізації маршрутизації трафіку з метою зменшення затримок, втрат пакетів і підвищення рівня використання ресурсів мережі та забезпечення необхідних показників якості обслуговування.

**Об'єктом дослідження** дисертаційної роботи є методи та моделі адаптивного управління мережевими ресурсами, маршрутизації й оптимізації трафіку в телекомунікаційних мережах, спрямовані на підвищення ефективності функціонування мережі та забезпечення необхідних показників якості обслуговування кінцевих користувачів.

**Предметом дослідження** дисертаційної роботи є процес адаптивного управління мережевими ресурсами, маршрутизацією та розподілом трафіку в телекомунікаційних мережах з метою забезпечення показників якості обслуговування QoS/QoE при наданні послуг кінцевим користувачам.

У дисертаційній роботі проведено теоретичні та експериментальні дослідження процесів управління телекомунікаційними мережами, розроблено математичні моделі та методи адаптивного розподілу мережевих ресурсів і оптимізації маршрутизації трафіку, а також виконано оцінювання ефективності запропонованих рішень.

На основі проведеного аналізу сформульовано вимоги до методів і моделей управління телекомунікаційними мережами, які повинні забезпечувати ефективний розподіл мережесих ресурсів, адаптацію до змін трафікового навантаження та підтримку заданих показників якості обслуговування кінцевих користувачів.

Особлива увага приділена дослідженню адаптивних методів розподілу пропускної здатності, алгоритмів оптимальної маршрутизації трафіку та механізмів забезпечення показників якості обслуговування QoS/QoE в умовах змінного навантаження мережі.

**Метою дослідження** є підвищення ефективності функціонування телекомунікаційних мереж за рахунок мінімізації затримок передавання, втрат пакетів і перевантажень каналів шляхом створення математичних моделей, алгоритмів багатокритеріальної маршрутизації та адаптивного керування пропускною здатністю.

В процесі досягнення зазначеної мети та вирішення наукового завдання у роботі одержано основні наукові результати:

**Вперше розроблено** метод адаптивного управління мережевими ресурсами телекомунікаційної мережі, який базується на замкненому контурі моніторингу стану мережі та динамічній адаптивній зміні пропускної здатності каналів зв'язку.

**Удосконалено** математичну модель функціонування телекомунікаційної мережі, яка, на відміну від відомих, дозволяє врахувати закони розподілу імовірностей завантаженості маршрутів через визначення вагових коефіцієнтів пакетів з урахуванням якості обслуговування.

**Вперше розроблено** метод багатокритеріальної оптимальної маршрутизації трафіку та пріоритезації критичних даних, який ґрунтується на використанні комбінованої вагової функції каналів та механізму багатошляхового розподілу потоків.

**Удосконалено** метод керування телекомунікаційними мережами, який, на відміну від існуючих забезпечує практичне впровадження запропонованих методів у сучасних IP/MPLS та програмно-керованих мережах.

Дисертаційна робота присвячена створенню науково обґрунтованих методів і моделей управління процесами функціонування телекомунікаційних мереж, що забезпечують підвищення ефективності використання пропускної здатності, зменшення затримок і втрат пакетів і покращення показників якості обслуговування кінцевих користувачів шляхом застосування адаптивного керування і багатокритеріальної оптимізації трафіку.

*У першому розділі* здійснено всебічний аналіз сучасних телекомунікаційних технологій та методів управління мережами, зокрема механізмів QoS/QoE, алгоритмів маршрутизації, балансування навантаження та розподілу пропускної здатності, проведено порівняльну оцінку існуючих рішень і виявлено їх обмеження в умовах змінного трафіку та зростання вимог до якості сервісів, що дозволило сформулювати наукову проблему, мету та задачі дослідження.

*У другому розділі* розроблено комплекс математичних моделей функціонування телекомунікаційної мережі, зокрема модель використання пропускної здатності каналів, модель формування затримок і втрат пакетів, а також багатокритеріальну модель оцінювання якості обслуговування, на основі яких сформульовано задачу оптимізації розподілу ресурсів та маршрутизації трафіку у вигляді зваженого графа.

*У третьому розділі* запропоновано адаптивний метод управління мережевими ресурсами на основі замкненого контуру моніторингу стану мережі та метод оптимальної маршрутизації трафіку з використанням багатокритеріальної вагової функції каналів, що дозволяє здійснювати динамічний розподіл трафіку, балансування навантаження та підвищення ефективності використання пропускної здатності телекомунікаційної мережі.

*У четвертому розділі* представлено реалізацію розроблених методів у вигляді алгоритмів і програмної моделі, наведено результати імітаційного моделювання та експериментальних досліджень, виконано порівняльний аналіз ефективності запропонованих і існуючих підходів, а також визначено вплив адаптивного управління на показники функціонування телекомунікаційної мережі.



Отримані результати свідчать, що застосування розроблених адаптивних методів управління та оптимальної маршрутизації трафіку дозволяє зменшити перевантаження каналів зв'язку, підвищити пропускну здатність мережі та забезпечити стабільне функціонування телекомунікаційної системи в умовах змінного трафікового навантаження.

Дисертаційну роботу виконано відповідно до напрямку науково-дослідних робіт: «Розробка методу оцінки завадостійкості когерентного прийому дискретного сигналу з багатопозиційною фазовою маніпуляцією в умовах впливу нефлуктаційних завад» (№ державної реєстрації 0126U001466, ДУІКТ, м. Київ), «Розроблення алгоритмів для забезпечення функціональної стійкості інтелектуальних систем при прийнятті рішень» (№ державної реєстрації 0125U000865, ДУІКТ, м. Київ).

Результати наукових досліджень прийняті та впроваджені в: ТОВ «ТЕСЕМ», Інститут прикладних систем управління НАН України, а також у науково-технічну діяльність та начальний процес Інституту телекомунікацій ДУІКТ.

**Ключові слова:** телекомунікаційна мережа, управління мережею, мережеві ресурси, маршрутизація трафіку, адаптивне управління, оптимізація, математична модель, алгоритми управління, якість обслуговування, пропускну здатність, затримка передачі, втрати пакетів, балансування навантаження, телекомунікаційні послуги, моделювання мережі.

## СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

*Наукові праці, в яких опубліковані основні наукові результати дисертації.*

### *Статті у наукових фахових виданнях*

1. **Брезіцький С.М., Герасимчук В.С.** (2024). Інтегральні методи оцінки продуктивності пакетних мереж: аналіз та формування. *Наукові записки ДУІКТ*. №2(6). С. 37 – 47.

DOI: [10.31673/2786-8362.2024.025073](https://doi.org/10.31673/2786-8362.2024.025073)

2. **Вікторія Ключєва, Владислав Герасимчук, Ігор Данилов.** (2024). Типові математичні моделі, які застосовуються в автоматизованих системах організаційного управління. *SMART TECHNOLOGIES: Industrial and Civil Engineering*. Issue 2(15). С. 4 – 9.

DOI: [10.32347/st.2024.2.1101](https://doi.org/10.32347/st.2024.2.1101)

3. **Брезіцький С.М., Герасимчук В.С.** (2025). Методика формування інтегральних показників якості роботи мережевих з'єднань в мережах з комутацією пакетів. *Наукові записки ДУІКТ*. №1(7). С. 90 – 96.

DOI: [10.31673/2786-8362.2025.016443](https://doi.org/10.31673/2786-8362.2025.016443)

4. **Пепа Ю.В., Герасимчук В.С.** (2026). Інтеграція AI в процес управління телекомунікаційними послугами на прикладі VoIP. *Прикладні системи управління та робототехніка: кол. моногр. / За заг. ред. С.О. Довгого.* – Київ, ІПСУ НАН України. С. 185 – 187.

ISBN 978-617-8261-54-2

DOI: [10.35668/978-617-8261-54-2](https://doi.org/10.35668/978-617-8261-54-2)

5. **Герасимчук В.С., Пепа Ю.В.** (2026) Моделі вибору оптимального трафіку передачі даних в інфокомунікаційних мережах. *Телекомунікаційні та інформаційні технології*. №1(90). С. 59 – 68.

DOI: [10.31673/2412-4338.2026.019007](https://doi.org/10.31673/2412-4338.2026.019007)

## Матеріали й тези наукових конференцій

1. **Герасимчук Владислав Сергійович, Федчук Володимир Сергійович** (2023) Перспективи розвитку та впровадження новітніх технологій оптоволоконного зв'язку. *V Всеукраїнська науково-практична конференція «Telecommunication: Problems and Innovation»*. Збірник тез (м. Київ, 20 грудня 2023 р.). С. 43 – 44.

[https://duikt.edu.ua/uploads/n\\_12108\\_23987482.pdf#page=43](https://duikt.edu.ua/uploads/n_12108_23987482.pdf#page=43)

2. **В.С. Герасимчук** (2025) Оцінка якості надання телекомунікаційних та інформаційних послуг. *Науково-практична конференція «Актуальні проблеми безпеки інформаційно-комунікаційних систем»*. Збірник тез (м. Київ, 5 листопада 2025 р.). С. 45 – 46.

[https://duikt.edu.ua/uploads/p\\_2779\\_72486260.pdf](https://duikt.edu.ua/uploads/p_2779_72486260.pdf)

3. **Крикун Юлія Віталіївна, Герасимчук Владислав Сергійович** (2025) Розподілена біометрія: автентифікація без центрального зберігання. *III Всеукраїнська науково-технічна конференція «Технологічні горизонти: дослідження та застосування інформаційних технологій для технологічного прогресу України і світу»*. Збірник тез (м. Київ, 18 листопада 2025 р.). С. 260 – 261.

[https://duikt.edu.ua/uploads/p\\_2779\\_90287546.pdf](https://duikt.edu.ua/uploads/p_2779_90287546.pdf)

4. **Герасимчук Владислав Сергійович** (2025) Методика розрахунку транзакційних послуг. *III Міжнародна науково-практична конференція «Сучасні аспекти діджиталізації та інформатизації в програмній та комп'ютерній інженерії»*. Збірник тез (м. Київ, 4 - 6 грудня 2025 р.). С. 394 – 396.

[https://duikt.edu.ua/uploads/p\\_2779\\_50068458.pdf](https://duikt.edu.ua/uploads/p_2779_50068458.pdf)

5. **Vladyslav HERASYMCHUK, Yuriy PEPA** (2026). Algorithm for Synthesizing Background Network Traffic. *The V International Conference on Emerging «Technology Trends on the Smart Industry and the Internet of Things “TTSIIT-2026”»*. Ukraine-Iraq-Poland (Kyiv, Ukraine, January 29 - 30). pp. 37 – 40.

<https://knute.edu.ua/file/Mjk1NzE=/32f194b7b0fbe0d3633338a2027c8f6c.pdf>

6. **Olha MARCHUK, Vladyslav HERASYMCHUK** (2026). Interference Immunity of MIMO Systems Under Different Probability Distributions of Fading.

*The V International Conference on Emerging «Technology Trends on the Smart Industry and the Internet of Things “TTSIIT-2026”». Ukraine-Iraq-Poland (Kyiv, Ukraine, January 29 - 30). pp. 130 – 132.*

<https://knute.edu.ua/file/Mjk1NzE=/32f194b7b0fbe0d3633338a2027c8f6c.pdf>

## ABSTRACT

Herasymchuk V.S. Methods for improving the efficiency of management in telecommunication networks when providing services to the end user. – Qualifying scientific work submitted as a manuscript.

The dissertation for the degree of Doctor of Philosophy in specialty 172 – Telecommunications and Radio Engineering (field of knowledge 17 – Electronics and Telecommunications). – State University of Information and Communication Technologies, Kyiv, 2026.

The dissertation is devoted to the development of methods, models, and algorithms for improving the efficiency of telecommunication network management when providing services to end users through adaptive allocation of network resources and optimization of traffic routing in order to reduce transmission delays, packet losses, and increase network resource utilization while ensuring the required quality of service indicators.

***The object of research*** is the process of adaptive management of network resources, routing, and traffic distribution in telecommunication networks in order to ensure quality of service QoS/QoE indicators when providing services to end users.

***The subject of research*** is methods and models of adaptive management of network resources, routing, and traffic optimization in telecommunication networks aimed at improving network performance efficiency and ensuring the required quality of service indicators for end users.

The dissertation presents theoretical and experimental studies of telecommunication network management processes. Mathematical models and methods for adaptive allocation of network resources and optimization of traffic routing have been developed, and the efficiency of the proposed solutions has been evaluated.

Based on the conducted analysis, requirements for methods and models of telecommunication network management were formulated, which should ensure efficient allocation of network resources, adaptation to changing traffic load conditions, and support of specified quality of service indicators for end users.

Special attention is paid to the study of adaptive bandwidth allocation methods, optimal traffic routing algorithms, and mechanisms for ensuring quality of service QoS/QoE indicators under dynamic network load conditions.

*The purpose of the research* is to improve the efficiency of telecommunication network operation by minimizing transmission delays, packet losses, and channel congestion through the development of mathematical models, multicriteria routing algorithms, and adaptive bandwidth management methods.

In the process of achieving this goal and solving the scientific problem, the following main scientific results were obtained:

*For the first time*, a method has been developed for the adaptive management of telecommunications network resources, based on a closed-loop network status monitoring system and dynamic, adaptive adjustment of communication channel bandwidth.

A mathematical model of telecommunications network operation *has been refined*; unlike existing models, it allows for the consideration of probability distribution laws governing route utilization by determining packet weighting coefficients that take service quality into account.

*For the first time*, a method has been developed for multi-criteria optimal traffic routing and prioritization of critical data, based on the use of a combined channel weighting function and a multi-path flow distribution mechanism.

A method for managing telecommunications networks *has been improved*; unlike existing methods, it enables the practical implementation of the proposed methods in modern IP/MPLS and software-defined networks.

The dissertation is devoted to the development of scientifically grounded methods and models for managing telecommunication network operation processes, ensuring improved bandwidth utilization, reduced delays and packet losses, and enhanced quality of service indicators through adaptive control and multicriteria traffic optimization.

*The first chapter* provides a comprehensive analysis of modern telecommunication technologies and network management methods, including QoS/QoE mechanisms, routing algorithms, load balancing, and bandwidth allocation, identifies their limitations, and formulates the research problem, goal, and objectives.

*The second chapter* develops a set of mathematical models describing telecommunication network operation, including bandwidth utilization models, delay and packet loss formation models, and a multicriteria quality of service evaluation model, forming the basis for traffic routing optimization using a weighted graph representation.

*The third chapter* proposes an adaptive network resource management method based on closed-loop network monitoring and an optimal traffic routing method using a multicriteria channel weight function, enabling dynamic traffic distribution, load balancing, and improved network resource utilization.

*The fourth chapter* presents the implementation of the proposed methods in the form of algorithms and a software model, provides simulation and experimental results, compares the proposed and existing approaches, and evaluates the impact of adaptive management on telecommunication network performance.

The obtained results demonstrate that the application of the developed adaptive management and optimal routing methods reduces channel congestion, increases network throughput, and ensures stable telecommunication system operation under variable traffic load conditions.

This dissertation was completed in accordance with the research area: «Development of a method for evaluating the noise immunity of coherent reception of a discrete signal with multi-position phase modulation under the influence of non-fluctuating noise» (State Registration No. 0126U001466, SUICT, Kyiv), «Development of algorithms to ensure the functional stability of intelligent systems during decision-making» (State Registration No. 0125U000865, SUICT, Kyiv).

The results of scientific research have been adopted and implemented at: «TSPG LLC», the Institute of Applied Control Systems of the National Academy of Sciences of Ukraine, as well as in the scientific and technical activities and educational process of the Institute of Telecommunications at SUICT.

**Keywords:** telecommunication network, network management, network resources, traffic routing, adaptive management, optimization, mathematical model, management algorithms, quality of service, bandwidth, transmission delay, packet loss, load balancing, telecommunication services, network modeling.

## LIST OF PUBLICATIONS OF THE APPLICANT

Scientific works in which the main scientific findings of the dissertation have been published.

### Articles in scientific journals

1. **Brezitskyi S.M., Herasymchuk V.S.** (2024). Intehralni metody otsinky produktyvnosti paketnykh merezh: analiz ta formuvannia. *Naukovi zapysky SUICT*. No 2(6). pp. 37 – 47.

DOI: [10.31673/2786-8362.2024.025073](https://doi.org/10.31673/2786-8362.2024.025073)

2. **Viktoriia Kliuieva, Vladyslav Herasymchuk, Ihor Danylov.** (2024). Typovi matematychni modeli, yaki zastosovuiutsia v avtomatyzovanykh systemakh orhanizatsiinoho upravlinnia. *SMART TECHNOLOGIES: Industrial and Civil Engineering*. Issue 2(15). pp. 4 – 9.

DOI: [10.32347/st.2024.2.1101](https://doi.org/10.32347/st.2024.2.1101)

3. **Brezitskyi S.M., Herasymchuk V.S.** (2025). Metodyka formuvannia intehralnykh pokaznykiv yakosti roboty merezhevykh ziednan v merezhakh z komutatsiieiu paketiv. *Naukovi zapysky SUICT*. No 1(7). pp. 90 – 96.

DOI: [10.31673/2786-8362.2025.016443](https://doi.org/10.31673/2786-8362.2025.016443)

4. **Pepa Y.V., Herasymchuk V.S.** (2026). Intehratsiia AI v protses upravlinnia telekomunikatsiinymy posluhamy na prykladi VoIP. *Prykladni systemy upravlinnia ta robototekhnika: kol. monohr. / Za zah. red. S.O. Dovhoho.* – Kyiv, IPSU NAN Ukraine. pp. 185 – 187.

ISBN 978-617-8261-54-2

DOI: [10.35668/978-617-8261-54-2](https://doi.org/10.35668/978-617-8261-54-2)

5. **Herasymchuk V.S., Pepa Y.V.** (2026) Modeli vyboru optymalnoho trafiku peredachi danykh v infokomunikatsiinnykh merezhakh. *Telekomunikatsiini ta informatsiini tekhnolohii*. No 1(90). pp. 59 – 68.

DOI: [10.31673/2412-4338.2026.019007](https://doi.org/10.31673/2412-4338.2026.019007)



### Proceedings and abstracts of academic conferences

1. **Herasymchuk Vladyslav Serhiiovych, Fedchuk Volodymyr Serhiiovych** (2023) Perspektyvy rozvytku ta vprovadzhennia novitnikh tekhnolohii optovolokonnoho zviazku. *V Vseukrainska naukovo-praktychna konferentsiia «Telecommunication: Problems and Innovation»*. Zbirnyk tez (Kyiv, December 20). pp. 43 – 44.

[https://duikt.edu.ua/uploads/n\\_12108\\_23987482.pdf#page=43](https://duikt.edu.ua/uploads/n_12108_23987482.pdf#page=43)

2. **V.S. Herasymchuk** (2025) Otsinka yakosti nadannia telekomunikatsiinykh ta informatsiinykh posluh. *Naukovo-praktychna konferentsiia «Aktualni problemy bezpeky informatsiino-komunikatsiinykh system»*. Zbirnyk tez (Kyiv, November 5). pp. 45 – 46.

[https://duikt.edu.ua/uploads/p\\_2779\\_72486260.pdf](https://duikt.edu.ua/uploads/p_2779_72486260.pdf)

3. **Krykun Yuliia Vitaliivna, Herasymchuk Vladyslav Serhiiovych** (2025) Rozpodilena biometriia: avtentyfikatsiia bez tsentralnoho zberihannia. *III Vseukrainska naukovo-tekhnichna konferentsiia «Tekhnolohichni horyzonty: doslidzhennia ta zastosuvannia informatsiinykh tekhnolohii dlia tekhnolohichnoho prohresu Ukrainy i svitu»*. Zbirnyk tez (Kyiv, November 18). pp. 260 – 261.

[https://duikt.edu.ua/uploads/p\\_2779\\_90287546.pdf](https://duikt.edu.ua/uploads/p_2779_90287546.pdf)

4. **Herasymchuk Vladyslav Serhiiovych** (2025) Metodyka rozrakhunku transaktsiinykh posluh. *III Mizhnarodna naukovo-praktychna konferentsiia «Suchasni aspekty didzhitalizatsii ta informatyzatsii v prohramnii ta kompiuternii inzhenerii»*. Zbirnyk tez (Kyiv, December 4 - 6). pp. 394 – 396.

[https://duikt.edu.ua/uploads/p\\_2779\\_50068458.pdf](https://duikt.edu.ua/uploads/p_2779_50068458.pdf)

5. **Vladyslav HERASYMCHUK, Yuriy PEPA** (2026). Algorithm for Synthesizing Background Network Traffic. *The V International Conference on Emerging «Technology Trends on the Smart Industry and the Internet of Things “TTSIIT-2026”»*. Ukraine-Iraq-Poland (Kyiv, Ukraine, January 29 - 30). pp. 37 – 40.

<https://knute.edu.ua/file/Mjk1NzE=/32f194b7b0fbe0d3633338a2027c8f6c.pdf>

6. **Olha MARCHUK, Vladyslav HERASYMCHUK** (2026). Interference Immunity of MIMO Systems Under Different Probability Distributions of Fading. *The V International Conference on Emerging «Technology Trends on the Smart Industry*

*and the Internet of Things “TTSIIT-2026”*». Ukraine-Iraq-Poland (Kyiv, Ukraine, January 29 - 30). pp. 130 – 132.

<https://knute.edu.ua/file/Mjk1NzE=/32f194b7b0fbe0d3633338a2027c8f6c.pdf>

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ .....	18
ВСТУП.....	19
РОЗДІЛ 1 .....	32
ТЕЛЕКОМУНІКАЦІЙНІ ПОСЛУГИ ТА СПОСОБИ ЇХ МОДЕЛЮВАННЯ.....	32
1.1 Шляхи розвитку телекомунікаційних послуг .....	32
1.2 Асинхронний режим роботи мереж .....	34
1.3. Специфікація якості обслуговування та метрика складності системи.....	37
1.4 Ентропія втрат бітів в пакетах даних.....	42
1.5 Сукупний коефіцієнт втрат бітів через затримку та впливи відношення сигнал/шум .....	44
1.6. Нечіткі атрибути характеристик втрат бітів в пакетах.....	48
1.7. Лінгвістичний опис нечітких правил.....	52
1.8 Межа допуску викликів: чіткі та нечіткі аспекти.....	59
1.9 Опис об'єкта дослідження .....	61
Висновки до розділу 1 .....	63
РОЗДІЛ 2 .....	65
МОДЕЛІ І МЕТОДИ МОДЕЛЮВАННЯ .....	65
ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ.....	65
2.1. Методи моделювання телекомунікаційних систем .....	65
2.2. Математичний апарат для моделювання інформаційно-комунікаційних систем.....	68
2.3 Класифікація послуг і сервісів за рівнем взаємодії з користувачем .....	75
2.4. Мікросервісна архітектура надання послуг користувачеві .....	80
2.5 Моделі оцінювання часових характеристик комбінованих послуг .....	85
2.6 Моделі та методи розрахунку затримок під час надання сеансових послуг користувачам .....	90
2.7 Математичний апарат для обчислення часових характеристик сеансових послуг .....	97

	16
Висновки до розділу 2 .....	99
РОЗДІЛ 3 .....	101
МОДЕЛЬ ДЛЯ ОЦІНКИ ХАРАКТЕРИСТИК ТРАНЗАКЦІЙНИХ ПОСЛУГ .....	101
ІЗ ПАРАЛЕЛЬНОЮ ОБРОБКОЮ ЗАПИТІВ.....	101
3.1. Транзакційні послуги, як сценарії реалізації робочих процесів .....	101
3.2. Паралельне опрацювання запитів за методом Fork-Join.....	109
3.3 Моделювання робочого процесу з використанням математичного апарату мереж Джексона .....	113
3.4. Обчислення параметрів затримки систем довільного типу.....	119
3.5 Оцінка часових затримок у послідовності локальних транзакцій.....	122
3.6. Оцінювання стану послідовних мереж систем G/M/1 .....	124
Висновки до розділу 3 .....	130
РОЗДІЛ 4 .....	132
МЕТОДИ ОЦІНЮВАННЯ ХАРАКТЕРИСТИК .....	132
ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ.....	132
4.1. Особливості етапів моделювання при розробці інформаційних систем.....	132
4.2 Алгоритм розрахунку ймовірно-часових характеристик телекомунікаційної системи .....	134
4.3. Результати моделювання .....	139
4.4. Обговорення результатів .....	148
4.5. Оцінка похибок вимірювань та розрахунків .....	150
4.6. Методи ефективного управління мережевими ресурсами .....	159
Висновки до розділу 4 .....	162
ВИСНОВКИ.....	164
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	167
ДОДАТКИ .....	186
ДОДАТОК А .....	187
ДОДАТОК Б.....	191
ДОДАТОК В.....	194
ДОДАТОК Г .....	195

ДОДАТОК Д.....	199
ДОДАТОК Е.....	202

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ВВХ – вхідно-вихідні характеристики

ІМ – імітаційне моделювання

ЙЧХ – ймовірно-часові характеристики

МСА – мікросервісна архітектура

СМО – система масового обслуговування

ФРВ – функція розподілу вірогідності

ЩРВ – щільність розподілу вірогідності

ATM – Asynchronous Transfer Mode

CAC – Connection Admission Control

CDV – Cell Delay Variation

CLR – Cell Loss Ratio

NFV – Network Function Virtualization

SDN – Software-Defined Networking

SNR – Signal-to-Noise Ratio

QoE – Quality of Experience

QoS – Quality of Services

## ВСТУП

Сучасна телекомунікаційна послуга розглядається як складна система, що характеризується стохастичними властивостями та нечіткими критеріями продуктивності. Як приклад такої системи проаналізовано телекомунікації на основі асинхронного режиму передачі (ATM). Для цієї технології розроблено процедуру контролю допуску викликів (CAC), що ґрунтується на методі нечіткого виведення та враховує складні схеми з'єднань у сервісі ATM. Застосування такого підходу дозволяє уникнути надмірно консервативного характеру розподілу ресурсів. У дослідженні введено поняття параметра складної системи, який використовується як нечітка змінна в розроблених стратегіях виведення. Окрему увагу приділено практичним аспектам та перевагам запропонованого підходу.

Сучасні телекомунікаційні системи цілком обґрунтовано можна віднести до класу складних систем [47, 68, 127]. Масштабність застосованих технологій, різноманітність сервісних вимог та велика кількість профілів застосувань зумовлюють неминучий характер взаємодії між підсистемами, що входять до складу сучасної телекомунікаційної інженерії. Ця взаємодія, як правило, визначає, яким чином ресурси однієї підсистеми витрачаються під час обміну з іншими у складний спосіб.

Параметри якості обслуговування (QoS), що пов'язані з телекомунікаційною системою, виступають метрикою її складності. Така метрика дозволяє оцінити глобальну ефективність наданої послуги у співвідношенні до витрат. Оцінка параметрів QoS або їхнє формальне визначення під час узгодження з постачальником послуг доступу до певного класу з'єднань у мережі потребує адекватного моделювання складності, притаманної телекомунікаційній системі. Відповідне моделювання має враховувати як просторову стохастичність взаємодіючих ресурсів, так і часову динаміку інформаційних потоків між кінцевими об'єктами телекомунікаційного каналу.

Дослідження [22, 136, 144] свідчать, що такі просторово-часові аспекти у сучасному телекомунікаційному середовищі часто мають нечіткий характер. Тобто

процес передачі інформації мережею відображає невизначеність значень і розмитість меж змінних, що беруть у ньому участь. Відтак, параметри QoS у телекомунікаціях можуть бути відображені у нечіткій області. Відповідно, характеристики продуктивності, а також зроблені на їх основі висновки можуть бути більш адекватно визначені у нечітких множинах. Крім того, будь-яке керування системою може реалізовуватися за допомогою механізму нечіткого виведення. Інакше кажучи, параметри складної системи, що характеризують її ефективність, можуть бути використані як нечіткі змінні для формування критерію прийняття керівних рішень.

У контексті сучасних телекомунікаційних систем [35-37] актуальність стохастичних чинників, нечітких характеристик та складності (як у просторовій, так і в часовій інтерпретаціях) виникає головним чином з таких причин:

1. Гетерогенність типів трафіку (передача голосу, відео та даних).
2. Різноманітність фізичних середовищ передавання (мідні проводи, оптичні волокна та бездротові засоби).
3. Різні версії комутації, які з'єднують лінії з буферами різного розміру.
4. Часові коливання попиту на трафік (навантаження у години пік та у періоди спаду).
5. Тимчасові варіації швидкості передавання у бітах за секунду від джерела (трафік із змінною бітовою швидкістю).
6. Залежні від часу потреби у пропускній здатності (передача мультимедійного контенту).
7. Синхронна та асинхронна передача пакетизованої інформації [168].
8. Типи пакетизації бітів (змінного розміру пакети або фіксовані за розміром комірки).
9. Специфіка протоколів, зокрема контроль допуску викликів САС, управління перевантаженнями та планування пріоритетів обслуговування.
10. Статистичне мультиплексування пакетів із пріоритетами трафіку при конфліктах.



11. Співвідношення сигнал/шум (SNR) та показники бітових помилок (виявлення та корекція помилок).

12. Статистичні аспекти кількості інформації, що передається різними джерелами (фактор навантаження як функція часу).

13. Мобільність кінцевих пристроїв (системи безпроводового зв'язку).

14. Типи та режими реалізації сигналізації для встановлення й завершення з'єднань.

15. Орієнтовані на з'єднання та безз'єднувальні конфігурації мережі.

16. Географічні масштаби: глобальні, міжміські (WAN), міські (MAN) та локальні (LAN) мережі.

17. Безпека мережі (вимоги до шифрування та дешифрування).

Загальні характеристики мережі (зокрема її розмірність за критерієм зв'язності та обсяг оброблюваного трафіку), а також стохастична природа її функціонування як функції часу зумовлюють реалістичну складність телекомунікаційної системи.

Для повного відображення інтуїтивного поняття складності телекомунікаційних систем необхідно репрезентувати її у багатовимірній моделі. Така багатовимірність має охоплювати сукупність показників, кожен із яких відображає певний аспект категорії послуг, цілей QoS та дескрипторів трафіку. Оскільки багато з цих показників можуть бути корельованими, модель повинна враховувати зростання оцінюваної складності внаслідок взаємодії між цими параметрами.

Метрики QoS у чіткій (crisp) формі є лише примітивними показниками. Вони самі по собі не відображають загальних та кореляційних властивостей параметрів, а також не відображають нечіткості змінних. У кращому випадку їх представляють у ймовірнісних нормах, які характеризують виключно стохастичні властивості.

Щоб комплексно врахувати стохастичні чинники, нечіткі характеристики та загальну складність системи, необхідна інтеграція таких підходів: урахування просторово-часової динаміки системи, описаної в площині ентропії або

інформаційної теорії [48, 74]; задання нечітких характеристик змінним, що беруть участь у динаміці [77, 80].

Метою даної роботи є показати, як складність телекомунікаційної системи може бути аналітично змодельована з урахуванням інформаційно-теоретичних (ІТ) підходів та нечітких властивостей. Зокрема, розглянуто комутаційний етап у системах асинхронного режиму передачі АТМ. Моделювання виконано з урахуванням двох параметрів QoS: (і) варіацій затримок передавання комірок. Такі коливання затримок (CDV) або джитер призводять до відкидання комірок, які перевищують допустиму межу затримки, що є особливо критичним у перевантаженому трафіку; додаткових втрат комірок (пакетизованих сегментів бітів з інформацією) унаслідок бітових помилок, спричинених шумовим каналом.

Продуктивність телекомунікаційної мережі відносно заданих цілей QoS може бути оцінена у термінах профілю складності, що визначається стохастичними характеристиками та нечіткими атрибутами відповідної технології. У цій роботі показано, як багатовимірність примітивних показників продуктивності мережі можна звести до єдиного параметра складності (з нечіткими властивостями), який дозволяє встановити верхню межу продуктивності. Згадана стратегія передбачає поєднання інформаційно-теоретичного підходу та нечіткої логіки.

Використовуючи виведений параметр складності, показано, як побудувати механізм нечіткого виведення, у якому цей параметр виступає як нечітка змінна. На основі цього механізму розроблено алгоритм для реалізації контролю допуску викликів САС у мережах АТМ. Іншими словами, створено механізм нечіткого виведення, що оцінює складні схеми з'єднань у мережі та формує нечіткі норми використання ресурсів. Такий підхід знижує ризик надмірно консервативного розподілу ресурсів у процесі реалізації САС. Ефективність цього алгоритму підтверджується результатами моделювання для типового середовища АТМ.

Таким чином, робота складається з двох частин. У розділі 2 виведено часово-залежну метрику складності  $s(t)$  за допомогою інформаційно-теоретичного підходу для телекомунікаційних систем АТМ. Статистичні межі для гарантій з'єднання (тобто параметрів QoS), які вимагає користувач від оператора мережі,

виражено через запропоновану метрику складності. Стратегія моделювання ґрунтується на тому, що потік  $i$  АТМ-комірок із інформаційним навантаженням розглядається як система масового обслуговування (СМО). На відміну від традиційних моделей черг, де використовуються часові характеристики надходження та очікування комірок, у даній роботі розглядається потік ентропії (інформаційного змісту у сенсі Шеннона), що переноситься чергою [138, 139]. Відповідно визначено метрику складності  $s(t)$  та виражено її через інформаційний зміст за Шенноном; часову динаміку  $s(t)$  сформульовано на основі стохастичного диференціального числення; нечіткі властивості  $s(t)$  введено за допомогою розширеного принципу Заде та отримано нечітке рівняння Фоккера–Планка [85, 86], яке описує динаміку  $s(t)$ .

Використання ІТ-підходу дозволяє безпосередньо оцінювати втрати інформації, що виникають через статистику втрат комірок. Тобто вплив втрат комірок, які знижують обсяг переданої інформації, враховано через введений параметр складності. Розглядаються втрати як унаслідок переповнення буфера або черги у комутаторі (або мультиплексорі) АТМ, так і втрати, спричинені некоригованими бітовими помилками через обмежене SNR (що відображає скінченний рівень шуму у системі).

В роботі, на основі стратегії виведено алгоритм САС та продемонстровано його ефективність за допомогою результатів моделювання.

#### **Актуальність теми дослідження.**

Стрімкий розвиток інформаційно-комунікаційних технологій, широке впровадження мереж наступного покоління (NGN), технологій 5G/6G, хмарних обчислень та мікросервісних архітектур зумовлюють значне зростання обсягів трафіку та підвищення вимог до якості надання телекомунікаційних послуг. Сучасні телекомунікаційні мережі повинні забезпечувати високий рівень QoS [104] та якості сприйняття користувачем (QoE) [101], мінімальні затримки передавання даних, низьку ймовірність втрат пакетів і ефективне використання обмежених мережевих ресурсів [6].

Особливістю сучасних телекомунікаційних систем є їх складна структура, яка включає велику кількість взаємопов'язаних компонентів, реалізованих на основі розподілених і мікросервісних архітектур. У таких системах процеси обробки запитів мають стохастичний характер, а параметри функціонування мережі можуть змінюватися в широких межах залежно від інтенсивності трафіку, кількості користувачів і характеристик мережевого середовища. Це значно ускладнює процес управління [164, 171] телекомунікаційними мережами та потребує застосування ефективних математичних моделей і методів адаптивного керування.

Існуючі методи управління мережевими ресурсами та маршрутизації трафіку, як правило, базуються на статичних або частково адаптивних підходах і не повною мірою враховують динамічний характер сучасних телекомунікаційних послуг, особливо транзакційних і комбінованих сервісів, що реалізуються у мікросервісному середовищі. Це призводить до неефективного використання пропускної здатності мережі, виникнення перевантажень, збільшення затримок передавання та погіршення показників якості обслуговування.

У зв'язку з цим актуальним є розроблення нових математичних моделей і методів управління телекомунікаційними мережами, які враховують стохастичний характер процесів обробки трафіку, особливості функціонування мікросервісних архітектур, паралельну обробку запитів і багатокритеріальні показники якості обслуговування. Використання таких моделей і методів дозволяє підвищити ефективність використання мережевих ресурсів, зменшити затримки передавання та втрати пакетів, а також забезпечити необхідний рівень QoS/QoE для кінцевих користувачів.

Таким чином, тема дисертаційної роботи є актуальною, оскільки спрямована на вирішення важливого науково-технічного завдання підвищення ефективності управління телекомунікаційними мережами шляхом розроблення математичних моделей і адаптивних методів управління ресурсами та маршрутизації трафіку, що має важливе значення для розвитку сучасних телекомунікаційних систем і технологій.

### **Ступінь розробленості теми.**

У сучасному інформаційному просторі спостерігається помітне зростання кількості й різновидів телекомунікаційних послуг, розробники яких разом зі спільнотою науковців активно публікують праці, що присвячені дослідженню якості обслуговування багатокористувацьких сервісів [129, 130], аналізу архітектурних рішень із підтримкою QoS [102-104], створенню складених телекомунікаційних послуг [100, 128] та вивченню можливостей застосування наявних рішень [46, 49, 73]. Напрацювання в цих напрямках можуть бути корисними для впровадження і використання телекомунікаційних послуг у різних організаціях.

Значний внесок у розвиток теорії телекомунікаційних систем, теорії телетрафіку, математичного моделювання мереж і забезпечення якості обслуговування зробили такі українські науковці, як О.І. Лисенко [79, 90], В.В. Поповський [12, 69], О.І. Романов [14], А.І. Семенко [15, 21], І.В. Стрілковська [16], А.Г. Ложковський [17-19], М.В. Захарченко [15, 21], В.Є. Бондаренко [23, 24], С.О. Кравчук [25-27], Л.О. Уривський [28, 29], О.М. Лебедєв [30], Л.Н. Беркман [13, 171, 172, 168] та інші.

Численні статті, наукові публікації та дисертації присвячені вивченню продуктивності та підвищенню ефективності телекомунікацій у різних сферах їхнього використання. Однак у наявних дослідженнях недостатньо уваги приділено окремим аспектам, таким як аналіз вимог до архітектурних рішень, порівняння різних сценаріїв реалізації процесів надання послуг, а також визначення якісних характеристик за допомогою математичного й імітаційного моделювання (ІМ). Це ускладнює ухвалення обґрунтованих рішень під час проєктування телекомунікаційних послуг.

Теоретичні й практичні інженерні дослідження частково висвітлюють окремі питання, які лежать в основі цієї дисертаційної роботи. Зокрема, у роботі [25] автори запропонували методику оцінювання сценаріїв розгортання архітектури слабо пов'язаних програмних компонентів, засновану на автоматизованому тестуванні продуктивності з використанням навантажувального тестування, орієнтованого на специфіку домену.

У дослідженні [58] із застосуванням графових моделей детально проаналізовано механізми взаємодії між слабо пов'язаними послугами в контексті робочих процесів. У роботі [62] обговорено принципи масштабування веб-сервісів, а також розглянуто проблеми, які вирішуються завдяки цьому підходу. Серед таких проблем – обробка великої кількості одночасних запитів, управління обмеженою пропускною здатністю та оптимізація затримок під час операцій із читання й запису даних. У роботах [63, 64] представлено метод оцінювання часу відгуку веб-сервісів для визначення їхньої продуктивності. Автори планують використовувати цей метод під час проєктування обчислювальних систем.

Ця дисертаційна робота є суттєвим внеском у наукову літературу, присвячену аналізу телекомунікаційних послуг. Вона зосереджена на вивченні часових характеристик і продуктивності в умовах різних архітектурних підходів, зокрема мікросервісної архітектури. Використання комбінації аналітичного та ІМ дає змогу отримати точні результати для різноманітних сценаріїв функціонування системи. Ці результати можуть слугувати основою для прийняття обґрунтованих рішень щодо управління якісними та кількісними показниками на етапах проєктування та експлуатації телекомунікаційних послуг, що підкреслює важливість роботи для розвитку цієї галузі.

**Об'єктом дослідження** дисертаційної роботи є процес адаптивного управління мережевими ресурсами, маршрутизацією та розподілом трафіку в телекомунікаційних мережах з метою забезпечення показників якості обслуговування QoS/QoE при наданні послуг кінцевим користувачам.

**Предметом дослідження** дисертаційної роботи є методи та моделі адаптивного управління мережевими ресурсами, маршрутизації й оптимізації трафіку в телекомунікаційних мережах, спрямовані на підвищення ефективності функціонування мережі та забезпечення необхідних показників якості обслуговування кінцевих користувачів.

#### **Мета та завдання дисертації.**

**Мета** – підвищення ефективності функціонування телекомунікаційних мереж за рахунок мінімізації затримок передавання, втрат пакетів і перевантажень

каналів шляхом створення математичних моделей, алгоритмів багатокритеріальної маршрутизації та адаптивного керування пропускнуою здатністю.

Для досягнення цієї мети визначено такі завдання:

1. Провести аналіз сучасного стану методів управління телекомунікаційними мережами, моделей розподілу мережевих ресурсів і маршрутизації трафіку, а також визначити їх переваги, недоліки та обмеження в умовах змінного трафікового навантаження.

2. Розробити математичну модель телекомунікаційної мережі, яка дозволяє описати процеси передавання трафіку, розподілу мережевих ресурсів і оцінювання показників якості обслуговування з урахуванням стохастичних характеристик трафіку.

3. Удосконалити математичний апарат оцінювання стану телекомунікаційної мережі шляхом врахування багатокритеріальних показників якості обслуговування, таких як затримка передавання, рівень завантаження каналів і ймовірність втрат пакетів.

4. Розробити метод адаптивного управління мережевими ресурсами телекомунікаційної мережі, який забезпечує ефективний розподіл пропускнуої здатності каналів зв'язку з урахуванням поточного стану мережі.

5. Розробити метод багатокритеріальної оптимальної маршрутизації трафіку, який забезпечує балансування навантаження, підвищення ефективності використання мережевих ресурсів і покращення показників якості обслуговування.

6. Провести ІМ та експериментальні дослідження запропонованих моделей і методів, а також оцінити практичну ефективність запропонованих методів і визначити можливість їх використання в сучасних телекомунікаційних мережах.

7. Розробити алгоритми та програмні засоби реалізації запропонованих методів управління телекомунікаційними мережами.

**Наукове завдання** полягає в розробці та дослідженні математичних моделей і адаптивних методів управління процесами розподілу мережевих ресурсів і маршрутизації трафіку в телекомунікаційних мережах, які враховують стохастичний характер трафіку, багатокритеріальні показники якості

обслуговування та забезпечують підвищення ефективності функціонування телекомунікаційної мережі.

### **Наукова новизна результатів дослідження:**

1. Уперше розроблено метод адаптивного управління мережевими ресурсами телекомунікаційної мережі, який, на відміну від існуючих, базується на використанні замкненого контуру моніторингу стану мережі та механізму динамічного перерозподілу пропускної здатності каналів зв'язку з урахуванням поточного трафікового навантаження, що дозволяє підвищити ефективність використання мережесих ресурсів і зменшити затримки передавання та втрати пакетів.

2. Удосконалено математичну модель функціонування телекомунікаційної мережі, яка, на відміну від відомих, враховує багатокритеріальні показники якості обслуговування (затримку передавання, рівень завантаження каналів, імовірність втрат пакетів) при формуванні вагових коефіцієнтів каналів зв'язку, що забезпечує більш адекватне відображення реального стану мережі та підвищує ефективність процесів управління і маршрутизації трафіку.

3. Уперше розроблено метод багатокритеріальної оптимальної маршрутизації трафіку, який відрізняється використанням комбінованої вагової функції каналів зв'язку та механізму багатошляхового розподілу потоків даних, що дозволяє здійснювати балансування навантаження, зменшувати перевантаження окремих сегментів мережі та підвищувати ефективність використання пропускної здатності телекомунікаційної мережі.

4. Набули подальшого розвитку математичні моделі та алгоритмічні засоби управління телекомунікаційними мережами, які, на відміну від існуючих, забезпечують можливість адаптивного управління процесами розподілу мережесих ресурсів і маршрутизації трафіку в умовах змінного навантаження та складної структури сучасних телекомунікаційних систем, що дозволяє підвищити інтегральні показники якості обслуговування кінцевих користувачів.

**Теоретична значущість роботи** полягає у розвитку математичного апарату моделювання та аналізу процесів функціонування телекомунікаційних мереж



шляхом розроблення нових і вдосконалення існуючих математичних моделей, методів і алгоритмів адаптивного управління мережевими ресурсами та маршрутизації трафіку з урахуванням багатокритеріальних показників якості обслуговування.

Запропоновані в дисертаційній роботі математичні моделі, що базуються на представленні телекомунікаційної мережі у вигляді зваженого графа та враховують стохастичний характер трафіку, дозволяють більш точно описувати процеси функціонування сучасних телекомунікаційних систем. Розроблені методи адаптивного управління ресурсами та багатокритеріальної оптимізації маршрутизації трафіку розширюють теоретичні основи управління телекомунікаційними мережами та створюють наукову базу для подальших досліджень у галузі підвищення ефективності функціонування інформаційно-комунікаційних систем.

**Практична значущість роботи** полягає у можливості використання розроблених математичних моделей, методів і алгоритмів адаптивного управління мережевими ресурсами та оптимальної маршрутизації трафіку для підвищення ефективності функціонування сучасних телекомунікаційних мереж.

Запропоновані методи дозволяють здійснювати динамічний розподіл мережеских ресурсів з урахуванням поточного стану мережі та характеристик трафіку, що забезпечує зменшення затримок передавання, зниження ймовірності втрат пакетів і підвищення ефективності використання пропускну здатності каналів зв'язку. Це сприяє покращенню показників QoS/QoE та забезпеченню стабільного функціонування телекомунікаційних систем.

Розроблені математичні моделі та алгоритми можуть бути використані при проектуванні, аналізі та оптимізації телекомунікаційних мереж, зокрема IP/MPLS-мереж, програмно-керованих мереж (SDN), мереж наступного покоління NGN, а також у системах, що використовують мікросервісні архітектури.

### **Методологія та методи дисертаційного дослідження.**

Методологічною основою дисертаційного дослідження є сучасні положення теорії телекомунікаційних мереж, теорії масового обслуговування, теорії графів,

теорії ймовірностей і математичної статистики, теорії оптимізації, а також методи математичного та ІМ складних телекомунікаційних систем.

У процесі виконання дисертаційної роботи використано комплекс взаємопов'язаних теоретичних і експериментальних методів дослідження, зокрема:

- методи системного аналізу – для дослідження процесів функціонування телекомунікаційних мереж, аналізу існуючих методів управління мережевими ресурсами та визначення напрямів підвищення ефективності управління;

- методи теорії графів – для формалізації структури телекомунікаційної мережі у вигляді зваженого графа та побудови математичних моделей процесів маршрутизації трафіку;

- методи теорії масового обслуговування – для моделювання процесів обробки трафіку, оцінювання часових характеристик обслуговування запитів і аналізу ефективності функціонування телекомунікаційних систем;

- методи теорії ймовірностей і математичної статистики – для опису стохастичних характеристик телекомунікаційного трафіку, аналізу випадкових процесів і оцінювання характеристик функціонування мережі;

- методи математичного моделювання – для розроблення моделей функціонування телекомунікаційної мережі, оцінювання показників якості обслуговування та аналізу ефективності управління мережевими ресурсами;

- методи багатокритеріальної оптимізації – для розроблення методів оптимальної маршрутизації трафіку та адаптивного розподілу мережесих ресурсів з урахуванням кількох критеріїв ефективності;

- методи імітаційного моделювання – для експериментального дослідження функціонування телекомунікаційної мережі, перевірки адекватності розроблених моделей і оцінювання ефективності запропонованих методів;

- методи алгоритмічного та програмного моделювання – для реалізації розроблених моделей і методів у вигляді програмних засобів і проведення чисельних експериментів.

Застосування зазначених методів дозволило розробити математичні моделі та ефективні методи управління телекомунікаційними мережами, виконати їх

теоретичне обґрунтування та експериментальну перевірку, а також підтвердити достовірність отриманих результатів.

**Основні положення, що виносяться на захист:**

1. Математична модель телекомунікаційної мережі у вигляді зваженого графа, яка, на відміну від існуючих моделей, враховує багатокритеріальні показники якості обслуговування, зокрема затримку передавання, рівень завантаження каналів і ймовірність втрат пакетів, що забезпечує більш точне оцінювання стану мережі та створює основу для ефективного управління мережевими ресурсами.

2. Метод адаптивного управління мережевими ресурсами телекомунікаційної мережі, який базується на використанні замкненого контуру моніторингу стану мережі та механізму динамічного перерозподілу пропускної здатності каналів зв'язку, що дозволяє підвищити ефективність використання мережевих ресурсів і покращити показники якості обслуговування кінцевих користувачів.

3. Метод багатокритеріальної оптимальної маршрутизації трафіку, який відрізняється використанням комбінованої вагової функції каналів зв'язку та механізму багатошляхового розподілу потоків даних, що забезпечує балансування навантаження між каналами, зменшення перевантажень мережі та підвищення ефективності використання пропускної здатності.

4. Математичні моделі та алгоритмічні засоби управління телекомунікаційними мережами, які забезпечують адаптивний розподіл мережевих ресурсів і оптимізацію маршрутизації трафіку в умовах змінного навантаження та дозволяють підвищити ефективність функціонування телекомунікаційної мережі.

5. Результати ІМ та експериментальних досліджень, які підтверджують ефективність запропонованих методів і моделей та демонструють їх переваги порівняно з існуючими підходами щодо підвищення ефективності використання мережевих ресурсів і забезпечення показників якості обслуговування.

## РОЗДІЛ 1

### ТЕЛЕКОМУНІКАЦІЙНІ ПОСЛУГИ ТА СПОСОБИ ЇХ МОДЕЛЮВАННЯ

#### 1.1 Шляхи розвитку телекомунікаційних послуг

Фундаментальною основою для досягнення зразкових показників якості сучасних сервісів [57] виступає детальний превентивний аналіз та математичне моделювання мережевих процесів ще на ранніх стадіях їх проектування. Такий підхід дозволяє суттєво мінімізувати потенційні ризики та оптимізувати системну архітектуру ще до моменту її практичного впровадження. В основі цього наукового методу лежать засади класичної теорії масового обслуговування [18], розвиток якої нерозривно пов'язаний з іменами таких видатних вчених, як James F. Kurose [34], William Stallings [33] та Andrew S. Tanenbaum [31, 32].

Центральна проблематика даного дослідження охоплює не лише питання стабільності базової інфраструктури, а й гарантування високого рівня якості прикладних сервісів, що функціонують на її основі. Лише комплексний розгляд системної архітектури та додатків дозволяє повною мірою реалізувати потенціал новітніх мережевих технологій [12, 125]. Важливо враховувати, що сучасна архітектура програмних продуктів перебуває у прямій залежності від технічних вимог: з посиленням запитів до безпеки, надійності та продуктивності складність системних структур зростає експоненційно.

Для подолання цих викликів ключового значення набувають специфічні архітектурні аспекти, які необхідно враховувати при розробці:

- **Горизонтальне масштабування:** забезпечення здатності системи до розширення у розподілених структурах центрів обробки даних;
- **Висока надійність:** гарантування безперебійної роботи для критично важливих функцій та сервісів;
- **Архітектурна гнучкість:** використання уніфікованих протоколів та інтерфейсів для безшовної зовнішньої інтеграції;
- **Автономність компонентів:** розділення системи на окремі модулі для підвищення загальної відмовостійкості.

Еволюція мережевої інфраструктури, що супроводжується активною розробкою концепцій 6G [117] та ініціатив 2030 під егідою ITU [120-123], ETSI [87] та IETF [88] і диктує нові, більш суворі умови для забезпечення QoS. Ці питання активно досліджувалися такими вченими, як Roy B. Jackson [135], Dimitri P. Bertsekas [173], О.В. Лемешко [52], Т.М. Лебеденко [52] та Л.Н. Беркман [69]. Розвиток технологій створює необхідні умови для впровадження інноваційних додатків, функціонування яких раніше було неможливим через існуючі технічні обмеження. При цьому такі додатки можуть виступати як самостійні сервіси, так і бути частинами складної багаторівневої взаємодії компонентів. Враховуючи вищевикладене, пріоритетним завданням даної роботи є розробка спеціалізованого алгоритму для проєктування телекомунікаційних систем, який би враховував складну взаємозалежність між архітектурними рішеннями та експлуатаційними вимогами.

Подальший аналіз телекомунікаційних сервісів планується здійснювати із застосуванням моделі абстрактної соціальної мережі [167], структурну схему якої наведено на рис. 1.1.

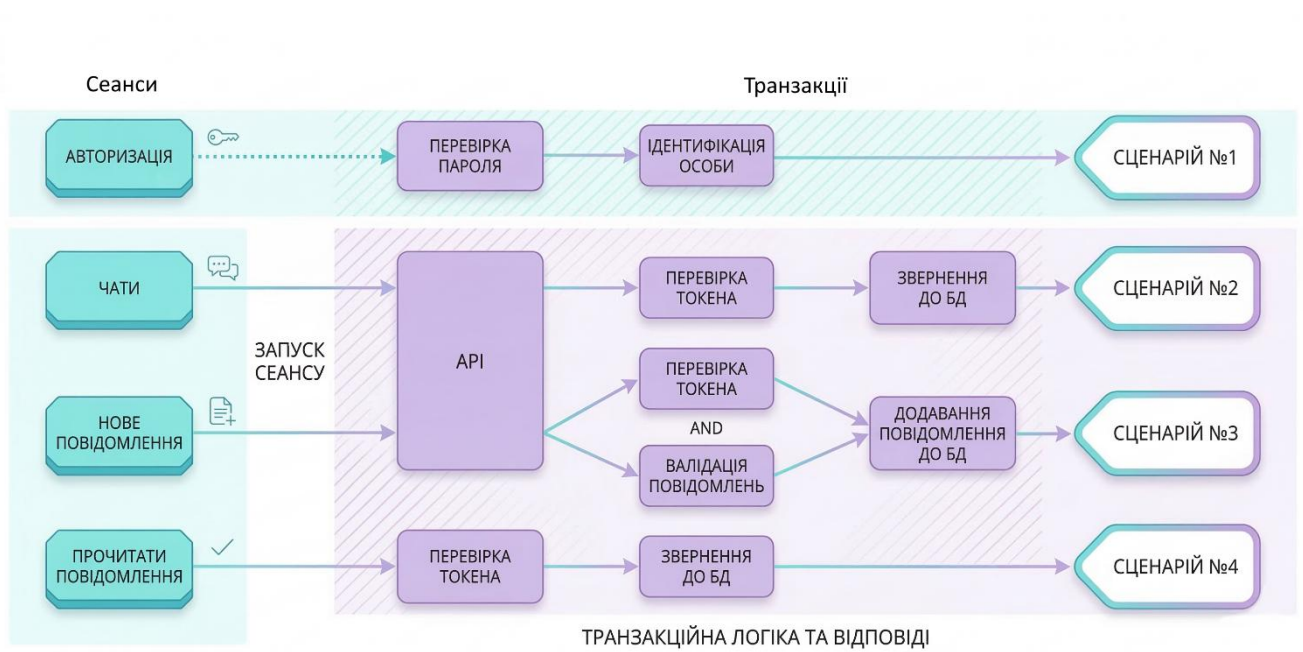


Рис. 1.1 – Функціональна схема абстрактної соціальної мережі

Для проєктування релевантної функціональної моделі та реалізації подальшого наукового пошуку обґрунтованим є вибір методології Domain-Driven

Design (DDD) [38, 78]. Ця концепція, орієнтована на пріоритетність предметної області, дозволяє сформувавши цілісну репрезентацію системи, яка базується на реальних процесах, фундаментальних поняттях та логічних правилах функціонування досліджуваного об'єкта.

Застосування DDD-підходу дозволяє трансформувати складну архітектуру в упорядковану структуру за рахунок наступних кроків:

- Логічна декомпозиція: поділ системи на автономні функціональні модулі;
- Встановлення зв'язків: чітке визначення механізмів та протоколів взаємодії між окремими блоками;
- Ідентифікація доменів: окреслення специфічного призначення кожного компонента системи.

Така систематизація не лише забезпечує глибоке розуміння архітектурних принципів роботи всього комплексу, а й створює умови для оптимізації його функціонування. У кінцевому підсумку це дозволяє досягти цільових показників якості обслуговування QoS та підвищити загальну ефективність інфокомунікаційного рішення.

## **1.2 Асинхронний режим роботи мереж**

ATM-мережа (Asynchronous Transfer Mode) – це високошвидкісна комутаційна телекомунікаційна технологія, призначена для інтегрованої передачі голосу, відео та даних з гарантованими параметрами QoS. ATM є платформою, рекомендованою для підтримки широкосмугових інтегрованих систем цифрових мереж (B-ISDN) [40, 56].

Ця технологія, розроблена для узгодженого обслуговування телекомунікаційних потоків різнорідних повідомлень, що надходять від різних джерел, таких як дані, голос і відео, в межах однієї мережі [67].

Характерні риси змішаного трафіку ATM такі:

- значні відмінності у швидкостях окремих потоків (у бітах за секунду);
- чутливість до затримок ізохронного трафіку, такого як відео та голос (порівняно з передаванням даних);

- втрати комірок під час їх проходження мережею;
- забезпечення віртуального з'єднання для комірок (пакети фіксованого розміру по 53 байти) для передавання мережею.

Коли затримка, яку відчуває комірка, стає надмірною, протокол ATM дозволяє відкинути цю комірку. Частка таких відкинутих комірок визначається параметром QoS, відомим як коефіцієнт втрат комірок (CLR). Оскільки причини та визначальні чинники цих затримок мають переважно статистичний характер, параметр затримки змінюється випадковим чином у часі. Таке статистично змінне відхилення затримки комірок позначається параметром варіація затримки комірки CDV і є одним із показників QoS.

Типове мультиплексування комірок від різних джерел і комутація викликів ATM проілюстровані на рис. 1.2.

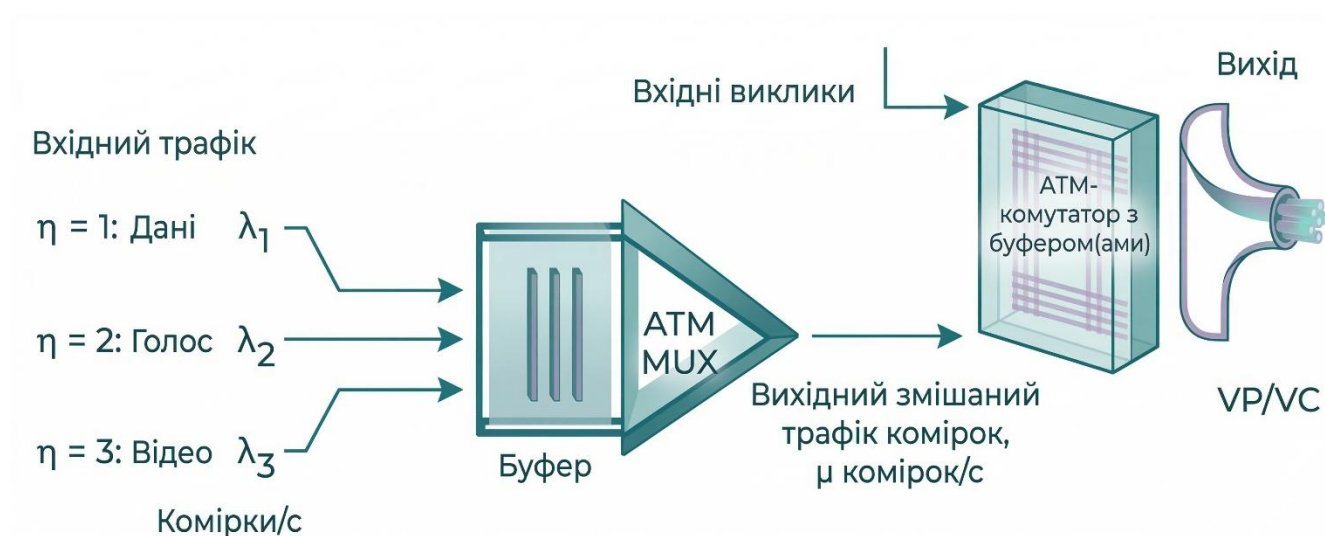


Рис. 1.2 – Мультиплексування та комутація в ATM:

$\eta = 1, 2, 3$  – вхідні комірки з різних типів джерел на стороні ініціації виклику

Окрім надмірної варіації затримки передавання комірок (CTD), що перевищує допустиму межу й призводить до втрат, існують також бітові помилки в заголовку комірки, які виявляються, але не можуть бути виправлені, що також змушує відкидати комірки. Перший тип втрат зумовлений випадковістю трафіку, другий – зміною бітового шаблону в заголовку комірки через наявність шуму (обмеженого рівнем SNR у каналі передавання). Межа CLR визначає гранично

допустиму кількість втрат комірок у каналі, яка може бути витримана без істотного погіршення прозорості інформації.

У цій роботі запропоновано об'єднати втрати комірок через CDV та втрати через шум в один параметр у площині інформаційної теорії. Відповідна верхня межа для результативної варіації CTD може бути визначена за ймовірністю «хвоста»  $A$  у розподілі ймовірностей CDV. Нехай CLR, що відповідає джитеру CDV, дорівнює CLR1, а CLR, зумовлений бітовими помилками через шум, – CLR2. Тоді загальна межа CLR має визначатися спільним впливом CLR1 і CLR2. Завдання поєднання CDV із CLR1 та SNR із CLR2 може бути розв'язане за допомогою інформаційно-теоретичних підходів.

Частина цієї роботи присвячена використанню метрики складності для розробки алгоритму SAC за умов нечіткого функціонування. Коли надходить запит на виклик, комутатор АТМ має вирішити, чи забезпечуватиметься необхідна QoS, якщо виклик буде прийнято. Якщо комутатор визначає, що необхідний QoS можна гарантувати, виклик приймається, інакше – відхиляється. Така стратегія відома як SAC. По суті, це набір правил або алгоритмів, які визначають, чи слід прийняти новий виклик [72].

Існує набір ключових параметрів QoS, що характеризують вихідний трафік і необхідні для вирішення завдань SAC. На основі цих параметрів, визначених користувачем, SAC може встановити тип необхідної послуги. Наприклад, з'єднання може бути надано на основі сервісу із доступною швидкістю (available-bit-rate service), що забезпечується мережею. У цьому випадку з'єднання отримує ту швидкість, яка є доступною на момент встановлення.

Широкосмуговий виклик [71], наприклад, може бути мультимедійним [15] і включати кілька компонентів, таких як статичне зображення, відео, голос і текст. Кожен компонент зазвичай вимагає окремого з'єднання. Надання з'єднання (що відповідає набору визначених параметрів QoS) означає присвоєння віртуального ідентифікатора каналу для тієї категорії комірок, що належить до виклику. Передача цих комірок відбувається по відповідним віртуальним каналам. Набір віртуальних каналів формує віртуальний шлях, який ідентифікується за допомогою



віртуального ідентифікатора шляху. Кожному віртуальному з'єднанню призначається пропускна здатність (смуга), відповідно до запиту користувача. Це зазвичай виконується під час процедури встановлення з'єднання за допомогою САС. У цьому процесі також визначаються параметри на кшталт максимального CTD/CDV та/або CLR, які можуть бути дозволені для конкретного з'єднання залежно від вимог користувача.

Рішення про те, чи дозволяти виклик, ухвалюється на основі характеристик трафіку та доступності ресурсів мережі для обробки додаткового навантаження без погіршення умов обслуговування існуючих викликів. Іншими словами, додаткове навантаження, спричинене новим викликом, не повинно впливати на продуктивність обслуговування вже активних з'єднань.

У літературі описано декілька алгоритмів САС [43, 54]. САС є програмною функцією у комутаторі [20, 51], яка виконується під час встановлення виклику, коли формується віртуальний канал або віртуальний шлях. Запит приймається лише в тому випадку, якщо QoS для всіх існуючих з'єднань залишиться гарантованим після його прийняття.

### **1.3. Специфікація якості обслуговування та метрика складності системи**

У даному дослідженні специфікація QoS телекомунікаційної системи (наприклад, мережі ATM) подається у вигляді єдиного параметра з урахуванням стохастичних та нечітких характеристик параметрів системи.

У відповідному моделюванні затримка кінець-у-кінець або CTD для  $i$ -ої комірки вважається зумовленою випадковими факторами, що виникають через чергування та буферизацію (мультиплексування) у межах мережі, а також рівнем шуму (у термінах SNR). Як зазначалося раніше, вважається, що параметр SNR у мережі призводить до появи бітових помилок, які формують скінченний коефіцієнт втрат комірок CLR2, тоді як буферна індукована варіація затримки комірок CDV зрештою спричиняє відкидання комірок на рівні, що визначається CLR. Таким чином, існують два явища, які зрештою формують ефективний CLR у системі. Зазвичай відкидання комірок через CDV визначається обсягом використаної

буферизації, тоді як відкидання через SNR зумовлюється переважно фізичним середовищем передавання. Наприклад, якщо фізичним середовищем є безпроводове передавання, то кількість бітових помилок буде відносно більшою, ніж у телекомунікаціях на основі волоконно-оптичних ліній.

У даній роботі запропоновано метрику складності, яка виводиться на основі принципу максимальної ентропії [27] у поєднанні з нечіткими характеристиками, пов'язаними зі стохастичними властивостями втрат комірок. Такі втрати охоплюють як вплив SNR, так і характеристики CDV. У межах отриманої метрики стає можливим розділяти наскрізну передачу мережі як просту чи складну у нечіткому просторі. Межа розподілу встановлює допустимий рівень CLR і визначає верхню межу для реалізації CAC. Відповідний аналіз подано в наступному розділі з використанням евристичного підходу до розглянутої проблеми.

Нехай  $0 \leq p(i) \leq 1$  позначає ймовірність появи  $i$ -ої комірки в АТМ-передачі. Нехай  $i = 1, 2, \dots, M$  представляє комірки, для яких оцінюється наскрізна продуктивність в АТМ-лінії. Тоді аксіоматична вимога для ймовірності полягає в тому, що:

$$\sum_{i=1}^M p(i) = 1,$$

а середнє значення

$$\sum_{i=1}^M i p(i) \leq B_0.$$

Тут  $B_0 > 0$  відображає обмежувальне значення для ансамблевого середнього, що визначається межею допустимої продуктивності трафіку. Для заданого набору параметрів, які характеризують трафік віртуального каналу, параметр ентропії [27, 133] (інформація за Шенноном) для моментів появи комірок задається наступним чином:

$$H(p) = - \sum_{i=1}^M p(i) \ln[p(i)] \text{ нат.} \quad (1.1)$$

Для оцінки ентропійних показників у системі, що охоплює  $M$  активних комірок-учасників, доцільно використовувати підхід, описаний у фаховій літературі [16, 21]. Математична модель ентропії будується на основі параметра складності  $s$ .

Для підвищення наочності розрахунків та структурування алгоритму, розділимо загальний вираз на дві складові. Спочатку визначимо середньостатистичний показник стану системи (математичне очікування)  $\bar{k}$ :

$$\bar{k} = \frac{s}{1-s} - \frac{(M+1)s^{M+1}}{1-s^{M+1}}.$$

Враховуючи отримане значення  $\bar{k}$ , підсумкове рівняння ентропії  $H(s, M)$  для випадку  $s < 1$  набуває наступного вигляду:

$$H(s, M) = \ln \left[ \frac{1-s^{M+1}}{1-s} \right] - \bar{k} \ln(s). \quad (1.2)$$

У ситуаціях, коли аналіз проводиться через обернений параметр складності  $p$  (де  $p \leq 1$ ), вираз трансформується для врахування специфіки розподілу ймовірностей у системі. Параметр  $s$  у даному контексті виступає ключовим індикатором структурної складності об'єкта дослідження [24].

Надані співвідношення стосуються максимізованого ентропійного функціонала, який ґрунтується на ймовірності наявності  $i$  неупорядкованих (з точки зору продуктивності) підмножин у просторі комірок загального розміру  $M$ . Тобто,  $p(i)$  описує розподіл ймовірностей, який максимізує ентропію  $H(p)$  неупорядкованості, пов'язаної з подіями появи комірок, і неявно відображає статистику відповідних погіршень продуктивності, таких як джитер і/або бітові помилки. Явно величина  $s$  у рівнянні (1.2) визначається як  $\exp(-b)$ , де  $b$  – це параметр Лагранжа, що використовується при максимізації ентропійного функціонала [61].

Ентропія неупорядкованих підмножин є додатною функцією, яка зростає монотонно відносно  $M$  для всіх значень  $s > 0$ . Вона також зростає монотонно зі збільшенням  $s$ , якщо  $s < 1$ , але зменшується монотонно зі збільшенням  $s$ , якщо  $s > 1$ .

Коефіцієнт  $s$  можна розглядати як міру складності, пов'язаної з АТМ-передачею, що зазнає втрат комірок унаслідок джитера затримки та SNR-залежних бітових помилок у наскрізному з'єднанні. Припустимо, що середнє значення  $B_0$  (визначене як обмеження для статистики  $M$ ) задане через  $s$  як  $B_0 = E[s, M]$ .

Для даної послідовності комірок  $s$  відображає рівень очікуваного CLR. Таким чином, якщо  $s = 0$ , то комірки, що передаються в межах з'єднання кінець-у-кінець, утворюють просту підсистему з очікуваним значенням  $E[0, M] = 0$ .

Протилежна крайня ситуація відповідає  $s \rightarrow \infty$ , у такому випадку система стає цілком складною, і  $E[-\infty; M] = M$ , що означає надмірно великі втрати інформації, оскільки втрати охоплюють усі  $M$  комірок.

Іншими словами, коли кількість неупорядкованих підмножин втрат комірок у популяції розміру  $M$  прямує до нескінченності, складність відображає весь простір джитерних комірок та/або бітових помилок. Відповідне очікуване значення  $E[s, \infty]$  може бути виведене за допомогою таких співвідношень [19].

У межі  $M \rightarrow \infty$  функції, що пов'язані з  $(M, s)$ , стають неаналітичними при  $s = 1$ . Явно це означає, що

$$E[s, M]|_{M \rightarrow \infty} = \begin{cases} \frac{s}{1-s}; & s < 1; \\ \frac{M}{2}; & s > 1; \\ \frac{p}{1-p}; & s = \frac{1}{p} \geq 1 \end{cases} \quad (1.3)$$

та  $E[s, \infty] = E[s, M]$  для малих значень  $s$  та помірно великих  $M$ .

Відповідно,

$$\begin{aligned} H[s, M]|_{M \rightarrow \infty} &= -\ln(1-s) - \frac{s \ln(s)}{1-s}; & 0 \leq s < 1; \\ &= -\ln(1-\rho) - \frac{\rho \ln(\rho)}{1-\rho}; & s = \frac{1}{\rho} \geq 1 \end{aligned} \quad (1.4 \text{ а, б})$$

і  $H(s, M) \rightarrow H(s, \infty)$  для малих значень  $s$ . Відповідні висновки, що стосуються цих алгоритмічних викладок, такі:

1. Для дуже малих рівнів CLR при  $s \ll 1$  очікуваний рівень втрат комірок майже не залежить від кількості  $M$  задіяних комірок;

2. Для дуже великих рівнів CLR при  $s \gg 1$  очікуваний рівень втрат комірок визначається числом задіяних комірок  $M$ ;

3. Характеристичне значення  $s = 1$  розділяє систему на просту або складну щодо рівня інформаційних втрат.

Проста, або мала величина CLR, при зростанні до більш високого рівня призводить до того, що продуктивність усієї системи починає визначатися як складна. Тобто малі значення втрат комірок можна розглядати як квазіавтономні (прості) підмножини, але при їх значному зростанні система стає складною з точки зору продуктивності, що оцінюється за параметрами CDV та/або SNR.

Для моделювання цього аспекту коефіцієнт складності  $s$  можна записати як функцію від  $M$ . Зокрема, поблизу  $s = 1$  нехай  $s = (1 - \varphi)$ , де  $\varphi = (A/M) + 0$  при  $M \rightarrow \infty$ , і константа  $A$  залишається сталою. Використовуючи розклад в ряд Тейлора [17, 19] для рівняння (1.4 а) при  $s = 1$ , отримаємо:

$$E[s, M] \approx \frac{M \left(1 - \frac{A}{3}\right)}{2 \left(1 - \frac{A}{2} + \frac{A^2}{6} + \dots\right)}. \quad (1.5)$$

Дане наближення дозволяє з високою точністю оцінити стан складності  $s$  для множини  $M$  елементів при заданому параметрі  $A$ . Використання такої форми запису замість складних нескінченних рядів суттєво спрощує подальший обчислювальний процес, зберігаючи при цьому необхідний рівень достовірності результатів.

Відповідна ентропія, виведена з рівняння (1.2), має вигляд:

$$H(s, M) \approx \ln(M + 1). \quad (1.6)$$

Цей результат дозволяє визначити коефіцієнт складності втрат комірок у функціональній формі такого типу:

$$s = \exp\left(-\frac{A}{M}\right) \quad (1.7)$$

при  $s < 1$  та  $A > 0$  (тобто коли система вважається простою).

Використовуючи експоненційну форму  $s$ , задану рівнянням (1.7), можна отримати такі результати при  $A > 0$ :

$$s(e^{-A/M}) = \frac{M}{A} [1 - e^{-A}] + \frac{1}{2} [1 - e^{-A}] + \vartheta\left(\frac{A}{M}\right); \quad (1.8)$$

$$E[e^{-A/M}, M] = MF(A) - G(A) + \vartheta\left(\frac{A}{M}\right); \quad (1.9)$$

$$H[e^{-A/M}, M] = \ln M - U(A) + \vartheta\left(\frac{A}{M}\right), \quad (1.10)$$

де

$$\begin{aligned} F(A) &= \frac{1}{A} - \frac{1}{e^A - 1}; \\ G(A) &= \frac{1}{2} + \frac{1}{(e^A - 1)} - \frac{Ae^A}{(e^A - 1)^2}; \\ U(A) &= AF(A) + \ln\left(\frac{1}{A}(1 - e^{-A})\right). \end{aligned}$$

У критичній точці  $s = 1$  та в її околі середнє значення продуктивності системи відображає екстенсивну властивість відносно можливої кількості пошкоджених комірок  $M$ , а схильність до пошкодження прямо пропорційна  $\ln(M)$ , тобто інформаційному вмісту  $M$  комірок.

Коли  $s \neq 1$ , експоненціальний закон може бути модифікований як  $s = \exp(A/M)$ , у такому випадку

$$E[e^{A/M}, M] = \{1 - F(A)\}M - G(A) + \vartheta\left(\frac{A}{M}\right); \quad (1.11)$$

$$H[e^{A/M}, M] = \ln M + U(A) + \vartheta\left(\frac{A}{M}\right). \quad (1.12)$$

Ці алгоритмічні міркування можна належним чином застосувати для використання параметра складності  $s$  як узгодженої міри ступеня параметра QoS (наприклад, втрат комірок) у наскрізному АТМ-з'єднанні. Відповідні деталі наведено далі.

#### 1.4 Ентропія втрат бітів в пакетах даних

Припустимо, що статистика CDV  $\delta_i$  трафіку (ідентифікованого як трафік від джерела  $\eta$ ) має гаусівський характер. Тоді функцію від  $\delta_\eta$ , а саме  $p(\delta_\eta)$ , можна записати так:

$$p(\delta_\eta) = \frac{1}{\sqrt{2\pi\sigma_{\delta_\eta}}} \exp \left[ -\frac{(\delta_\eta - \mu_{\delta_\eta})^2}{2\sigma_{\delta_\eta}^2} \right], \quad (1.13)$$

де  $\sigma_{\delta_\eta}$  та  $\mu_{\delta_\eta}$  позначають відповідно середнє значення і стандартне відхилення  $\delta_\eta$ .

Максимальна ентропія відносно гаусівської статистики визначається як:

$$(H_{max})_\eta = \frac{1}{2} \ln \left[ 2\pi e \left( \sigma_{\delta_\eta}^* \right)^2 \right], \quad (1.14)$$

де  $\sigma_{\delta_\eta}^*$  – це  $\sigma_{\delta_\eta}$ , яке нормалізоване відносно параметра часу, наприклад, загального часу передавання  $M$  комірок,  $T$  секунд.

Як було детально пояснено раніше, критичний перехід від простого до складного стану, що визначається  $M$  комірками, які надходять із швидкістю  $\lambda_\eta$  комірок/секунду, відбувається тоді, коли параметр складності  $s \rightarrow 1$ . Використовуючи рівняння (1.6), відповідна максимальна ентропія, пов'язана з передаванням  $M$  комірок із швидкістю  $\lambda_\eta$ , задається виразом:

$$(H_{max})_\eta = \ln \left[ \frac{M+1}{\lambda_\eta T} \right]. \quad (1.15)$$

Таким чином, поєднавши рівняння (1.14) та (1.5), отримуємо, що:

$$\ln \left[ \frac{M+1}{\lambda_\eta T} \right] = \ln \left[ 2\pi e \left( \sigma_{\delta_\eta}^* \right)^2 \right]^{1/2}$$

або ж

$$\frac{M+1}{\lambda_\eta} = \sqrt{2\pi e} \left( \sigma_{\delta_\eta} \right)_{max}.$$

Іншими словами, для множини з  $M$  комірок, для яких оцінюється затримка кінець-у-кінець, максимальна ентропія, пов'язана із затримкою  $\delta_\eta$ , має стандартне відхилення, яке визначається як:

$$\sigma_{\delta_\eta} = \frac{M+1}{\sqrt{2\pi e} \lambda_\eta} \text{ сек.} \quad (1.16)$$

Відповідний параметр дисперсії визначає верхню межу допустимого джитера затримки комірок, після перевищення якої мережа починає відкидати комірки, як зазначалося раніше.

### 1.5 Сукупний коефіцієнт втрат бітів через затримку та впливи відношення сигнал/шум

Беручи до уваги дві можливі причини втрат комірок, кінцева ймовірність  $A$ , яку можна вивести з рівняння (1.13), встановлює верхню межу для сукупного CLR; при цьому  $A$  можна розділити на  $A_1$  та  $A_2$ , що відповідають CLR1 та CLR2 відповідно. Тобто функціональні співвідношення, які необхідно визначити, мають вигляд:

$$\alpha_1 \Leftrightarrow CLR1 \Rightarrow CTD1; \quad (1.17 \text{ а})$$

$$\alpha_2 \Leftrightarrow CLR2 \Rightarrow CTD2, \quad (1.17 \text{ б})$$

де  $CTD1$  та  $CTD2$  є параметрами, які спільно визначають сукупну верхню межу для CTD та, відповідно, призводять до встановлення (сукупної) верхньої межі для CLR.

Проблема пов'язання  $\alpha_1$  та CLR1 полягає в тому ж самому, що й у встановленні кореляції дисперсії  $\delta$ , а саме  $(\sigma_\delta)^2$ , із CLR1. Тобто дисперсія CDV (у найгіршому випадку), яка відповідає CTD1, перевищує максимальне значення, що дорівнює фіксованій затримці  $F$  плюс  $\sigma_\delta$ , з імовірністю не більшою, ніж  $\alpha_1$ . Це визначає максимальну межу для CTD, тобто CTD1.

Аналогічно, якщо припустити, що втрати комірок через бітові помилки, зумовлені SNR (представлені як CLR2), відповідають еквівалентній затримці CTD2, тоді необхідно встановити співвідношення між обмежувальними параметрами  $\alpha_2$  (для CTD2) та CLR2.

Відповідну методику наведено в [41, 42] та узагальнено як результуючий параметр втрат комірок, що може бути визначений із CLR1 та CLR2, який зрештою встановлює межу, яку можна використати для формування критерію SAC.

Стосовно джерела  $\eta$ , припустимо, що ймовірність втрат комірок  $P_{x\eta}$  визначається як параметр погіршення наскрізної продуктивності трафіку внаслідок



асинхронного мультиплексування та/або перевантаження, що спричиняє CDV, а також  $P_{\varepsilon\eta}$ , яке позначає ймовірність бітової помилки, що призводить до відповідних стирань комірок через скінченне SNR. Стратегія реалізації необхідних співвідношень, визначених у рівнянні (1.17), базується на таких евристичках.

Параметр CLR1, що стосується втрат комірок (зумовлених CDV через мультиплексування або перевантаження в ATM-каналах), може бути виражений відповідною ймовірністю стирання комірок  $P_{x\eta}$ , яка спричиняє втрату середнього інформаційного вмісту на одиницю смуги пропускання, пов'язаної з  $M$  комірками, що формують ATM-трафік [53]. Ненульове значення  $P_{x\eta}$  можна розглядати як еквівалентну ймовірність помилки, індуковану фактором пошкодження  $C_{x\eta}$ . Тобто втрати комірок через надмірний CDV, які виникають унаслідок механізму мультиплексування чи перевантаження, можна трактувати як такі, що спричинені деякою пошкоджувальною сутністю (аналогічною до шуму), присутньою у потоці трафіку.

Таким чином, з точки зору теорії цифрових комунікацій, яка застосовується до бітових помилок, індукованих шумом,  $P_{x\eta}$  відносно  $\eta$  – го джерела ( $\eta = 1, 2, \dots, N$ ) може бути визначене експоненційним співвідношенням:

$$P_{x\eta} = k_m \exp(-C_{x\eta}),$$

де  $k_m$  – це константа, що залежить від схеми модуляції (наприклад,  $k_m = 1/2$  при частотній маніпуляції).

У цьому співвідношенні фактор пошкодження  $C_{x\eta}$  виступає як параметр стирання, що задає межу для CDV, перевищення якої визначає верхню межу. Це співвідношення є аналогом параметра SNR.

Чим більше джерел активні у певний момент часу, тим меншу швидкість обслуговування отримує кожне з них, оскільки зростає конкуренція. Загальна швидкість обслуговування, яку надає MUX, залежить від стану черги через кількість комірок, що змагаються за мультиплексорний сервіс. Відповідно до інформаційно-теоретичної моделі багатодоступної системи, розглянутої в [59], і використовуючи базові положення рівняння (1.6), можна показати, що середні

втрати інформації  $H_\eta(L_\eta)$  для  $H$ -го джерела, визначені за одиницю часу на мультиплексорі, можуть бути задані як:

$$H_\eta(L_\eta) = \ln[M_T L_\eta(m_\eta) + 1] \text{ нат/(на одиницю смуги пропускання)}, \quad (1.18)$$

де  $L_\eta(\cdot)$  позначає очікувану кількість комірок, відкинутих на мультиплексорі внаслідок затримки, спричиненої чергою та конфліктами;  $M_T$  – це загальна кількість комірок від усіх  $N$  активних джерел, що обробляються за одиницю часу.

Крім того,  $m_\eta$  відноситься до комірок у  $\eta$ -му потоковому трафіку.

Середній час затримки  $\Delta_{1\eta}$  (що представляє втрату комірок) може бути отриманий із відомої формули Літтла [143], тобто, коли

$$\begin{aligned} M_T L_\eta(m_\eta) &\gg 1; \\ \Delta_{1\eta} &\approx \frac{M_T L_\eta(m_\eta)}{\lambda_\eta} \text{ сек}, \end{aligned} \quad (1.19)$$

де індекс 1 відповідає погіршенню трафіку, визначеному через  $CLR1$ .

Очікувані втрати інформації, спричинені втратами комірок для  $\eta$ -го джерела, як зазначено у рівнянні (1.18), тотожно дорівнюють ентропії, що пов'язана з CDV. Отже,

$$\ln \left[ \frac{M_T L_\eta(m_\eta)}{\lambda_\eta} \right] \equiv \ln \left( 2\pi e \sigma_{\Delta_{1\eta}}^2 \right)^{1/2} \quad (1.20 \text{ a})$$

або

$$\sigma_{1\eta} = \frac{\Delta_{1\eta}}{(2\pi e)^{1/2}}, \quad (1.20 \text{ б})$$

яке представляє стандартне відхилення статистики, що описує еквівалентність CDV втрат комірок, пов'язаних із  $\eta$ -м джерелом. Слід зазначити, що як середнє значення  $\Delta_{1\eta}$ , так і стандартне відхилення  $\sigma_{1\eta}$  задаються на одиницю смуги пропускання магістральної передачі та визначаються за одиницю часу.

Додаткова затримка  $\sigma_{2\eta}$  виникає внаслідок втрат комірок через невірні бітові помилки, спричинені скінченним значенням  $SNR_\eta$ , що існує у  $\eta$ -му каналі від джерела ATM до MUX. Її можна явно задати через ймовірність помилки:

$$P_{E\eta} = k_m \exp(-SNR_\eta).$$

Тобто, враховуючи  $P_{E\eta}$  або його показник  $SNR_{\eta}$ , відповідний CLR, а саме CLR2, при зв'язку з CTD зумовить відповідну середню затримку  $\Delta_{2\eta}$  з дисперсією  $\sigma_{2\eta}^2$ . Параметри  $\Delta_{2\eta}$  та  $\sigma_{2\eta}$  можуть бути визначені подібною процедурою, наведеною у розділі 3.4, замінюючи  $P_{X\eta}$  на  $P_{E\eta}$  і  $C_{X\eta}$  на SNR. Знову ж таки, індекс 2 у  $\Delta$  та  $\sigma$  відповідає умовам, визначеним CLR2.

Отже, підсумовуючи, специфікація сумарного CTD для АТМ-системи щодо  $\eta$ -го джерела може бути представлена у вигляді:

1. Фіксована затримка  $F$ , що виникає через поширення сигналу, затримку, спричинену системою комутації та/або процесами, а також через фіксовані компоненти.

2. Випадкова затримка, що виникає внаслідок асинхронної передачі з буферизацією вхідних комірок з різних джерел (з різною смугою пропускання та бітовими швидкостями). У цьому випадку маємо найгірший сценарій або максимальне значення, визначене  $\Delta_{1\eta}$  та  $\sigma_{1\eta}$ .

3. Випадкова затримка, викликана втратами комірок через невірні бітові помилки, спричинені скінченням значенням SNR, що існує в АТМ-каналі. Тут максимальні значення визначаються  $\Delta_{2\eta}$  та  $\sigma_{2\eta}$ .

Сумарний ефект мультиплексування/перевантаження та невірних бітових помилок може бути описаний єдиною гауссовою статистикою [132] CDV із середнім значенням  $\Delta_{\eta} = (\Delta_{1\eta} + \Delta_{2\eta})$  та стандартним відхиленням  $\sigma'_{\eta} = \left[ (\sigma_{1\eta})^2 + (\sigma_{2\eta})^2 \right]^{(1/2)}$ .

Таким чином, можна встановити відповідну ефективну межу CLR. Ймовірність  $\alpha$  визначається верхньою межею, заданою квантилем, що дорівнює  $2/3\sigma'_{\eta}$ .

Таким чином, для заданого набору специфікацій щодо очікуваної кількості втрачених одиниць  $(L_{X\eta}; L_{E\eta})$ , ймовірності значень помилок  $(P_{X\eta}; P_{E\eta})$  швидкості передачі  $\lambda_{\eta}$  і  $m_{\eta}$  для  $\eta$ -го трафіку, гаусівська функція щільності ймовірності може

бути побудована для  $\delta_\eta$ , коли сумарно  $M_T$  комірок (від усіх джерел) надходить на вхід до MUX.

$$p(\delta_\eta) = \frac{1}{\sqrt{2\pi}\rho_\eta} \exp \left\{ - \left[ \frac{(\delta_\eta - \Delta_\eta)^2}{2\rho_\eta^2} \right] \right\}. \quad (1.21)$$

### 1.6. Нечіткі атрибути характеристик втрат бітів в пакетах

Посилаючись на попередній виклад, у діапазоні значень  $s \leq 1$ , де АТМ-система вважається простою (з огляду на те, що сумарний показник CLR є нижчим за певну верхню межу), метрика складності відносно ентропії визначається рівнянням (1.6). Звідси випливає, що

$$\Phi(s) = \exp[H(s)] = \left[ 1 - \frac{A}{\ln(s)} \right] \quad \text{при } s = 1, \text{ коли } A > 0. \quad (1.22)$$

Далі, у термінах  $M(s)$ , рівняння (1.16), (1.10) та (1.12) можуть бути записані відповідно так:

$$H(s) = \ln(M + 1) = \ln[\Phi(s)] \quad \text{при } s = 1; \quad (1.23 \text{ а})$$

$$H(s) = \ln(M) + U(A) \approx \ln[\Phi(s)] + U(A) \quad \text{при } s < 1; \quad (1.23 \text{ б})$$

$$H(s) = \ln(M) + U(A) \approx \ln[\Phi(s)] + U(A) \quad \text{при } s > 1 \quad (1.23 \text{ в})$$

з наближеннями, що враховуються при  $M \rightarrow \infty$ .

Для трьох випадків, наведених у рівнянні (1.23), можна коректно задати таке диференціальне рівняння:

$$\frac{d^2 H(s)}{ds^2} + \left[ \frac{dH(s)}{ds} \right]^2 = \frac{\Phi''(s)}{\Phi'(s)}. \quad (1.24)$$

Для проведення кількісної оцінки ентропії (або інформаційної міри за Шенноном), що корелює з показником складності  $s$ , застосовується математичний апарат, який визначає функціональну залежність  $H(s)$  у межах заданої області. Проте специфіка сучасних телекомунікаційних систем [70] вимагає впровадження методів нечіткої логіки для обробки інтервальних значень у межах нелінійних диференціальних моделей. У цьому контексті ми спираємося на принцип екстенсивності за Заде [65], що дозволяє інтегрувати нечіткі атрибути в архітектуру системи, де динаміка втрат описується через згаданий параметр  $s$ .

Процес переходу від детермінованих моделей до гнучких сценаріїв передбачає наступні кроки:

- ідентифікація ступеня належності для кожного компонента універсальної множини;
- перехід від точкових значень до інтервальних діапазонів невизначеності;
- використання лінгвістичного градування для змінних  $x$  та  $y$  за допомогою функцій належності.

На початковому етапі ми розглядаємо загальне алгебраїчне представлення рівняння (1.24), яке встановлює зв'язок між чіткими множинами  $s_i$  (незалежні змінні) та  $H_i(s)$  (залежні результати). Хоча класичний підхід дозволяє однозначно ідентифікувати елементи множини, реальні умови експлуатації мереж диктують потребу врахування факторів невизначеності. Саме тому змінні  $x$  та  $y$  трансформуються у нечіткі аналоги  $x^f$  та  $y^f$ , де індекс  $f$  сигналізує про нечіткий характер характеристик. Така градація дозволяє адекватно відображати розмах значень у межах природної мови та специфічного контексту системи.

Під час адаптації рівняння (1.24) до умов нечіткості вхідна послідовність  $x_i$  розглядається як набір інтервальних даних. Шляхом розширення формули Тейлора для інтервальних обчислень, ми отримуємо вихідний результат у формі нечіткого відображення  $y_i^f(x_i)$ . Це означає, що кожна  $i$ -та компонента результуючої множини описується відносно граничних станів усіх векторів на визначеному інтервалі  $x \in [x^L; x^H]$ :

$$y_i^f(x_i) = y_i^f(x_L) + \sum_{j=1}^{k-1} [\psi^{(j-1)} y_i^f(x_L)] \frac{x_i^j}{j!}. \quad (1.25)$$

Математична структура рівняння (1.25) базується на алгебраїчній композиції складників, кожен з яких оперує інтервальними величинами. Процес підсумовування в межах інтервальної арифметики неминуче призводить до формування так званої (ширини результату), що є прямим наслідком кумулятивного накладання окремих доданків. Характерною особливістю такої арифметики є динамічна зміна розмаху значень:

- зростання вхідних параметрів  $x_i^f$  провокує пропорційне розширення інтервального показника  $y_i^f$ ;
- зменшення значень аргументів  $x_i^f$ , навпаки, сприяє звуженню діапазону результуючої величини  $y_i^f$ .

Що стосується рівняння (1.24), то його подання в нечіткій області здійснюється за допомогою векторної форми запису. У такій інтерпретації всі змінні та коефіцієнти розглядаються як інтервали. При цьому вводиться параметр  $k$ , що визначає загальну чисельність складників з інтервальною невизначеністю, тоді як індекс вказує на порядок досліджуваної системи. Така формалізація дозволяє перейти до більш гнучкого опису динаміки процесів у нечіткому домені:

$$\frac{d^2 Y^f(x, I)}{dx^2} + \left( \frac{dY(x, I)}{dx} \right)^2 = g(x), \quad (1.26)$$

де векторна межа  $y_i^f$  визначається як  $Y^f(X_{L,H}) = Y_{L,H}^f$ . Крім того,  $I$  позначає векторний набір параметрів, значення яких є інтервалом, а  $I = [I_1, I_2, \dots, I_k]$ .

Оскільки рівняння (1.24) є диференціальним рівнянням другого порядку (нелінійним), його порядок  $n$  дорівнює 2. Явний розв'язок цього нечіткого диференціального рівняння (1.26), а саме  $y^f(x) = fF(x)$ , наведено буде далі.

Загалом, значення меж CLR, необхідних для САС, не можуть бути визначені з детермінованою точністю. Це зумовлено притаманними нечіткими атрибутами та стохастичною природою складної телекомунікаційної передачі, про що вже йшлося раніше.

Невизначеність меж CLR, отже, вимагає застосування схеми нечіткого виводу, яку комутатор АТМ може використовувати для полегшення допуску виклику, щоб забезпечене з'єднання гарантувало значення CLR, яке не перевищує верхньої межі.

У роботах Уехари та Хірої [174, 175] було запропоновано нечіткий САС для АТМ-мереж, у якому гарантія обмежується лише значенням CLR; а що стосується CDV, то відповідні наслідки компенсуються за допомогою згладжування трафіку.

У межах поточного дослідження ключовим індикатором деградації трафіку через втрати комірок CLR виступає узагальнений параметр складності  $s$ . Цей показник є інтегральним, оскільки він одночасно враховує деструктивний вплив джиттера CDV та відношення сигналу до шуму (SNR або бітові помилки) на якість передачі даних у каналах ATM. Важливою особливістю є можливість трактування параметра  $s$  як нечіткої величини, що дозволяє більш гнучко описувати стан системи.

Спираючись на значення складності, що визначається наскрізною продуктивністю ATM-каналу, було розроблено спеціалізовану схему управління допуском з'єднань САС. Головна мета цього механізму – суворе дотримання встановлених межових значень CLR. В основі запропонованого алгоритму САС лежить апарат нечіткого виводу, який використовує розрахунки складності, що базуються на інформаційно-теоретичних (ІТ) міркуваннях.

Фактично, параметр  $s$  акумулює в собі ентропійні показники (інформацію за Шенноном), які безпосередньо пов'язані з втратою інформаційних одиниць у мережі. Для практичної реалізації процедури САС на основі цього параметра я пропоную використовувати два взаємодоповнюючі аналітичні підходи:

#### **Формування бази продукційних правил (*if-then*):**

1. Цей метод [23] передбачає побудову матриці імплікацій, яка пов'язує набір дескрипторів трафіку та профілів мережевих ресурсів із відповідним класом належності за метрикою  $s$ ;
2. Для кожного індивідуального  $i$ -го виклику система генерує лінгвістичний висновок, який може частково перекриватися з суміжними значеннями;
3. На фінальному етапі застосовується математична операція дефазифікації, що дозволяє трансформувати нечіткий результат у конкретне числове значення, на основі якого приймається остаточне рішення про допуск запиту в систему.

#### **Ідентифікація розподілу у нечіткому домені:**

1. Даний підхід [60] орієнтований на встановлення верхньої межі для параметра  $s$  за допомогою спеціалізованих алгоритмів;

2. Визначений поріг слугує фундаментальним критерієм для САС у ситуаціях, коли комутатор отримує множинні запити на встановлення зв'язку одночасно;

3. Це гарантує, що будь-яке прийняте з'єднання буде функціонувати в межах заданих лімітів CLR та пов'язаного з ним показника CDV.

Детальний опис процедурних кроків, а також результати проведених імітаційних експериментів, що підтверджують ефективність обраної стратегії, я виклала у наступних підрозділах.

### **1.7. Лінгвістичний опис нечітких правил**

В основі моделювання сучасних телекомунікаційних мереж лежить розуміння того, що залежність коефіцієнта втрат комірок CLR від інтенсивності навантаження або кількості активних сесій є виражено нелінійною. Це зумовлено значними флуктуаціями в паттернах генерації трафіку (зокрема для сервісів CBR та VBR), де швидкість передачі може суттєво варіюватися в межах одного класу. Статистичні бази даних CLR, отримані шляхом багаторазових замірів, демонструють високий рівень дисперсії, що вимагає особливого підходу до визначення верхньої межі втрат для коректної роботи механізмів управління допуском САС.

Для вирішення цієї задачі в науковій літературі [161, 162] розглядаються кілька методологічних підходів [112, 119, 141], кожен з яких має свої обмеження:

**1. Штучні нейронні мережі (ШНМ):** Ефективні для виявлення загальних тенденцій, проте їхнє навчання базується на усереднених даних, що заважає адекватному відображенню нечіткого профілю реальних значень CLR.

**2. Апроксимаційні моделі (межа Чернова):** Базуються на аналізі активних та пасивних періодів генерації трафіку. Попри свою корисність, вони часто не враховують об'єм буфера комутатора та не розрізняють індивідуальні вимоги до CLR для різних з'єднань, що знижує статистичний виграш від мультиплексування.



**3. Непараметричні методи:** Враховують лише пікові та середні показники, що часто призводить до надмірної консервативності оцінок і зниження загальної ефективності САС.

Як перспективну альтернативу я пропоную паралельний підхід, заснований на апараті нечіткої логіки та використанні інтегрального параметра складності  $s$ . Цей параметр дозволяє цілісно оцінити вплив таких факторів, як переповнення буфера CDV та бітові помилки, спричинені SNR. У межах цього підходу для регулювання з'єднань використовується система нечітких правил формату *if-then*, де на основі обчисленого значення  $s$  визначається клас належності  $M(s)$ , що стає базисом для прийняття рішення САС.

Для лінгвістичного опису процедури оцінки складності наскрізного з'єднання (індекс  $i$ ) ми вводимо два ключові функціональні атрибути:

**Змінна  $X_i$ .** Характеризує властивості джерела трафіку, зокрема кількісні та якісні показники різноманітності швидкостей передачі бітів (наприклад, сукупність голосових та відеоданих у мультимедійному потоці).

**Змінна  $Z_i$ .** Відображає стан доступності та поточного розподілу мережевих ресурсів, необхідних для обслуговування даного трафіку.

Таким чином, індивідуальний профіль складності  $s_i$  формується як похідна функція від стану вхідного потоку та ресурсних можливостей:  $s_i \in (X_i, Z_i)$ . Сукупність змінних  $\{X_i, Z_i, s_i\}$  утворює єдину нечітку множину, для якої задаються лінгвістичні норми опису (що деталізовано у табл. 1.1 - 1.3). Загалом, такий алгоритм САС має бути адаптивним як до стандартних, так і до пікових режимів навантаження, забезпечуючи гнучкість у встановленні порогових значень допуску та можливість динамічного перебронювання ресурсів для максимізації мультиплексорного виграшу.

Таблиця 1.1

Лінгвістичні норми для змінної  $X_i$ 

Набір $X_i$ . Тип джерела	Лінгвістична ідентифікація
Надзвичайно змінні складові бітрейти та сплески	EX
Значно змінні складові бітрейти та сплески	LX
Помірно змінні складові бітрейти та сплески	MX
Постійні складові бітрейти	CX

Таблиця 1.2

Лінгвістичні норми для змінної  $Z_i$ 

Набір $Z_i$ . Профіль ресурсів	Лінгвістична ідентифікація
Великий обсяг буферів та згладжування трафіку плюс фізичне середовище з високим SNR	RA
Помірний обсяг буферів та згладжування трафіку плюс фізичне середовище з високим SNR	RB
Великий обсяг буферів та згладжування трафіку, але фізичне середовище з низьким SNR	RC
Помірний обсяг буферів та згладжування трафіку, але фізичне середовище з низьким SNR	RD
Малий обсяг буферів та згладжування трафіку плюс фізичне середовище з високим SNR	RE
Малий обсяг буферів та згладжування трафіку плюс фізичне середовище з низьким SNR	RF

Лінгвістичні норми для змінної  $s_i$ 

Набір $s_i$ . Тип системної складності	Лінгвістична ідентифікація
Надзвичайно проста система	ES
Дуже проста система	HS
Дещо проста система	SS
Дещо складна система	SC
Дуже складна система	HC
Надзвичайно складна система	EC

Кожна нечітка змінна множини  $X_i, Z_i, s_i$ , ідентифікована цими лінгвістичними описами, може бути віднесена до певного класу належності за допомогою функції належності  $\mu(\cdot) \in [0; 1]$ , як показано на рис. 1.3. Вибір функції належності  $\mu(\cdot)$  є довільним. На рис. 1.3 використовується трикутна функція. В літературі [96, 111] було застосовано також інші функції для представлення  $\mu(\cdot)$ .

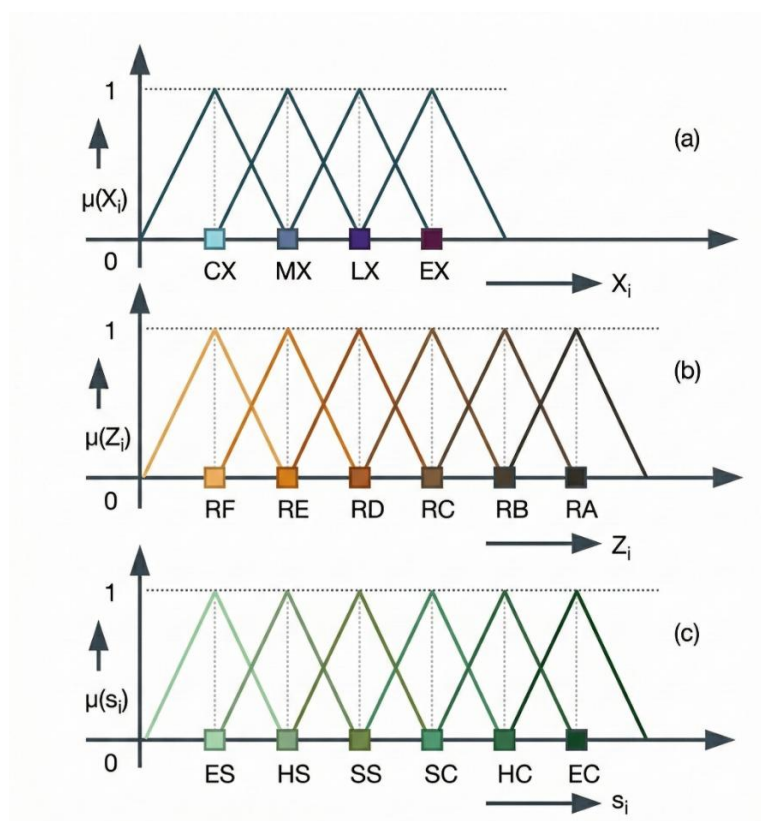


Рис. 1.3 – Графічне представлення класу належності та перекриття елементів множини  $X_i, Z_i, s_i$

З точки зору лінгвістичних описів  $X_i, Z_i$ , ідентифікованих раніше, ступінь складності  $s_i$  слід оцінювати, незважаючи на неоднозначні перекривні риси лінгвістичних норм, зображених на рис. 1.3. З цією метою слід спочатку встановити набір правил якщо-то *if-then* для чітких змінних. Основні логічні конструкції представлені наступним чином:

**Правило 1:** У сценарії, коли вхідне джерело класифікується як тип EX, а відповідний йому профіль доступних ресурсів ідентифікується як RF, результуючий показник складності системи встановлюється на рівні ES.

**Правило 2:** Якщо аналізоване джерело належить до категорії CX, а структура профілю ресурсів відповідає специфікації RA, то згідно з алгоритмом рівень складності визначається як ES.

**Правило 3:** У разі ідентифікації джерела за типом MX за умови використання ресурсного профілю моделі RD, підсумкове значення складності приймає параметр SS.

Даний набір правил є відкритим, що дозволяє масштабувати систему та впроваджувати додаткові логічні умови залежно від специфіки нових типів транзакційних послуг або зміни конфігурації мережевих ресурсів.

Правила 1–3 значною мірою є інтуїтивними, але за ними стоїть експертне обґрунтування. Наприклад, правило 1 відображає крайні випадки щодо характеристик джерела та обмежень ресурсів. За таких умов досягнення необхідного рівня продуктивності наскрізного з'єднання з урахуванням вимог QoS щодо CDV/CLR є досить обмеженим. Відповідно, пов'язана складність оголошується як надзвичайна.

Усі правила типу якщо-то, що стосуються чіткого набору  $X_i, Z_i, s_i$ , можна стисло подати у вигляді таблиці пошуку. Приблизно мовні описи, подані в таблицю на рис. 1.4, можна розділити на чітку дихотомію класів – простих і складних, за допомогою симетрично розділеної межі (подвійної товстої лінії) у таблиці. Верхня множина цієї межі є складною (при  $s_i > 1$ ), а нижня множина є простою (при  $s_i > 1$ ). Тобто для лінгвістичних елементів верхньої множини значення  $s_i$

визначаються згідно з рівнянням (1.4 а) при  $s_i < 1$ ; а для нижньої множини значення  $s_i$  відповідають рівнянню (1.4 б) при  $s_i \geq 0$ .

Тепер можна розглянути таблицю рішень, що відповідає нечітким аспектам. Тобто, процедуру побудови таблиці рішень, яка враховує перекриваючі атрибути кожного з наборів  $X_i$ ,  $Z_i$  та  $s_i$ . Відповідно до карти з таблиці на рис. 1.4, можна побудувати нечітку таблицю рішень, як показано в таблиці на рис. 1.4.

$Z_i$ $X_i$	RA	RB	RC	RD	RE	RF
EX	SC	HC	HC	EC	HC	EC
LX	HS	SS	SS	SC	SC	HC
MX	HS	SS	SS	SS	SC	HC
CX	ES	ES	HS	SS	HC	HC

Рис. 1.4 – Таблиця рішень щодо  $s_i$  для заданих  $X_i, Z_i$ : представлення у чіткому домені (товста лінія межі визначає симетричну дихотомію простих і складних сутностей; для верхньої множини виконується умова  $s_i < 1$ , а для нижньої множини –  $s_i \geq 1$ )

Процедура, використана при побудові таблиці на рис. 1.5, є наступною. Спочатку записуються лінгвістичні ідентифікації  $X_i$  та  $Z_i$  у їхніх перекриваючих формах (узгоджених з їхніми профілями належності, зображеними на рис. 1.4). Далі, для кожного перетину  $X_i$  та  $Z_i$ , відповідна лінгвістична ідентифікація складності системи береться з таблиці і позначається. Для кутових та крайових елементів, що мають лише одного сусіда, відсутній сусід позначається як нульовий з використанням зірочок.

Лінгвістичне подання відносин типу якщо–то між змінними  $X_i^f, Z_i^f$  проти  $s_i^f$ , як зазначено в таблиці на рис. 1.5, відображає перекриваючий атрибут наслідку  $s_i^f$ . Тобто, поєднання  $X_i^f$  та  $Z_i^f$  активує множину перекриваючих наслідкових імплікацій.

$Z_i^{(f)}$ $X_i^{(f)}$	*-RA-RB	RA-RB-RC	RB-RC-RD	RC-RD-RE	RD-RE-RF	RE-RF-*
* EX LX	* ● HS HC	* ● SS SC HC	* ● SS HC HC	* ● SC EC HC	* ● SC EC EC	* ● HC EC HC
EX LX MX	* ● HS SC	● ● SS HS HC	● ● SS SS HC	● ● SS SS EC	● ● SC SC HC	● ● HC SC EC
LX MX CX	* ● ES HS	● ● ES HS SS	● ● HS SS SS	● ● SS SS SC	● ● HC SC SC	● ● HC SC HC
MX CS *	* ● ES HS	● ● HS ES SS	● ● ES HS SS	● ● HC HS SS	● ● HC SS SC	● ● HC HC HC

Рис. 1.5 – Нечітка таблиця рішень щодо  $s_i^f$  за умов  $X_i^f$  та  $Z_i^f$   
(● – це дефазифіковане значення, яке потрібно визначити)

Припустимо, що множина  $(X_i^f, Z_i^f)$  позначає конкретний трафік, який вимагає допуску виклику на комутаторі. Аналогічно, існують й інші трафіки, кожен із власним універсумом дискурсу  $(X_i^f, Z_i^f)_{j,k,l,\dots}$ , які також запитують з'єднання на тому ж комутаторі. Чи буде даний виклик допущено, чи ні, вирішується шляхом визначення імплікативного вектору  $(s_i^f)_{j,k,l,\dots}$ , що стосується цього виклику.

У таблиці цей імплікативний вектор  $s_i^f$  для  $i$ -го трафіку (або виклику) розташовується в області сусідства як множина з чотирьох нечітких змінних, представлених через лінгвістичні ідентифікації. Із перекривних атрибутів  $s_i^f$  має бути оцінено змістовний показник, щоб відобразити фактичну складність  $i$ -го виклику. Це дозволить використати такий показник для прийняття рішення щодо критерію САС в умовах нечіткого дискурсу.

### 1.8 Межа допуску викликів: чіткі та нечіткі аспекти

Як зазначалося раніше, у практичних механізмах САС АТМ-комутатор перетворює запит на з'єднання, що сигналізується через інтерфейс користувач–мережа, у відповідну категоризацію послуг [89], опис трафіку та цільові параметри QoS.

Категоризація послуг визначається через параметр відношення пікової до середньої швидкості передавання бітів, пов'язаний із різними типами послуг (такими як голос, відео чи дані), що обробляються. Кожна з цих категоризованих послуг може бути позначена як окремий тип виклику з визначеними вимогами до трафіку або дескрипторами (наприклад, пікова швидкість передавання комірок) та/або цільовими параметрами QoS, такими як CLR, CTD та CDV.

На рис. 1.6 (а) показано чітку межу допуску викликів (crisp call admission boundary) для викликів, позначених індексом  $i = 1, 2, 3, \dots, K$ . Для кожного виклику CLR розглядається як параметр QoS, що підлягає контролю.

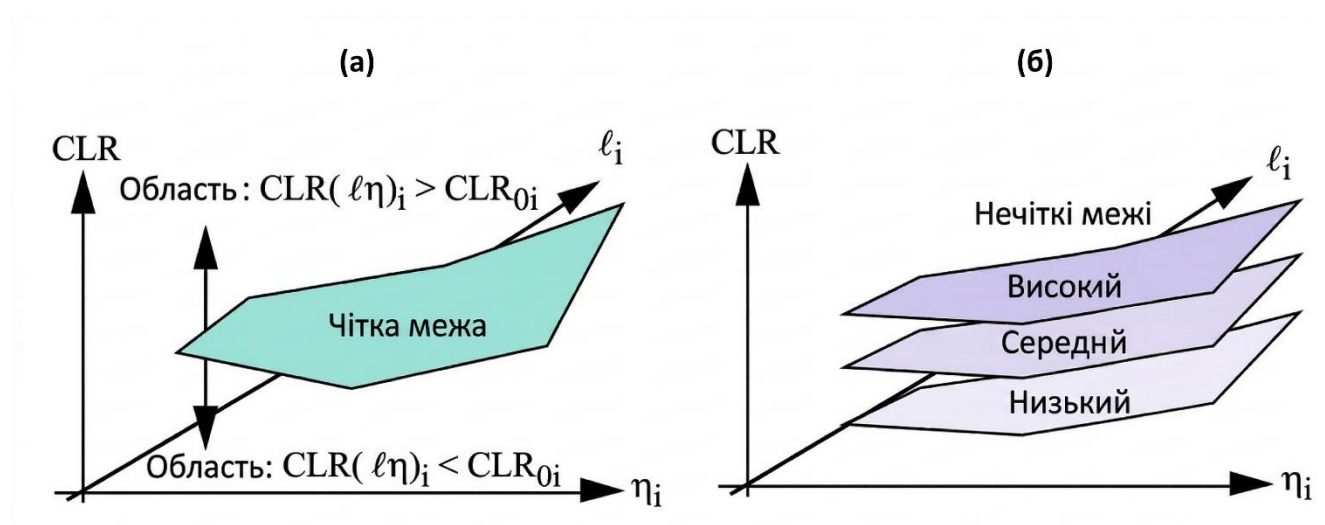


Рис. 1.6 – Поверхня межі допуску виклику для  $i$ -го виклику:

(а) – чітка межа; (б) – перекривні нечіткі межі

Припустимо, що кожен виклик охоплює три різні типи послуг, позначені як  $\eta_1, \eta_2$  та  $\eta_3$ . Виклик являє собою мультиплекс цих послуг. Далі, рівень втрат комірок для  $i$ -го трафіку з певною комбінацією  $\eta_1, \eta_2$  та  $\eta_3$  позначається як

$CLR(\ell_i) = (\ell_1\eta_1; \ell_2\eta_2; \ell_3\eta_3)$ , де  $\ell_1, \ell_2, \ell_3$  є відносними частками трафіку, що належать відповідним категоріям послуг.

Нехай потрібна цільова величина  $CLR$  задається як верхня межа  $CLR_0$ . Тоді можна поділити карту  $(\ell_1\eta_1; \ell_2\eta_2; \ell_3\eta_3)$  на дві області:

$$CLR(\ell_i) > CLR_{0i}; \quad CLR(\ell_i) < CLR_{0i}.$$

Відповідна гранична поверхня між цими двома областями є чіткою межею допуску викликів. Коли поєднання кількостей підключених категоризованих викликів перевищує цю межу, вузол комутації АТМ відхиляє запити на встановлення викликів.

З огляду на раніше викладені причини, якщо висновок щодо допуску виклику робиться на основі чіткої межі, як описано, правило прийняття рішення є жорстким і найчастіше призводить до консервативного підходу в розподілі доступних ресурсів. Тобто використання ресурсів стає неефективним і економічно невиправданим.

Крім того, спираючись на попередні обговорення, доцільніше розглядати межу допуску викликів як нечітку, із лінгвістичними атрибутами високий, низький або середній у контексті прийняття рішення щодо допуску виклику з бажаним значенням  $CLR_{0i}$ . Іншими словами, рис. 1.6 (а) можна переробити так, щоб відобразити нечіткість режимів прийняти/відхилити, як показано на рис. 1.6 (б).

Таким чином, для заданої комбінації  $(\ell_1\eta_1; \ell_2\eta_2; \ell_3\eta_3)$ , що формує  $i$ -й виклик, асоціюється нечітке відношення втрат комірок:

$$(f)CLR(\ell_1) = (f)(\ell_1\eta_1; \ell_2\eta_2; \ell_3\eta_3),$$

де верхній індекс  $f$  явно позначає нечіткий атрибут. Клас належності  $CLR(l_i)$  до високого, середнього або низького режиму визначається функцією належності:  $\mu^{(f)}CLR(l_i) = [1; 0]$ .

У цьому дослідженні [169] критерій допуску викликів задається подібною нечіткою процедурою, але з певною модифікацією. Межа допуску викликів будується у термінах параметра складності  $s_i$  (замість CLR або будь-яких інших об'єктивних параметрів, таких як CTD чи CDV). Причина (як зазначалося раніше)



полягає в тому, що параметр складності  $s_i$  виводиться з міркувань теорії інформації та цілісно враховує втрати інформації, яких зазнає певний набір  $(\ell_1\eta_1; \ell_2\eta_2; \ell_3\eta_3)$  у результаті глобальних погіршень трафіку, викликаних втратами комірок через CDV та/або SNR.

Таким чином, на рис. 1.6 (б), який ілюструє нечіткі межі допуску викликів, їх можна модифікувати для позначення перекривних меж складності системи – від простої до складної – і визначити через лінгвістичні ідентифікатори від ES до EC. Отже, відповідний нечіткий висновок дає змогу ухвалювати рішення щодо допуску викликів без надмірного або недостатнього використання доступних ресурсів. Тому погіршення, спричинене втратою інформації через CDV та/або SNR, враховується через параметр складності  $s$ .

## 1.9 Опис об'єкта дослідження

У сучасному нормативному полі послуга розглядається як специфічний інструмент генерування цінності для споживача, що сприяє досягненню конкретних результатів без необхідності прямого управління витратами чи ризиками. З технічної ж точки зору, сервіс часто трактують як програмний модуль, що функціонує за принципом «запит-відповідь» і має чітко визначені параметри QoS.

У телекомунікаційній галузі сервіси [94, 97] трансформувалися у багаторівневі екосистеми, що інтегрують інформаційні та комунікаційні технології. Їхня основна функція – створення середовища для безперешкодного обміну даними та доступу до цифрових ресурсів [45]. Сюди належать:

- Традиційний зв'язок [76, 109]: мобільні та фіксовані мережі;
- Хмарні рішення [91, 115, 116]: віддалені обчислення та зберігання даних;
- Комунікаційні платформи [92, 93]: відео зв'язок, месенджери, електронна пошта;
- Інфраструктурні послуги [95]: забезпечення доступу до мережі Інтернет.

Глибокий аналіз таких сервісів дозволяє не лише оцінити їхню поточну продуктивність і надійність, а й визначити потенціал адаптивності системи до мінливих запитів користувачів і критичних навантажень.

Ключове завдання цього дослідження – розробка спеціалізованого алгоритму для оцінювання ймовірнісно-часових характеристик ( $P - T$  показників) складних, комбінованих послуг. Мова йде про поєднання сеансових та транзакційних моделей у системах із слабким зв'язком між архітектурними елементами, що є критично важливим для сучасних розподілених мереж [118].

Структуру дисертаційного дослідження можна представити, як на рис. 1.7.

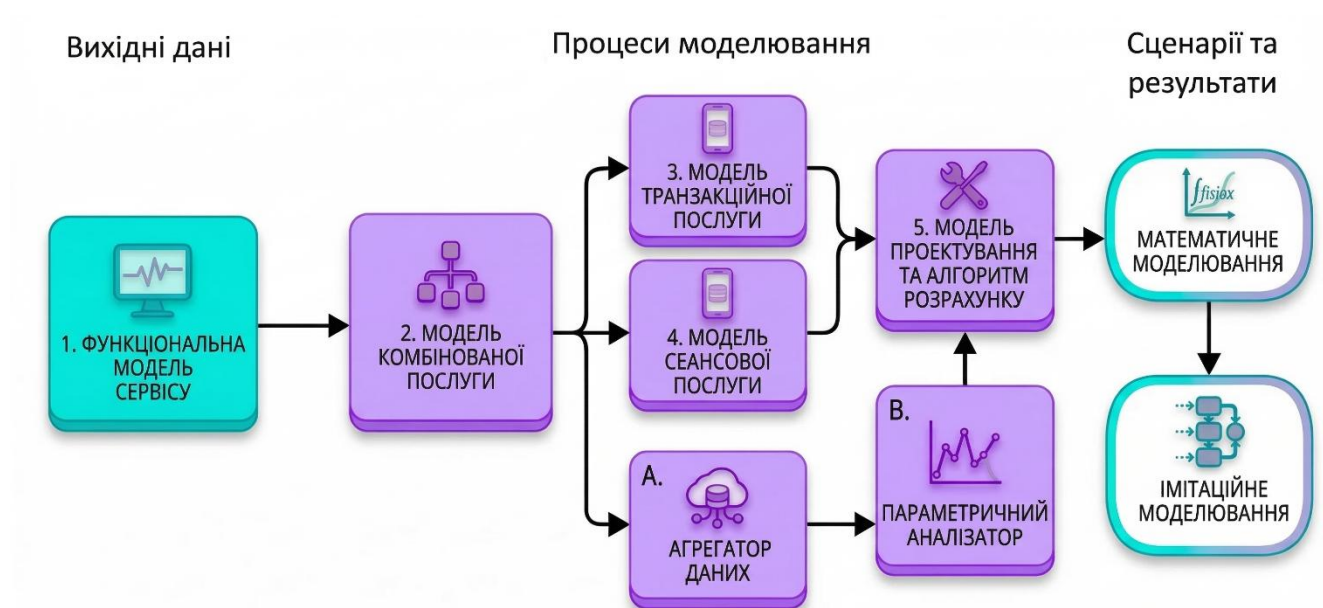


Рис. 1.7 – Структура дисертаційної роботи

Для досягнення поставленої мети в роботі передбачається виконання таких основних завдань:

1. Провести аналіз сучасного стану методів управління телекомунікаційними мережами, моделей розподілу мережевих ресурсів і підходів до маршрутизації трафіку, а також визначити їхні переваги, недоліки та наявні обмеження в умовах змінного трафікового навантаження.

2. Розробити математичну модель телекомунікаційної мережі, яка дасть змогу описати процеси передавання трафіку, розподілу мережевих ресурсів і оцінювання показників якості обслуговування з урахуванням стохастичного характеру трафіку.

3. Удосконалити математичний апарат оцінювання стану телекомунікаційної мережі шляхом урахування багатокритеріальних показників якості обслуговування, зокрема затримки передавання, рівня завантаження каналів і ймовірності втрат пакетів.

4. Розробити метод адаптивного управління мережевими ресурсами телекомунікаційної мережі, який забезпечуватиме ефективний розподіл пропускної здатності каналів зв'язку з урахуванням поточного стану мережі.

5. Розробити метод багатокритеріальної оптимальної маршрутизації трафіку, спрямований на балансування навантаження, підвищення ефективності використання мережеских ресурсів і покращення показників якості обслуговування.

6. Провести ІМ та експериментальне дослідження запропонованих моделей і методів, а також оцінити практичну ефективність запропонованих підходів і визначити можливості їх використання в сучасних телекомунікаційних мережах. Окремо передбачається розроблення алгоритмів і програмних засобів для реалізації запропонованих методів управління телекомунікаційними мережами.

Відповідно до наведеної схеми та логіки дослідження, у роботі доцільно застосувати комбінований підхід, який поєднує математичні та імітаційні методи моделювання на всіх основних етапах. Використання такого підходу дає можливість більш повно та точно оцінити часові характеристики телекомунікаційних послуг, враховуючи різноманіття архітектурних рішень і особливості функціонування мережі.

### **Висновки до розділу 1**

Сучасні телекомунікаційні послуги висувають високі вимоги, що зумовлено стрімким зростанням функціональних можливостей у розробці сервісів і веб-продуктів. Це обумовило необхідність створення високоякісних послуг QoS, що вимагає використання методів моделювання для прогнозування та забезпечення готового продукту для користувачів.

У цьому розділі отримані такі результати:

1. Проведено аналіз методів і підходів до моделювання слабозв'язаних телекомунікаційних послуг. Визначено математичні інструменти теорії масового обслуговування, які будуть використані в подальшому аналізі.

2. Запропоновано класифікацію телекомунікаційних послуг, яка допомагає у виборі відповідних методів моделювання.

3. Встановлено, що мікросервісна архітектура (МСА) є оптимальним рішенням для побудови сучасних високонавантажених сервісів. Особливості цієї архітектури будуть використані для розробки моделей, а також введено поняття робочого процесу обробки запитів у слабозв'язаних архітектурних елементах.

4. Розгляд оцінки ймовірісно-часових характеристик систем і використання запропонованих методів моделювання визначено ключовими елементами цієї дисертації.

Наступні етапи роботи будуть спрямовані на розробку моделей і методик їх застосування для оцінки ймовірісно-часових характеристик слабозв'язаних послуг.

## РОЗДІЛ 2

### МОДЕЛІ І МЕТОДИ МОДЕЛЮВАННЯ ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ

#### 2.1. Методи моделювання телекомунікаційних систем

На етапах проєктування, розроблення та подальшого вдосконалення телекомунікаційних систем моделювання відіграє одну з ключових ролей [166]. Його застосування дає змогу не лише попередньо оцінити функціональні можливості майбутньої системи, а й виявити потенційні проблеми ще до впровадження реального рішення. Завдяки цьому зменшуються часові витрати на створення системи, підвищується обґрунтованість технічних рішень і забезпечується краща якість кінцевого продукту.

У загальному вигляді вплив змін входних параметрів на узагальнений показник функціонування системи може бути поданий так:

$$\Delta Y = \sum_{i=1}^n w_i \Delta x_i, \quad (2.1)$$

де  $\Delta Y$  – відносна зміна узагальненого показника функціонування системи;  $\Delta x_i$  – відносна зміна  $i$ -го входного параметра;  $W_i$  – ваговий коефіцієнт, що відображає ступінь впливу відповідного параметра на підсумковий результат.

Як вихідний показник можуть розглядатися продуктивність, надійність, доступність або інші характеристики системи, а до входних змінних можуть належати обсяг ресурсів, інтенсивність обробки запитів, затримка реакції чи інші параметри функціонування.

Для дослідження телекомунікаційних послуг і мережевих процесів використовується сукупність різних методів моделювання. Вибір конкретного підходу визначається метою дослідження, необхідною точністю результатів, складністю системи та наявними обчислювальними ресурсами. Найбільш поширеними є такі методи:

**Математичне моделювання.** Цей підхід [157] ґрунтується на побудові аналітичного опису процесів, що відбуваються в телекомунікаційній системі. За

допомогою математичних рівнянь можна досліджувати закономірності проходження трафіку, оцінювати пропускну здатність мережі, визначати показники надійності та аналізувати залежність параметрів якості обслуговування від навантаження. Перевагою такого методу є можливість отримання формалізованих залежностей і проведення теоретичного аналізу без створення фізичного прототипу.

**Імітаційне моделювання.** Імітаційні моделі [81, 158] дозволяють відтворювати поведінку телекомунікаційної системи в умовах, максимально наближених до реальної експлуатації. За їх допомогою можна досліджувати функціонування мережі при різних сценаріях навантаження, тестувати реакцію системи на зміни параметрів, аналізувати продуктивність окремих вузлів і перевіряти ефективність керуючих рішень. Такий підхід є особливо корисним у випадках, коли аналітичний опис системи є надто складним або спрощення можуть призвести до втрати суттєвих особливостей її поведінки.

**Аналіз даних.** В умовах сучасних телекомунікаційних мереж формується значний обсяг статистичної та експлуатаційної інформації [55, 66], яка може бути використана для дослідження стану системи. Аналіз даних дає можливість виявляти закономірності у споживанні послуг, визначати характер зміни навантаження, знаходити типові сценарії роботи мережі та формувати основу для подальшого прогнозування й оптимізації.

**Методи машинного навчання.** Ці методи [82, 106] дедалі активніше застосовуються для побудови прогнозних моделей у телекомунікаційній сфері. На основі накопичених даних вони дають змогу передбачати зміну попиту на послуги, виявляти приховані тенденції в роботі мережі, прогнозувати перевантаження окремих сегментів і формувати керуючі впливи для підвищення якості обслуговування користувачів.

**Експертні методи.** У низці випадків для оцінювання стану мережі та вибору оптимальних рішень доцільно залучати знання фахівців [39, 110, 113] предметної області. Експертні підходи особливо корисні тоді, коли формалізація всіх процесів

є ускладненою, а також під час прийняття рішень щодо проєктування, модернізації чи адаптації телекомунікаційної системи до нових умов функціонування.

Слід зазначити, що на практиці окремі методи не часто використовуються ізольовано. Значно частіше ефективний результат забезпечує поєднання кількох підходів (рис. 2.1), оскільки кожен із них має власні сильні сторони. Математичні моделі забезпечують формальний опис і виявлення основних закономірностей, тоді як імітаційні моделі дають змогу перевірити ці закономірності в умовах, близьких до реальної експлуатації мережі. Саме тому для оцінювання ймовірно-часових характеристик систем доцільно застосовувати гібридну імітаційно-аналітичну модель, яка об'єднує переваги обох підходів.

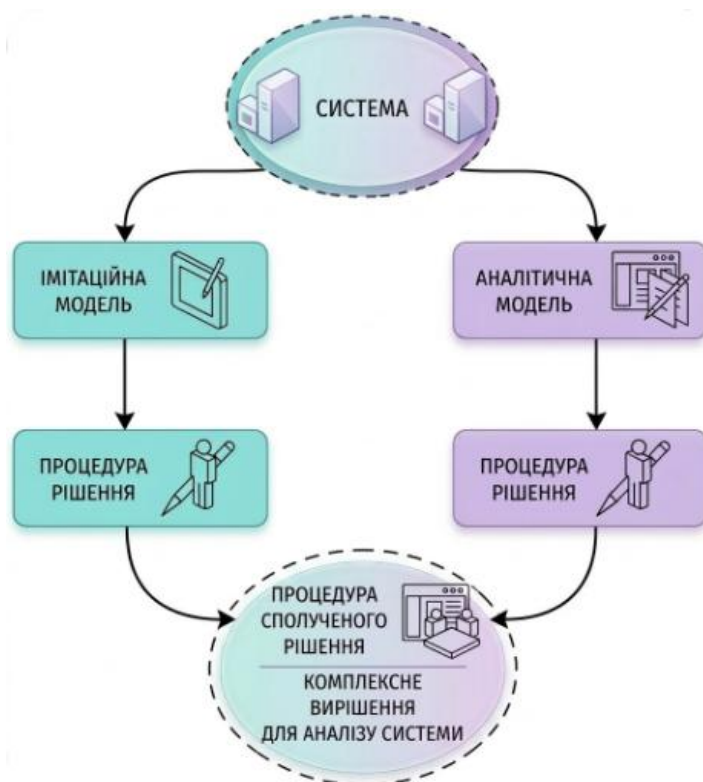


Рис. 2.1 – Комбінована модель, що поєднує імітаційний та аналітичний підходи

Окремої уваги заслуговує підхід Model-Driven Architecture (MDA) [40, 105], або модельно-орієнтована архітектура. Його сутність полягає у використанні абстрактних моделей як основи для подальшої реалізації системи. У межах цього підходу первинно формується модель предметної області, після чого на її основі можуть бути згенеровані програмні компоненти, конфігурації або інші елементи

реалізації. Такий спосіб організації розроблення підвищує узгодженість між етапами аналізу, проєктування та програмної реалізації.

У сучасному інформаційному середовищі підхід MDA набуває особливої актуальності, оскільки орієнтація на типові рішення, шаблони та повторно використовувані моделі сприяє скороченню строків розроблення й підвищує рівень автоматизації процесів. Для складних телекомунікаційних систем це є суттєвою перевагою, адже дозволяє знизити трудомісткість створення нових рішень і водночас підвищити їхню структурну впорядкованість.

Таким чином, моделювання телекомунікаційних систем доцільно розглядати як багаторівневий інструмент дослідження, що охоплює аналітичний опис, імітацію реальної поведінки, обробку емпіричних даних, інтелектуальне прогнозування та експертну підтримку прийняття рішень. У межах даної роботи найбільш обґрунтованим є застосування комбінованого підходу, який поєднує математичний апарат, бібліотеку імітаційних моделей та принципи модельно-орієнтованого проєктування для оцінювання ймовірно-часових характеристик (ЙЧХ) телекомунікаційної системи.

## **2.2. Математичний апарат для моделювання інформаційно-комунікаційних систем**

Інформаційно-комунікаційні системи, у межах яких одночасно обробляється значна кількість різнорідних запитів, доцільно описувати засобами теорії СМО [75, 131]. Такий підхід дає змогу подати телекомунікаційні послуги як сукупність взаємопов'язаних вузлів, між якими циркулюють потоки заявок. Для формального опису подібних структур використовують положення теорії телетрафіка, класичні моделі СМО та апарат мереж Джексона [137, 140]. У найзагальнішому випадку система розглядається як відкрита мережа обслуговування зі зворотними переходами між окремими елементами (рис. 2.2).



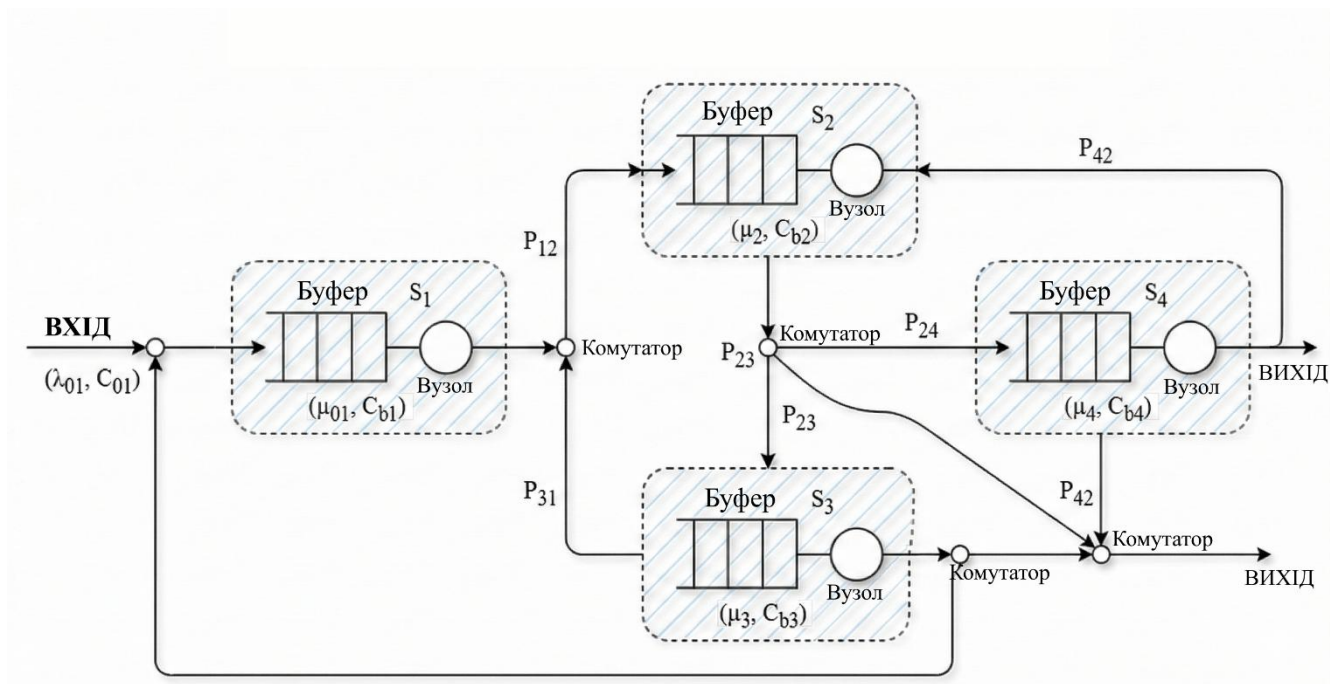


Рис. 2.2 – Приклад моделі відкритої СМО

Під час побудови такої моделі враховують набір параметрів, що визначають як поведінку окремих вузлів, так і зв'язки між ними. Зокрема,  $\mu_i$  задає інтенсивність обслуговування в  $i$ -му вузлі,  $\lambda_i$  характеризує інтенсивність потоку, що передається з вузла  $i$  до вузла  $j$ ,  $C_{ij}$  описує варіацію процесу надходження між вузлами, а  $C_{bi}$  – варіацію процесу обслуговування у вузлі  $i$ . Додатково враховують загальну кількість вузлів у системі  $Y$  та матрицю маршрутних імовірностей  $R = (p_{ij})$ , у якій  $p_{ij}$  визначає ймовірність переходу заявки від вузла  $i$  до вузла  $j, i = 1 \dots Y; j = 1 \dots Y$ .

У спрощеному варіанті вважають, що переміщення запиту між елементами мережі не потребує часу. Якщо ж затримка передавання між вузлами є суттєвою, до моделі додають окремий обслуговуючий елемент типу  $G/G/\infty$ , який відтворює процес транспортування сигналу або повідомлення. Це дозволяє точніше подати часову структуру обміну даними між комутаційними та обчислювальними компонентами.

Щоб зменшити складність дослідження, послідовність взаємопов'язаних операцій можна перетворити на еквівалентний укрупнений вузол із власним законом обслуговування. Такий підхід спрощує первинний аналіз усієї системи, а

далі дає змогу деталізувати модель, розбиваючи її на менші підсистеми без втрати загальної логіки функціонування. Фактично кожен елемент у подальшому може розглядатися за принципом «чорної скриньки», коли важливими є його вхідні та вихідні характеристики, а не внутрішня реалізація.

Після такого перетворення окремий вузол мережі можна подати як автономну СМО з параметрами надходження, обслуговування, черги та керування (рис. 2.3).

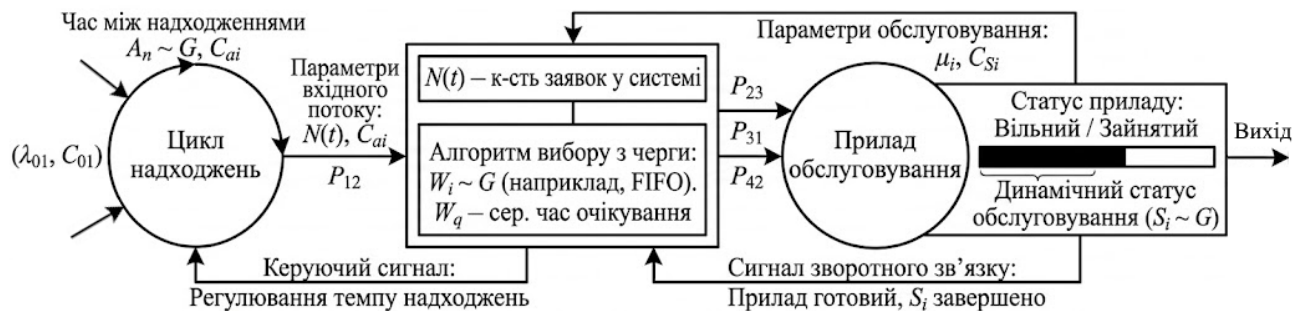


Рис. 2.3 – Приклад комплексної відкритої мережі СМО (Схема G/G/1)

У розглядуваній моделі інтервали між надходженнями запитів описуються довільним законом розподілу  $G$  [142]. Після прибуття заявка, за наявності вільного обслуговуючого пристрою, негайно переходить до обслуговування. Якщо ж пристрій зайнятий, запит накопичується в буфері очікування. Основними характеристиками вузла в такій моделі є наступними.

Сама  $\lambda_i$  визначає інтенсивність вхідного потоку до вузла  $i$ ,  $C_{ai}$  характеризує коефіцієнт варіації процесу надходження,  $\mu_i$  задає інтенсивність обслуговування у вузлі  $i$ , причому  $t_i = 1 / \lambda_i$  відповідає середньому часу обслуговування заявки, а  $C_{si}$  відображає коефіцієнт варіації тривалості обслуговування.

Щоб мережа функціонувала у стаціонарному режимі, для кожного вузла має виконуватися умова завантаження:

$$\rho_i = \lambda_i / \mu_i < 1. \quad (2.2)$$

Ця умова є базовою для подальшого аналізу, оскільки перевищення граничного рівня завантаження призводить до необмеженого зростання черг і погіршення показників якості обслуговування.

Отримання точних аналітичних рішень для мереж загального вигляду пов'язане зі значними труднощами, тому на практиці широко застосовують наближені методи оцінювання продуктивності. До найбільш уживаних підходів належать дифузійні наближення [150], метод середніх значень [151], оперативний аналіз [154, 155] і методи декомпозиції [153].

Дифузійні моделі дають змогу описувати випадкові коливання стану системи в часі та аналізувати стохастичну динаміку її функціонування. Метод середніх значень орієнтований на визначення усереднених показників, зокрема середнього часу очікування, середньої тривалості обслуговування та середньої кількості заявок у системі. Оперативний аналіз ґрунтується на використанні поточних даних про стан мережі для швидкого оцінювання її продуктивності. Методи декомпозиції базуються на поділі складної багатовузлової системи на простіші елементи, для яких можливо побудувати зручні аналітичні апроксимації.

У межах декомпозиційного підходу мережу подають як сукупність взаємодіючих вузлів, а її аналіз виконують поетапно. Спочатку досліджують передавання потоку між сусідніми вузлами, після чого визначають параметри потоку, що надходить до наступного елемента. Для послідовного ланцюга систем інтенсивність потоку на вході наступного вузла може бути подана співвідношенням

$$\lambda_{i+1} = \lambda_i. \quad (2.3)$$

Квадрат коефіцієнта варіації вхідного потоку для наступного вузла визначають наближено так:

$$C_{a_{i+1}}^2 \approx 1 + \rho_i^2 (C_{s_i}^2 - 1) + (1 - \rho_i^2) (C_{a_i}^2 - 1). \quad (2.4)$$

У наведеному співвідношенні слід урахувати такі положення:

- $C_{a_i}^2$  є квадратом коефіцієнта варіації (SCV) вхідного потоку попереднього вузла;
- $\rho_i$  характеризує коефіцієнт завантаження вузла;
- за малого навантаження, коли  $\rho \rightarrow 0$ , вираз переходить до співвідношення  $C_d^2 \rightarrow C_{a_i}^2$ ;

- це означає, що вплив процесу обслуговування на структуру потоку поступово зменшується.

Окремим випадком декомпозиції є розщеплення потоку, коли загальний потік заявок поділяється на кілька підпотоків [10, 163]. Такий підхід доцільний у системах, де різні категорії трафіку обслуговуються за окремими правилами або маршрутизуються різними гілками мережі. У цьому разі інтенсивність вихідного потоку визначають як

$$\lambda_j = \lambda_i p_{ij}. \quad (2.5)$$

Квадрат коефіцієнта варіації вихідного потоку при цьому можна оцінити співвідношенням

$$C_j^2 \approx 1 + p_{ij}(C_i^2 - 1). \quad (2.6)$$

Зворотною операцією є об'єднання кількох потоків в один сумарний потік. Така ситуація має місце тоді, коли декілька джерел навантаження використовують спільний канал передавання або спільний ресурс обслуговування. Для цього випадку сумарна інтенсивність потоку визначається як

$$\lambda_{out} = \sum_{i=1}^N \lambda_i. \quad (2.7)$$

Квадрат коефіцієнта варіації результуючого потоку можна оцінити виразом

$$C_{out}^2 \approx 1 + \sum_{i=1}^N \left( \frac{\lambda_i}{\lambda_{out}} \right)^2 (C_i^2 - 1). \quad (2.8)$$

Таким чином, операції поділу та об'єднання потоків становлять основу моделювання маршрутизації запитів у телекомунікаційних мережах і дозволяють оцінювати розподіл навантаження [5, 170] між окремими ресурсами системи.

Практична реалізація методу декомпозиції [3], як правило, передбачає поділ мережі на підсистеми, окремий аналіз кожної з них, апроксимацію необоротних процесів стаціонарними еквівалентами, визначення часових характеристик та подальше спрощення загальної задачі до набору базових операцій. Завдяки цьому вдається отримувати прийнятні оцінки продуктивності за рахунок комутації пакетів, навіть для складних багатовузлових структур.

Для аналізу часових характеристик системи використовують функцію розподілу випадкової величини  $T$ , яка задається співвідношенням:

$$F(t) = P(T \leq t), \quad (2.9)$$

де  $k$ -й момент цієї випадкової величини визначають як:

$$E[T^k] = k \int_0^{\infty} t^{k-1} (1 - F(t)) dt. \quad (2.10)$$

На основі отриманих моментів може бути обчислений коефіцієнт варіації ( $c$ ), який дозволяє оцінити ступінь нерівномірності часових процесів у системи:

$$c = \frac{\sqrt{E[T^2] - (E[T])^2}}{E[T]}. \quad (2.11)$$

Одержані співвідношення використовують для оцінювання інтегральних показників функціонування мережі. Одним із ключових є середній час реакції системи  $\bar{T}$ , який складається із затримки в черзі та часу безпосереднього обслуговування:

$$\bar{T} = \bar{W} + \overline{t_{\text{обсл}}}, \quad (2.12)$$

де  $\bar{W}$  – середня затримка в черзі;  $\overline{t_{\text{обсл}}}$  – середній час обслуговування.

Ефективність системи та часові показники безпосередньо залежать від її завантаженості  $\rho$ . Для одноканальних систем інтенсивність використання ресурсу визначається як:

$$\rho = \lambda \cdot \overline{t_{\text{обсл}}}, \quad (2.13)$$

де  $\lambda$  – інтенсивність вхідного потоку заявок.

Для стабільної роботи системи необхідно дотримуватися умови  $\rho < 1$ .

На початковому етапі класифікацію підсистем доцільно подати у вигляді таких положень:

- Усі підсистеми спочатку розглядають як такі, що належать до загального класу  $G/G/v$  [146] (довільний вхідний потік, довільне обслуговування,  $v$  каналів).
- Надалі, залежно від типу вхідного потоку та характеру обслуговування, застосовують уточнену класифікацію Кендалла [145]:

- $M$  (Markovian) – відповідає Марківським процесам [156] (експоненціальний розподіл);

- $D$  (Deterministic) – позначає детерміновані (регулярні) процеси [152];
- $H_k$  (Hyperexponential) – відповідає гіперекспоненціальному розподілу  $k$ -го порядку [160].

Для комплексного аналізу будь-якої системи застосовується закон Літгла:  $L = \lambda \cdot \bar{T}$ . Проте найбільш критичним для інженерного аналізу є розуміння нелінійної природи затримок.

Графічна інтерпретація завантаженості.

Якщо розглянути модель  $M/M/1$  [145], характеристики черги виражаються через завантаженість  $\rho$ :

- Середня кількість заявок у черзі ( $\bar{Q}$ ):

$$\bar{Q} = \frac{\rho^2}{1 - \rho}.$$

- Середній час очікування  $\bar{W}$ :

$$\bar{W} = \frac{\rho \cdot \overline{t_{\text{обсл}}}}{1 - \rho}.$$

На побудованому графіку промодельована залежність  $\bar{W}$  від  $\rho$ , де можна виділити три зони (рис. 2.4):

1. Зона низького навантаження ( $\rho < 0.5$ ): Затримки зростають лінійно і майже непомітно.
2. Зона насичення ( $0.5 < \rho < 0.8$ ): Починається помітне зростання черги.
3. Зона деградації ( $\rho > 0.8$ ): Час очікування зростає катастрофічно (асимптотично наближаючись до нескінченності при  $\rho \rightarrow 1$ ).

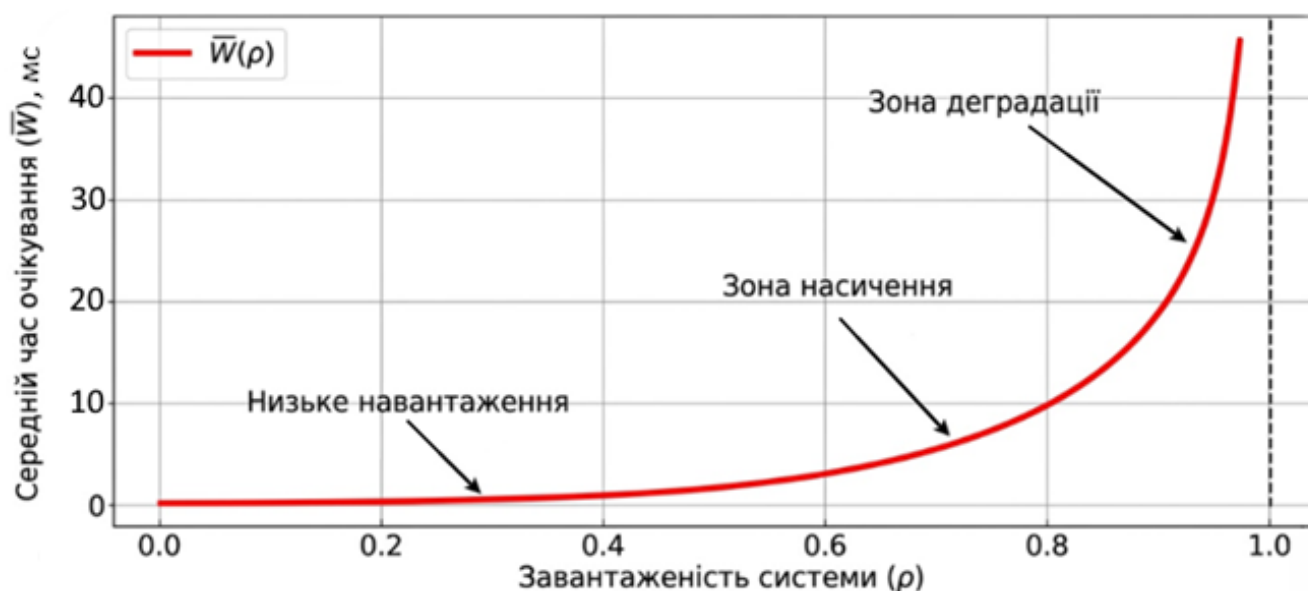


Рис. 2.4 – Залежність середнього часу очікування  $\bar{W}$  від завантаженості системи  $\rho$  (модель  $M/M/1$ )

Отже, математичний апарат СМО дає змогу не лише формально описати структуру інформаційно-комунікаційної системи, а й визначити параметри, необхідні для подальшого прогнозування її функціонування. Саме тому перед вибором методики дослідження доцільно класифікувати вид послуги, встановити характеристики потоків запитів, параметри обслуговування та особливості взаємодії між елементами мережі.

### 2.3 Класифікація послуг і сервісів за рівнем взаємодії з користувачем

Класифікація послуг дає змогу обрати відповідну стратегію моделювання для визначення якісних характеристик залежно від їх призначення, типу інформації, технологій, архітектурних рішень тощо. В табл. 2.1 наведено основні категорії поділу.

## Сервіси та їх характеристики

Критерій класифікації	Категорія сервісів	Характеристика	Приклади
<b>За якістю обслуговування (QoS)</b>	(Critical Services) Критично важливі послуги	Потребують найвищої якості обслуговування та максимальної доступності	Екстрені виклики, системи керування трафіком і безпекою, медичні сервіси у реальному часі
	(Real-time Interactive Services) Часокритичні послуги	Вимагають мінімальних затримок і високої доступності	Голосові виклики, онлайн-ігри, відеоконференції,
	(Interactive Services) Інтерактивні послуги	Забезпечують надійну передачу даних із достатньою швидкістю	Мультимедіа, відеострімінг, онлайн-чати
	(Transactional Services) Транзакційні послуги	Орієнтовані на продуктивність і надійність, а не на затримки	Електронна комерція, онлайн-банкінг, цифрові транзакції
	(Bulk Transfer Services) Масові послуги	Допускають значні затримки, мають низькі вимоги до QoS	Передача великих файлів, пакетний обмін даними
<b>За функціональним призначенням</b>	Комунікаційні сервіси	Забезпечують передачу голосових, текстових і відеоданих	Мобільний зв'язок, VoIP, електронна пошта
	Послуги безпеки	Спрямовані на захист даних і систем	Антивірусні сервіси, системи кібербезпеки
	Інтернет-сервіси	Забезпечують доступ до онлайн-контенту та сервісів	Пошукові системи, стрімінг медіа, соціальні мережі, e-commerce



Критерій класифікації	Категорія сервісів	Характеристика	Приклади
	Аналітичні та BI-сервіси	Виконують аналіз і обробку великих обсягів даних	Data Mining, Big Data-платформи, BI-системи
За типом використовуваних технологій	Кабельний зв'язок	Передача даних через фізичні лінії зв'язку	Ethernet, оптоволоконні мережі
	Бездротовий зв'язок	Передача даних без фізичного з'єднання	Wi-Fi, мобільні мережі
	Супутниковий зв'язок	Використовує супутники для глобального покриття	Супутниковий Інтернет, навігаційні системи
	Хмарні технології	Доступ до обчислювальних ресурсів через Інтернет	IaaS, PaaS, SaaS
	Інтернет речей (IoT)	Мережа взаємодіючих фізичних пристроїв	Розумний дім, смарт-міста
	Віртуальна та доповнена реальність	Забезпечують імерсивну взаємодію з користувачем	Ігри, навчальні тренажери, конференції
	Блокчейн	Розподілені реєстри для зберігання транзакцій	Криптовалюти, смарт-контракти
	Штучний інтелект	Автоматизована обробка даних та прийняття рішень	Машинне навчання, експертні системи

<b>За архітектурними рішеннями</b>	Монолітна архітектура	Усі компоненти інтегровані в єдину систему	Класичні корпоративні застосунки
	Сервісно-орієнтована архітектура (SOA)	Незалежні сервіси взаємодіють через мережу	Корпоративні інформаційні системи
	Подієво-орієнтована архітектура	Обмін інформацією через події та повідомлення	Системи обробки подій у реальному часі
	Безсерверна архітектура (Serverless)	Масштабування та управління інфраструктурою виконує провайдер	Хмарні функції, FaaS-платформи

У межах даної роботи головний акцент робиться на особливостях МСА [124], оскільки саме вона дозволяє найбільш ефективно аналізувати, моделювати та оцінювати високонавантажені сервіси. Вибір МСА зумовлений її здатністю вирішувати проблеми зростаючої складності сучасних систем. Зокрема, використання мікросервісів оптимізує витрати та управління у веб-проектах, а також гарантує високу адаптивність у сфері промислової автоматизації. Для того щоб надалі успішно моделювати ІКМ, ми запроваджуємо підхід із розділенням послуг. Вони будуть класифікуватися на сеансові та транзакційні, виходячи з того, як саме користувач взаємодіє з системою.

Взаємодію користувача із системою можна класифікувати за тривалістю та цільовим призначенням, виділивши транзакційні (TR) та сеансові (SE) послуги.

Транзакційна модель [126] базується на принципі короткочасної дії: система отримує запит, обробляє дані (наприклад, проводить платіж в електронній комерції або пересилає файл) і завершує операцію.

На противагу цьому, сеансові послуги передбачають тривале утримання користувача в системі. Контакт може не перериватися днями, як це відбувається у додатках для спілкування чи онлайн-іграх. Відповідно, архітектура SE-послуг є

складнішою, оскільки вимагає постійного збереження поточного стану сесії та регулярних процедур аутентифікації. Варто також зазначити, що в телекомунікаційній галузі сеансовий підхід є базовим для створення неперервних ліній зв'язку, завдяки чому функціонують багатокористувацькі відеоконференції та IP-телефонія.

Сеансові послуги SE [107] зазвичай передбачають тривалу взаємодію користувача із системою, наприклад, для обміну повідомленнями у додатках або онлайн-іграх, де кінцевий користувач може залишатися активним у системі протягом годин чи навіть днів. У таких додатках широко використовуються механізми авторизації, автентифікації [8], збереження стану сесії та обробки запитів на основі поточного стану.

Основна відмінність між сеансовими та транзакційними послугами полягає у способах забезпечення QoS. По-перше, час обслуговування  $T_{SE}$  безпосередньо залежить від дій користувача, тобто від тривалості, яку він витрачає на отримання послуги. У межах цієї роботи не враховуються зміни у поведінці користувача, зокрема тимчасові обмеження, які можуть впливати на сприйняття якості послуги QoE. У випадку транзакційних систем час обслуговування запитів визначається характеристиками самої системи, такими як продуктивність та надійність. Покращення цих показників можливе завдяки вдосконаленню апаратного та програмного забезпечення.

По-друге, масштабованість та продуктивність системи по-різному впливають на ці два типи послуг. Висока якість клієнтського сервісу залежить від раціонального використання обчислювальних ресурсів і здатності інфраструктури до адаптивного розширення. Для транзакційних послуг критично важливими є швидкість обробки та мінімізація помилок під час виконання запитів.

Час обслуговування транзакційного запиту у порівнянні з тривалістю користувацької сесії є незначним і наближається до нуля. Тому при моделюванні сеансових послуг є обґрунтованим ігнорування зайнятості ресурсів транзакційними запитами (TR) за таких умов.

Переваги мікросервісного підходу [124] яскраво проявляються під час декомпозиції системи, що значно покращує взаємодію між її компонентами. Наприклад, для побудови високонавантажених систем сеансові послуги раціонально розбивати на ізольовані вузли: окремо для автентифікації, підтримки сесій чи роботи з поштою. Водночас транзакційні послуги можуть формувати або єдиний мікросервіс для збереження даних, або цілу групу сервісів, поділених за бізнес-завданнями. Однак для аналізу ефективності таких архітектурних рішень досі немає єдиної моделі, яка б охоплювала всі типи послуг. Оптимальним виходом є застосування математичного апарату теорії масового обслуговування (ТМО) [131]. Як свідчить успішний досвід розрахунку ресурсів для локацій в ігрових сервісах, методи ТМО дають змогу з високою точністю визначати ймовірнісні та часові характеристики як для сеансових, так і для транзакційних процесів.

#### **2.4. Мікросервісна архітектура надання послуг користувачеві**

Сьогодні концепція МСА є одним із найбільш досліджуваних напрямків у проектуванні програмного забезпечення [147]. Багато уваги приділяється як порівнянню МСА з класичними монолітними системами, так і розробці стратегій міграції на мікросервіси. З технічної точки зору, під мікросервісом розуміють слабо зв'язаний, автономний процес, що функціонує в чітко обмеженому контексті та обмінюється даними з іншими компонентами за допомогою повідомлень.

Перехід на МСА став логічною відповіддю на виклики, пов'язані з розробкою та підтримкою масштабних розподілених веб-додатків (рис. 2.5). Практика доводить високу ефективність цього підходу в абсолютно різних сферах:

- системи інтернет-банкінгу;
- платформи для обробки великих масивів даних (Big Data);
- системи управління метаданими користувачів;
- інфраструктура для вбудованих систем;
- рішення для інтеграції перехресних сервісів.

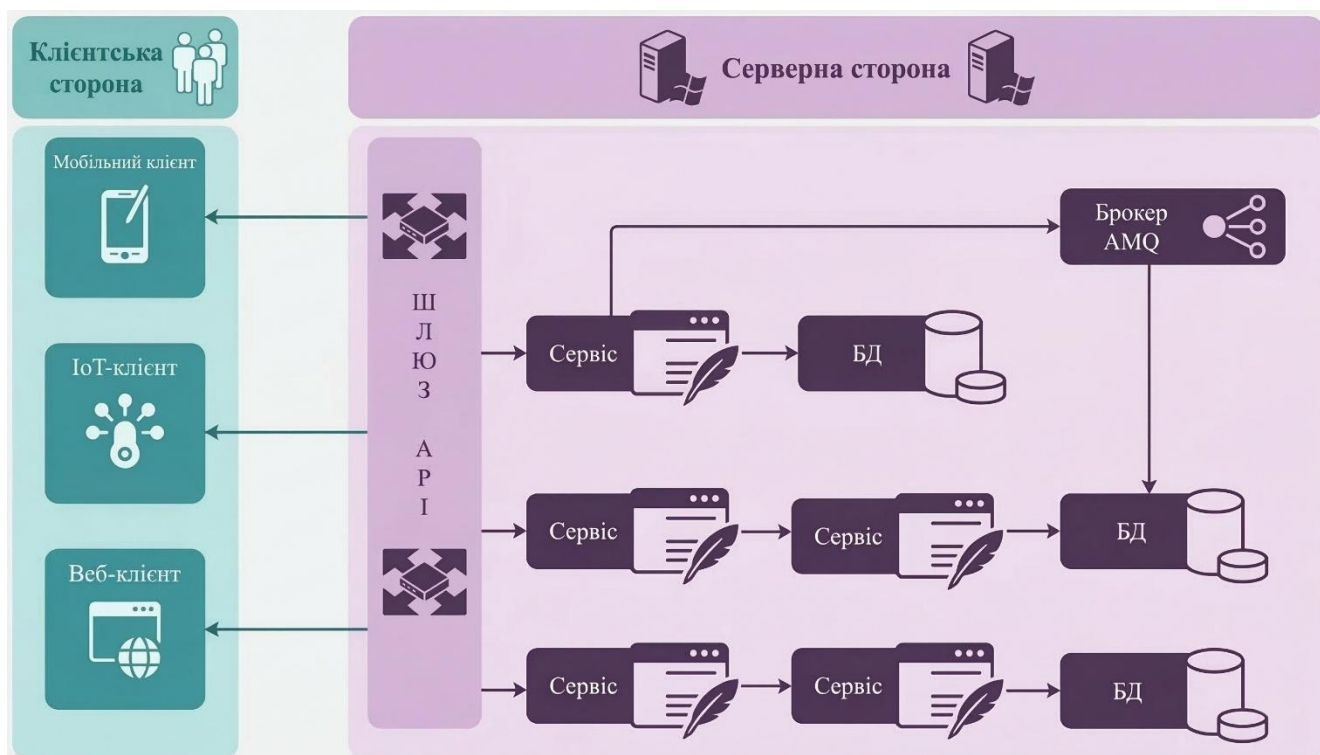


Рис. 2.5 – Типовий приклад мікросервісної архітектури

Проектування в умовах МСА кардинально відрізняється від традиційних підходів і вимагає застосування специфічних хмарних шаблонів. Їх впровадження не лише підвищує загальну продуктивність, а й суттєво покращує якісні показники системи. У контексті нашого дослідження, орієнтованого на оцінку слабо зв'язаних сервісів, ми пропонуємо розділяти послуги на сеансові та транзакційні (про що детально йтиметься далі), і будувати архітектуру з урахуванням характеру взаємодії клієнта з системою [167].

Як самостійна одиниця розгортання, мікросервіс відповідає за конкретну бізнес-функцію і взаємодіє з іншим середовищем виключно через стандартизовані інтерфейси. Хоча строгих обмежень щодо розміру мікросервісу не існує, він повинен відповідати п'яти базовим критеріям:

1. Автономність: застосунок декомпонується на невеликі незалежні модулі, кожен з яких вирішує власне завдання;
2. Обмежений контекст: функціонал сервісу суворо локалізований у межах визначеної бізнес-області;

3. Незалежне розгортання: оновлення чи масштабування одного компонента відбувається без зупинки або впливу на всю систему загалом;

4. Слабка зв'язність: комунікація між компонентами здійснюється переважно через API-шлюзи;

5. Ізоляція ресурсів: кожен мікросервіс керує власною базою даних та кодовою базою, незалежно від інших.

Для забезпечення високої доступності (High Availability) часто застосовують концепцію еквівалентних мікросервісів. Це сервіси, які виконують абсолютно ідентичну функцію, але можуть бути реалізовані різними методами чи технологіями. У разі відмови одного компонента, його задачі миттєво перехоплює еквівалентний. Крім того, це дозволяє розгортати систему в різних хмарних середовищах та оптимізувати часові затримки шляхом паралельної обробки запитів і балансування навантаження. Детальний математичний апарат паралельної обробки буде наведено у третьому розділі. Вибір конкретного еквівалентного сервісу для виконання задачі зазвичай базується на показниках QoS.

Складні додатки вимагають координації десятків мікросервісів, для чого створюються автоматизовані робочі процеси [20, 27]. Один сервіс може бути задіяний одразу в кількох таких процесах (рис. 2.6).

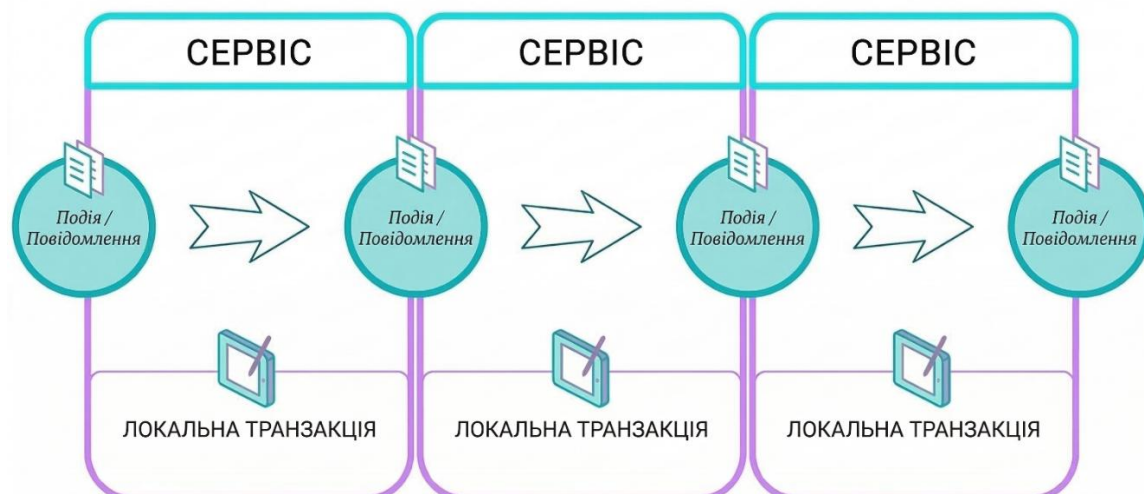


Рис. 2.6 – Шаблон проєктування робочого процесу в МСА

Управління послідовністю кроків транзакції, де кожен крок обробляється окремим сервісом, може бути реалізовано за двома основними паттернами [24, 27]:

**1. Централізована логіка (Оркестратор)** У цьому сценарії [114] створюється окремий керуючий компонент – координатор (або оркестратор), як вказано на рис. 2.7. Він чітко знає послідовність дій і віддає команди відповідним мікросервісам. Наприклад, якщо на етапі обробки завдання одним із сервісів виникає помилка, оркестратор ініціює процедуру скасування (компенсації) попередніх змін у сервісах [84]. Перевагою підходу є легкість управління бізнес-логікою. Недоліком – наявність єдиної точки відмови: при пікових навантаженнях оркестратор може стати вузьким місцем системи.



Рис. 2.7 – МСА з централізованою логікою взаємодії (оркестрацією): сценарій успішного виконання транзакції (етапи 1, 2 – отримання запиту; етапи 3-10 – послідовне делегування задач мікросервісам та отримання підтверджень; етапи 11, 12 – повернення успішного результату клієнту).

**2. Розподілена комунікація через події (Хореографія)** Альтернативний підхід [83] базується на подійно-орієнтованій архітектурі. Тут немає єдиного центру управління. Кожен мікросервіс знає лише про свою зону відповідальності і реагує на події, що надходять через подійну шину даних (брокер повідомлень). Процес виглядає так: Сервіс виконує свою дію і публікує подію у брокер; сервіс, який «слухає» брокер, отримує цю подію і починає свою частину роботи. Цей

підхід дозволяє сервісам діяти максимально автономно та усуває проблему єдиного вузла відмови, що робить його ідеальним для високонавантажених систем (рис. 2.8).



Рис. 2.8 – МСА з розподіленою подійно-орієнтованою комунікацією (хореографією): сценарій успішної обробки замовлення (взаємодія здійснюється асинхронно шляхом публікації подій та підписки на них через подійну шину)

Для ефективного управління мікросервісами у великих проєктах (наприклад, на платформах e-commerce з десятками окремих сервісів для кошика, обробки платежів та логістики) обов'язково використовуються спеціалізовані інфраструктурні фреймворки. Оскільки ручна координація такої кількості компонентів є неможливою, ці інструменти дозволяють автоматизувати маршрутизацію потоків даних та гарантують загальну надійність системи. Крім того, вони здатні динамічно масштабувати ресурси: миттєво виділяти додаткові сервери в моменти пікових навантажень та автоматично вивільняти їх під час спаду активності користувачів.

Щоб заздалегідь оцінити ефективність та масштабованість різних конфігурацій архітектури під реальним навантаженням, використовують емпіричні профілі робочих ситуацій. Вони являють собою детальні моделі тестування, які відтворюють реальну поведінку користувачів і дозволяють виявити потенційні вузькі місця системи. Головною метрикою успішності під час таких випробувань



виступає загальна частка послуг, що були виконані без системних збоїв та із суворим дотриманням вимог масштабованості (зокрема, щодо лімітів часу відгуку). Вона розраховується за формулою:

$$\hat{S}(\lambda) = \sum_{j=1}^n w_j c_j(\lambda), \quad (2.14)$$

де  $n$  – загальна кількість мікросервісів, що беруть участь у процесі;  $w_j$  – ваговий коефіцієнт  $j$ -го сервісу (наприклад, частка запитів, що на нього надходить, або рівень його критичності для бізнесу), причому

$$\sum_{j=1}^n w_j = 1;$$

тоді  $c_j(\lambda)$  – бінарна функція-індикатор успішності  $j$ -го сервісу. Вона приймає значення 1, якщо середній час відгуку сервісу  $x_j(\lambda)$  не перевищує допустимого SLA-порогу  $\Gamma_j(\lambda_0)$  і 0 – якщо ліміт часу перевищено.

Для того, щоб цей підхід став повноцінно універсальним, необхідно розробити комплексну модель мережі черг для обчислення показника  $\hat{S}(\lambda)$ . Саме це завдання буде вирішуватися в рамках даної дисертаційної роботи при моделюванні процесів обслуговування транзакційних систем [24].

## 2.5 Моделі оцінювання часових характеристик комбінованих послуг

У сучасних умовах одна ТК-інфраструктура часто виступає базою для надання цілого спектра різноманітних послуг. Їхня тісна інтеграція дозволяє формувати так звані «комбіновані послуги», які значно розширюють функціональні можливості та підвищують зручність для кінцевого споживача. Найкращим архітектурним паттерном для реалізації таких рішень визнано МСА. Завдяки декомпозиції додатку на ізольовані та автономні компоненти (кожен з яких відповідає за свою вузьку задачу), МСА гарантує гнучкість управління, спрощує комунікацію між вузлами та дає змогу оперативно модернізувати систему під мінливі запити користувачів [27, 28].

Проте при проєктуванні комбінованих систем критично важливо враховувати архітектурну специфіку: транзакційні функції та сеансові сервіси висувають абсолютно різні вимоги до продуктивності, надійності та здатності до масштабування. Саме тому для забезпечення стабільної та ефективної роботи доцільно виділяти під кожну функцію відокремлені інфраструктурні ресурси.

Щоб зробити запропоновані методи оцінки максимально універсальними, було побудовано абстрактну модель взаємодії сервісів. Її візуалізацію наведено нижче на рис. 2.9.

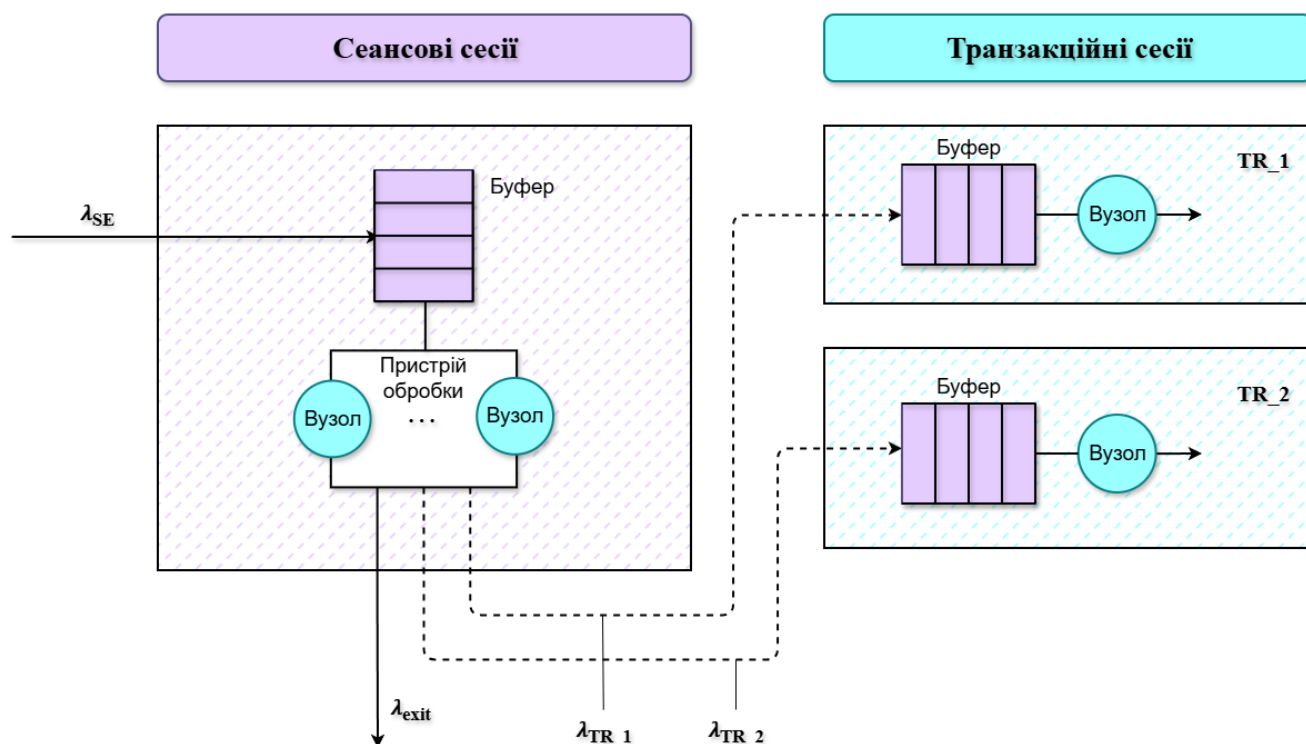


Рис. 2.9 – Функціональна схема взаємодії між сеансовою та транзакційною послугами

Алгоритм взаємодії між сеансовою та транзакційною складовими можна описати у вигляді такого сценарію:

1. Спочатку генерується вхідний потік клієнтських запитів із певною інтенсивністю  $\lambda_{SE}$ , який спрямовується на вузол обслуговування;
2. Далі цей запит або переходить у стан очікування (буферизації), або ж ініціює відкриття користувацької сесії в певний момент часу  $t_1$ ;

3. Починаючи від моменту  $t_1$ , система починає генерувати виклики до транзакційної підсистеми. Таким чином, активність транзакційних послуг є прямою похідною від наявності та функціонування клієнтських сесій.

Зважаючи на кардинальні розбіжності у вимогах до якості та функціональних параметрах, намагатися об'єднати сеансові та транзакційні процеси в єдину модель є нераціональним кроком. Цей висновок наочно підтверджується під час аналізу часових характеристик системи.

Для кращого розуміння розглянемо часову шкалу окремої клієнтської сесії  $T_{SE}$ , яка представлена на рис. 2.10. Як показано на графіку, у межах цієї сесії ініціюються потоки завдань двох різних типів:  $T_{TR-1}$  та  $T_{TR-2}$ . Відповідні інтенсивності для цих потоків розраховуються як  $\lambda_{TR-1} = 1/t_{RI-1}$  та  $\lambda_{TR-2} = 1/t_{RI-2}$ .

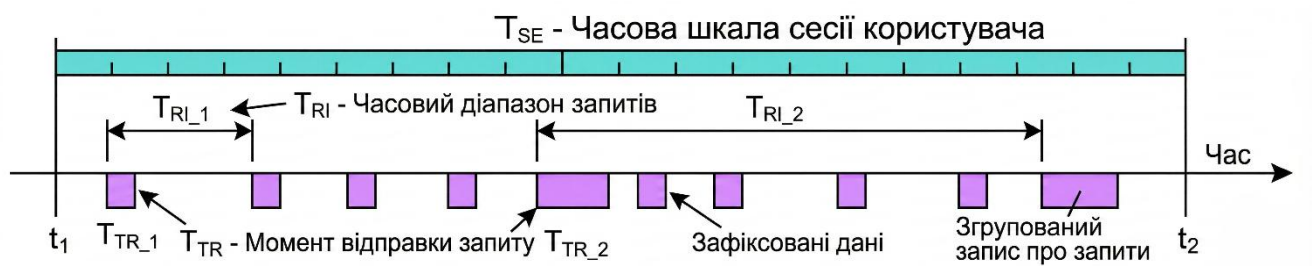


Рис. 2.10 – Представлення граничних значень розподілу часу надання меж послуг типів  $T_r$  та  $S_e$

Для оцінки загального потоку запитів, що генерується множиною активних сесій (кількість яких фіксується як  $\bar{K}$ ) та спрямовується до вузла обробки транзакцій TR, застосовується двоетапний підхід.

Першим кроком є визначення базового навантаження на користувацький сесійний вузол. Цей параметр, позначений як  $A_{SE}$ , можна виразити через класичне співвідношення інтенсивностей  $\left(\frac{\lambda}{\mu}\right)$ . Проте на практиці його частіше розраховують як добуток середньої тривалості одного сеансу  $\overline{t_{SE}}$  та частоти надходження нових звернень від користувачів сервісу  $\lambda_{SE}$ :

$$A_{SE} = \overline{t_{SE}} \cdot \lambda_{SE}.$$

Маючи показник навантаження, можна обчислити фінальну сумарну швидкість надходження запитів –  $\lambda'_{TR_i}$ . Вона визначається шляхом множення знайденого коефіцієнта  $A_{SE}$  на  $\lambda_{TR_i}$  (інтенсивність, з якою конкретний сеансовий вузол  $S_i$  самостійно генерує звернення до транзакційного вузла). Отже, підсумкова залежність набуває такого вигляду:

$$\lambda'_{TR_i} = \lambda_{TR_i} \cdot A_{SE}.$$

Описаний вище механізм ілюструє класичний випадок, коли під час сесії мікросервіс звертається до власної, ізольованої бази даних. Проте на практиці часто зустрічається інший підхід: кілька різних мікросервісів можуть паралельно надсилати запити до єдиної спільної БД. Цю конфігурацію проілюстровано нижче на рис. 2.11.

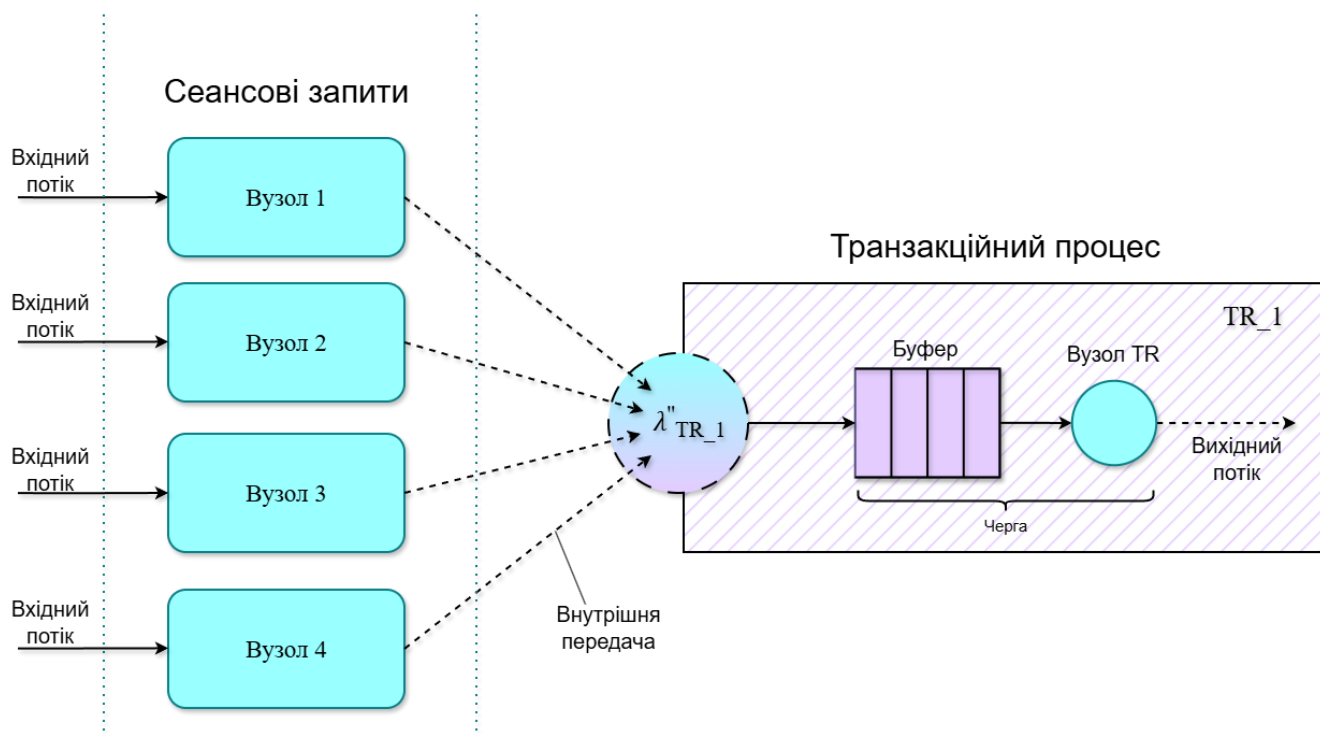


Рис. 2.11 – Функціональна схема передачі запитів від послуги SE до послуги TR

Часто виникають сценарії, за яких мікросервіси мають відмінну тривалість сесій  $\overline{t_{SE-j}}$  та генерують запити до вузла транзакцій з неоднаковою інтенсивністю  $\lambda_{TR_j}$ . У такому разі для визначення сукупного вхідного потоку до підсистеми TR доцільно застосувати поетапний розрахунок.

Спочатку визначається частковий потік для кожного окремого  $j$ -го вузла користувацької послуги. Він є добутком інтенсивності звернень цього вузла  $\lambda_{j,TR}$  на його специфічний рівень навантаження  $A_j$ :

$$\lambda_{part,j} = \lambda_{j,TR} \cdot A_j.$$

Після цього обчислюється загальний потік запитів на вході  $\lambda''_{TR}$  шляхом агрегування всіх отриманих часткових значень. Підсумовування здійснюється за всією множиною сеансових послуг, загальна кількість яких дорівнює  $Y_{SE}$ :

$$\lambda''_{TR} = \sum_{j=1}^{Y_{SE}} \lambda_{part,j}.$$

Щодо побудови повноцінної імітаційної моделі такого телекомунікаційного сервісу, то вона базується на трьох ключових групах параметрів, які формалізуються у вигляді векторів. Першочергово враховуються глобальні впливи, або ж зовнішні фактори, об'єднані у вектор  $C = \{c_1, c_2, \dots, c_n\}$  (сюди входять діюча система пріоритетів та загальна інтенсивність на вході). Наступною складовою є вектор внутрішніх архітектурних параметрів  $B = \{b_1, b_2, \dots, b_n\}$ , який деталізує конфігурацію системи. Він містить інформацію про кількість робочих вузлів  $Y$  та пристроїв обробки  $v_i$ , інтенсивності їх обслуговування  $\mu_i$ , а також ймовірності відмови чи втрати запитів  $p_{відм,i}$  для кожного  $i \in Y$ . Кінцева оцінка якості обслуговування та результативності роботи сервісу формується за допомогою третьої групи – вектору вихідних характеристик  $H = \{h_1, h_2, \dots, h_n\}$ .

З огляду на вищевикладене, вирішення поставленої наукової задачі зводиться до створення двокомпонентного математичного апарату. Перша модель має визначати характеристики сеансових послуг, аналізувати їхні ймовірно-часові параметри та оцінювати створюване ними навантаження на транзакційну підсистему. Друга математична модель повинна описувати якісні характеристики самих транзакційних послуг. У кінцевому підсумку це дозволить інтегрувати обидві моделі в єдиний алгоритм для комплексного розрахунку параметрів комбінованих, слабо зв'язаних послуг.

## 2.6 Моделі та методи розрахунку затримок під час надання сеансових послуг користувачам

Поняття сесійного сервісу [24, 27] в сучасній ІТ-індустрії впливає з базового визначення сеансу в обчислювальних середовищах, що описує активну фазу взаємодії між користувачем та програмним забезпеченням або ресурсами системи.

У нормативній базі ці послуги розглядаються крізь призму телекомунікацій. Кожен документ адаптує термінологію під конкретні задачі галузі. Наприклад, згідно з ITU-T E.800 [120], під сеансовою послугою розуміють комплекс дій, спрямованих на ініціацію, підтримку та завершення зв'язку між кількома точками, включно з обміном даними в межах цього циклу. Інші джерела акцентують увагу на функціональному аспекті: це сервіс, що надає клієнту доступ до певних можливостей чи системних ресурсів протягом обмеженого часу [165], припускаючи інтеграцію з іншими зовнішніми системами.

Цікаво, що в бібліотеці ITIL [176] чітко сформульована дефініція сеансових послуг відсутня. Проте в межах управління рівнем сервісу (SLM) для моніторингу їхньої якості зазвичай використовують такі показники, як стабільність (доступність) та швидкість реакції (час відгуку).

У межах даного дослідження для аналізу сеансових послуг виділимо наступні часові та ймовірнісні параметри:

- **Рівень завантаження вузлів:** демонструє ступінь використання потужностей окремих елементів у розподіленій мережі;
- **Час очікування в черзі:** середній період від надходження запиту до початку його фактичного опрацювання;
- **Ризик затримки сесії:** вірогідність виникнення паузи перед тим, як система допустить користувача до роботи;
- **Локальна інтенсивність трафіку:** частка загального потоку запитів, що припадає на конкретний вузол;
- **Повна тривалість перебування:** час від моменту авторизації/входу до завершення роботи;

- **Пропускна здатність:** гранична кількість операцій або запитів, які інфраструктура здатна виконати за одиницю часу.

Практичне втілення таких послуг у веб-сегменті охоплює широкий спектр рішень, починаючи від інструментів для миттєвого зв'язку [95[]], таких як Telegram чи WhatsApp, що забезпечують текстовий чат у реальному часі, та спеціалізованих сервісів відео зв'язку на кшталт Google Meet, орієнтованих на групову онлайн-комунікацію. Важливу роль також відіграє цифрова телефонія (VoIP) [4, 98], яка дозволяє передавати голосові дані через мережеві протоколи, та ігрові онлайн-проекти, де безперервна взаємодія між учасниками потребує мінімальних затримок. Окрім повсякденних потреб, сеансові сервіси активно використовуються в професійних сферах: у моніторингових системах та відеоспостереженні, що функціонують у режимі live, в автоматизації виробництва для контролю процесів безпосередньо в момент їх перебігу, а також у телемедицині, де фахівці отримують можливість проводити віддалені консультації та навіть складні операції в реальному часі.

Впровадження таких технологічних рішень суттєво оптимізує роботу фахівців, підвищуючи ефективність сервісу та кінцеві результати діяльності.

Ефективний аналіз якісних показників нових багатокористувацьких сервісів у телекомунікаціях неможливий без чіткої ідентифікації базових детермінант системи. Головною перешкодою тут зазвичай стає хаотичність дій споживачів, що суттєво ускладнює розрахунок часових параметрів. Для подолання цієї невизначеності доцільно використовувати комбіновані методи: від ретроспективного вивчення попередніх масивів запитів до аналізу функціонування аналогічних систем або моделювання на основі даних, отриманих під час адаптації живого сервісу.

Процес розробки моделі для оцінки обслуговуваного навантаження передбачає сегментацію серверних потужностей. Це дозволяє деталізовано описати функціональну структуру послуги та виділити ключові вхідні дані, як-от часову динаміку надходження запитів, вірогідність переміщення активності користувача

між різними вузлами, а також тривалість перебування запиту в кожній точці обробки.

У науковій літературі [109] панує підхід, за якого динаміка переходів між серверами (зокрема в ігрових середовищах) апроксимується ланцюгами Маркова [149]. Додатково пропонуються специфічні алгоритми підготовки до випробувань, спираючись на поведінкові паттерни користувачів. Що стосується математичного опису процесів, то логічним буде використовувати широкий спектр статистичних розподілів: від класичного експоненціального та Вейбулла [143] до складніших моделей на кшталт розподілу Парето [158], оберненого Гауссівського [151] чи змішаних експоненціальних варіантів [148].

Практична реалізація вимірювань зазвичай базується на двох стратегіях:

- глибокий аналіз мережевого трафіку та трасувань;
- інсталяція спеціалізованого софту на стороні клієнта для прецизійного фіксування тривалості кожної сесії.

Для аналізу статистичних закономірностей використовують математичну модель розподілу Парето, де інтегральна функція  $F(x)$  та відповідна їй щільність ймовірності  $f(x)$  мають наступний вигляд:

$$F(x) = P(X \leq x) = 1 - \left(\frac{x_{min}}{x}\right)^\alpha \quad \text{коли } x \geq x_{min};$$

$$f(x) = \begin{cases} \frac{\alpha x_{min}^\alpha}{x^{(\alpha+1)}} & x \geq x_{min}; \\ 0 & x < x_{min}. \end{cases} \quad (2.15)$$

Для оцінки статистичного розсіювання та центральних тенденцій випадкової величини, що описується даним законом, застосовують апарат моментів та розрахунок дисперсії [132]:

**Математичне сподівання  $n$ -го порядку:** Для знаходження початкового моменту  $n$ -го порядку використовується вираз:

$$E[X^n] = \frac{\alpha \cdot x_{min}^n}{\alpha - n}.$$

**Показник дисперсії:** Розрахунок варіативності (дисперсії) здійснюється за формулою:



$$\text{Var}(X) = \frac{\alpha \cdot x_{\min}^2}{(\alpha - 1)^2(\alpha - 2)}.$$

Для аналізу часових меж сесій у сучасних дослідженнях часто застосовують спеціалізовані статистичні моделі. Зокрема, одним із ефективних методів є декомпозиція розподілу Вейбулла–Гнеденка [132] через його представлення у вигляді сукупності трьох випадкових складників. Такий підхід, поряд із використанням гіперекспоненціальної апроксимації, дозволяє значно спростити математичний опис складних систем під час оцінювання їхніх якісних показників.

Закон розподілу Вейбулла–Гнеденка для випадкової величини  $X \sim W(k, \lambda)$  описується через наступний математичний апарат:

**Ймовірнісна функція розподілу:** Показує накопичену ймовірність того, що тривалість сесії не перевищить значення  $x$ :

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^k\right) \quad \text{коли } x \geq 0.$$

**Функція щільності розподілу:** Визначає концентрацію ймовірності в околі точки  $x$ :

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}.$$

**Математичне сподівання  $n$ -го порядку:** Дозволяє розрахувати статистичні моменти випадкової величини:

$$E[X^n] = \lambda^n \cdot \Gamma\left(1 + \frac{n}{k}\right).$$

У наведених виразах використовуються наступні параметри:  $k$  – коефіцієнт форми (shape parameter), де  $k > 0$ ;  $\lambda$  – коефіцієнт масштабу (scale parameter), де  $\lambda > 0$ ;  $\Gamma(z)$  – Ейлерова гамма-функція, що визначається як невластний інтеграл:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt.$$

В основі пропонованої моделі лежить припущення, що тривалість користувацької активності – як у масштабах усієї інфраструктури, так і на рівні її окремих сегментів – підпорядковується закону експоненціального розподілу. Такий вибір є базовим для подальшої розробки з кількох стратегічних міркувань:

**1. Аналітична гнучкість на етапі дизайну.** Використання експоненціальних моделей суттєво спрощує математичні операції. Це дозволяє ще до збору реальної статистики детально дослідити ймовірісно-часові показники системи та локалізувати потенційні критичні зони (вузькі місця) у мережі масового обслуговування.

**2. Відсутність ефекту пам'яті.** Ця властивість означає, що шанси на завершення обробки запиту в найближчий момент не корелюють із тим, скільки часу цей запит уже провів у черзі.

**3. Універсальність.** Даний тип розподілу є класичним інструментом для опису стохастичних процесів, таких як випадкове надходження нових завдань або вихід обладнання з ладу.

Архітектура розроблюваного сервісу розглядається крізь призму математичного апарату мереж Джексона. Це дозволяє представити всю систему як цілісну сукупність взаємопов'язаних вузлів масового обслуговування. Безпосередню роботу із сесіями виконують окремі модулі (позначені як  $S_i$ ).

Кожен такий функціональний блок функціонує як багатоканальна структура з необмеженою ємністю черги. Обробка вхідних потоків здійснюється за принципом FIFO (First In – First Out) [27], де пріоритет надається запитам у порядку їхнього надходження.

Для візуалізації логіки сервісу використовується абстрактна схема (рис. 2.12). Вона включає генератор вхідного навантаження, який ініціює запити, та відповідні модулі обробки.

Окремо виділено так званий нульовий блок, який виконує роль зовнішнього середовища. Хоча він не є частиною внутрішньої обчислювальної мережі, його наявність необхідна для моделювання повноцінної взаємодії системи з навколишнім цифровим простором та фіксації завершених операцій.

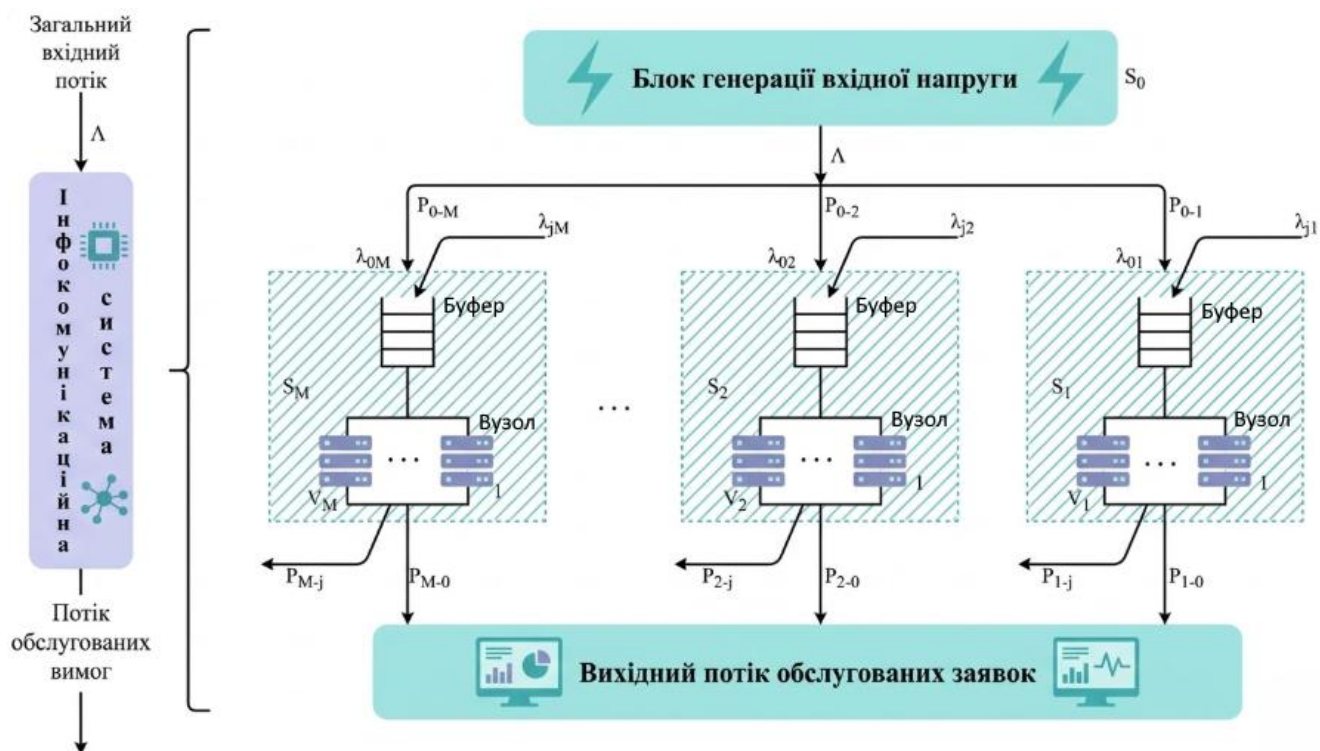


Рис. 2.12 – Узагальнена модель сеансового багатокористувацького сервісу

В основу моделювання покладено концепцію замкнених мереж масового обслуговування, де імітація активності користувацьких сесій реалізована через виокремлену підсистему (рис. 2.13). Такий підхід, описаний у раніше згаданому джерелі, дозволяє розглядати сесії як динамічні джерела запитів, що циркулюють у системі.

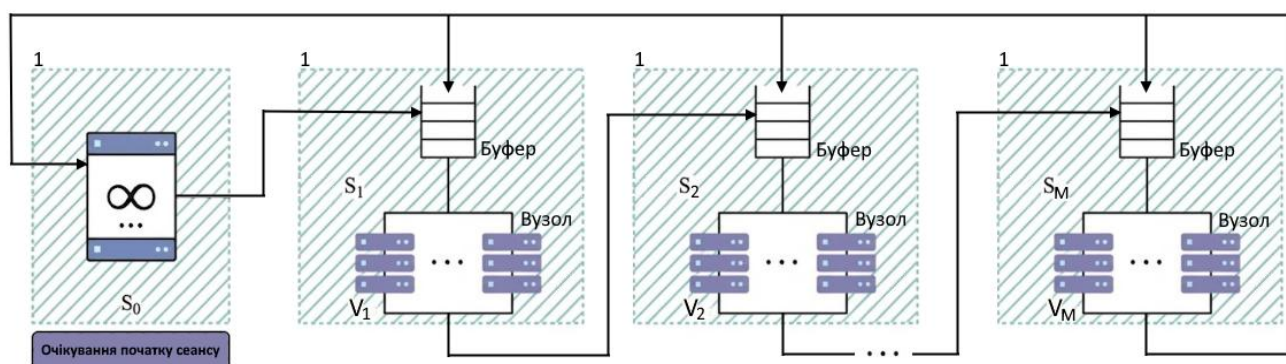


Рис. 2.13 – Альтернативна модель обслуговування сервісів

У межах побудованої моделі [2] передбачається, що зовнішнє середовище (назвемо його базовим вузлом  $E_0$ ) функціонує як буферна зона, що імітує паузи користувачів між їхніми активними діями в системі. Хоча математично така

абстракція припускає нескінченну циркуляцію запитів, на практиці популяція сесій є динамічною: старі завершуються, а нові постійно генеруються. Обрана топологія мережі масового обслуговування виявилася найбільш оптимальною для проведення подальших аналітичних розрахунків.

Процес маршрутизації виглядає наступним чином. Із зовнішнього середовища  $E_0$  генерується сумарний потік вхідних запитів з інтенсивністю  $\gamma$ . Цей потік розгалужується і з певною вірогідністю потрапляє до одного з  $N$  функціональних кластерів системи. Коли користувацька сесія досягає обраного кластера (позначимо його як  $C_i$ ), вона займає один із  $k$  доступних потоків обробки на середній час  $\tau_i$ .

Після того, як вузол завершує свою частину роботи, сесія з імовірністю  $\theta_{i,j}$  переходить до наступного кластера  $C_j$  для продовження обслуговування. Сукупність усіх можливих маршрутів між  $N$  кластерами системи формує стохастичну матрицю переходів  $\Theta$ :

$$\Theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \dots & \theta_{1,N} \\ \theta_{2,1} & \theta_{2,2} & \dots & \theta_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{N,1} & \theta_{N,2} & \dots & \theta_{N,N} \end{bmatrix}.$$

Використання такої матриці дозволяє деталізувати розрахунок сумарного навантаження на кожен окремий вузол мережі. Замість класичного згорнутого рівняння Джексона [159], вхідну інтенсивність для  $i$ -го кластера  $v_i$  доцільно декомпонувати на дві складові:

**Зовнішнє навантаження** (частка первинного трафіку, що надходить безпосередньо з  $E_0$ ):

$$v_i^{ext} = \gamma \cdot \theta_{0,i}.$$

**Внутрішнє транзитне навантаження** (запити, що рециркулюють з інших кластерів):

$$v_i^{int} = \sum_{j=1}^N v_j \cdot \theta_{j,i}.$$

Відповідно, повна інтенсивність потоку, що надходить на кластер  $C_i$ , є сумою цих компонентів:

$$v_i = v_i^{ext} + v_i^{int}.$$

Оскільки ми моделюємо відкриту систему, сценарій «нескінченного зациклення» сесії є неприпустимим. Тому для кожного вузла  $C_i$  обов'язково існує ймовірність остаточного виходу із системи (завершення сеансу), яка вираховується як доповнення до одиниці:

$$\theta_{i,out} = 1 - \sum_{j=1}^N \theta_{i,j}.$$

Для опису безперервних Марківських переходів між станами системи використовуються локальні інтенсивності  $q_{\alpha,\beta}$ . Швидкість переходу системи залежить від інтенсивності обслуговування  $\mu_i$  на конкретному вузлі та ймовірності подальшого маршруту:

- Інтенсивність внутрішнього переміщення:  $q_{i,j} = \mu_i \cdot \theta_{i,j}$ ;
- Інтенсивність виходу з системи:  $q_{i,out} = \mu_i \cdot \theta_{i,out}$ .

Такий підхід до декомпозиції параметрів є критично важливим для точного моделювання часових затримок у складних багаторівневих сервісах.

## 2.7 Математичний апарат для обчислення часових характеристик сеансових послуг

Для визначення середньої тривалості перебування користувача в системі  $T_{total}$ , що працює в стаціонарному режимі, використовується адитивна модель. Загальний час сесії в системі  $S$  розраховується як сума витрат часу на кожному окремому вузлі, з урахуванням частоти звернень до них:

### 1. Загальний час перебування:

$$T_{total} = \sum_{i=1}^Y a_i \cdot (w_i + \tau_i),$$

де  $a_i$  – середня кількість візитів до  $i$ -го вузла за одну сесію;  $w_i$  – очікування в черзі на обслуговування;  $\tau_i$  – чистий час отримання сервісу (обслуговування).

Сумарний час перебування запиту на вузлі можна виділити в окрему змінну  $u_i = w_i + \tau_i$ .

**2. Баланс інтенсивності звернень:** Коефіцієнти візитів  $a_i$  визначаються через систему лінійних рівнянь, що описують маршрутизацію всередині системи:

$$a_i = p_{0i} + \sum_{j=1}^Y a_j \cdot p_{ji} \quad \text{причому } i = \overline{1, Y}.$$

де  $p_{ji}$  – ймовірність переходу від вузла  $j$  до вузла  $i$ , а початкова умова для входу в систему зазвичай приймається як  $a_0 = 1$ .

### **Розрахунок інтенсивності та параметрів обслуговування**

Знаючи загальну інтенсивність вхідного потоку в систему  $\Lambda$ , можна обчислити локальну інтенсивність для кожного конкретного вузла:

$$\lambda_i = \Lambda \cdot a_i.$$

Ці показники є базою для отримання таких метрик, як завантаженість вузлів  $\rho_i$ , середня довжина черг  $L_i$  та затримки.

**Модель інтенсивності обслуговування:** Якщо вузол має кілька паралельних каналів обслуговування  $v$ , загальна інтенсивність  $\mu_n$  змінюється залежно від кількості заявок у системі:

$$\mu_n = \begin{cases} n\mu & \text{при } 0 \leq n \leq v; \\ v\mu & \text{при } n \geq v, \end{cases}$$

де  $\mu$  – продуктивність одного пристрою.

### **Показники ефективності (Формула Літтла та стан системи)**

Для оцінки середнього часу очікування  $\bar{W}$  використовується фундаментальна залежність Літтла, що пов'язує довжину черги та інтенсивність потоку:

$$\bar{W} = \frac{\bar{L}}{\lambda}.$$

**Умови стабільності:** Система вважається стаціонарною (черга не зростає до нескінченності), якщо виконується умова:

$$A_i < v_i \Rightarrow \lambda_i / \mu_i < v_i.$$

Або через коефіцієнт завантаження:

$$\rho_i = \frac{\lambda_i \cdot t_{obs}}{v_i} < 1.$$

**Ймовірнісні характеристики черги:** Середня довжина черги  $\bar{L}$  для багатоканальної системи розраховується через ймовірність застання всіх каналів зайнятими  $P_v$ :

$$\bar{L} = P_v \cdot \frac{A \cdot v}{(v - A)^2}.$$

Ймовірність того, що в системі немає жодного запиту  $P_0$ , визначається як:

$$P_0 = \left[ \sum_{x=0}^v \frac{A^x}{x!} + \frac{A^v}{v!} \cdot \frac{A}{v - A} \right]^{-1}.$$

Після чого знайти  $P_v$  можна за формулою:  $P_v = \frac{A^v}{v!} \cdot P_0$ .

У сучасних соцмережах тривалість обслуговування сильно варіюється. Короткі запити (текст) та тривалі (стрімінг відео) створюють нерівномірне навантаження, що вимагає динамічного регулювання ресурсів вузлів.

## Висновки до розділу 2

Отже, для стабільного стану системи характерним є значне переважання часу обробки запитів над сумарним часом затримки. Однак у випадку сеансових послуг, як і в мережах масового обслуговування, виникає необхідність вирішення проблем, пов'язаних із високим навантаженням на вузли. Для цього рекомендовано використовувати математичні та імітаційні методи моделювання, поєднання яких дозволяє знизити трудовитрати, як було описано у фінальному розділі цієї глави.

У цьому розділі дисертаційної роботи досягнуто таких результатів:

1. Розроблено та формалізовано методологію класифікації інформаційно-комунікаційних (ІК) послуг за рівнем взаємодії з користувачем:

- сеансові послуги – забезпечують функціонування комунікаційних сесій користувачів;

- транзакційні послуги – надають можливість виконання системних операцій;

- комбіновані послуги – поєднують функції сеансових і транзакційних послуг в одному сервісі, забезпечуючи користувачеві повну функціональність.

2. Досліджено специфіку взаємодії між різними типами послуг у рамках єдиних архітектурних рішень, а також визначено підхід до їхнього моделювання.

3. Розроблено імітаційні та аналітичні моделі для оцінювання часових характеристик сеансових послуг веб-сервісів.

Проведений аналіз сеансових послуг створює передумови для подальшого вирішення завдань, а саме розробки математичного інструментарію для оцінки часових характеристик (ЧХ) робочих процесів транзакційних послуг.



## РОЗДІЛ 3

### МОДЕЛЬ ДЛЯ ОЦІНКИ ХАРАКТЕРИСТИК ТРАНЗАКЦІЙНИХ ПОСЛУГ ІЗ ПАРАЛЕЛЬНОЮ ОБРОБКОЮ ЗАПИТІВ

#### 3.1. Транзакційні послуги, як сценарії реалізації робочих процесів

Будь-яка транзакція за своєю суттю є логічно завершеною одиницею робочого процесу, що охоплює одну або декілька маніпуляцій із даними. Для забезпечення цілісності системи кожна операція повинна відповідати базовому набору властивостей, відомих як ACID (Atomicity, Consistency, Isolation, Durability) [31]:

- **Атомарність (Atomicity)** – реалізує принцип неподільності: транзакція або виконується в повному обсязі, або не виконується взагалі. Якщо хоча б одна мікрооперація всередині ланцюжка зазнає невдачі, система ініціює повний відкат (rollback) до вихідного стану.

- **Узгодженість (Consistency)** – забезпечує трансформацію бази даних з одного стабільного стану в інший. Будь-яка транзакція, що завершилася успішно, має залишити дані у валідному стані, не порушуючи встановлених правил цілісності та бізнес-логіки.

- **Ізольованість (Isolation)** – забезпечує незалежність паралельних процесів. Проміжні результати транзакції, що ще не завершилася, залишаються невидимими для інших запитів, що запобігає виникненню конфліктів при одночасному доступі до спільних ресурсів.

- **Стійкість (Durability)** – гарантує, що після успішного підтвердження транзакції всі внесені зміни стають постійними. Навіть у випадку раптового апаратного збою чи перезавантаження вузла, дані не можуть бути втрачені.

ACID та її властивості представлено на рис. 3.1.

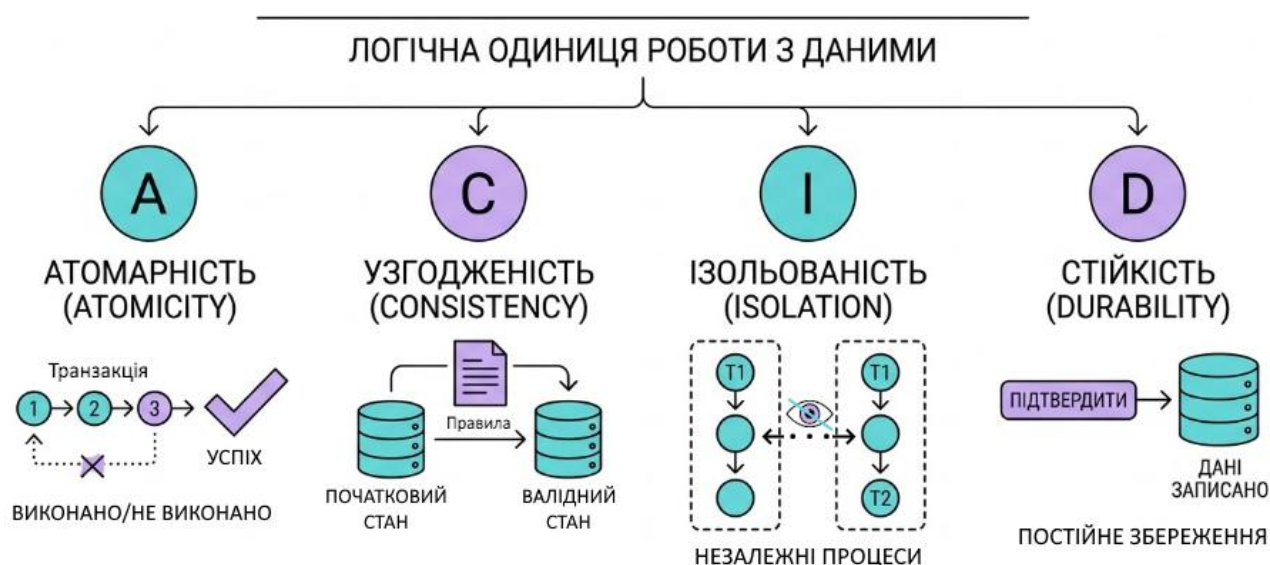


Рис. 3.1 – ACID властивості транзакцій

Дотримання цих властивостей робить мікросервісну архітектуру ідеальним вибором для реалізації таких послуг, де кожна окрема транзакція замикається в межах свого мікросервісу.

### Показники якості та продуктивності

При проектуванні транзакційних систем доцільно зосередити увагу на параметрах якості обслуговування та впровадженні методів мінімізації затримок при обробці складних операцій у DATA-центрах. Це дозволяє забезпечити стабільну роботу програмно-апаратного середовища за умов високої інтенсивності запитів.

Ефективність таких систем оцінюється за наступним комплексом характеристик:

1. **Часові параметри:** сюди входить час перебування запиту в черзі та системі (очікуване значення та дисперсія), чистий час обслуговування, а також загальний час реалізації всього робочого процесу.

2. **Пропускна здатність:** здатність системи обробляти максимально можливу кількість запитів за фіксований проміжок часу.

3. **Експлуатаційні метрики:** рівень надійності (ймовірність успіху) та мінімізація часу простою системи, що безпосередньо впливає на її коефіцієнт корисної дії.

## Організація та моделі робочих процесів (Workflow)

Робочий процес у контексті транзакційних систем – це чіткий сценарій, за яким задача проходить крок за кроком через різні сервіси. У межах системного проєктування Workflow [24, 27] розглядається як універсальна обчислювальна модель для стандартизації обміну даними в мережі.

Типовий життєвий цикл обробки задачі складається з таких фаз:

- **Ініціація:** прийом вхідного запиту;
- **Валідація:** верифікація даних на коректність;
- **Екзекуція:** безпосередня обробка;
- **Фіксація та завершення:** підтвердження результатів та закриття операції.

Залежно від архітектурних вимог, виділяють кілька базових моделей виконання процесів:

**1. Послідовна модель (Sequential)** Кроки виконуються строго один за одним у детермінованому порядку. Наступний етап не розпочнеться, поки не завершиться попередній, що забезпечує лінійну передбачуваність стану даних. Це значно спрощує діагностику помилок, проте загальний час виконання дорівнює сумі тривалостей усіх етапів.

**2. Паралельна модель (Parallel)** Дозволяє одночасно виконувати кілька незалежних етапів обробки. Система переходить до фінальної стадії лише після того, як будуть отримані та успішно синхронізовані результати від усіх паралельних гілок. Такий підхід суттєво скорочує загальний час відповіді (latency), що критично важливо для систем із високим навантаженням.

**3. Модель на основі правил (Rules-Based)** Логіка переходів у цій моделі є динамічною і залежить від виконання певних умов або атрибутів вхідних даних. Це дозволяє системі адаптивно обирати оптимальний шлях обробки «на льоту», підвищуючи гнучкість сценаріїв обслуговування.

**4. Гібридна модель (Hybrid)** Комбінує вищезгадані підходи для досягнення максимальної ефективності у складних сценаріях. Певні блоки задач виконуються паралельно для швидкості, тоді як критичні операції з базою даних залишаються строго послідовними для гарантування ACID-властивостей.

Вибір конкретної моделі Workflow визначається характером задачі, складністю бізнес-логіки та жорсткістю вимог до часу відгуку та відмовостійкості системи.

### Графічне моделювання транзакційних сценаріїв

Етап проектування сервісу, що базується на складному робочому процесі, обов'язково включає візуалізацію сценаріїв його виконання. Найбільш ефективним інструментом для цього є UML-діаграми [23], які дозволяють декомпонувати систему на окремі компоненти. Такий підхід узгоджується з сучасними методиками, де UML-моделі дій використовуються для створення гнучких композицій веб-сервісів із можливістю повторного використання їхніх елементів.

Прикладом такої моделі є сценарій авторизації користувача (рис. 3.2).

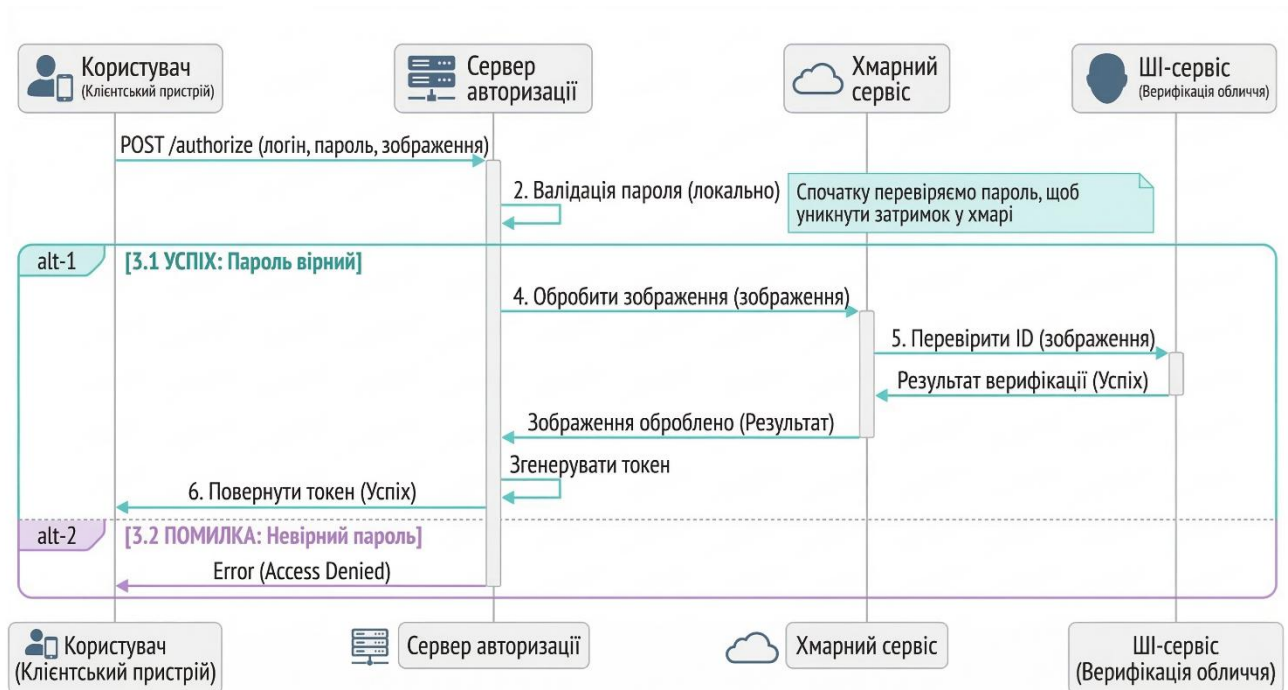


Рис. 3.2 – Процес авторизації

**Логіка роботи хмарного сервісу авторизації:** Розглянута система виконує ідентифікацію особи на основі аналізу фотографій. Оскільки обробка зображень методами машинного навчання (ML) [4, 116] потребує значних обчислювальних потужностей та хмарних ресурсів, алгоритм роботи був оптимізований для зниження витрат:

1. **Початковий запит:** Користувач надсилає ідентифікатор та пароль.

**2. Попередня фільтрація:** Першим етапом є серверна перевірка пароля. Це дозволяє відсіяти некоректні запити ще до звернення до високовартісних хмарних ML-моделей, мінімізуючи мережеві затримки.

**3. Розгалуження сценарію:** Якщо пароль невірний, система одразу відмовляє в доступі. У разі успіху – активується гілка обробки зображення та верифікація особи за біометричними даними.

**4. Фіналізація:** Після підтвердження ідентичності клієнту повертається авторизаційний токен.

Такий розподіл функцій дозволяє не лише описати структуру, а й підготувати базу для імітаційного та математичного аналізу, деталі якого будуть розкриті пізніше.

### **Ключові принципи організації робочих процесів**

Для ефективного функціонування в межах транзакційних систем робочі процеси повинні володіти наступним набором характеристик:

- **Високий рівень автоматизації:** Виключення ручного введення даних мінімізує людський фактор, прискорює обробку та підвищує загальний коефіцієнт корисної дії.

- **Автономність та ізолюваність:** Незалежність окремих потоків забезпечує стійкість системи [11] – критична помилка в одному процесі не призводить до каскадного збою всієї структури.

- **Адаптивність та гнучкість:** Система повинна легко інтегрувати нові функції або змінювати існуючі алгоритми відповідно до нових вимог.

- **Горизонтальна масштабованість:** Здатність архітектури нарощувати потужність при збільшенні обсягів транзакцій без втрати продуктивності.

### **Підходи до забезпечення надійності та математичне моделювання**

Розвиток інформаційних систем нерозривно пов'язаний з математичною формалізацією їхніх процесів. Для складних веб-структур часто використовується апарат напівмарківських процесів, що дозволяє ідентифікувати вузькі місця та оцінювати параметри QoS.

**Еквівалентність та реплікація мікросервісів:** Для максимізації відмовостійкості доцільно впроваджувати механізми еквівалентних вузлів. Якщо основний мікросервіс виходить з ладу, його роль миттєво бере на себе аналогічний компонент, що може працювати в іншому хмарному середовищі.

Загальна схема впровадження таких дублюючих вузлів представлена в структурі реплікації (рис. 3.3).

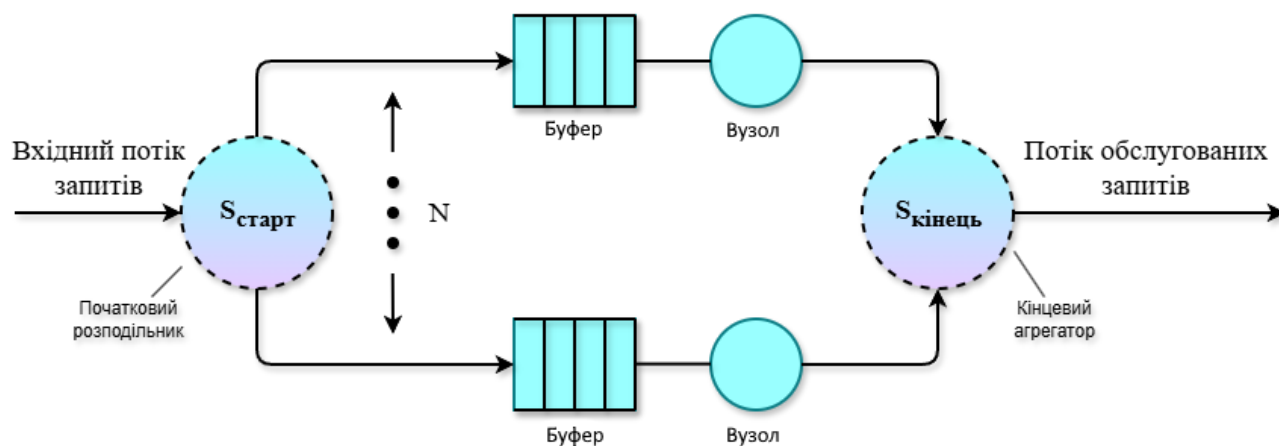


Рис. 3.3 – Загальна схема реплікації

Позначимо:  $S_{\text{старт}}$  – початковий вузол,  $S_{\text{кінець}}$  – кінцевий вузол, а  $N$  являє собою масштабовані шляхи та визначає ступінь реплікації завдань у початковому вузлі.

Сценарій відмовостійкості (рис. 3.4) передбачає, що у разі невдалої обробки в одній СМО система ініціює додаткову спробу в резервній гілці.

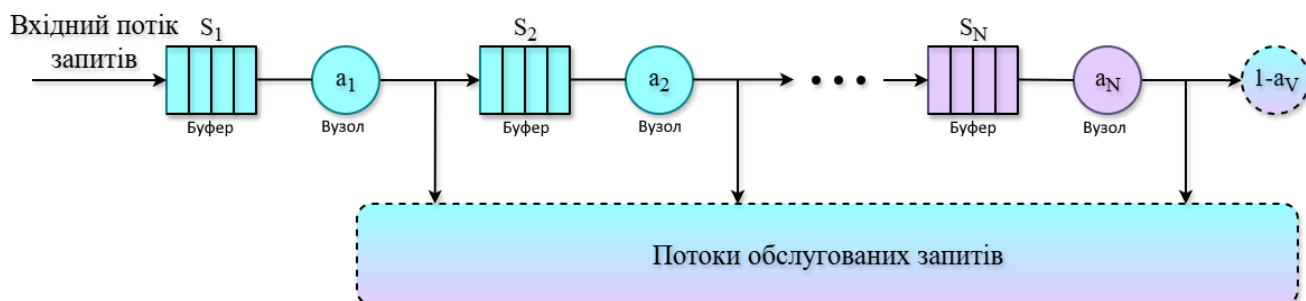


Рис. 3.4 – Сценарій реплікації з відмовостійкістю

**Особливості аналітичного оцінювання:** При побудові математичних моделей часто виникає дилема точності розрахунків. Існуючі підходи [10, 24, 27]

іноді ігнорують час, витрачений на невдалу обробку (час відмови вузла), або прирівнюють його до часу успішного виконання.

У цій роботі акцентується увага на тому, що неврахування часу відмови при аналізі еквівалентних гілок суттєво викривляє результати, оскільки ресурси, витрачені на хибний шлях, можуть значно відрізнятись від параметрів успішної транзакції.

### **Аналітична оцінка часових показників у відмовостійких структурах**

Для систем, що реалізують сценарії послідовного резервування (failover) [15, 24], ключовим показником є математичне сподівання часу відповіді. Якщо система активує вузли один за одним у разі відмови попереднього, сумарний час обслуговування  $T_{total}$  можна оцінити за наступною залежністю:

$$T_{total} = \frac{\sum_{i=1}^N p_i \cdot (\sum_{k=1}^i \tau_k \cdot \prod_{v=1}^{i-1} (1 - p_v))}{1 - \prod_{i=1}^N (1 - p_i)},$$

де  $N$  – загальна кількість доступних вузлів;  $p_i$  – показник надійності  $i$ -го елемента;  $\tau_i$  – тривалість обробки запиту на відповідному вузлі.

**Врахування затримок на сповіщення про збої:** На практиці процес ідентифікації помилки не є миттєвим. Невдала спроба обслуговування може завершитися швидше за успішну (переривання) або, навпаки, затягнутися через очікування тайм-ауту чи формування звіту про помилку. Для більш точного моделювання систем, що мають понад дві еквівалентні гілки, доцільно використовувати розширену формулу:

$$T = \frac{\sum_{i=1}^N \left( p_i \cdot \prod_{v=1}^{i-1} (1 - p_v) \cdot (\tau_i + \sum_{k=1}^{i-1} \tau_{err,k}) \right)}{1 - \prod_{i=1}^N (1 - p_i)},$$

де  $\tau_{err,k}$  – часові витрати на формування сигналу про відмову в  $k$ -му вузлі.

Загальна надійності такої мережевої структури  $P_{sys}$  визначається як імовірність успіху хоча б одного з  $N$  вузлів:

$$P_{sys} = 1 - \prod_{i=1}^N (1 - p_i).$$

### **Стратегія спекулятивного паралелізму**

Іншим підходом до підвищення продуктивності є спекулятивне виконання [24]. У цьому сценарії система не чекає відмови першого вузла, а ініціює обробку транзакції одночасно на всіх  $k$  гілках (рис. 3.5). Кінцевим результатом стає відповідь від того сервісу, який завершив обробку першим.

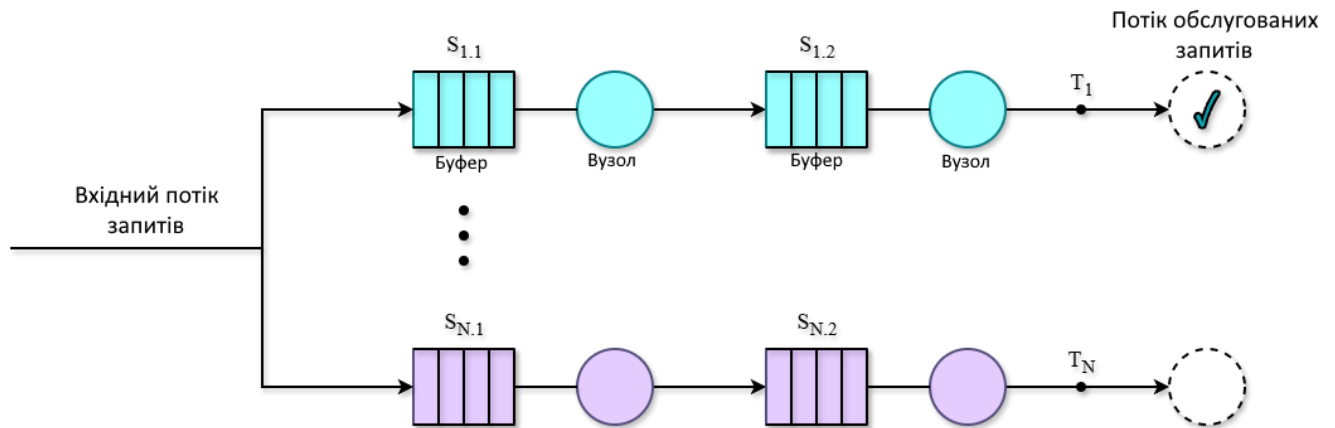


Рис. 3.5 – Схема сценарію спекулятивного паралелізму в реплікації

Як тільки одна з гілок  $S_{1,i}$  демонструє успішний фініш, робота інших гілок  $S_{j,i}$ , де  $j$  – кількість решти процесів) припиняється. Час відгуку в такій моделі залежить від результативності спроб:

$$T = \begin{cases} \min(T_{succ}) \text{ при } |T_{succ}| > 0 & \text{(найшвидший успіх);} \\ \max(T_{fail}) \text{ при } |T_{fail}| = 0 & \text{(відмова всіх гілок),} \end{cases}$$

де  $\tau_{succ}$  – набір значень часу успішної обробки в активованих гілках;  
 $\tau_{fail}$  – тривалість невдалого обслуговування.

**Порівняння з моделлю Fork-Join:** Описаний паралелізм має спільні риси з класичною моделлю Fork-Join [162], проте відрізняється логікою завершення. Якщо у Fork-Join система чекає на завершення абсолютно всіх потоків для синхронізації, то у спекулятивній моделі достатньо першого валідного результату. Це дозволяє використовувати аналогічний математичний апарат для оцінки характеристик, але з коригуванням на правило першої відповіді, що значно знижує середні затримки в хмарних середовищах.



### 3.2. Паралельне опрацювання запитів за методом Fork-Join

#### Концепція паралелізму та реплікації запитів

У сучасних розподілених системах [20, 24] паралельна обробка базується на клонуванні вхідного запиту на  $N$  незалежних вузлів. Такий підхід зазвичай переслідує дві стратегічні цілі:

**1. Дистрибутивне опитування:** Одночасне звернення до кількох автономних сервісів для збору агрегованої відповіді. Критичною особливістю тут є залежність: якщо бодай один вузол не здатний виконати завдання, весь процес деактивується, а запит маркується як необроблений.

**2. Декомпозиція завдань (Sharding):** Розподіл масиву даних на дрібніші сегменти, які опрацьовуються паралельно на різних обчислювальних потужностях. Це дозволяє радикально скоротити час виконання транзакції.

**Оцінка часу обробки при декомпозиції:** Якщо завдання розділене порівну між  $N$  ресурсами, ідеалізований час обробки  $\bar{\tau}$  можна визначити як:

$$\bar{\tau} = \frac{\tau_{total}}{N},$$

де  $\tau_{total}$  – час виконання задачі одним вузлом без розпаралелювання.

#### Архітектура Fork-Join та синхронізація

Для реалізації складних сценаріїв обслуговування використовується математичний апарат Fork-Join [164]. Він забезпечує збереження цілісності (атомарності) транзакцій у розгалужених мережах [24, 161].

Тоді паралельну обробку сценаріїв в такій архітектурі можна представити як на рис. 3.6.

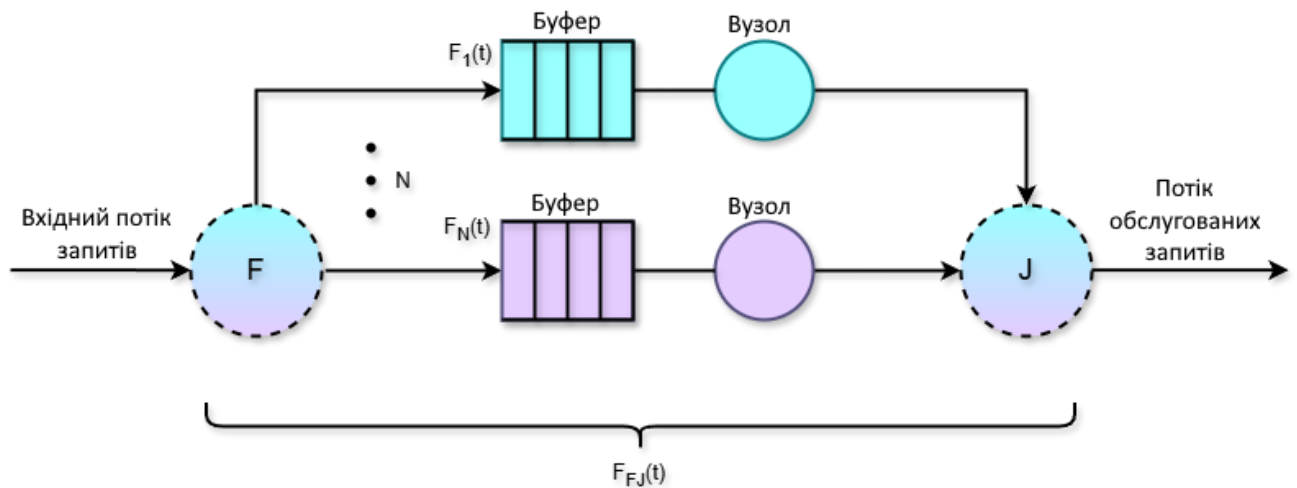


Рис. 3.6 – Схема паралельної обробки запитів

#### Алгоритм функціонування:

- **Стадія Fork:** Вхідний потік потрапляє в абстрактний блок розділення, де запит миттєво тиражується або сегментується на  $N$  ідентичних фрагментів, що спрямовуються до відповідних підсистем.
- **Стадія Обробки:** Кожен вузол  $1 \dots N$  СМО.
- **Стадія Join:** Після завершення обробки всі репліки потрапляють у чергу синхронізації. Об'єднання в єдину відповідь відбувається лише тоді, коли всі фрагменти успішно опрацьовані.

**Часові характеристики синхронізації:** У такій моделі загальний час перебування запиту в системі  $T_{sync}$  визначається за правилом «найповільнішої ланки» для успіху та «найшвидшої ланки» для відмови:

$$T_{sync} = \begin{cases} \min(\tau_{succ,i}), & \text{якщо гілка успішна;} \\ \max(\tau_{fail,k}), & \text{якщо виникла помилка,} \end{cases}$$

де  $\tau_{succ}$  – час валідної обробки в кожній активній гілці;  $\tau_{fail}$  – момент фіксації першого збою, що зупиняє весь процес.

#### Переваги імплементації та складність моделювання

Використання паттерна Fork-Join при проєктуванні телекомунікаційних структур (рис. 3.7) надає ряд переваг:

- **Оптимізація ресурсів:** Максимальне завантаження доступних вузлів через паралельне обслуговування черг.

- **Масштабованість:** Можливість легкого додавання нових підсистем для обробки великих обсягів даних.
- **Динамічна адаптивність:** Система гнучко реагує на зміну інтенсивності вхідного потоку, перерозподіляючи навантаження.



Рис. 3.7 – Паттерн Fork-join у телекомунікаційних структурах та його переваги

Попри очевидні плюси, математичне моделювання Fork-Join структур є нетривіальним завданням. Воно потребує врахування стохастичної природи потоків, теорії ймовірностей та методів числового аналізу для точної оцінки QoS в мікросервісних середовищах.

### Методи апроксимації для складних мереж обслуговування

Для аналізу мереж типу  $G/G/1$  за різних режимів інтенсивності (від низького до критичного навантаження) доцільно застосовувати методи, що базуються на інтерполяції середнього часу перебування заявки [149, 155]. Основна ідея полягає у використанні граничних значень параметрів системи для отримання наближеного результату.

Для систем із Марківським вхідним потоком та експоненціальним обслуговуванням  $G/M/1$  [27, 152, 155] апроксимоване значення часу  $E[W]_M$  набуває вигляду:

$$E[W]_M \approx \left[ H_M + \left( \sum_{i=1}^M \binom{M}{i} (-1)^{i-1} \sum_{m=1}^i \binom{i}{m} \frac{(m-1)!}{i^{m+1}} - H_M \right) \frac{\lambda_{in}}{\mu} \right] \frac{1}{\mu - \lambda_{in}},$$

де  $\lambda_{in}$  – інтенсивність входу;  $H_M$  – гармонічне число порядку  $M$ .

Альтернативні підходи [27, 156, 157] дозволяють розглядати системи з довільним законом розподілу M/G/1 або використовувати апарат мереж Петрі [20] для опису паралельних структур. У класичному випадку для мережі, що складається з двох послідовних вузлів M/M/1, точне значення часового показника  $T_2$  визначається через базовий час  $T_1$ :

$$T_2 = \frac{12 - \psi}{8} T_1, \quad \text{де} \quad \psi = \frac{\lambda_{in}}{\mu_{srv}}.$$

### Граничні оцінки та наближені обчислення

У межах дослідження часових характеристик важливо встановити верхню та нижню межі для показника  $T_N$ . Верхня границя базується на припущенні про повну незалежність фрагментів запиту, тоді як нижня відповідає ідеальній системі без черг:

$$H_N \cdot \frac{1}{\mu_{srv}} \leq T_N \leq H_N \cdot T_1.$$

Беручи до уваги логарифмічну швидкість зростання гармонійного ряду  $O(\ln N)$ , для систем із кількістю вузлів  $N \geq 2$  отримано наближений вираз:

$$T_N \approx \left( \frac{H_N}{H_2} + \frac{4}{11} \left( 1 - \frac{H_N}{H_2} \right) \psi \right) T_2.$$

Функція розподілу для мереж із реплікацією  $m$  описується через коефіцієнти  $c_i$ :

$$F_{sys}(t) \approx \sum_{i=1}^N (c_i - c_{i+1}) \prod_{j=1}^i F_{node,j}(t).$$

Ці коефіцієнти розраховуються на основі ймовірності перебування запиту в черзі  $p_{queue}$ , що дозволяє враховувати стохастичну природу системи.

### Аналіз паралельної обробки автономними сервісами

При розрахунку вхідно-вихідних характеристик (VBX) для систем із двома паралельними автономними вузлами використовуються складні аналітичні вирази

для функції розподілу вірогідності (ФРВ) та щільності розподілу вірогідності (ЩРВ). Для моделі M/M/1 ці функції мають вигляд:

$$F_{Fork2}(s) = \frac{(\lambda_{in} - s - \mu) \left( -\lambda_{in} + 4\mu + \frac{4\mu(-2\lambda_{in} + s + 2\mu)}{s} \right)}{4\mu \cdot (\lambda_{in} - s - \mu)(2\mu + s - 2\lambda_{in})} + \dots f_{Fork2}(s) =$$

$$= \frac{8\mu^3 - 16\lambda_{in}\mu^2 + (\lambda_{in}s + 8\lambda_{in}^2)\mu - \lambda_{in}^2s}{4\mu \cdot (\mu + s - \lambda_{in})(2\mu + s - 2\lambda_{in})}.$$

Попри традиційне застосування моделі Fork-Join до елементарних систем типу M/M/1, її математичний інструментарій залишається ефективним і для складніших конфігурацій, таких як M/G/1 або G/G/1. Проте для систем із високим коефіцієнтом реплікації  $N > 2$  або нетиповими розподілами аналітичні розрахунки стають занадто громіздкими.

У таких випадках я вважаю за доцільне переходити до ІМ. Саме тому в межах даної роботи було розроблено програмну модель сценарію Fork-Join, яка дозволяє гнучко налаштовувати функціонал для  $N$  вузлів і отримувати точні часові метрики без виведення надскладних формул.

### 3.3 Моделювання робочого процесу з використанням математичного апарату мереж Джексона

Ключовим завданням цієї наукової роботи є проєктування комплексу моделей, що дозволяють обчислювати ВВХ для сервісів транзакційного типу. Аналіз проводиться з урахуванням специфіки робочих процесів у системах, де застосовується механізм паралельної обробки вхідних запитів, що детально розбирався у попередніх розділах.

На рис. 3.8 продемонстровано структуру функціональної моделі, яка описує операційний цикл обслуговування послідовного стеку локальних транзакцій. Ця топологія мережі базується на використанні ідентичних (еквівалентних) вузлів обслуговування. Для отримання релевантних часових характеристик системи необхідно виконати поетапний аналіз динаміки переміщення заявок між окремими компонентами ланцюга.

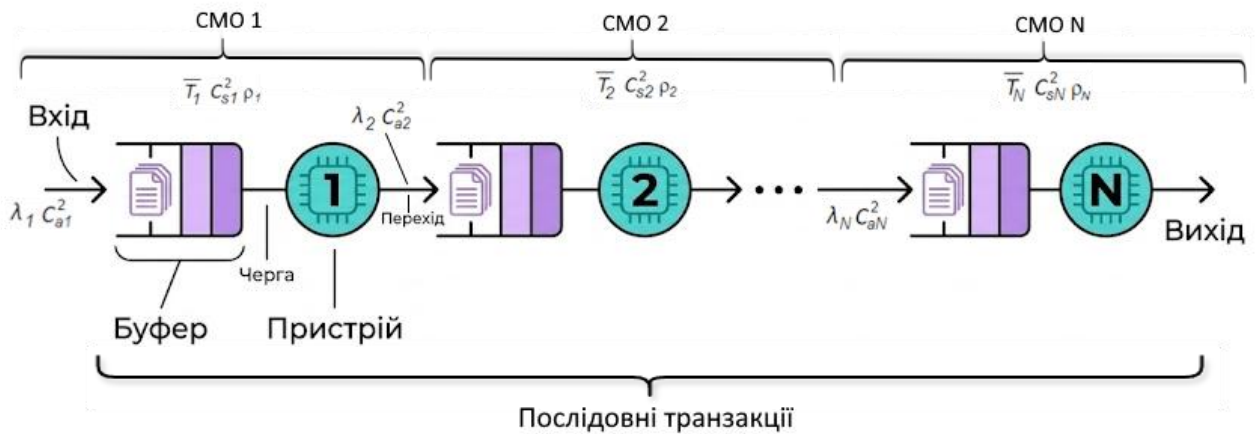


Рис. 3.8 – Модель обслуговування послідовного стеку локальних транзакцій

Для визначення інтегрального показника затримки  $T_s$ , що виникає впродовж усього шляху від моменту надходження до завершення обробки у заданій послідовності СМО, використовується метод адитивного підсумовування локальних затримок  $\bar{T}_i$  для кожного конкретного вузла.

Розрахункову модель можна представити наступним чином.

Визначимо загальний час очікування в системі як кумулятивний ефект від роботи всіх  $N$  ланок:

$$T_{\Sigma} = \sum_{i=1}^N \bar{T}_i,$$

де параметр  $i$  варіюється в межах  $i \in \{1, 2, \dots, N\}$ .

Для спрощення математичного опису та подальшого аналізу такої послідовності вузлів доцільно застосовувати принцип зчеплення (агрегації) [25, 51, 56], як це ілюструє рис. 3.9. Фактично, складний ланцюг послідовних СМО трансформується в одну узагальнену (еквівалентну) одиницю. Такий вузол описується за допомогою ФРВ, що характеризує повний час проходження через систему:

$$F_x(t) = P(T \leq t).$$

Це дозволяє значно знизити обчислювальну складність при збереженні точності моделювання часових параметрів системи.

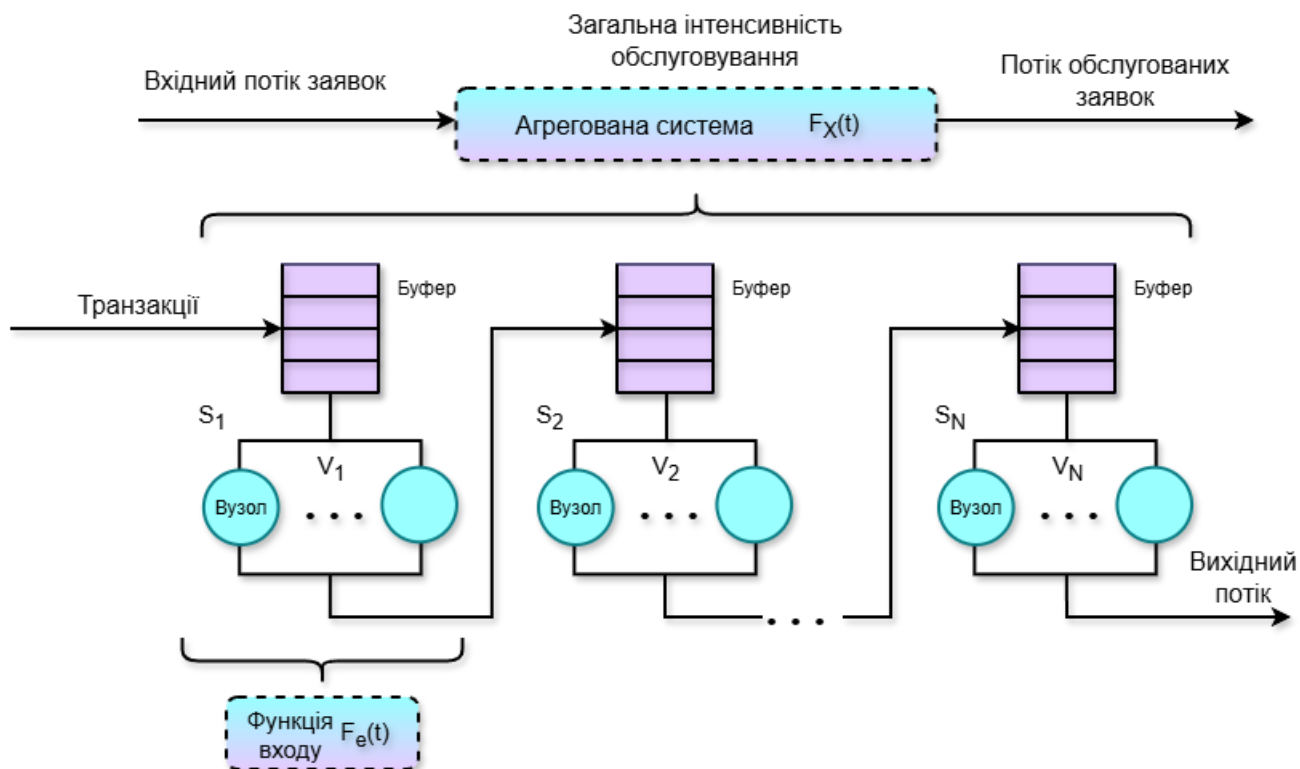


Рис. 3.9 – Робочий процес системи, модель ланцюга локальних транзакцій

Для реалізації повноцінного аналізу часової динаміки мережі, що передбачає обчислення математичного очікування та значень варіативності (дисперсії), першочерговим завданням є ідентифікація закону, за яким розподіляється час перебування транзакції в системі (визначення щільності ймовірності).

Спираючись на класичні наукові підходи [12, 24, 27] до моделювання подібних процесів, результуючий вираз для ФРВ часового інтервалу, необхідного для подолання всього каскаду мікросервісних компонентів, можна отримати шляхом послідовної операції згортки індивідуальних показників кожного вузла системи:

$$F_{\Sigma}(t) = F_{e1}(t) * F_{e2}(t) * \dots * F_{eN}(t) = \prod_{i=1}^N F_{ei}(t).$$

У даному контексті  $F_{ei}(t)$  характеризує ймовірність того, що обробка запиту в межах окремого  $i$ -го сегмента мережі  $i = 1, \dots, N$  буде завершена до моменту часу  $t$ .

Для кожного ізольованого елемента мережевої структури повна функція розподілу часу перебування формується як композиційне поєднання (згортка) двох часових складових: періоду очікування в черзі та безпосередньо етапу виконання операції:

$$F_e(t) = \int_0^t F_w(t-x) dF_s(x).$$

У символічному вигляді ці операції записуються як:

$$F_e(t) = F_w(t) * F_s(t).$$

Аби спростити процедуру математичних перетворень над згортками, доцільно перейти від часової області до простору комплексних зображень. Для цього використовується апарат перетворення Лапласа-Стілтєса [147]. Відповідне зображення  $f(s)$  для вихідних функцій та їхніх щільностей визначається через наступний інтегральний вираз:

$$f(s) = \int_0^\infty e^{-st} dF(t) \quad \text{при} \quad \operatorname{Re}(s) \geq 0.$$

Якщо розглядати базові системи з експоненціальним розподілом типу М/М/1, то аналітичні моделі для часу очікування та загального періоду знаходження в системі набувають такого вигляду:

**Для стадії очікування:**

$$F_w(t) = 1 - \frac{\lambda}{\mu} e^{-(\mu-\lambda)t}.$$

**Для повного циклу в вузлі:**

$$F_e(t) = 1 - e^{-(\mu-\lambda)t}.$$

Використовуючи отриману функцію  $F_x(t)$ , ми можемо вивести фундаментальні моменти розподілу для всього операційного циклу:

**Математичне очікування часу перебування:**

$$M[T] = \int_0^\infty (1 - F_x(t)) dt. \quad (3.1)$$

**Дисперсія (варіація) часових характеристик:**

$$D[T] = \int_0^\infty t^2 dF_x(t) - (M[T])^2. \quad (3.2)$$



## Аналіз розгалужених топологій та використання спільних обчислювальних ресурсів

При переході до складніших конфігурацій вкрай важливо враховувати мінливість інтенсивності потоків (рис. 3.10).

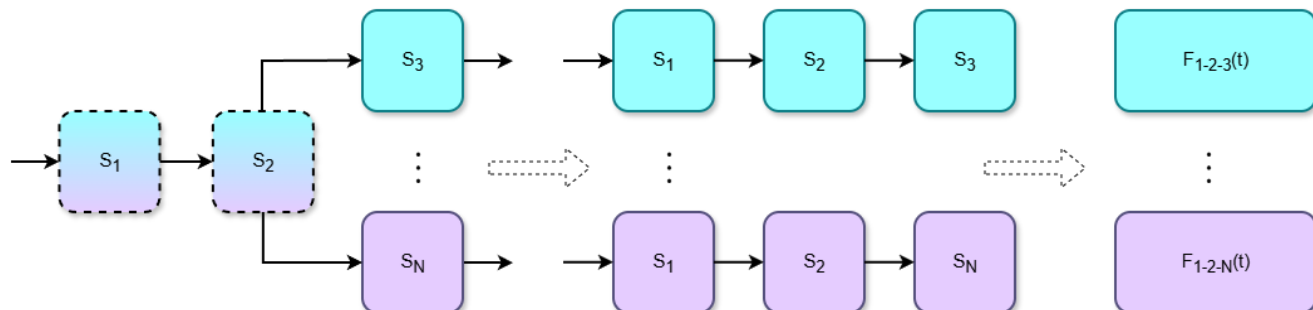


Рис. 3.10 – Модель розгалуження сценаріїв робочих процесів системи

Сумарна інтенсивність  $\lambda_j$ , що спрямовується до певного вузла  $j$  від попереднього вузла  $i$ , визначається за допомогою матричного оператора ймовірностей переходів  $R = |p_{ij}|$ :

$$\overrightarrow{\lambda_{total}} = \overrightarrow{\lambda_{in}} \cdot R.$$

У розгалужених мережах зображення Лапласа-Стілтєса для функціонального стану вузла  $j$  може бути представлено рекурсивною моделлю:

$$f_{ij}(s) = \frac{p_{ij}f_i(s)}{1 - (1 - p_{ij})f_i(s)}.$$

Модель сценарію спільного використання ресурсів робочими процесами, на відміну від рис. 3.10, дещо ускладнюється (рис.3.11).

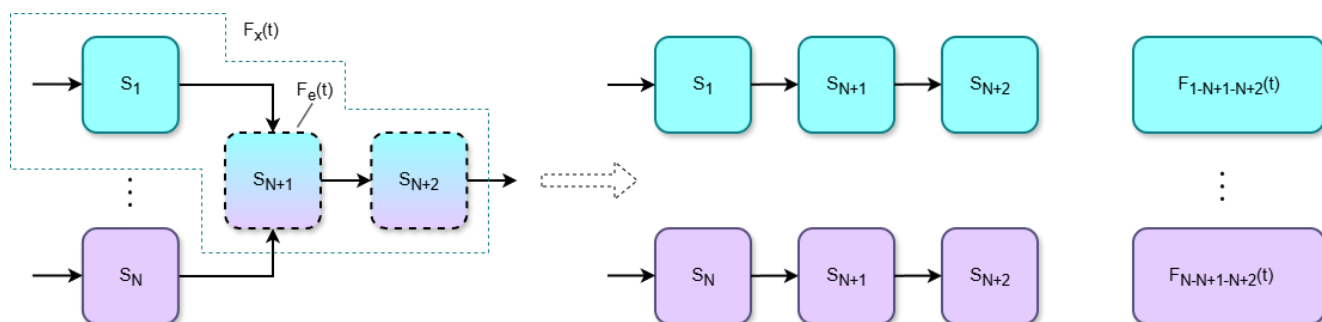


Рис. 3.11 – Модель сценарію спільного використання ресурсів робочими процесами

Коли один системний ресурс задіяний для паралельного обслуговування кількох незалежних потоків транзакцій, інтегральна функція розподілу  $F(t)$  для такого агрегованого процесу обчислюється як:

$$F(t) = 1 - \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \left[ (1 - F_1(t)) \int_t^\infty (1 - F_2(u)) du + (1 - F_2(t)) \int_t^\infty (1 - F_1(u)) du \right].$$

### Динаміка вихідних потоків та їхні властивості

Одним із ключових аспектів аналізу послідовних СМО є вивчення того, як змінюється структура потоку на виході. Позначимо через  $d(t)$  щільність розподілу пауз між заявками після обробки у вузлі  $i$ , а через  $D(s)$  – її відповідник у просторі Лапласа-Стілтєса. Загальна модель для  $D(s)$  базується на зважуванні станів активності та простою вузла:

$$D(s) = (1 - \rho)D(s)_{\text{вільний}} + \rho D(s)_{\text{зайнятий}}$$

або, за допомогою параметрів вхідного трафіку  $I'(s)$  та характеристик продуктивності  $B(s)$ :

$$D(s) = (1 - \rho)I'(s)B(s) + \rho B(s).$$

Згідно з відомими теоретичними положеннями [18, 131] про властивості вихідних потоків, для стаціонарних систем формату М/М/ν вихідний трафік зберігає Пуассонівський характер з незмінним параметром інтенсивності  $\lambda$ . В операторній формі це виглядає так:

$$D(s) = \frac{\lambda}{s + \lambda}.$$

Відповідно, у часовому вимірі розподіл інтервалів залишається експоненціальним:

$$D(t) = 1 - \exp(-\lambda t) \quad \text{при } t \geq 0.$$

Ця унікальна властивість дозволяє декомпонувати загальну мережу на окремі автономні сегменти, що суттєво полегшує розрахунки. Для складніших систем типу G/G/1 цей принцип стає основою для розробки апроксимаційних методик, які будуть деталізовані далі.

### 3.4. Обчислення параметрів затримки систем довільного типу

Раніше ми розглядали готові математичні апарати для аналізу систем типу  $M/M/1$ , де процеси очікування описуються досить просто. Однак для систем загального виду  $G/G/1$ , які набагато краще відображають реальні умови експлуатації, доцільно застосовувати методику, що базується на фундаментальних працях з теорії масового обслуговування. Зокрема, у відомих дослідженнях у цій галузі [141, 144] було виведено інтегральне рівняння Ліндлі [142, 143], яке дозволяє оцінювати затримки транзакцій для будь-яких значень аргументу:

$$W(y) + W_-(y) = \begin{cases} \int_{-\infty}^y W(y-u)dC(u) & \text{при } y \geq 0; \\ 0 & \text{при } y < 0, \end{cases}$$

де функція  $C(u)$  виступає інтегральним законом розподілу випадкової величини  $u$ .

Для знаходження розв'язку цього рівняння найчастіше обирають метод спектрального аналізу. При цьому зображення за Лапласом [147] для функції розподілу  $W_-(y)$  ми позначимо як  $\Phi_-(s)$ , а для  $W(y)$  – відповідно  $\Phi_+(s)$ . Остання величина безпосередньо корелює з показником густини розподілу періоду очікування  $W(s)$  через наступне співвідношення:

$$\Phi_+(s) = \frac{W^*(s)}{s}.$$

Оскільки тривалість очікування за своєю природою є невід'ємною величиною, права частина математичної моделі повинна містити операцію згортки між функцією розподілу часу очікування та густиною випадкової величини. Таким чином, отримуємо вираз:

$$\Phi_+(s) + \Phi_-(s) = \Phi_+(s) \cdot C^*(s).$$

При цьому результат перетворення Лапласа-Стільтєса для функції  $C(u)$  можна представити як добуток двох складових:

$$C^*(s) = A^*(-s) \cdot B^*(s).$$

Отже, для аналітичного опису поведінки системи в області збіжності  $Re(s) < D$ , шукана функція  $\Phi_-(s)$  набуває вигляду:

$$\Phi_{-}(s) = \Phi_{+}(s)[A^{*}(-s)B^{*}(s) - 1].$$

Шляхом алгебраїчних перетворень це рівняння можна представити через відношення допоміжних аналітичних функцій  $\Psi_{+}(s)$  та  $\Psi_{-}(s)$ :

$$\Phi_{-}(s) = \Phi_{+}(s) \frac{\Psi_{+}(s)}{\Psi_{-}(s)}.$$

Тут важливо зазначити, що  $\Psi_{+}(s)$  є аналітичною за параметром  $s$  у правій півплощині  $Re(s) > 0$  і не має там нульових значень, тоді як  $\Psi_{-}(s)$  зберігає аналітичність в області  $Re(s) < D$ .

У науковій літературі [27, 152, 155], що присвячена цій проблематиці, наведено чимало прикладів розрахунків для систем M/M/1 та G/M/1. Як практичну ілюстрацію для нашого дослідження ми оберемо модель  $H_2/H_2/1$  (система з гіперекспоненціальним розподілом другого порядку [155]). Такий підхід дозволяє нам у подальшому розробити алгоритм для ланцюга послідовних вузлів, використовуючи апроксимаційні методи [155, 158] для опису інтервалів між надходженнями транзакцій.

Вирази для перетворення Лапласа-Стільтєса, що описують процеси обслуговування та надходження запитів за гіперекспоненціальним законом  $H_2$ , мають таку структуру:

**Для потоку надходжень:**

$$A(s) = \frac{\lambda_1 p}{\lambda_1 + s} + \frac{\lambda_2 (1 - p)}{\lambda_2 + s},$$

де  $p$  – ймовірнісна вага першої експоненціальної складової з інтенсивністю  $\lambda_1$ .

**Для процесу обслуговування (аналогічно):**

$$B(s) = \frac{\mu_1 q}{\mu_1 + s} + \frac{\mu_2 (1 - q)}{\mu_2 + s},$$

де  $q$  визначає ймовірність обробки за першим сценарієм з інтенсивністю  $\mu_1$ .

Проводячи процедуру розкладу згідно з раніше визначеними виразами, ми прийдемо до дробово-раціональної форми:

$$\frac{\Psi_+(s)}{\Psi_-(s)} = \left( \frac{\lambda_1 p}{\lambda_1 - s} + \frac{\lambda_2(1-p)}{\lambda_2 - s} \right) \left( \frac{\mu_1 q}{\mu_1 + s} + \frac{\mu_2(1-q)}{\mu_2 + s} \right) - 1.$$

Після зведення до спільного знаменника та групування членів, отримуємо:

$$\frac{\Psi_+(s)}{\Psi_-(s)} = \frac{s(s^3 - as^2 - bs - c)}{(\lambda_2 - s)(\lambda_1 - s)(\mu_1 + s)(\mu_2 + s)}.$$

Для обчислення параметрів цієї моделі використовуються коефіцієнти:

- $a = \lambda_1 + \lambda_2 - \mu_1 - \mu_2$ ;
- $b$  та  $c$  розраховуються як комбінації добутків інтенсивностей та ймовірнісних ваг  $p$  і  $q$  (повні формули наведено у розрахунковому блоці).

Щоб знайти остаточний розв'язок, необхідно ідентифікувати корені та полюси функції  $\Psi_+(s)$ , які розташовані в лівій півплощині. Для спрощеного випадку системи G/M/1 результат можна представити у вигляді:

$$W(t) = 1 - \left(1 - \frac{x}{\mu}\right) e^{-xt} \quad \text{при } t \geq 0,$$

де допоміжний параметр  $x$  визначається через граничне значення  $x = \mu \cdot \lim_{s \rightarrow 0} \frac{\Psi_+(s)}{s}$ .

В умовах критичного навантаження системи  $\rho \rightarrow 1$ , модель G/G/1 може бути апроксимована наступним виразом:

$$A^*(-s)B^*(s) - 1 \approx s(s - s_0) \frac{\sigma_a^2 + \sigma_b^2}{2},$$

де  $\sigma_a^2$  та  $\sigma_b^2$  відображають дисперсію часових інтервалів надходження та обслуговування відповідно.

Середній час очікування  $\bar{W}$  у такому режимі становить:

$$\bar{W} = \frac{\sigma_a^2 + \sigma_b^2}{2(1 - \rho)} \cdot \frac{1}{t}.$$

У сучасних наукових розробках [156, 158] також пропонуються теоретико-інформаційні моделі [142, 143] для систем M/G/1, де знання про густину ймовірності обмежене лише кількома початковими моментами. Такі апроксимації дозволяють вважати прості системи типу M/M/1 цілком адекватним наближенням

для складніших випадків при відомих перших моментах розподілу. У цьому разі густина часу очікування записується як:

$$f_w(t) = (1 - \rho)\delta(t) + \lambda(1 - \rho)e^{-(1-\rho)\mu t}.$$

Подальша розробка алгоритмів для розрахунку характеристик систем загального виду вимагає розгляду методів апроксимації для тих випадків, де функції ймовірності не задані у явному вигляді або мають складні обмеження для інтегрування.

### 3.5 Оцінка часових затримок у послідовності локальних транзакцій

Враховуючи попередні теоретичні викладки, розрахунок параметрів затримки базується на застосуванні оператора перетворення Лапласа-Стілтєса до ФРВ інтервалів між надходженнями та безпосередньої тривалості обслуговування. Даний методологічний підхід є універсальним для аналізу систем формату G/G/1, де закони розподілу вхідного трафіку та часу обробки можуть бути довільними. Проте, при проєктуванні багаторівневого ланцюга локальних транзакцій виникає специфічна складність: необхідність точної ідентифікації ФРВ часових проміжків між заявками, що надходять до кожного наступного сегмента  $j$  (де  $j = 2, \dots, Y$ ) у загальній послідовності.

На стадії моделювання вихідного трафіку для серійної структури вузлів  $D(s)$ , про яку йшлося в попередніх підрозділах, критично важливо визначити характеристику  $I'(s)$ . Цей параметр представляє собою образ Лапласа-Стілтєса, як ФРВ часу очікування чергового запиту, що фактично описує тривалість простою обчислювального ресурсу. У загальному випадку для систем типу G/G/1 отримання прямого аналітичного розв'язку цієї задачі є проблематичним. Це зумовлено тим, що в таких моделях тривалість періоду пасивності системи знаходиться у прямої статистичній залежності від динаміки попереднього циклу зайнятості.

На відміну від класичних систем M/M/1, де вхідний потік має властивість відсутності післядії (незалежності), системи загального виду характеризуються внутрішніми кореляційними зв'язками. Через це єдиного універсального розв'язку

для них не існує. Точні математичні моделі вдається побудувати лише для окремих випадків, таких як M/M/1, або специфічних конфігурацій типу M/G/1 та G/D/1.

Для вузлів класу M/G/1, які передбачають довільний характер тривалості обслуговування  $B(t)$ , ґрунтовний аналіз було проведено у фундаментальних галузевих працях. Дана група моделей є досить широкою і може охоплювати як системи з детермінованим часом обробки M/D/1, так і прості експоненціальні моделі M/M/1. Результируючі показники для таких систем, що враховують варіативність процесів обслуговування, визначаються за допомогою модифікованих виразів Полячека-Хінчина [27, 156].

Зокрема, середню кількість транзакцій у системі (довжину черги)  $N$  можна подати як суму коефіцієнта завантаження та нормованої дисперсії:

$$N = \rho + \frac{\rho^2 \cdot (1 + C_s^2)}{2 \cdot (1 - \rho)}.$$

Відповідно, інтегральний показник тривалості очікування  $W$  розраховується через середній час обслуговування  $\bar{T}$  наступним чином:

$$W = \left[ \frac{\rho \cdot (1 + C_s^2)}{2 \cdot (1 - \rho)} \right] \bar{T}.$$

У цих рівняннях  $C_s$  виступає як коефіцієнт варіації часових витрат на обробку запиту. Для певних конфігурацій систем (наприклад, при гіперекспоненціальній апроксимації) цей коефіцієнт може бути деталізований через параметри форми  $p$ :

$$C_s = \sqrt{\frac{2p(1-p) + (1-2p)^2}{2p(1-p)}}.$$

Важливою перевагою систем типу M/G/1 є те, що період їхнього простою описується експоненціальним законом. Це дозволяє отримати точне зображення Лапласа-Стілтєса для вхідного потоку наступного вузла, що значно спрощує аналіз всього ланцюга. Відповідний вираз для образу Лапласа-Стілтєса, як ФРВ періоду пасивності має вигляд:

$$I'(s) = \frac{\mu}{s + \mu}.$$

Для повноцінної реалізації методу згорток при оцінці затримок у системі M/G/1 використовуються операторні форми рівняння Полячека-Хінчина. Вони дозволяють знайти зображення Лапласа-Стілтєсса для:

**Тривалості перебування в черзі:**

$$w(s) = \frac{s \cdot (1 - \rho)}{s - \lambda + \lambda \cdot \beta(s)}.$$

**Загальної затримки транзакції у вузлі СМО:**

$$\epsilon(s) = \frac{s \cdot (1 - \rho) \cdot \beta(s)}{s - \lambda + \lambda \cdot \beta(s)}.$$

Тут параметр  $\beta(s)$  відповідає перетворенню Лапласа-Стілтєсса, як ФРВ часу безпосереднього обслуговування, а  $\lambda$  та  $\rho$  відображають інтенсивність трафіку та рівень утилізації вузла відповідно.

Завдяки тому, що окремі сегменти M/G/1 у межах єдиного транзакційного ланцюга можна вважати стохастично незалежними, загальний час проходження запиту через усю систему успішно обчислюється за допомогою процедури послідовної згортки індивідуальних часових характеристик кожного вузла.

### 3.6. Оцінювання стану послідовних мереж систем G/M/1

Під час математичного опису мереж, де вхідний трафік базується на довільних імовірнісних законах, дослідники стикаються із серйозною проблематикою ідентифікації параметрів вихідного потоку. Навіть якщо процес безпосередньої обробки у вузлі підпорядковується класичному експоненціальному розподілу, структура потоку заявок після виходу із системи зазнає складних трансформацій.

Спрощена гіпотеза про те, що вихідний трафік можна апроксимувати аналогічним експоненціальним законом (де інтенсивність просто прирівнюється до середнього значення інтервалів надходження), є методологічно помилковою. Таке припущення не враховує динамічну мінливість системи та реальні умови її експлуатації, що в кінцевому підсумку призводить до суттєвих розбіжностей у результатах.



Моделі класу G/M/1 (рис. 3.12), які працюють із нерегулярним характером надходження запитів, являють собою складні обчислювальні сценарії. Оскільки для них не існує універсальних аналітичних інструментів для отримання точних розв'язків, у сучасній науковій практиці [27, 99, 152] було вироблено кілька альтернативних підходів. Зокрема, ці методики зосереджені на знаходженні перетворень Лапласа-Стілтєса для функцій розподілу періодів пасивності (простою) системи, що дозволяє досягти необхідної точності через апроксимаційні алгоритми.

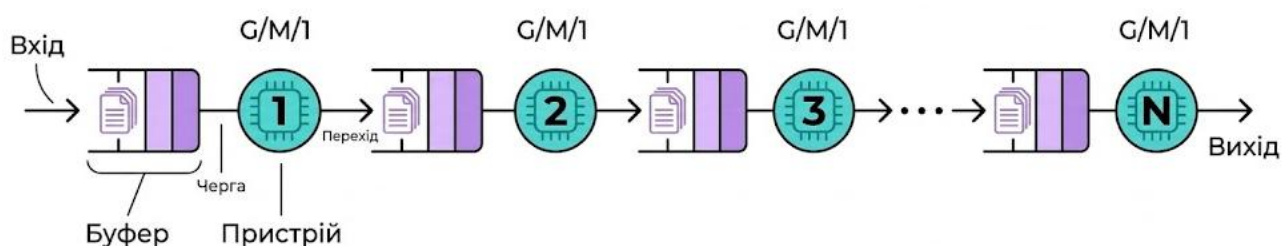


Рис. 3.12 – Послідовна ланцюгова система G/M/1

Як вхідні характеристики системи було визначено вхідний потік як гіперекспоненціальний розподіл шостого порядку  $H_6$  з експоненціальним часом обслуговування. Оцінка часу очікування обслуговування була отримана за допомогою методу, описаного в розділі 2 і приведена на рис. 3.13.

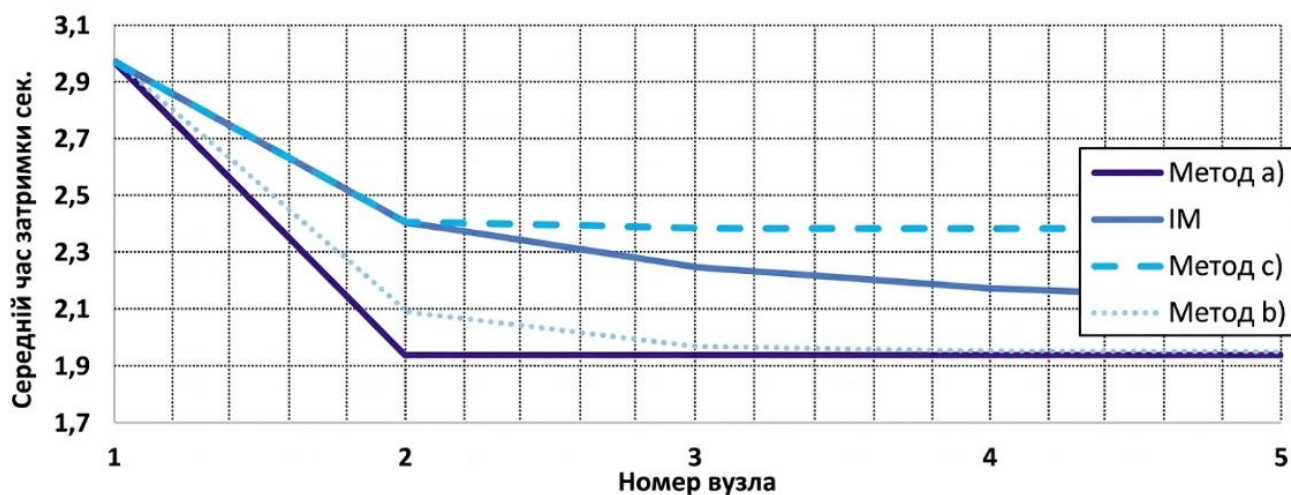


Рис. 3.13 – Середній час очікування на кожному вузлі в залежності від обраного методу апроксимації

Графічна інтерпретація результатів моделювання представлена на рис. 3.13 дозволяє зробити важливі висновки щодо точності існуючих підходів:

а) **Експоненціальне наближення для  $I'(s)$ .** Аналіз демонструє, що використання класичного експоненціального розподілу для апроксимації параметрів затримки має критичний недолік. Даний метод практично ігнорує специфіку потоків, що характеризуються високою дисперсією (варіативністю). Це призводить до того, що під час розрахунку часових показників очікування на другому вузлі системи виникає суттєва похибка, яка може сягати близько 20%.

б) **Моделювання періоду пасивності на основі інтервалів надходження.** Для опису вихідного потоку  $D(s)$  можна застосувати комбіновану модель, що враховує стан завантаженості системи  $\rho$ . Результируючий вираз для перетворення Лапласа-Стілтєсса для вихідної ФРВ формується як зважена сума характеристик вхідного потоку  $A(s)$  та параметрів обслуговування  $B(s)$ :

$$D(s) = B(s) \cdot [(1 - \rho) \cdot A(s) + \rho].$$

У випадках, коли вхідний потік описується гіперекспоненціальним законом (що є поширеним сценарієм для складних мереж), значення  $A(s)$  визначається згідно з відповідними аналітичними виразами для цього класу розподілів.

в) **Апроксимаційне припущення щодо залишкового часу.** У фаховій літературі [155, 158] часто використовується підхід, де тривалість періоду простою ототожнюється із залишковим значенням часового проміжку між вхідними запитами. Логічне обґрунтування цього методу базується на стохастичній властивості відновлення: оскільки тривалі інтервали часу між подіями займають фізично більші сегменти на часовій осі, ймовірність того, що випадково обраний момент часу  $t$  потрапить саме у довгий проміжок, є значно вищою, ніж потрапляння у короткий інтервал.

Виходячи з цієї концепції, перетворення Лапласа-Стілтєсса для періоду бездіяльності  $I'(s)$  можна представити у вигляді диференційованої моделі, що залежить від порогу завантаження системи:

$$I'(s) = \begin{cases} \frac{1 - I'_a(s)}{s\bar{x}} & \text{якщо } \rho \leq 0.5; \\ \frac{1 - B(s)}{s\bar{m}} & \text{якщо } \rho > 0.5. \end{cases}$$

Для коректного використання цієї моделі введемо такі параметри:

- $\bar{m}$  – середній час, що витрачається на безпосередню обробку запиту;
- $B(s)$  – образ Лапласа-Стільтєса для ФРВ тривалості обслуговування;
- $I'_a(s)$  – апроксимаційний показник для зображення Лапласа-Стільтєса

залишкового часу, що розраховується через параметр  $\epsilon$ :

$$I'_a(s) = \frac{\epsilon}{\epsilon + s}.$$

При цьому коефіцієнт  $\epsilon$  визначається через взаємозв'язок середнього інтервалу надходження  $\bar{t}$  та середнього залишкового часу  $\bar{x}$ :

$$\epsilon = \frac{1}{\bar{x}} = \frac{1}{\bar{t} - \bar{x}}.$$

Така структура розрахунків дозволяє більш гнучко адаптувати модель до реальних умов експлуатації системи, враховуючи як низькі, так і високі рівні навантаження на обчислювальні вузли.

Дослідження продемонструвало, що розглянутий метод забезпечує досить високу кореляцію значень на другому етапі (вузлі) мережевої структури. Проте, при подальшому русі транзакцій по ланцюгу (від 2-го до  $n$ -го вузла), здатність даної моделі точно відображати трансформацію вихідного трафіку суттєво знижується. Це призводить до прогресуючого зростання відносної похибки обчислень, що, у свою чергу, робить довгострокове прогнозування поведінки черг у системі вкрай складним завданням.

Порівняльна оцінка інтегральної затримки в ланцюгу СМО при варіюванні рівня навантаження для трьох альтернативних стратегій моделювання наведена на відповідному графіку (рис. 3.14).



Рис. 3.14 – Оцінка середньої затримки для всієї СМО  
за трьома методами моделювання

Аналізуючи результати оцінки точності, що представлені на рис. 3.15, можна констатувати, що показники методу «с» мають тенденцію до зближення з даними, отриманими шляхом ІМ.

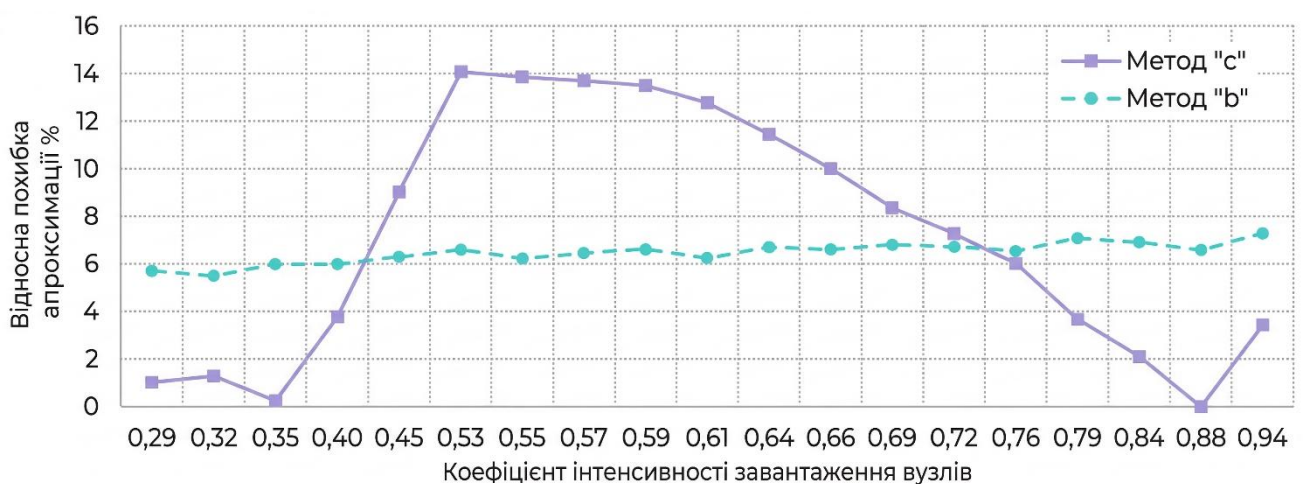


Рис. 3.15 – Відносна оцінка апроксимації двох методів моделювання  
щодо результатів ІМ

Проте існують критичні зони завантаження  $\rho$ , де рівень похибки апроксимації долає бар'єр у 10%. Цікавим є той факт, що при наближенні системи до стану насичення  $\rho \rightarrow 1$ , відносний рівень помилки починає зменшуватися. Це явище знаходить своє математичне обґрунтування у раніше виведеному співвідношенні для граничних станів мережі:

$$\Delta_{err} = f(\rho) \rightarrow \min \quad \text{коли} \quad \rho \rightarrow 1.$$

Що стосується методики «b», то результати спостережень показують дещо іншу картину. Тут похибка апроксимації демонструє відносну стабільність, постійно коливаючись у діапазоні 6 – 8%. Попри таку стабільність, цей підхід не здатний забезпечити ідеальну точність при довільних рівнях інтенсивності трафіку.

Враховуючи отримані дані, жоден із запропонованих методів поодиночки не є абсолютним еталоном. Однак комплексний аналіз дозволив виявити ефективну стратегію: використання методів «b» та «c» як інструментів для визначення нижньої та верхньої меж імовірних значень цільової величини. Таке коридорне оцінювання наочно представлено на графіках похибок.

Особливий інтерес становлять результати вивчення еволюції вхідного потоку в процесі його проходження через каскад обчислювальних вузлів у транзакційній системі. Динаміка швидкості збіжності для гіперекспоненціального вхідного трафіку залежно від коефіцієнта завантаження  $\rho$  відображена на рис. 3.16.

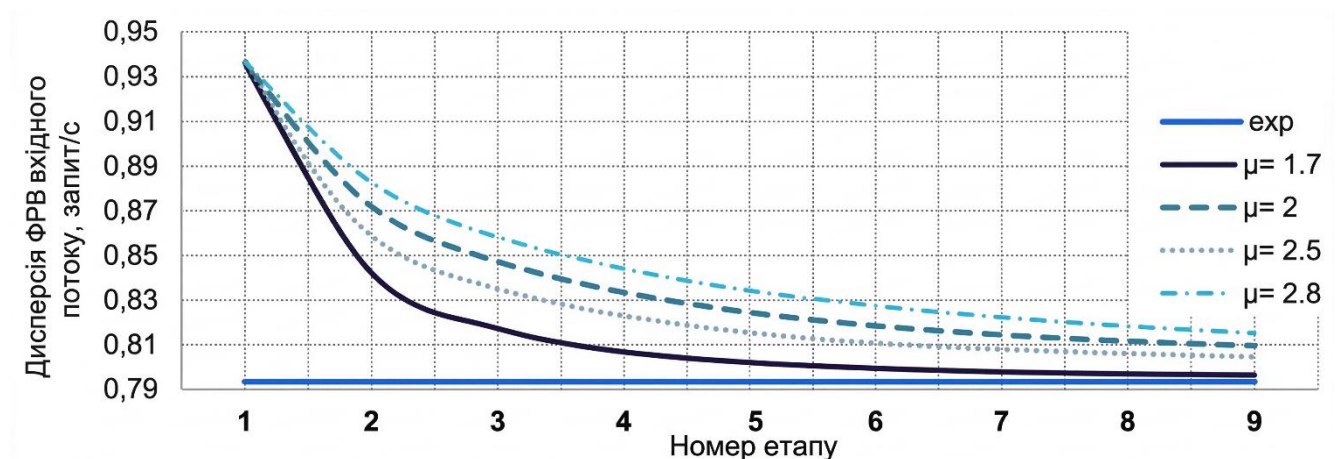


Рис. 3.16 – Дисперсія ФРВ часу між надходженням запитів до кожного вузла в мережі при різних інтенсивностях обслуговування

Згідно з отриманими кривими, зі зростанням інтенсивності навантаження на систему швидкість адаптації форми вхідного потоку до класичного експоненціального розподілу, характерного для моделей G/M/1, суттєво зростає. Головним каталізатором цього процесу є рівень утилізації ресурсів, що математично підтверджується через вирази для моментів розподілу.

Аналогічні висновки були отримані й для інших типів статистичних розподілів, таких як закони Парето, Гамма-розподіл та логнормальний закон.

Вимірювання для цих випадків (рис. 3.17) при  $\mu \sim const$  та  $x \sim const$  підтверджують загальну закономірність.

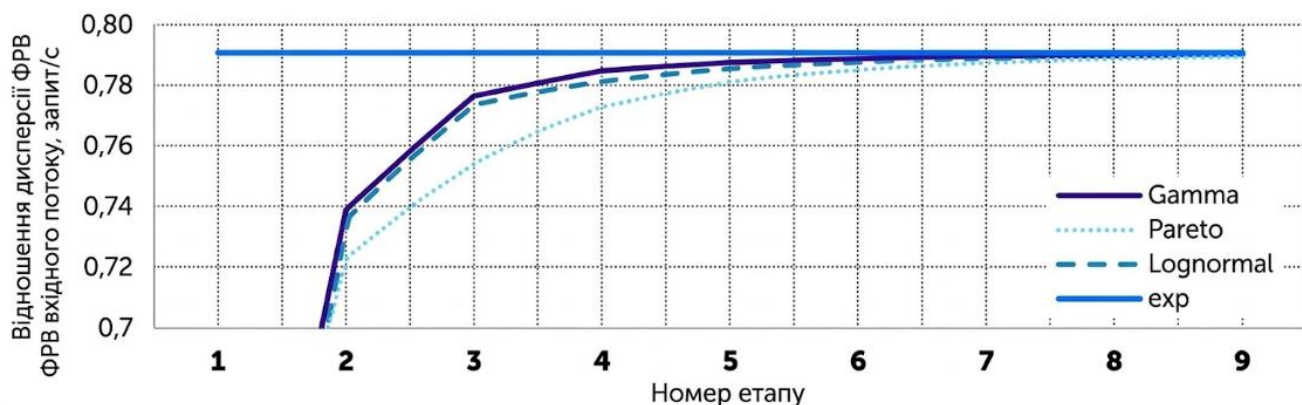


Рис. 3.17 – Дисперсія ФРВ часу між надходженнями запитів до кожного вузла в мережі для різних типів законів надходження

Як можна побачити з фінальної візуалізації, кожна наступна ітерація обробки трафіку на вузлах СМО призводить до того, що стохастичні характеристики потоку в моделях типу G/M/1 дедалі більше наближаються до параметрів ідеального експоненціального потоку. Це явище нормалізації потоку всередині мережі дозволяє використовувати спрощені апроксимаційні моделі на глибоких рівнях ієрархії системи без суттєвої втрати точності.

### Висновки до розділу 3

У межах цього розділу було сформовано підхід до оцінювання часових параметрів обробки запитів у слабозв'язаних транзакційних сервісах, що базується на застосуванні математичного апарату згортки. В основі запропонованого методу лежить аналітичний апарат, який виходить за межі оперування виключно середніми величинами під час аналізу мереж масового обслуговування із послідовною структурою вузлів, де відбуваються процеси розгалуження та агрегації потоків. Завдяки інтеграції моделі Fork-Join підхід дозволяє коректно оцінювати системи з паралельним виконанням операцій у межах одного циклу обробки.

Варто зауважити, що наявні методи прецизійного обчислення часових показників зазвичай обмежені рамками класичних експоненціальних моделей М/М/1. Однак для оцінки характеристик ланцюгів систем довільного типу G/G/1 необхідно впровадження методів апроксимації, розроблених і формалізованих у цьому розділі, а саме:

1. Методи апроксимації неекспоненціальних розподілів з високими коефіцієнтами варіації, що дозволяють подолати складнощі при аналітичному моделюванні та математичних розрахунках послідовних мереж СМО.

2. Методи апроксимації для оцінки характеристик вихідного потоку, обробленого в вузлах розглянутих мереж систем довільного типу G/G/1, для яких немає загального точного рішення для оцінки вихідного потоку при всіх значеннях навантаження.



## РОЗДІЛ 4

### МЕТОДИ ОЦІНЮВАННЯ ХАРАКТЕРИСТИХ ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ

У цій роботі була розроблена методика, що поєднує результати створення аналітичних математичних моделей та ІМ для підвищення ефективності процесу проектування інформаційних систем.

Методика ґрунтується на використанні апроксимаційних формул, які дозволяють наближено моделювати поведінку запитів у робочих процесах. Ці формули були розроблені на основі досліджень, наведених у 3 розділі цієї дисертації. Апроксимаційні формули дають змогу оцінювати різні параметри та характеристики запитів, що сприяє аналізу і визначенню вимог до робочих процесів як проєктованих, так і вже функціонуючих систем.

#### **4.1. Особливості етапів моделювання при розробці інформаційних систем**

Процес розробки та впровадження інфокомунікаційних систем і сучасних послуг зв'язку [166] передбачає кілька ключових фаз моделювання (рис. 4.1).

На стадії попереднього планування та концептуалізації доцільно використовувати аналітичне моделювання. Цей підхід базується на апроксимаційних обчисленнях з використанням стаціонарних ординарних систем типу М/М/1 (рис. 4.1). Такий рівень абстракції дозволяє:

- Провести попередню оцінку робочих параметрів майбутньої системи;
- Виявити потенційні вузькі місця (bottlenecks) у структурі виділених операційних процесів ще до початку розробки;
- Визначити критичні зони перевантаження, які обмежують загальну продуктивність мережі.



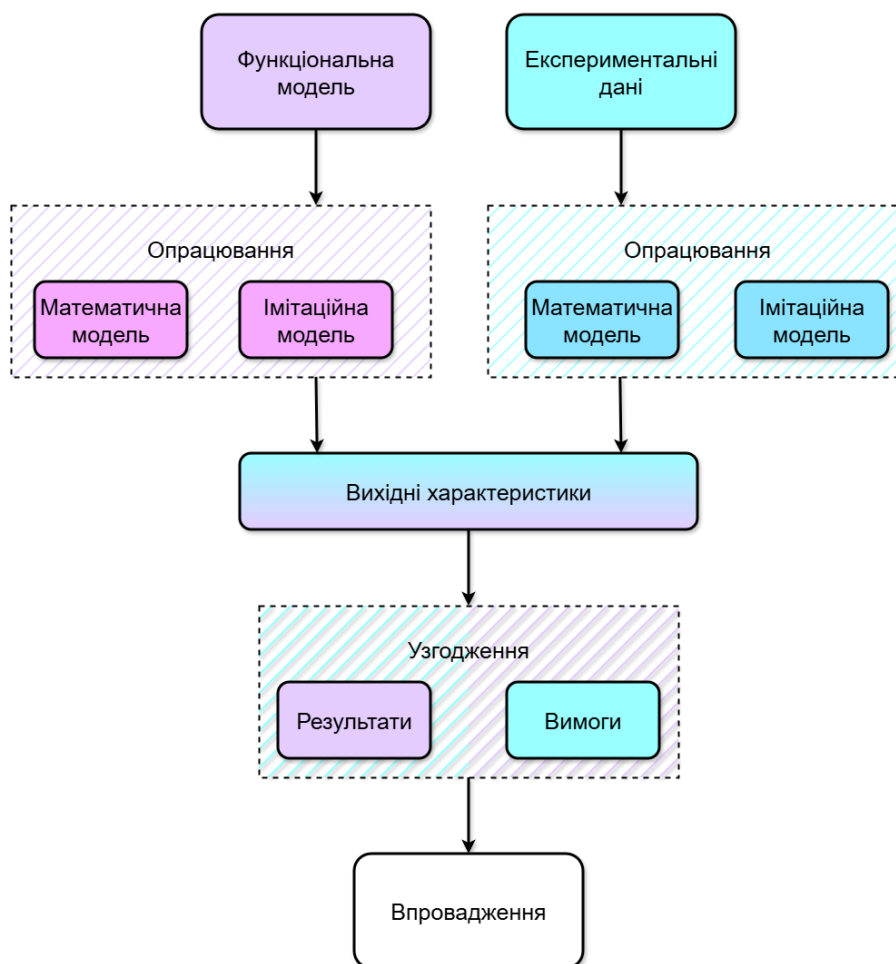


Рис. 4.1 – Етапи розробки телекомунікаційних послуг

Виявлення компонентів, що лімітують швидкість обробки даних, є фундаментальним кроком для забезпечення стабільної роботи та високої ефективності ІС. Після завершення математичного опису рекомендується переходити до ІМ. Використання даних, отриманих на базі функціональних та аналітичних моделей, дозволяє провести верифікацію математичного апарату та підтвердити коректність обраних алгоритмів.

При переході до детального проєктування, коли за основу беруться результати експериментальних досліджень реальних робочих процесів, характер вхідних потоків часто змінюється. Це призводить до необхідності розгляду систем загального виду  $G/G/1$ , де розподіли часових інтервалів відхиляються від експоненціальних і мають високу варіативність.

У таких сценаріях імітаційний підхід (рис. 4.2) демонструє свої ключові переваги:

**1. Висока точність:** Отримання характеристик системи (зокрема моментів та дисперсії вихідних потоків) без використання апроксимаційних допущень;

**2. Адаптивність:** Можливість тонкого налаштування параметрів моделі відповідно до специфіки реального трафіку;

**3. Наочність:** Проведення віртуальних випробувань у контрольованому середовищі.

Варто зауважити, що ІМ має певне обмеження: при роботі з надвеликою кількістю ітерацій виникають труднощі з одночасним отриманням усіх вихідних показників у реальному часі.



Рис. 4.2 – Переваги імітаційного підходу

Загалом, використання математичного та імітаційного інструментарію при створенні ІС дозволяє не лише прогнозувати поведінку системи, а й ефективно керувати її якісними та кількісними показниками ще до моменту фізичної реалізації проєкту.

#### **4.2 Алгоритм розрахунку ймовірно-часових характеристик телекомунікаційної системи**

Синтезуючи результати, отримані на попередніх етапах дослідження (розділи 1–3), пропонується наступна узагальнена методика [5, 7] для визначення часових та ймовірнісних параметрів функціонування системи:

**Етап 1. Формування вхідного масиву даних.** На початковій стадії проводиться збір базових характеристик об'єкта моделювання. Сюди входить ідентифікація множини зовнішніх чинників, що впливають на сервіс  $C = \{c_1, c_2, \dots, c_n\}$ , а також сукупності внутрішніх системних параметрів  $B = \{b_1, b_2, b_3, \dots, b_n\}$ .

**Етап 2. Моделювання інтерфейсу взаємодії послуг.** Здійснюється побудова матриці зв'язків між сеансовим  $SE$  та транзакційним  $TR$  рівнями сервісів. Результатом цього етапу є формування матриці  $R_{ji}^c$ , яка відображає інтенсивність переходів запитів від  $j$ -го сегмента сеансової послуги до  $i$ -го операційного процесу транзакційного типу.

**Етап 3. Аналіз продуктивності сеансового рівня.**

**3.1. Опис архітектури.** Визначається топологія сеансової послуги. Ключовим вихідним даним тут є матриця ймовірностей  $R_{ji}$ , що описує маршрутизацію запитів користувачів між вузлами відправлення  $j$  та прийому  $i$ .

**3.2. Розрахунок потоків.** Використовуючи значення загальної вхідної інтенсивності та сформовану матрицю переходів, за допомогою теореми Джексона обчислюється інтенсивність трафіку  $\lambda_i$  для кожного індивідуального вузла  $S_{ei}$ .

**3.3. Визначення показників якості.** На основі отриманих значень  $\lambda_i$  проводиться розрахунок якісних характеристик системи з використанням математичного апарату, деталізованого у формулах (2.12) – (2.15).

**Етап 4. Трансформація навантаження на транзакційний рівень.** Враховуючи розраховане навантаження користувацьких сесій  $A_{SE}$  та структуру переходів  $R_{ij}^c$ , виконується калькуляція вхідного потоку, що спрямовується з  $SE_j$  на транзакційні підсистеми  $TR_i$ .

- **4.1. Спеціальний випадок (Пуассонівський потік).** При надходженні найпростішого потоку рекомендується моделювати обробку транзакцій як систему класу  $M/M/1/\infty/N$ . У цьому разі враховується лімітована кількість джерел навантаження згідно з виразами (2.1) – (2.2).

- **4.2. Емпіричний підхід.** Якщо система має специфічні характеристики, параметри вхідного трафіку визначаються шляхом експериментальних замірів

[51, 62]. При складних законах розподілу застосовується ІМ або апроксимація через ступінчасті емпіричні функції.

**Етап 5. Детальна оцінка транзакційних процесів.** Маючи вхідну інтенсивність  $\lambda_i^{TR}$ , проводиться фінальний розрахунок робочих циклів.

- **5.1. Вибір сценарію.** Ідентифікується конкретний операційний сценарій транзакційної послуги.

- **5.2. Перевірка інтегрованості.** Аналізується можливість математичного інтегрування ФРВ вхідних параметрів. У разі виявлення неінтегрованих або ступінчастих функцій (отриманих експериментально), виконується їх апроксимація за раніше описаними методами.

- **5.3. Операційні кроки обробки:**

- **5.3.1.** Починаючи з першої ланки в ланцюзі транзакцій, визначаються перетворення Лапласа-Стілтєса для часу обслуговування  $B(s)$  та інтервалів між заявками  $A(s)$ .

- **5.3.2. Обчислення часу очікування.** Залежно від типу системи (М/М/1, М/Г/1 або Г/Г/1) застосовуються відповідні розрахункові моделі, включаючи спектральні методи для рівняння Ліндлі.

- **5.3.3. Динаміка вихідного потоку.** Для багатоланкових систем  $N > 1$  ідентифікується форма щільності вихідного трафіку  $D_i(s)$  (через апроксимацію або імітацію для складних випадків).

- **5.3.4. Формування повної ФРВ.** Визначається щільність часу перебування запиту в системі  $F_{ei}(t)$  через операцію згортки розподілів очікування та обслуговування.

- **5.3.5.** Кроки повторюються для всіх вузлів ланцюга  $i \leq A$ .

- **5.4. Інтегральний показник сценарію.** На основі згортки ЩРВ часу перебування та зворотного перетворення Лапласа визначається результуюча щільність  $F_x(t)$  для всього сценарію.

- **5.5. Фінальні метрики.** Обчислюються перший момент (математичне очікування) та дисперсія за допомогою виразів (3.1) – (3.2).

Візуалізація даної методики розроблена автором і представлена у формі SDL-діаграм (рис. 4.3 та рис. 4.4).

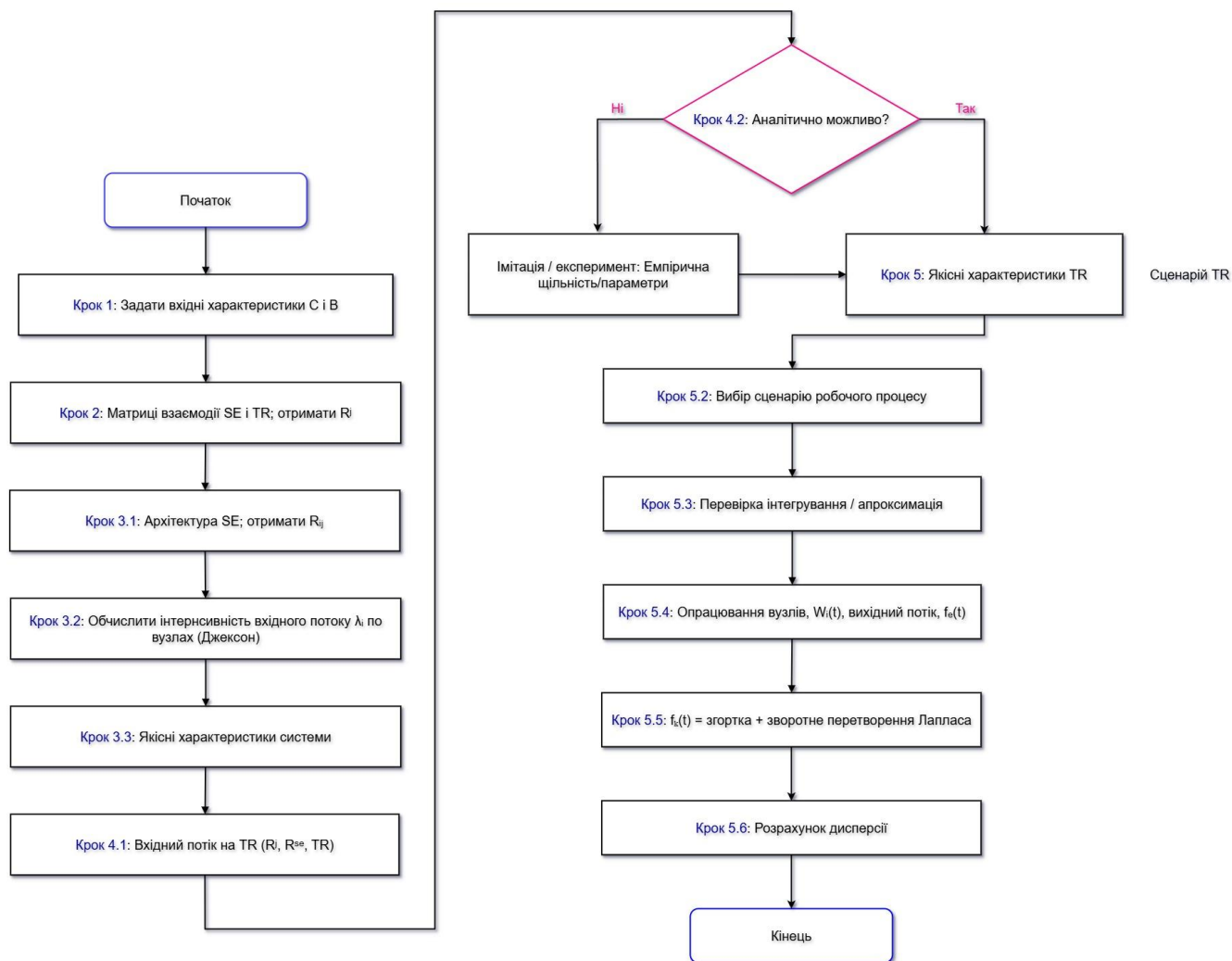


Рис. 4.3 – Алгоритм визначення якісних характеристик комбінованих послуг

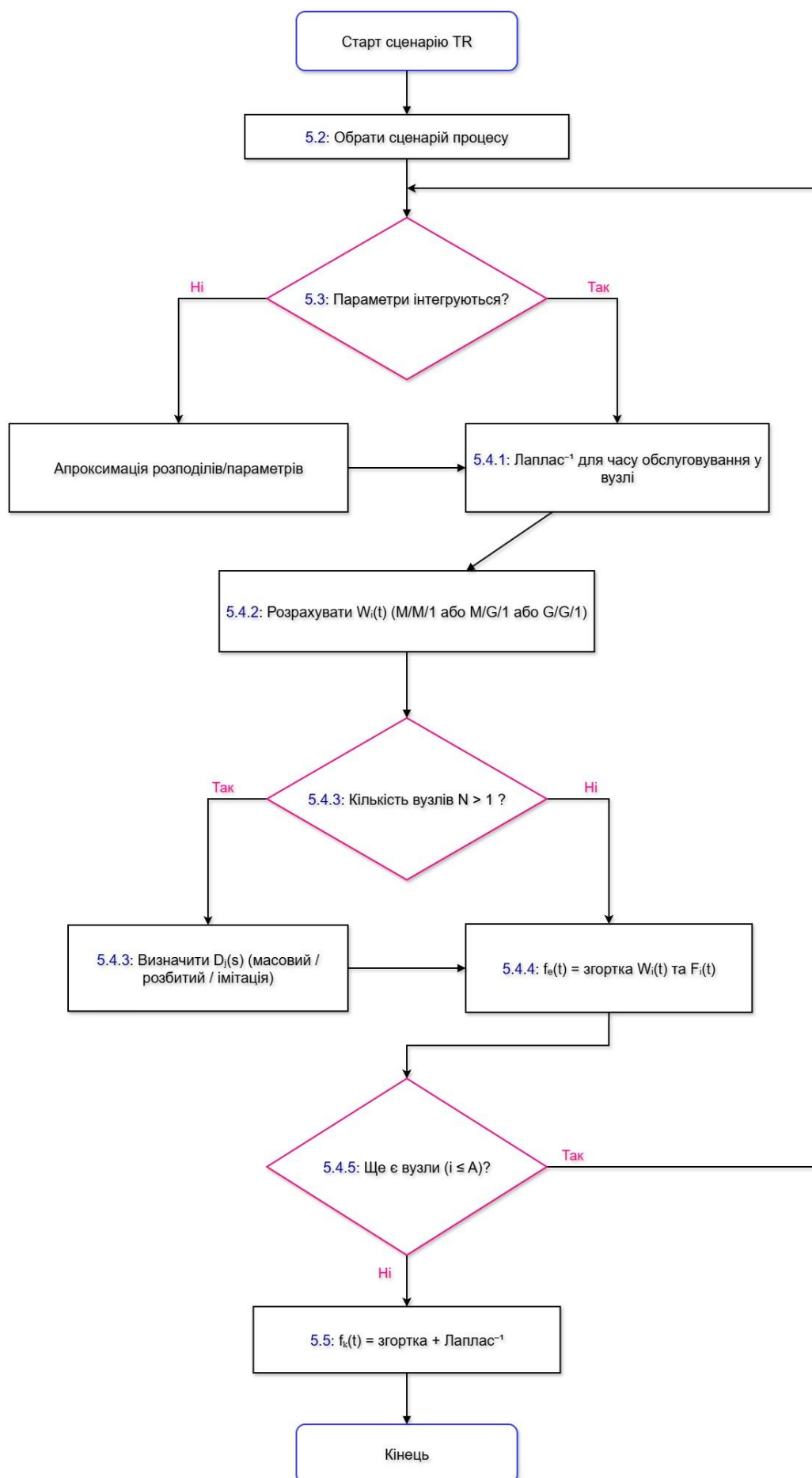


Рис. 4.4 – Методика розрахунку транзакційних послуг

### 4.3. Результати моделювання

Розглянемо нечітку процедура викликів VoIP [39] на АТМ-комутатор [77]. Припустимо, що схема мережі (рис. В.1 у додатках) розроблена з урахуванням того, що 24 різні виклики (згідно [44]) (описані в табл. Б.1 та табл. Б.2 у додатках) подаються на АТМ-комутатор. Критерій допуску виклику для  $i$ -го трафіку, наприклад, визначається значенням параметра  $s_i$ , отриманим однією з процедур дефазифікації (Підхід 1 або Підхід 2, тобто для двох абонентів на різних кінцях мережі), описаних у розділах вище. Відповідні параметри наведено в табл. 4.1 – табл. 4.4.

Для цілей моделювання припускається два набори даних із описами джерел  $X_i$  та профілями ресурсів  $Z_i$  для  $i$ -го виклику, які подані в табл. 1.1 та табл. 1.2.

Основні дослідження і практичні результати покажемо для голосової телефонії VoIP [4].

Набір даних I (профіль ресурсів  $Z_i$ )

Категорія послуг	Параметри	Пріоритет (якщо)
ЕХ	Відео: $\lambda_3 = 10 \text{ Mbps}$ (MPEG2); $BP = 3$ ; $\ell_3^* = 0.35$ ; $P_{NB} = 0.35$ ; $P_{NB} = 0.55$ . Голос: $\lambda_2 = 32 \text{ Kbps}$ (compressed voice); $P_{on} = 0.50$ ; $P_{off} = 0.20$ ; $\ell_3^* = 0.40$ . Дані: $\lambda_{1b} = 2 \text{ Mbps}$ (data file transfer); $\lambda_{2b} = 64 \text{ Kbps}$ ; $P_{LH} = 0.50$ ; $P_{LH} = 0.90$ ; $\ell_1^* = 0.30$ .	I: Відео  II: Голос  III: Дані
LX	Відео: $\lambda_3 = 10 \text{ Mbps}$ (broadband video retrieval); $BP = 5$ ; $\ell_3^* = 0.30$ ; $P_{NB} = 0.50$ ; $P_{NB} = 0.50$ . Голос: $\lambda_2 = 64 \text{ Kbps}$ (telephony); $P_{on} = 0.80$ ; $P_{off} = 0.10$ ; $\ell_2^* = 0.30$ . Дані: $\lambda_{1b} = 64 \text{ Kbps}$ (narrowband document retrieval); $\lambda_{1f} = 64 \text{ Kbps}$ ; $P_{LH} = 0.50$ ; $P_{HL} = 0.50$ ; $\ell_1^* = 0.40$ .	I: Відео  II: Голос  III: Дані
МХ	Відео: $\lambda_3 = 2 \text{ Mbps}$ (video phone); $BP = 1$ ; $\ell_3^* = 0.80$ . Голос: $\lambda_2 = 64 \text{ Kbps}$ (telephony); $P_{on} = 0.60$ ; $P_{off} = 0.40$ ; $\ell_2^* = 0.20$ .	I: Відео  II: Голос
СХ	Голос: $\lambda_2 = 32 \text{ Kbps}$ (compressed voice); $P_{on} = 0.50$ ; $P_{off} = 0.50$ ; $\ell_2^* = 0.40$ . Дані: $64 \text{ Kbps} = \lambda_{1b} = \lambda_{1f}$ (data on demand); $P_{LH} = 0.50$ ; $P_{HL} = 0.50$ ; $\ell_1^* = 0.60$ .	I: Голос  II: Дані



Набір даних I (типи трафіку (джерел)  $X_i$ )

Тип	Опис	Номінальні параметри (типові)
RA: (інтенсивна обробка трафіку та середній BER)	Великі буфери, середній SNR	Розмір буфера: 30; SNR: 3.5 дБ Інтенсивності трафіку: $\mu/\lambda_3 = 15$ ; $\mu/\lambda_{1h} \approx 75$
RB: (помірна обробка трафіку та високий BER)	Середні буфери, низький SNR	Розмір буфера: 10; SNR: 2 дБ Інтенсивності трафіку: $\mu/\lambda_3 = 1.5$ ; $\mu/\lambda_{1h} \approx 30$
RC: (інтенсивна обробка трафіку та високий BER)	Великі буфери, низький SNR	Розмір буфера: 30; SNR: 2.5 дБ Інтенсивності трафіку: $\mu/\lambda_3 = 5$ ; $\mu/\lambda_{1h} \approx 20$
RD: (низька обробка трафіку та середній BER)	Малі буфери, середній SNR	Розмір буфера: 5; SNR: 3 дБ Інтенсивності трафіку: $\mu/\lambda_3 = 2$ ; $\mu/\lambda_{1h} = 5$
RE: (низька обробка трафіку та низький BER)	Малі буфери, високий SNR	Розмір буфера: 3; SNR: 10 дБ Інтенсивності трафіку: $\mu/\lambda_3 = 5$ ; $\mu/\lambda_{1h} = 10$
RF: (низька обробка трафіку та високий BER)	Малі буфери, низький SNR	Розмір буфера: 2; SNR: 1.5 дБ Інтенсивності трафіку: $\mu/\lambda_3 = 3$ ; $\mu/\lambda_{1h} = 10$

Набір даних II (профіль ресурсів  $Z_i$ )

Категорія послуг	Параметри	Пріоритет (якщо)
ЕХ	Відео: $\lambda_3 = 90 \text{ Mbps}$ ; (MPEG2); $BP = 4$ ; $\ell_3^* = 0.35$ ; $P_{NB} = 0.50$ ; $P_{BN} = 0.20$ . Голос: $\lambda_2 = 64 \text{ Kbps}$ ; $P_{on} = 0.60$ ; $P_{off} = 0.40$ ; $\ell_2^* = 0.55$ . Дані: $\lambda_{1b} = 10 \text{ Mbps}$ ; $\lambda_{2b} = 64 \text{ Kbps}$ ; $P_{LH} = 0.20$ ; $P_{HL} = 0.85$ ; $\ell_1^* = 0.30$ .	I: Відео  II: Голос  III: Дані
LX	Відео: $\lambda_3 = 30 \text{ Mbps}$ ; $BP = 3$ ; $P_{NB} = 0.75$ ; $P_{BN} = 0.13$ . Голос: $\lambda_2 = 64 \text{ Kbps}$ ; $P_{on} = 0.76$ ; $P_{off} = 0.24$ ; $\ell_2^* = 0.40$ . Дані: $\lambda_{1b} = 6.3 \text{ Mbps}$ ; $\lambda_{1f} = 2.3 \text{ Mbps}$ ; $P_{LH} = 0.30$ ; $P_{HL} = 0.90$ ; $\ell_1^* = 0.30$ .	I: Відео  II: Голос  II: Дані
MX	Дані: $\lambda_{1f} = 10 \text{ Mbps}$ ; $BP = 1$ ; $\lambda_{1\ell} = 3 \text{ Mbps}$ ; $P_{on} = 0.70$ ; $\ell_2^* = 0.50$ . Голос: $\lambda_2 = 32 \text{ Kbps}$ ; $P_{on} = 0.60$ ; $P_{off} = 0.30$ ; $\ell_2^* = 0.50$ .	I: Дані  II: Голос
CX	Голос: $\lambda_{1b} = 64 \text{ Kbps}$ ; $P_{on} = 0.85$ ; $P_{off} = 0.15$ ; $\ell_2^* = 0.50$ . Дані: $\lambda_{1b} = 64 \text{ Kbps}$ ; $\lambda_{1f} = 500 \text{ Kbps}$ ; $P_{LH} = 0.30$ ; $P_{HL} = 0.80$ ; $\ell_1^* = 0.35$ .	I: Голос  II: Дані

Набір даних II (типи трафіку (джерела)  $X_i$ )

Тип	Опис	Номінальні параметри (типові)
RA: (інтенсивна обробка трафіку та середній BER)	Велика кількість буферів, наявне згладжування трафіку, великий SNR	Розмір буфера: 20; SNR: 15 дБ; Інтенсивності трафіку: $\mu/\lambda_3 = 4; \mu/\lambda_{1h} = 30$
RB: (помірна обробка трафіку та високий BER)	Середня кількість буферів, наявне згладжування трафіку, середній SNR	Розмір буфера: 15; SNR: 10 дБ; Інтенсивності трафіку: $\mu/\lambda_3 = 3; \mu/\lambda_{1h} = 15$
RC: (інтенсивна обробка трафіку та високий BER)	Велика кількість буферів, наявне згладжування трафіку, малий SNR	Розмір буфера: 20; SNR: 3 дБ; Інтенсивності трафіку: $\mu/\lambda_3 = 4; \mu/\lambda_{1h} = 30$
RD: (низька обробка трафіку та середній BER)	Середня кількість буферів, наявне згладжування трафіку, малий SNR	Розмір буфера: 10; SNR: 2 дБ; Інтенсивності трафіку: $\mu/\lambda_3 = 2; \mu/\lambda_{1h} = 10$
RE: (низька обробка трафіку та низький BER)	Невелика кількість буферів, наявне згладжування трафіку, середній SNR	Розмір буфера: 5; SNR: 7.5 дБ; Інтенсивності трафіку: $\mu/\lambda_3 = 3; \mu/\lambda_{1h} = 5$
RF: (низька обробка трафіку та високий BER)	Невелика кількість буферів, наявне згладжування трафіку, малий SNR	Розмір буфера: 5; SNR: 2 дБ; Інтенсивності трафіку: $\mu/\lambda_3 = 3; \mu/\lambda_{1h} = 5$

Відповідно до нечітких евристичних підходів щодо складності атрибутів передавання АТМ, як зазначено раніше, у моделюванні використовується стратегія допуску викликів. Вона передбачає призначення вихідного віртуального каналу або віртуального шляху у відповідь на запит на встановлення з'єднання на АТМ-комутаторі.

Допуск виклику забезпечує виконання узгоджених цілей QoS у гетерогенній системі вхідних викликів [18, 49]. Як зазначалося раніше, цілі QoS об'єднуються в єдиний параметр складності  $s$ .

Крім того, трафік, що асоціюється з кожним вхідним викликом, характеризується певними стохастичними атрибутами, а відповідні оцінки продуктивності вважаються нечіткими (fuzzy).

Припустимо, що вхідний виклик на АТМ-комутаторі відповідає певному набору категорій характеристик джерела та профілів ресурсів (топология мережі та маршрутизатори), як зазначено на рис. В.1 у додатках. Передбачається, що потоки комірок, запропоновані кожним джерелом, мультиплексуються на вході високошвидкісної лінії зв'язку з інтенсивністю  $\mu$  комірок за секунду.

Зірочка позначає нормалізацію відносно пікового значення, а  $(s_i^*)_{df}$  – це дефазифіковане значення, отримане з  $s_i^{(f)}$ . (А) Значення  $(s_i^*)_{df}$  подано залежно від типу/номера виклику, що відповідає простому, сірому та складному режимам. Верхня межа рішення про прийняття виклику  $\{\xi_i\}$  та нижня межа рішення про відхилення виклику  $\{\bar{\xi}_i\}$  із сірою зоною переходу між  $\xi_i$  та  $\bar{\xi}_i$ .

Тепер функцію САС щодо мультиплексованих комірок кожного вхідного виклику  $i$ -го типу можна інтерпретувати як відображення вектора стану  $\{s_i\}$  у вектор прийняття рішення  $\{\xi_i\}$ .

Це функціональне відображення розділяє простір станів на дві чіткі області – область прийняття та область відхилення (як показано на рис. 4.5. (А)), якщо вхідною змінною є  $s_i$ , а рішення  $\{\xi_i\}$  розглядаються у чітких форматах.

Однак, з огляду на зазначену раніше причину, вхідні дані  $\{X_i, Z_i\}$  можуть створювати нечіткі, перекривні області прийняття рішень під час реалізації процесу прийняття або відхилення запиту на з'єднання.

Отже, використовуючи нечітку множину  $\{s_i\} \in \{X_i, Z_i\}$ , що відповідає  $i$ -му виклику, процедура дефазифікації (Підхід 1 або Підхід 2) призводить до отримання дефазифікованого значення  $(s_i)^f$  або просто  $s_i$ .

У підході 1 це відповідає центроїдному значенню, а в підході 2 – модальному значенню  $s_{im}$ .

Використовуючи методи дефазифікації, дефазифіковані значення  $s_i$  отримано для 24 викликів ( $i = 1, 2, \dots, 24$ ), що відповідають наборам даних I та II (табл. 4.1 та табл. 4.4).

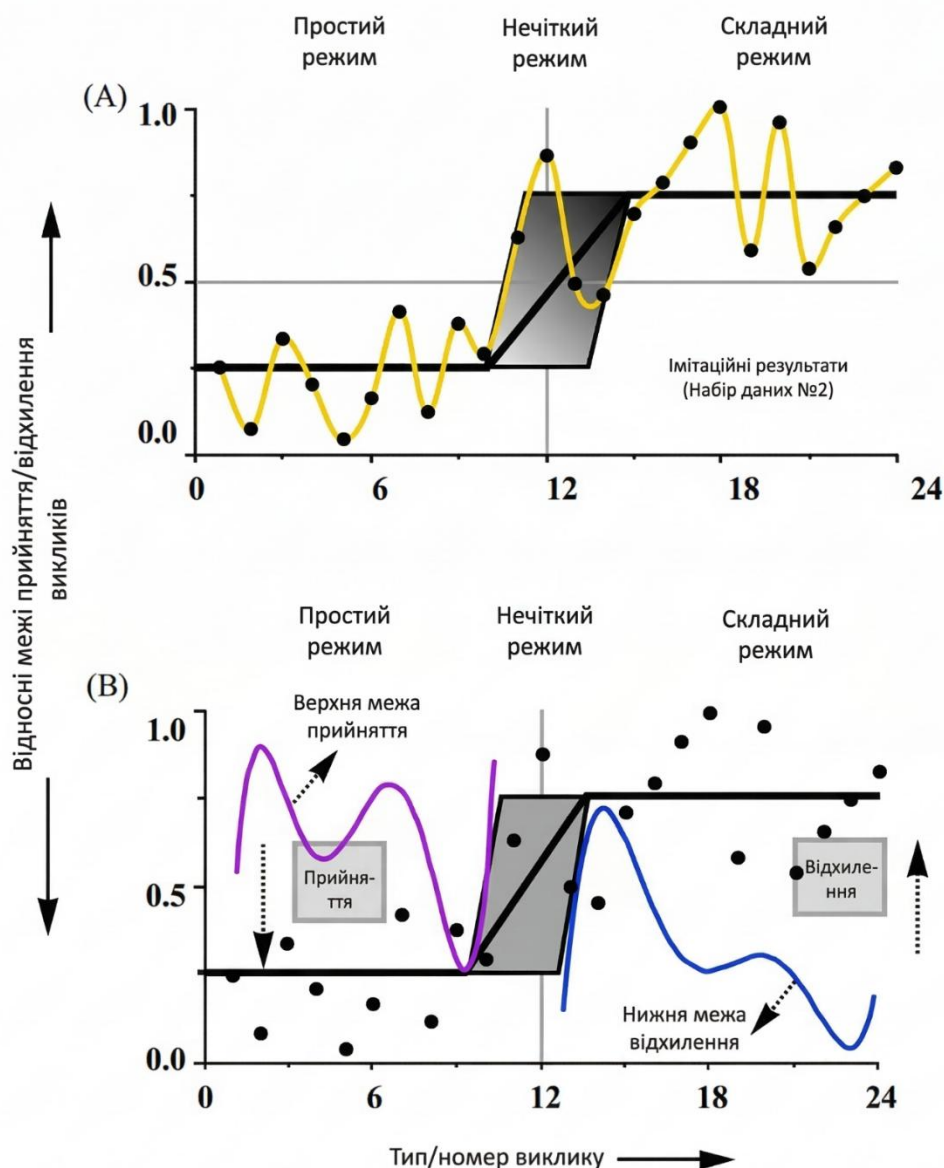


Рис. 4.5 – Імітаційні результати, що стосуються набору даних I (підхід 2)

Нижні індекси 1 та 2 у виразі  $s_i^*$  позначають 1 та 2 абонентів відповідно. Значення  $(s_i^*)_1$  та  $(s_i^*)_2$  подано залежно від типу/номера виклику для зображення простого, сірого та складного режимів.

Верхня межа рішення про прийняття виклику  $\{\xi_i\}$  та нижня межа рішення про відхилення виклику  $\{\xi_i\}$  із сірою зоною переходу між  $\xi_i$  та  $\bar{\xi}_i$ .

Оскільки допуск виклику відповідає умові, що  $s_i$  будь-якого  $i$ -го виклику відображається на заздалегідь визначений вектор прийняття рішення  $\{\xi_i\}$ , виклик дозволяється, якщо  $s_i \in \{\xi_i\}$  (рис. 4.5 (В)).

Вказане моделювання може виконуватись для будь-якої заданої групи викликів, які надходять до АТМ-комутатора для прийняття рішення про допуск; при цьому можна відповідно визначити графік пріоритетів прийому викликів згідно лінгвістичних атрибутів табл. Б.3 з додатків.

Розглянемо графічні результати набору даних І, подані на рис. 4.6 (А) та рис. 4.6 (В), що відповідають підходам 1 і 2 відповідно.

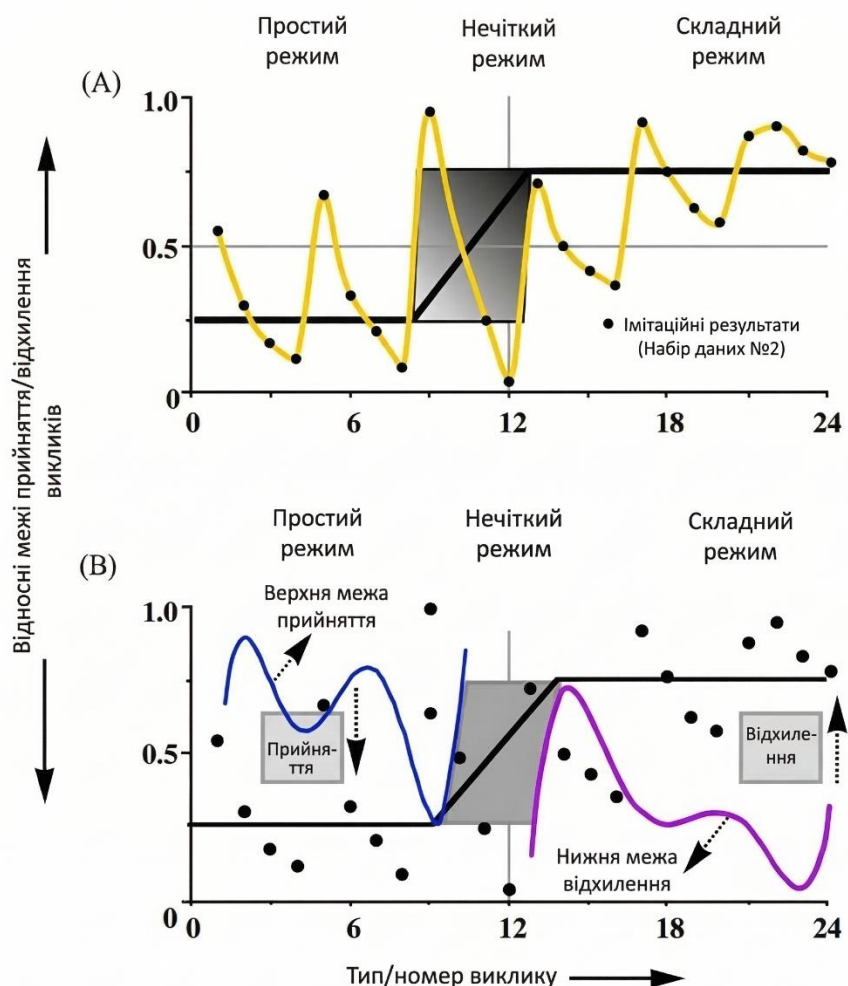


Рис. 4.6 – Імітаційні результати, що стосуються набору даних ІІ (підхід 1)

На рис. 4.6 (А) наведено дефазифіковані та нормалізовані значення  $s_i$ , тобто  $(s_i^*)_{df}$ , залежно від типу (номера) кожного виклику.

Ці значення демонструють дві чіткі кластери в крайніх областях, позначених як простий і складний режими. У середині спостерігається набір значень із перехідними характеристиками, тобто їхня належність до простого або складного режиму є досить нечіткою або невизначеною.

Значення  $(s_i^*)_{df}$ , подано залежно від типу/номера виклику для відображення простого, нечіткого та складного режимів. Верхня межа рішення про прийняття виклику  $\{\xi_i\}$  та нижня межа рішення про відхилення виклику  $\{\bar{\xi}_l\}$  із сірою зоною переходу між  $\xi_i$  та  $\bar{\xi}_l$ .

Тепер необхідно визначити межу всередині цього нечіткого режиму. Це робиться, як показано на рис. 4.6 (В). Спочатку встановлюється верхня межа (supremum) для простого (або приймального) режиму  $\xi_i$  і нижня межа (infimum) для складного (або відхиляючого) режиму  $\bar{\xi}_l$ . Ці межі визначаються шляхом розгляду значень  $s_i$  для лінгвістичних атрибутів табл. Б.3 з додатків.

Аналогічно, використовуючи Підхід 2, моделювання з набором даних І (як показано на рис. 4.5 (В)) дає змогу визначити межі за допомогою значень  $(s_i^*)_{2m}$  і  $(s_i^*)_{1m}$ , також – взаємно.

В обох підходах межі областей прийняття/відхилення викликів позначають перехідний нечіткий режим, як показано на рис. 4.5 та рис. 4.6.

Лінію, побудовану в межах цієї сірої області розриву, в обох випадках можна розглядати як межу розділення між  $\xi_i$  та  $\bar{\xi}_l$  (у цьому сірому режимі).

Таким чином, маючи встановлені граничні межі для прийняття або відхилення викликів, можна порівняти обидва підходи.

У табл. В.3 з додатків наведено кількість прийнятих викликів типу А та відхилених викликів типу R, підрахованих у межах, зазначених на рис. 4.5 (В) та рис. 4.6 (В).

Існують також виклики, що розташовані поблизу меж, які позначаються як A/BL або R/BL (де BL означає граничні випадки).

#### 4.4. Обговорення результатів

Наведені вище спостереження є цілком обґрунтованими.

Підхід 1 базується на центроїдному значенні нечіткої логіки, отриманої з сусідніх параметрів, тоді як Підхід 2 спирається на дихотомічну логіку, зосереджену навколо модального значення. Отже, у Підході 1 можна очікувати плавного переходу між простими та складними режимами, тоді як у Підході 2 – різких переходів між цими режимами.

Порівняно з цим, розмиті рішення, отримані за допомогою Підходу 1, є більш гнучкими, оскільки також враховують граничні випадки (рішення, які можуть бути прийняті лише в межах системної толерантності або за попередньо визначеним критерієм).

Підхід 2, навпаки, має схильність до жорстких (чітких) рішень, що може призвести до надмірно конкретизованих або, навпаки, недостатньо визначених результатів.

Враховуючи наведені спостереження, можна сформулювати такі висновки.

Лінгвістична класифікація викликів чітко поділяється на три категорії: простий тип, складний тип і нечіткий тип, що визначають їхню прийнятність у межах АТМ-комутатора.

Представлені підходи 1 та 2 класифікують ці виклики та групують їх у дві основні області: простий/допускний режим  $\xi_i$  та складний/відхиляючий режим  $\bar{\xi}_i$ . Крім того, вказується невизначена (нечітка) область, що лежить між режимами прийняття та відхилення, де розташована частина викликів.

Запропоновано процедуру [1] для визначення граничних меж – верхньої межі (supremum) для рішень про прийняття викликів і нижньої межі (infimum) для рішень про відхилення викликів у відповідних областях  $\xi_i$  та  $\bar{\xi}_i$ .

У межах нечіткої області формується перехідна межа, яка (за потреби) може примусово розділяти рішення як на прийняте або відхилене. Ця межа визначається шляхом згладжування асимптотичних розривів між верхньою та нижньою межами.

Встановивши такі граничні межі, отримані рішення за обома підходами – для наборів даних I та II – можуть бути узагальнені.



Ще одне цікаве узагальнення, яке можна зробити на основі рис. 4.5 та рис. 4.6, стосується нелінійного переходу від простого/допускового режиму до складного/відхиляючого через нечітку область, що лежить між ними.

Нелінійність цього типу виникає в результаті взаємодії між різними підсистемами або параметрами, які визначають продуктивність (тобто простий чи складний характер), з урахуванням профілю допустимості викликів. Взаємодія та пов'язана з нею невизначеність призводять до нечіткої нелінійності, яка описується нечітким нелінійним рівнянням – як було зазначено одним із авторів у роботі [2].

Дане дослідження пропонує надійну методику проєктування [9] для створення нечіткого механізму виведення (fuzzy inference engine), який може керувати САС у передаваннях даних в мережах АТМ.

У роботі використано концепцію складності (complexity) для опису експертних знань щодо перекривних характеристик параметрів трафіку [7]. Запропонована метрика нечіткої складності (fuzzy complexity metric), яка узгоджено поєднує фактори затримки, що виникають через заповнення буфера, та втрати комірок, спричинені SNR.

Крім того, у дослідженні наведено два підходи до дефазифікації. Ефективність і реалізація продемонстровані шляхом моделювання набору викликів, які мають характеристики практичного значення.

Запропонований механізм [5] здатен обробляти будь-який трафік із різними середніми швидкостями та змішаними застосунками, що зазвичай трапляються в АТМ-телекомунікаціях. Основною перевагою запропонованої стратегії є її простота та очевидність у реалізації.

Додатково система (рис. В.1 з додатків) тестувалась за алгоритмом відкриття шлюзу для надання послуги VoIP (рис. Г.1 з додатків). Також розроблено механізм контролю допуску викликів VoIP [4] на основі нечіткого оцінювання параметрів QoS (рис. Г.2 з додатків).

Алгоритм тестування мережі EVANGELION (згідно рис. В.1. з додатків) та побудови графіків був реалізований мові Python (Додаток Д). Отримано log-файл (Додаток Е) тестування мережі EVANGELION з відкритим шлюзом для VoIP в

режимі передачі відео- і аудіо- контенту через маршрутизатори (D-LINK DGS-1210-28P та ALCATEL Omni PCX 3EH08502AA).

Побудовано графік ІМ переданого трафіку для VoIP (рис. 4.7) від одного користувача іншому. Видно, що навантаження на тестову мережу приблизно рівномірне, за умов завантаження мережі іншими видами трафіку.

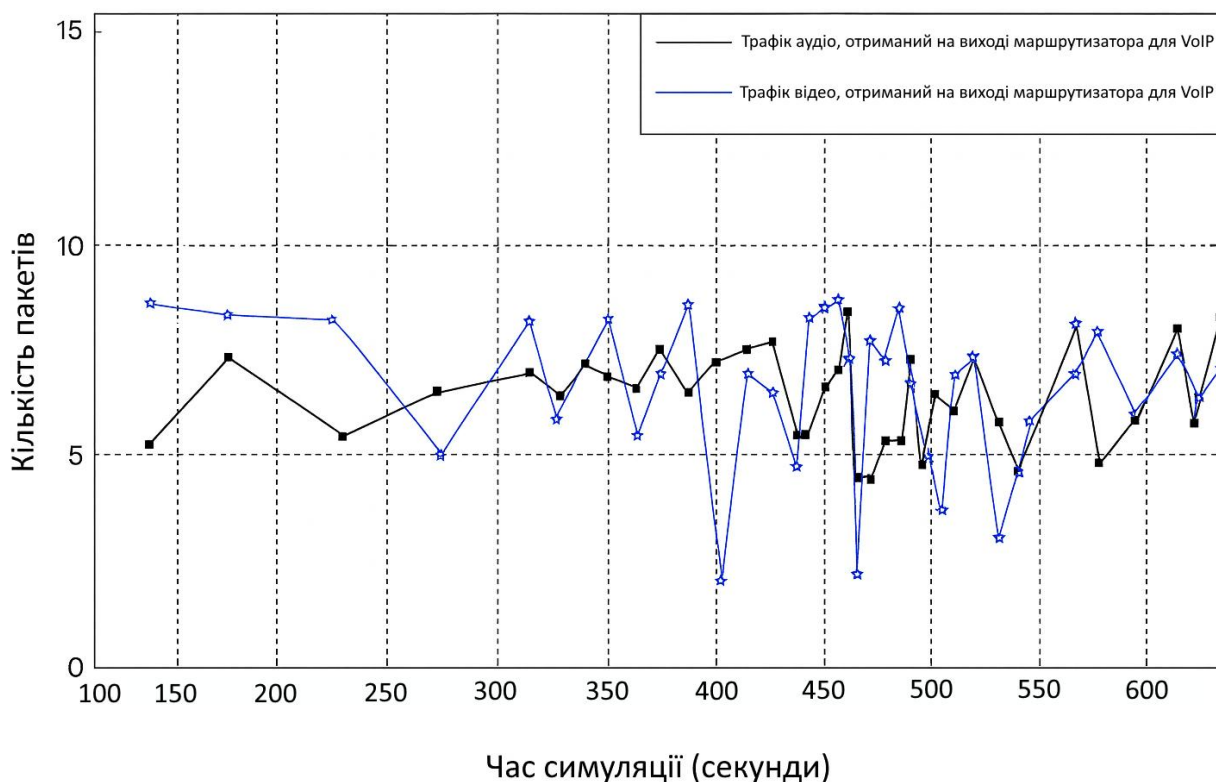


Рис. 4.7 – Трафік переданих пакетів VoIP в реальному часі

Також експериментально отримано графік переданих пакетів VoIP (рис. Г.3 з додатків) та графік затримки/втрати пакетів для VoIP (рис. Г.4 з додатків). Видно, що багато пакетів втрачаються, а також виділяється інтенсивність відео- трафіку за рахунок встановлених підвищених вимог до якості відео [10].

#### 4.5. Оцінка похибок вимірювань та розрахунків

Для оцінювання будь-якого методу вимірювань доцільно визначати інформаційну складність (ємність) об'єкта і кількість інформації, що отримується

про об'єкт. Для нестаціонарних об'єктів слід також урахувати швидкість застарівання інформації.

Отримана під час вимірювання інформація визначається зменшенням невизначеності наших уявлень про систему. Своєю чергою, для опису ступеня невизначеності використовується поняття ентропії системи.

Найбільш загальним є означення ентропії, запропоноване Р. Хартлі [143]:

$$H = k \ln M, \quad (4.1)$$

де  $M$  – статистична вага даного стану системи.

Коефіцієнт  $k$  визначається обраною шкалою вимірювання інформації. Для вимірювання інформації в бітах слід покласти  $k = 1 / \ln 2 \approx 1.442695$ . При  $k = 1$  одиницею вимірювання інформації є – нат.

У випадку неперервної залежності ймовірності від параметра можна скористатися означенням Шеннона [12]:

$$H_{\xi} = -k \int_X \rho(x) \ln(\rho(x)) dx, \quad (4.2)$$

де  $\rho(x)$  – щільність імовірності розподілу значення параметра;  $X$  – область визначення параметра.

Задача достовірності отримання певної кількості інформації при вимірюванні скалярної величини може розглядатися як задача непрямого вимірювання значення параметра об'єкта  $Q$ , при цьому слід враховувати динаміку роботи мережі і шуми вимірювання. На цьому рівні інкапсуляції задача оцінювання кількості інформації, отриманої під час вимірювань, є еквівалентною задачі оцінювання кількості інформації, отриманої під час вимірювання випадкової скалярної величини.

Припустимо, що задано стаціонарну скалярну випадкову величину  $x_0$ , і відома (яким-небудь чином оцінена) початкова щільність імовірності її розподілу  $\rho_0(x)$ . Нехай також за час  $\tau$  було проведено вимірювання величини  $x$ . Результатом цього вимірювання є нова оцінка щільності імовірності  $\rho(x, \tau)$ . Потрібно оцінити інформацію, отриману в цьому вимірюванні.

Апріорна ентропія  $H_0$  визначається виразом

$$H_0 = -k \int_{-\infty}^{\infty} \rho_0(x) \ln(\rho_0(x)) dx, \quad (4.3)$$

а поточна

$$H(\tau) = -k \int_{-\infty}^{\infty} \rho(x, \tau) \ln(\rho(x, \tau)) dx. \quad (4.4)$$

У вищенаведених виразах вважаємо, що  $0 \ln(0) = 0$ .

Інформація, отримана в результаті вимірювання, дорівнює різниці ентропій

$$I(\tau) = -(H(\tau) - H_0). \quad (4.5)$$

Якщо початкова невизначеність більша за отриману в результаті вимірювання, то кількість отриманої інформації є додатною. У протилежному випадку вимірювання не дало нової інформації порівняно з апіорною.

Насправді прямим результатом вимірювання є не щільність імовірності, а множина оцінок  $\bar{x}$  вимірюваної величини  $x_0$ . Для отримання параметрів розподілу необхідна або апіорна інформація про застосовуваний метод вимірювання, або додаткові вимірювання, призначені для отримання статистичних характеристик самого методу вимірювання.

Сама величина  $x_0$  є вимірювано недоступною, і під час вимірювання цієї величини виникають неминучі шуми вимірювання  $w(t)$  [22]. Під шумами вимірювання розуміють не тільки шуми, реально наявні в системі, а й помилки вимірювання, зумовлені способом вимірювання. Наприклад, якщо вимірювана величина піддається дискретизації, то похибку дискретизації також слід вважати шумом вимірювання. Своєю чергою, під час вимірювання напруги стрілочним вольтметром похибки вимірювання, зумовлені кінцевим часом реакції, нелінійністю механічних та електромагнітних частин приладу, роблять свій внесок у шум вимірювання.

Шуми вимірювання  $w(t)$  можуть бути адитивними, мультиплікативними, змішаними, а також параметричними, конкретний вигляд яких зумовлюється процесами, що відбуваються під час вимірювання:

- адитивний шум:  $x^w(t) = x_0 + w_a(t)$ ;
- мультиплікативний шум:  $x^w(t) = x_0 \cdot w_m(t)$ ;

- змішаний шум:  $x^w(t) = (x_0 + w_a(t)) \cdot w_m(t)$ .

У випадку безпосереднього вимірювання шум найчастіше має адитивну природу і зазвичай включає дві компоненти – випадкову та систематичну:

$$w_a(t) = w_r(t) + w_s.$$

Характеристики випадкової компоненти  $w_r(t)$  можуть бути оцінені за результатами вимірювань. Значення систематичної компоненти  $w_s$  або взагалі невідоме, або визначається апіорно (задається довільно для тестових задач або визначається фізикою процесу для реальних вимірювань).

Під час моделювання процесів вимірювання як  $w_a(t)$  часто використовують шум із нормальним розподілом, що характеризується середньоквадратичним відхиленням  $\sigma^w$  та часом автокореляції  $\tau^w$ . Рідше (наприклад, під час порогових вимірювань) для моделювання застосовують шум із рівномірним розподілом.

У тих випадках, коли вимірювана величина є параметром деякого процесу, шум часто має мультиплікативну природу [22]. Параметри такого шуму визначаються параметрами вхідного сигналу та характеристиками процесу. Під час дослідження систем і динамічних об'єктів реакцію сигналу похибки на вхідний сигнал доцільно описувати як шум мультиплікативної природи.

Вимірювання завжди включає в себе певний вид усереднення (фільтрації). З одного боку, таке усереднення визначається фізичними процесами, що відбуваються під час вимірювання (наприклад, механічною інерцією вимірювального приладу, сталими часу електромагнітних систем тощо). З іншого боку, додаткове усереднення застосовують для підвищення точності та надійності вимірювання. При цьому можуть використовуватися спеціальні методи фільтрації [30] за умови наявності відомостей про вид шумового впливу.

Оцінимо систематичну компоненту похибки вимірювання експерименту.

Припустимо тепер, що  $w_s \neq 0$ . У цьому випадку вираз (4.4) є непридатним для інформаційної оцінки, оскільки жодним чином не враховує зміщення розподілу відносно вимірюваного значення.

Розглянемо приклад з нормальним розподілом та ненульовою систематичною похибкою  $w_s \neq 0$ , тоді:

$$\sigma_x(\tau) \approx \sqrt{\sigma_w^2 \frac{\tau_w}{\tau} + w_s^2} \quad \text{при} \quad \lim_{\tau \rightarrow \infty} \sigma_x(\tau) = |w_s|. \quad (4.6)$$

Отримаємо оцінку ентропії:

$$\tilde{H}(\sigma_x) = \frac{1}{2 \ln 2} \ln(2\pi e \sigma_x^2). \quad (4.7)$$

Це означення збігається при  $w_s = 0$  і дає адекватні результати у протилежному випадку. Отримаємо:

$$\tilde{H}(\tau) = \frac{1}{2 \ln 2} \ln \left( 2\pi e \left( \sigma_w^2 \frac{\tau_w}{\tau} + w_s^2 \right) \right); \quad (4.8)$$

$$\tilde{I}(\tau) = H_0 - \tilde{H}(\tau) = \frac{1}{2 \ln 2} \ln \left( \frac{\sigma_{x0}^2}{\sigma_w^2 \frac{\tau_w}{\tau} + w_s^2} \right). \quad (4.9)$$

Ентропія не зменшуватиметься нескінченно, а відповідно прямуватиме до скінченної величини, обмеженої знизу (рис. 4.8):

$$\lim_{\tau \rightarrow \infty} \tilde{H}(\tau) = \frac{1}{2 \ln 2} \ln(2\pi e w_s^2). \quad (4.10)$$

Своєю чергою, кількість отриманої інформації не зростатиме нескінченно з плином часу, а прямуватиме до скінченної величини, визначеної способом вимірювання (рис. 4.8):

$$I_{s,lim} = \lim_{\tau \rightarrow \infty} I(\tau) = \frac{1}{2 \ln 2} \ln \left( \frac{\sigma_{x0}^2}{w_s^2} \right) = \frac{1}{\ln 2} \ln \left( \frac{\sigma_{x0}}{|w_s|} \right). \quad (4.11)$$

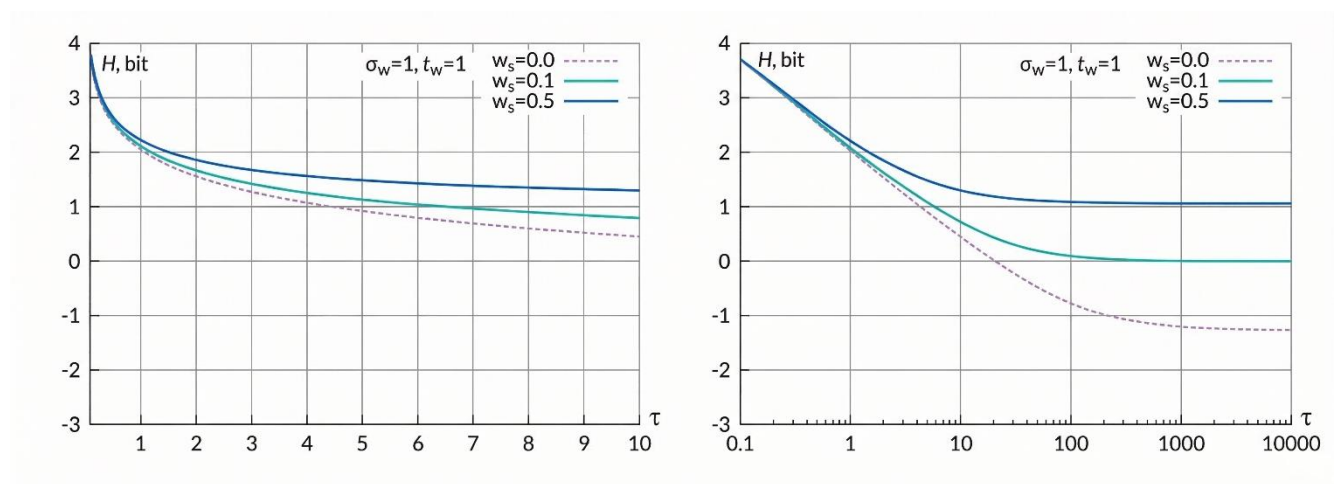


Рис. 4.8 – Залежність оцінки ентропії  $\tilde{H}$  від  $\tau$  за наявності систематичної похибки

Середня швидкість отримання інформації  $\bar{v}_i(\tau)$  (рис. 4.9) буде меншою, ніж у випадку відсутності систематичної похибки.

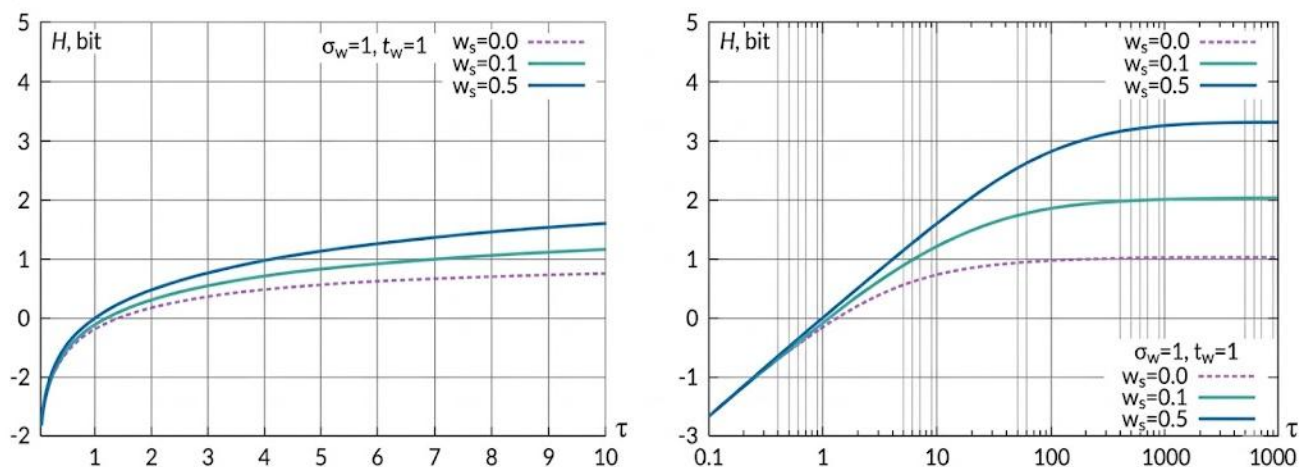


Рис. 4.9 – Залежність кількості отриманої інформації  $\tilde{I}$  від  $\tau$  у випадку наявності систематичної похибки

За заданих значень  $\sigma_{x0}$  та  $w_s$  величина  $I_{s,lim}$  визначає максимальну кількість інформації, яку можна отримати під час вимірювання, або, іншими словами, інформаційну ємність вимірювання. Отримання цієї кількості інформації потребує нескінченного часу.

Точкам перетину асимптот на логарифмічних графіках рис. 4.8 і рис. 4.9 відповідає таке значення часу вимірювання  $\tau_s$ , при якому внесок шумів і систематичної похибки є однаковим:

$$\sigma_w^2 \frac{\tau_w}{\tau_s} = w_s^2 \Rightarrow \tau_s = \tau_w \frac{\sigma_w^2}{w_s^2}. \quad (4.12)$$

Залежність середньої швидкості отримання інформації від  $\tau$  показана на рис. 4.10.

Час  $\tau_s$  у цьому випадку задає характерний часовий масштаб для даного процесу вимірювання. Тому є сенс увести безрозмірний час таким чином:

$$\tilde{\tau}_s = \frac{\tau}{\tau_s} = \frac{\tau w_s^2}{\tau_w \sigma_w^2}. \quad (4.13)$$

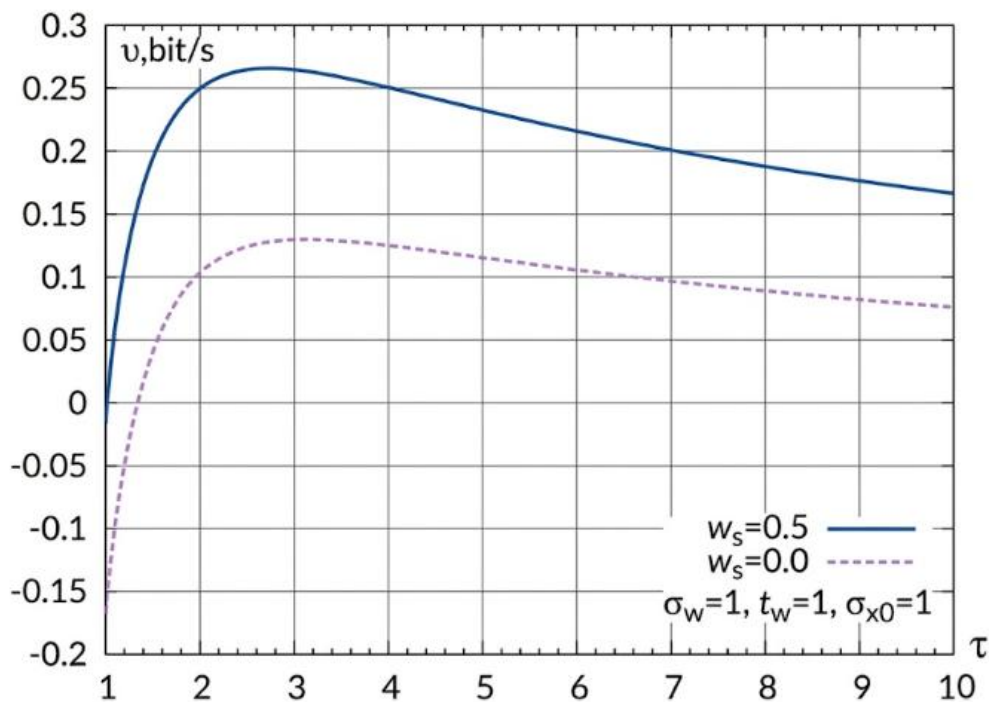


Рис. 4.10 – Залежність середньої швидкості отримання інформації від  $\tau$  у випадку наявності систематичної похибки  $w_s$

З використанням такого позначення запишемо:

$$I(\tilde{\tau}_s) = \frac{1}{2 \ln 2} \ln \left( \frac{\sigma_{x0}^2}{w_s^2 \left(1 + \frac{1}{\tilde{\tau}_s}\right)} \right). \quad (4.14)$$

Різниця між інформаційною ємністю та отриманою інформацією

$$I_{s,lim} - I(\tilde{\tau}_s) = \frac{1}{2 \ln 2} \ln \left( 1 + \frac{1}{\tilde{\tau}_s} \right). \quad (4.15)$$

визначає кількість інформації, яку можна отримати, виконуючи вимірювання від поточного моменту часу до нескінченності.

Ця величина дуже швидко зменшується з плином часу. Тому не має сенсу проводити вимірювання на часових інтервалах, довших за  $\tau_s$ , оскільки кількість інформації, отриманої після цього часу, є незначною – близько 0,5 біта.

Введемо безрозмірну величину  $E_s$ :

$$E_s = \sqrt{\frac{\tau_s}{\tau_0}} = \frac{\sigma_{x0}}{w_s}, \quad (4.16)$$



яка показує, наскільки доцільним є проведення вимірювання за заданих умов з урахуванням статичної похибки. Наприклад, при  $E_s \approx 1$  практично неможливо отримати більше інформації про вимірюваний об'єкт, ніж було відомо апіорно. Чим більшою є величина  $E_s$ , тим більше інформації можна отримати порівняно з апіорною:

$$I_{s,lim} = \frac{1}{\ln 2} \ln E_s. \quad (4.17)$$

Розглянемо нестационарну вимірювану величину.

Нехай вимірювана величина змінює своє значення з постійною швидкістю  $v_x = dx/dt$ . При цьому систематична похибка є незначною порівняно з динамічним зміщенням параметра, тоді

$$\sigma_x(\tau) \approx \sqrt{\sigma_w^2 \frac{\tau_w}{\tau} + v_x^2 \tau^2} \quad \text{при} \quad \lim_{\tau \rightarrow \infty} \sigma_x(\tau) = \infty. \quad (4.18)$$

Використовуючи означення (4.8), отримаємо

$$\tilde{H}(\tau) = \frac{1}{2 \ln 2} \ln \left( 2\pi e \left( \sigma_w^2 \frac{\tau_w}{\tau} + v_x^2 \tau^2 \right) \right); \quad (4.19)$$

$$\tilde{I}(\tau) = H_0 - \tilde{H}(\tau) = \frac{1}{2 \ln 2} \ln \left( \frac{\sigma_{x0}^2}{\sigma_w^2 \frac{\tau_w}{\tau} + v_x^2 \tau^2} \right). \quad (4.20)$$

У цьому випадку значення ентропії проходить, відповідно, через мінімум (рис. 4.11), а інформації – через максимум (рис. 4.12). Значення цього максимуму дорівнює

$$I_{d,max} = \frac{1}{2 \ln 2} \ln \frac{\sigma_{x0}^2}{\left( \frac{2v_x^2 \sigma_w^4}{\tau_w} \right)^{\frac{2}{3}} + \left( \frac{v_x \sigma_w^2 \tau_w}{2} \right)^{\frac{1}{3}}}. \quad (4.21)$$

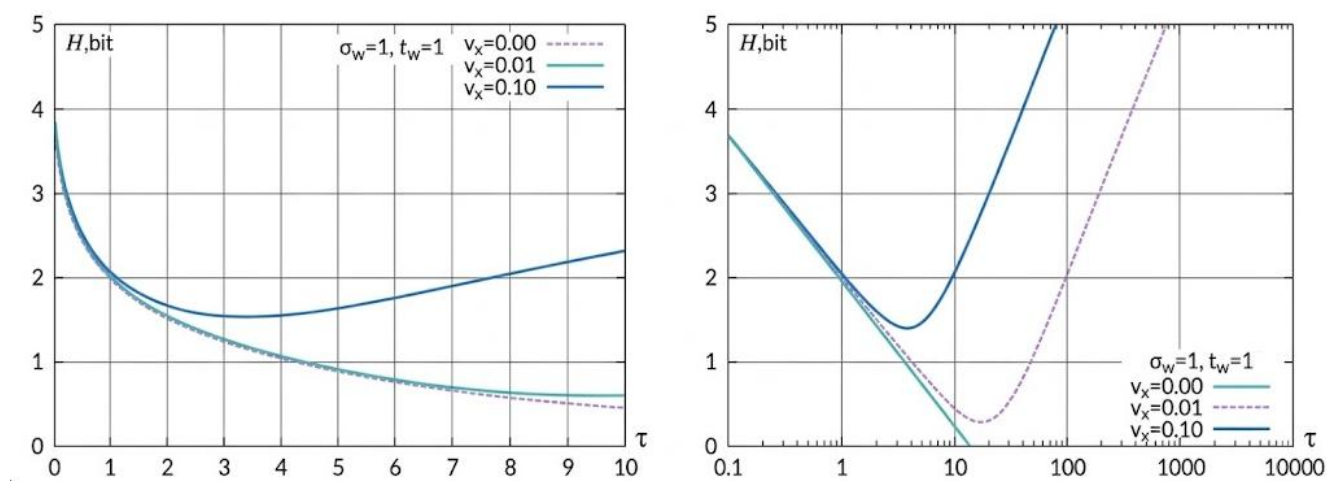


Рис. 4.11 – Залежність ентропії  $H$  від  $\tau$  у випадку нестационарного параметра

Максимум ентропії досягається при часі оцінювання, що дорівнює

$$\tau_d^* = \sqrt[3]{\frac{\sigma_w^2 \tau_w}{2v_x^2}}. \quad (4.22)$$

і тим самим визначаючи інформаційну ємність вимірювання для нестационарного параметра.

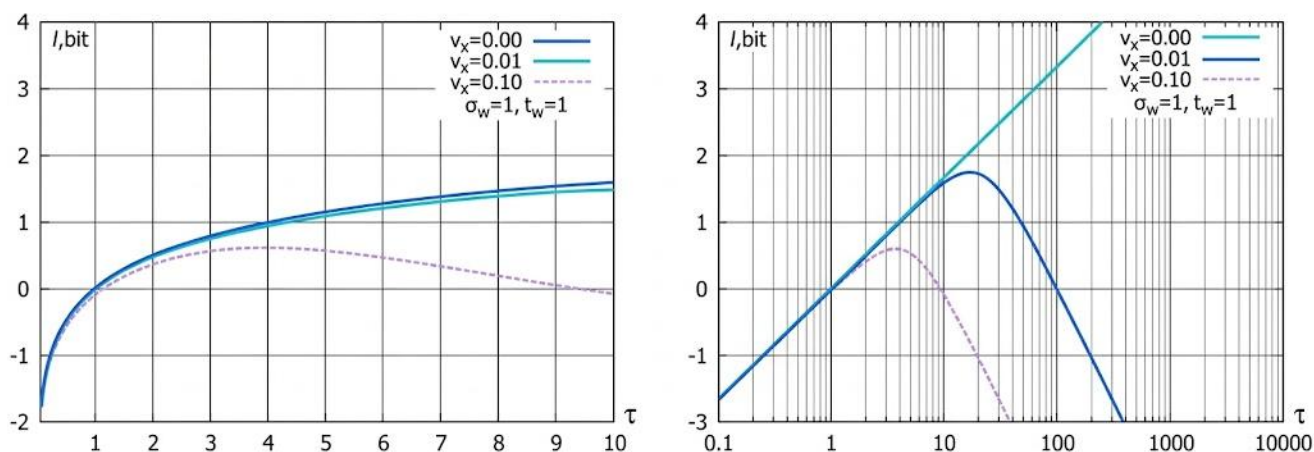


Рис. 4.12 – Залежність кількості отриманої інформації  $I$  від  $\tau$  у випадку нестационарності ідентифікованого параметра

Проводячи спостереження на часових інтервалах, більших за  $\tau_d^*$ , зменшується кількість відомої нам інформації про систему. Отже, для проведення вимірювань у системах з нестационарним параметром доцільно застосовувати

методи усереднення другого типу, причому час усереднення методу не повинен перевищувати  $\tau_d^*$ .

Безрозмірна величина  $\Xi_d$ :

$$\Xi_d = \sqrt{\frac{\tau_d^*}{\tau_0}} = \frac{\sigma_{x0}}{\sqrt[3]{2v_x\tau_w\sigma_w^2}} \quad (4.23)$$

показує, наскільки доцільним є проведення вимірювання за заданих умов з урахуванням нестационарності.

#### 4.6. Методи ефективного управління мережевими ресурсами

Сучасні телекомунікаційні мережі характеризуються високою динамічністю структури трафіку, зумовленою зростанням кількості мультимедійних сервісів [172], мобільних застосунків, хмарних обчислень та IoT-пристроїв [108]. Інтенсивність потоків даних є нестационарною та має виражений імпульсний характер, що проявляється у виникненні короточасних пікових навантажень і перевантаженні окремих сегментів мережі [134].

Застосування статичних політик керування ресурсами не забезпечує належного рівня якості обслуговування, оскільки фіксований розподіл смуги пропускання не враховує змін трафіку в реальному часі. Це призводить до:

- збільшення затримок передавання пакетів;
- зростання ймовірності втрат;
- погіршення показників QoS;
- неефективного використання пропускну здатності.

У зв'язку з цим виникає потреба в розробленні адаптивних методів керування [13], які дозволяють оперативно змінювати розподіл ресурсів відповідно до поточного стану мережі.

Як ефективний спосіб управління мережею пропонується адаптивний спосіб керування [26] з кільцевим контуром (рис. 4.13).



Рис. 4.13 – Адаптивний алгоритм управління ресурсами мережі

Алгоритм передбачає декілька етапів:

1. Моніторинг стану мережі (1.1. Вимірюємо затримку пакету; 1.2. Визначаємо джитер, 1.3. Оцінюємо втрати пакетів; 1.4. Визначаємо рівень завантаження каналів);

2. Оцінюємо QoS (2.1. Обчислюємо інтегральний показник QoS; 2.2. Виконуємо статистичний аналіз метрик; 2.3. Якщо QoS відповідає вимогам, то переходимо до кроку 8) [1];

3. Прогнозування навантаження (3.1. Застосовуємо модель прогнозування (ARIMA/ML); 3.2. Оцінюємо майбутню інтенсивність трафіку);

4. Оптимізація розподілу ресурсів (4.1. Формуємо цільову функцію, що враховує дані з попередніх етапів; 4.2. Враховуємо обмеження пропускної здатності; 4.3. Розв'язуємо задачу оптимізації) [5];

5. Адаптивне переналаштування мережі (5.1. Оновлюємо маршрути передачі пакетів; 5.2. Виконуємо перерозподіл смуги пропускання; 5.3. Налаштовуємо параметри черг QoS-політики) [9, 10];

6. Пріоритезація трафіку; (6.1. Розподіляємо трафік за типами сервісів; 6.2. Надаємо пріоритет критичним сервісам (VoIP, відео, дані)) [10];

7. Реалізація керуючого впливу (7.1. Передаємо керуючі команди до мережевих пристроїв; 7.2. Застосовуємо зміни через SDN-контролер або локальні алгоритми);

8. Перехід до наступного циклу (Повертаємося до кроку 1).

Запропонований автором алгоритм належить до класу адаптивних систем керування із замкненим контуром та забезпечує динамічну оптимізацію параметрів мережі на основі поточного стану та прогнозованого навантаження, що дозволяє підвищити ефективність використання ресурсів та гарантувати QoS. Цей алгоритм має наступні переваги:

- забезпечує динамічне балансування трафіку;
- підвищує відмовостійкість мережі;
- гарантує QoS для різних класів сервісів;
- створює основу для реалізації інтелектуальних мереж SDN/NFV [37, 67].

Після проведених досліджень можна оцінити отримані практичні результати дисертаційного дослідження. Вони дозволили:

- зменшити затримки при передачі пакетів на  $\approx 15$  %;
- знизити перевантаження в мережі (рис. Г.5 з додатків) на 30 %;
- збільшити пропускну здатність на 20 % за рахунок розбиття блоку пакетів на два підблоки;
- покращити показник QoS (рис. Г.6 з додатків) на  $\approx 33$  %;
- підвищити доступність мережі за рахунок пріоритетних сервісів (рис. Г.7 з додатків).

На відміну від класичних підходів [34, 45], розроблений метод базується на інтеграції моніторингу, прогнозування та оптимізації, що дозволяє досягти більш високої ефективності використання мережевих ресурсів в умовах змінного навантаження та дозволяє гарантувати QoS для різних класів сервісів із заданою точністю.

#### **Висновки до розділу 4**

Розроблений алгоритм дає змогу проводити розрахунок часових характеристик комбінованих телекомунікаційних послуг з можливістю застосування методів апроксимації для мереж систем довільного типу G/G/1.

Також були наведені приклади розрахунку транзакційної послуги з кількома сценаріями можливих варіантів реалізації робочих процесів обслуговування запиту, що наочно демонструє вплив моделювання на процес прийняття рішень для веб-сервісів.

Окрім методів математичного моделювання, були запропоновані підходи до інтеграції імітаційних моделей для оцінки характеристик складних мереж СМО. В рамках цієї роботи була створена бібліотека імітаційних моделей, для якої було запропоновано рішення для інтеграції апарату паралельної обробки запитів Fork-Join як окремий елемент структури з подальшим комбінуванням у різноманітні масштабні моделі.

Проведені практичні експерименти та отримані результати передачі трафіку VoIP за різних умов і режимів роботи мережі.

Також оцінені похибки експериментальних вимірювань і похибки моделювання.

## ВИСНОВКИ

У дисертаційній роботі вирішено актуальне наукове завдання розроблення моделей і методів оцінювання ймовірно-часових характеристик телекомунікаційних послуг і транзакційних робочих процесів у мікросервісних архітектурах, що дозволяє підвищити ефективність функціонування телекомунікаційних систем і забезпечити необхідні показники якості обслуговування кінцевих користувачів.

На основі проведених досліджень отримано такі основні результати та сформульовано відповідні висновки:

1. Проведено комплексний аналіз сучасних телекомунікаційних послуг і архітектур їх реалізації. Встановлено, що сучасні телекомунікаційні системи характеризуються високим рівнем складності, стохастичними властивостями та нечіткими параметрами якості обслуговування, що обумовлено гетерогенністю трафіку, змінністю навантаження, використанням мікросервісної архітектури та необхідністю забезпечення QoS/QoE. Показано, що існуючі підходи недостатньо враховують стохастичні, інформаційно-теоретичні та структурні особливості комбінованих телекомунікаційних послуг, що обґрунтовує необхідність розроблення нових моделей і методів їх аналізу.

2. Розроблено математичні моделі для оцінювання часових характеристик телекомунікаційних послуг, що підтримують користувацькі сеанси. Запропоновано підхід до моделювання телекомунікаційних послуг як СМО з урахуванням стохастичних параметрів надходження запитів і часу обслуговування. Використано математичний апарат мереж Джексона для аналізу характеристик багатокористувацьких систем, що дозволяє визначати середній час обслуговування, час очікування та інші показники продуктивності. Розроблені моделі забезпечують високу точність оцінювання часових характеристик і можуть бути застосовані для аналізу складних телекомунікаційних сервісів.

3. Уперше розроблено модель для оцінювання характеристик транзакційних телекомунікаційних послуг із паралельною обробкою запитів. Запропоновано



математичну модель на основі механізму Fork-Join, яка дозволяє враховувати паралельне виконання транзакцій і синхронізацію потоків. Розроблений підхід дозволяє оцінювати часові характеристики робочих процесів, визначати затримки виконання транзакцій і аналізувати ефективність функціонування систем із паралельною обробкою запитів.

4. Розроблено метод апроксимації для визначення характеристик систем довільного типу  $G/G/1$ . Запропонований метод дозволяє визначати ЙЧХ СМО загального типу без обмеження щодо типу розподілу процесів надходження та обслуговування. Це забезпечує можливість аналізу реальних телекомунікаційних систем, які не відповідають класичним припущенням експоненціального розподілу.

5. Удосконалено методи оцінювання часових характеристик комбінованих телекомунікаційних послуг. Розроблено математичний апарат для визначення часових характеристик послідовних і комбінованих СМО, що дозволяє враховувати складні сценарії взаємодії між компонентами мікросервісної архітектури. Запропоновані моделі забезпечують більш точну оцінку продуктивності систем і дозволяють прогнозувати їх поведінку в різних умовах навантаження.

6. Розроблено комплекс імітаційних моделей і програмних засобів для дослідження телекомунікаційних систем. Створено програмну реалізацію моделей із використанням MatLab і спеціалізованих математичних пакетів, що дозволяє проводити ІМ систем різної складності. Розроблений комплекс забезпечує можливість аналізу часових характеристик систем, перевірки аналітичних моделей і оцінювання ефективності різних архітектурних рішень.

7. Проведено експериментальні дослідження та ІМ запропонованих моделей і методів. Результати досліджень підтвердили адекватність і високу точність розроблених моделей. Встановлено, що запропоновані методи дозволяють отримувати результати з похибкою, що не перевищує допустимих значень, і можуть ефективно використовуватися для оцінювання характеристик телекомунікаційних систем у реальних умовах.

8. Доведено практичну ефективність запропонованих моделей і методів. Розроблені моделі та методи можуть використовуватися при проектуванні, аналізі

та оптимізації телекомунікаційних послуг і систем, що базуються на мікросервісній архітектурі. Їх застосування дозволяє підвищити ефективність управління телекомунікаційними ресурсами, забезпечити необхідний рівень QoS/QoE та покращити продуктивність телекомунікаційних систем.

Таким чином, у дисертаційній роботі вирішено наукове завдання розроблення моделей і методів оцінювання характеристик телекомунікаційних послуг і транзакційних робочих процесів у сучасних телекомунікаційних системах. Запропоновані моделі, методи та алгоритми забезпечують підвищення ефективності функціонування телекомунікаційних систем, дозволяють більш точно прогнозувати їх характеристики та можуть бути використані при проєктуванні, аналізі та оптимізації сучасних телекомунікаційних мереж і сервісів.

Отримані результати мають важливе теоретичне та практичне значення для розвитку телекомунікаційних систем, зокрема систем, побудованих на основі мікросервісної архітектури, та можуть бути використані при створенні сучасних інформаційно-комунікаційних технологій.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Брезіцький С.М., Герасимчук В.С. (2024). Інтегральні методи оцінки продуктивності пакетних мереж: аналіз та формування. Наукові записки ДУІКТ. №2(6). С. 37 – 47. DOI: 10.31673/2786-8362.2024.025073
2. Вікторія Ключєва, Владислав Герасимчук, Ігор Данилов. (2024). Типові математичні моделі, які застосовуються в автоматизованих системах організаційного управління. SMART TECHNOLOGIES: Industrial and Civil Engineering. Issue 2(15). С. 4 – 9. DOI: 10.32347/st.2024.2.1101
3. Брезіцький С.М., Герасимчук В.С. (2025). Методика формування інтегральних показників якості роботи мережевих з'єднань в мережах з комутацією пакетів. Наукові записки ДУІКТ. №1(7). С. 90 – 96. DOI: 10.31673/2786-8362.2025.016443
4. Пепа Ю.В., Герасимчук В.С. (2026). Інтеграція AI в процес управління телекомунікаційними послугами на прикладі VoIP. Прикладні системи управління та робототехніка: кол. моногр. / За заг. ред. С.О. Довгого. – Київ, ІПСУ НАН України. С. 185 – 187. ISBN 978-617-8261-54-2. DOI: 10.35668/978-617-8261-54-2
5. Герасимчук В.С., Пепа Ю.В. (2026) Моделі вибору оптимального трафіку передачі даних в інфокомунікаційних мережах. Телекомунікаційні та інформаційні технології. №1(90). С. 59 – 68. DOI: 10.31673/2412-4338.2026.019007
6. Герасимчук Владислав Сергійович, Федчук Володимир Сергійович (2023) Перспективи розвитку та впровадження новітніх технологій оптоволоконного зв'язку. V Всеукраїнська науково-практична конференція «Telecommunication: Problems and Innovation». Збірник тез (м. Київ, 20 грудня 2023 р.). С. 43 – 44.
7. В.С. Герасимчук (2025) Оцінка якості надання телекомунікаційних та інформаційних послуг. Науково-практична конференція «Актуальні проблеми безпеки інформаційно-комунікаційних систем». Збірник тез (м. Київ, 5 листопада 2025 р.). С. 45 – 46.

8. Крикун Юлія Віталіївна, Герасимчук Владислав Сергійович (2025) Розподілена біометрія: автентифікація без центрального зберігання. III Всеукраїнська науково-технічна конференція «Технологічні горизонти: дослідження та застосування інформаційних технологій для технологічного прогресу України і світу». Збірник тез (м. Київ, 18 листопада 2025 р.). С. 260 – 261.
9. Герасимчук Владислав Сергійович (2025) Методика розрахунку транзакційних послуг. III Міжнародна науково-практична конференція «Сучасні аспекти діджиталізації та інформатизації в програмній та комп'ютерній інженерії». Збірник тез (м. Київ, 4 - 6 грудня 2025 р.). С. 394 – 396.
10. Vladyslav HERASYMCHUK, Yuriy PEPA (2026). Algorithm for Synthesizing Background Network Traffic. The V International Conference on Emerging «Technology Trends on the Smart Industry and the Internet of Things “TTSIIT-2026”». Ukraine-Iraq-Poland (Kyiv, Ukraine, January 29 - 30). pp. 37 – 40.
11. Olha MARCHUK, Vladyslav HERASYMCHUK (2026). Interference Immunity of MIMO Systems Under Different Probability Distributions of Fading. The V International Conference on Emerging «Technology Trends on the Smart Industry and the Internet of Things “TTSIIT-2026”». Ukraine-Iraq-Poland (Kyiv, Ukraine, January 29 - 30). pp. 130 – 132.
12. Поповський В.В. (2018) Основи теорії телекомунікаційних систем: підручник. Харків: ХНУРЕ. 368 с. ISBN 978-966-659-233-3
13. Беркман Л.Н., Климаш М.М., Костік Б.Я., Олійник В.Ф., Остапенко І.А., Поповський В.В. та ін. (2009) Система управління сучасними телекомунікаційними мережами: монографія. К.: Зв'язок. С. 259 – 266 ISBN 978-966-2970-30-2
14. Романов О.І., Гришко Н.М., Нітенко А.С. (1999) Управління потоками мовних повідомлень на мережах зв'язку спеціального призначення (автореф.). 36 с.
15. Лунтовський А.О., Захарченко М.В., Семенко А.І. (2013) Мультисервісні мобільні платформи: монографія. К.: Одеса : ОНАЗ ім. О. С. Попова. 216 с.
16. Соловська І.М., Стрілковська І.В. (2013) Оценка характеристик качества обслуживания разнородного трафика в мультисервисной сети. Цифрові технології. Збірник наук. праць. №13. С. 34 – 41. УДК 621.391

17. Ложковський А.Г., Вербанов О.В., Колчар В.М. та ін. (2014) Методи проектування телекомунікаційних систем та мереж в умовах реального трафіка. Наукові праці ОНАЗ ім. О.С. Попова. №2. С. 54 – 63. УДК 621.39
18. Ложковський А.Г. (2010) Теорія масового обслуговування в телекомунікаціях. Одеса: ОНАЗ ім. О.С. Попова. С. 112 УДК 621.39
19. А. Г. Ложковський, В. Ю. Гордієнко. (2010) Функція розподілу станів системи масового обслуговування та ймовірність відмов для моделі типу  $ММ/D/m$ . Вісник Національного університету «Львівська політехніка». Сер. Радіoeлектроніка та телекомунікації. № 680. С.109 – 113.
20. Захарченко М.В., Вараксін О.О., Кононович В.Г., Вараксін С.О. (2008) Протоколи, термінальне обладнання та інформаційна безпека у мережах наступного покоління: навч. посібник; за ред. М.В. Захарченка. Одеса: ОНАЗ ім. О.С. Попова, С. 128 УДК 621.39: 681.324
21. Н.В. Захарченко, А.И. Семенко, В.В. Корчинский, Е.А. Домрачева (2016) Использование таймерных сигнальных конструкций для мониторинга сигналов аналогового датчика. Зв'язок. №6(124). С. 40 – 43. ISSN 2412-9070
22. Варфоломеева О. Г., Колченко Т. В. (2013) Нові підходи до управління телекомунікаційними мережами. Наукові записки УНДІЗ. №1(25). С. 41 – 44.
23. Бондаренко В.Є. (2016) Задача побудови замкнутих систем та її реалізація для синтезу живучих кільцевих транспортних телекомунікаційних мереж. Журнал «Телекомунікаційні та інформаційні технології». №1. С. 48 – 55. ISSN 2412-4338
24. Бондаренко В.Є. (2016) Модифікація методу аналізу ієрархій для прийняття рішень у складних системах. Наукові записки УНДІЗ. №1(41). С.20 – 25. УДК 51:62-505.72
25. Юнчик А.О., Кравчук С.О. (2011) Архітектурні рішення ретрансляцій в мережі LTE-Advanced. Механіка гіроскопічних систем. № 23. С. 92 – 101. УДК 621.39
26. Скляр М.О., Кравчук С.О. (2012) Динамічне управління радіоресурсом у гетерогенних мережах з точки зору виконання вимог QoS. Матер. 6-ї міжнар.

наук.-техн. конф. “Проблеми телекомунікацій”, 24 – 27 квітня, 2012 р. К.: НТУУ «КПІ». С. 82 – 84

27. Ільченко М.Ю., Кравчук С.О. (2017) Телекомунікаційні системи. Моногр. Київ: Наукова думка. 730 с. ISBN 978-966-00-1566-1

28. Уривський Л.О. Мошинська А. В. Осипчук С. О. (2022) Імітаційне моделювання систем і процесів у телекомунікаціях: навч. посіб. Київ КПІ ім. Ігоря Сікорського: КПІ ім. Ігоря Сікорського. 202 с.

29. Уривський Л.О., Шмігель Б.О. (2022) Оцінка ефективності використання ресурсів безпроводових каналів зв'язку. ВЧЕНІ ЗАПИСКИ Таврійського національного університету імені В.І. Вернадського Серія: Технічні науки. Том 33(72), №5. С. 27-32. DOI <https://doi.org/10.32782/2663-5941/2022.5/04>

30. Лебедев О.М., Ладик О.І. (2004) Цифрова техніка. К.: Політехніка. 320 с.

31. Tanenbaum A.S., Wetherall D.J. (2010) Computer Networks (5th Edition). Pearson Education. 960 p. ISBN-13: 978-0-13-212695-3

32. Tanenbaum A.S., van Steen M. (2006) Distributed Systems: Principles and Paradigms (2nd Edition). Pearson. 704 p. ISBN-13: 978-0132392273

33. Stallings W. (2017) Data and Computer Communications (10th Edition). PEARSON INDIA. 280 p. ISBN-13: 978-9332586932

34. Kurose J.F., Ross K.W. (2021) Computer Networking: A Top-Down Approach, Global Edition (8th). Pearson. 800 p. ISBN-13: 978-1292405513

35. Гранатуров В. М., Литовченко І. В. (2010). Управління послугами зв'язку. Навчальний посібник. Львів: ЛПІ. 254 с.

36. Kapiton A., Talybov R., Dziuban O., Kanivets B. (2023). Service Quality Management in Telecommunication Communication Networks: Need for User-Centric Methods. Scientific Journal. Control, Navigation and Communication Systems. №4. pp. 156 – 159. DOI: 10.26906/SUNZ.2023.4.156

37. Pidpalyi O. (2024). Optimising Service Quality and Network Efficiency by SDN and Broadband Access Integration. Information Technologies and Computer Engineering. Vol. 21, No. 3. pp. 29 – 42. DOI: 10.63341/vitce/3.2024.29

38. Е. М. Бовда, О. М. Гук, О. Г. Гаврилюк (2018). МЕТОДИКА ОЦІНКИ ЕФЕКТИВНОСТІ СИСТЕМИ УПРАВЛІННЯ ТЕЛЕКОМУНІКАЦІЙНИМИ МЕРЕЖАМИ. *Modern Information Technologies in the Sphere of Security and Defence* № 2(32) p. 125 – 134. ISSN 2410-7336

39. Demóstenes Z. Rodriguez, Renata L. Rosa, Rodrigo D. Nunes and Emmanuel T. Affonso. Assessment of Quality-of-Experience in Telecommunication Services. *International Journal of Digital Information and Wireless Communications (IJDIWC)* № 6(4). pp. 241 – 259. ISSN: 2225-658X

40. М. Medvetskyi, Mykola Beshley, A. Pryslupskyi. (2021) Метод управління якістю сприйняття послуг для програмно-конфігурованих мереж заснованих на намірах ICTEE. Volume 1, № 1. pp. 76 - 85. DOI: 10.23939/ict2021.01.076

41. Bulashenko A.V. (2020) Resource allocation for low-power devices of M2M technology in 5G networks. *KPI Science news*. Vol. 3. pp. 7 – 13. DOI: 10.20535/kpi-sn.2020.3.203863.

42. Nalapko O., Sova O., Shyshatskyi A. et al. (2021). Analysis of Methods for Increasing Efficiency of Dynamic Routing Protocols in Telecommunication Networks. *Technology Audit and Production Reserves*. № 5(2). pp. 44 – 48. DOI: 10.15587/2706-5448.2021.239096.

43. Mozhaev O., Kuchuk H., Kuchuk N., Mozhaiv M. (2017) Method for Calculating of E-Learning Traffic Peakedness. Четверта IEEE науково-практична конференції «Проблеми інфокомунікацій. Наука та технології». pp. 141-145.

44. Ліпенцев А. В., Думенкова, Т. Г. (2018). Підвищення ефективності надання послуг call-центрами. Ефективність державного управління. № 57. С. 13 – 25. DOI: 10.33990/2070-4011.57.2018.164590

45. Alkurd R., Abualhaol I., Yanikomeroglu H. (2020). Big-data-Driven and AI-Based Framework for Personalization in Wireless Networks. *IEEE Communications Magazine*. March 2020. pp. 18 – 24. DOI: 10.1109/MCOM.001.1900533

46. Radivilova T., Kirichenko L., Ageiev D., Bulakh V. (2020). Methods to Improve QoS by Accounting Secure Parameters. *Advances in Computer Science for*

Engineering and Education II. Advances in Intelligent Systems and Computing, Vol. 938. pp. 13 – 22. DOI: 10.1007/978-3-030-16621-2\_32

47. Толубко В. Б., Беркман Л. Н., Макаренко А. О., Заїка В. Ф., Березнюк А. В. (2019). Удосконалення процесів управління телекомунікаційними мережами за стандартом Telecommunications Management Network. Наукові записки УНДІЗ. №2(54). С. 5 – 13. ISSN 2518-7678

48. Козловський В.В., Туровський О.Л., Кулінський В.Д. (2020). Формалізація вимог до системи управління телекомунікаційними мережами. Проблеми інформатизації та управління, №64, С. 41 – 47. DOI: 10.18372/2073-4751.64.15148

49. Гороховський Є.П., Зіненко Ю.М., Крючкова Л.П., Борисенко І.І. (2016). Застосування методу модифікованих робочих характеристик для розрахунку показників якості систем управління телекомунікаційними мережами. Наукові записки УНДІЗ. №4(44). С. 5 – 12. ISSN 2518-7678

50. Лебеденко Т.М., Головешко М.В., Холодкова А.В. (2019). Дослідження методу активного управління чергами на інтерфейсах маршрутизаторів телекомунікаційних мереж. Системи управління, навігації та зв'язку, (розділ «Зв'язок, телекомунікації та радіотехніка»). № 4(56) С. 57 – 62. DOI: 10.26906/SUNZ.2019.4.057

51. Лебеденко Т.М., Головешко М.В., Холодкова А.В. (2019). Результати експериментального дослідження методу активного управління чергами на інтерфейсах телекомунікаційних мереж. Електронне наукове фахове видання – журнал «Проблеми телекомунікацій». №2(25). С. 37 – 55. УДК 621.391

52. О. В. Лемешко, А. С. Шаповалова, О. С. Єременко, М. О. Євдокименко, А. М. Хайлан. (2019) Математична модель швидкої перемаршрутизації з балансуванням навантаження та диференційованого обмеження трафіка в мережах SD-WAN. Системи управління, навігації та зв'язку. Том 4 № 56. С. 63 – 71. DOI: 10.26906/SUNZ.2019.4.063

53. Самохвалов Ю. Я., Бовда Е. М., Клименко В. М., Устинов Д. А. (2025). Управління балансуванням навантаження в гібридних SDN-мережах на основі



алгоритму ітераційного розслаблення з масштабуванням і округленням. Системи і технології зв'язку, інформатизації та кібербезпеки. ВІТІ. № 8. С. 153 – 162. DOI: 10.58254/viti.8.2025.13.153

54. А.А. Астраханцев, Л.С. Глоба, А.М. Давидюк, О.В. Сушко (2022). Дослідження ефективності алгоритмів машинного навчання для класифікації трафіка в мобільних мережах. Електронне наукове фахове видання–журнал «Проблеми телекомунікацій». №1(30). С. 3 – 17. DOI: 10.30837/pt.2022.1.01

55. Коляденко Ю.Ю., Чурсанов М.О. (2021). Аналіз ефективності методу вимірювання параметрів мережних елементів і сигналів в умовах зосереджених за спектром завад. Електронне наукове фахове видання–журнал «Проблеми телекомунікацій». №2(29). С. 36 – 44. DOI: 10.30837/pt.2021.2.04

56. Медвецький М., Бешля М., Прислупський Я., Бешля Г. (2021). Метод ініціації хендоверу в програмно-конфігурованій безпроводній мережі на основі показника якості сприйняття послуг. Інфокомунікаційні технології та електронна інженерія. №2(2). С. 1 – 10. DOI: 10.23939/ictee2021.02.001

57. Беркман Л.Н., Варфоломеева О.Г., Захаржевський А.Г., Твердохліб М.Г., Ярцев В.П. (2020). Управління якістю обслуговування в мережах NGN. Телекомунікаційні та інформаційні технології. №2(67). С. 4 – 12. DOI: 10.31673/2412-4338.2020.020412.

58. Воронець В., Пустовойтов П. (2024). Метод формування плану передачі пакетів при піковому навантаженні мережі, який знижує відгук. Системи управління, навігації та зв'язку, №1(75). С. 185 – 188. DOI: 10.26906/SUNZ.2024.1.185.

59. Ганзій В., Коваленко А., Ситник О. (2024). Аналіз методів управління процесами передачі даних та трафіком у мультисервісних комп'ютерних мережах. Системи управління, навігації та зв'язку. №1(75). С. 189 – 192. DOI: 10.26906/SUNZ.2024.1.189.

60. I. M. Syvolovskyi, V. P. Lysechko, O. S. Zhuchenko, O. M. Komar, V. V. Pastushenko (2024). Analysis of methods for organizing distributed telecommunication

systems using the paradigm of edge computing. Control, Navigation and Communication Systems. Vol.1 №75. pp. 206 – 210. DOI: 10.26906/SUNZ.2024.1.206

61. О.В. Лемешко, М.О. Євдокименко (2020). Вдосконалення потокової моделі маршрутизації в мультисервісній телекомунікаційній мережі із забезпеченням якості обслуговування. Системи озброєння і військова техніка. № 1(61). С. 31 – 43. DOI: 10.30748/soivt.2020.61.04

62. Я.І. Базуков, В.Х. Чакрян, А.А. Акулінічев, О.О. Мартинчук (2025). Лабораторне дослідження методу обслуговування черг, заснованого на класах, на маршрутизаторах IP-мереж. Електронне наукове фахове видання–журнал «Проблеми телекомунікацій». №2(37). С. 46 – 62. DOI: 10.30837/pt.2025.2.03

63. Євдокименко М.О., Кравченко Б.Д., Лукова-Чуйко Н.В. (2017). Аналіз проблеми забезпечення ефективного балансування навантаження в мережах MPLS-TE. Електронне наукове фахове видання–журнал «Проблеми телекомунікацій». №2(21). С. 32 – 41. УДК: 621.391

64. Євдокименко М.О., Шаповал М.М., Крепко А.В. (2020). Потокова модель маршрутизації з балансуванням навантаження на принципах TRAFFIC ENGINEERING з врахуванням ризиків інформаційної безпеки. Електронне наукове фахове видання–журнал «Проблеми телекомунікацій». №2(27). С. 32 – 42. DOI: 10.30837/pt.2020.2.03

65. Лемешко О.В., Єременко О.С. (2016). Динамічне представлення тензорної моделі багатошляхової QoS-маршрутизації у нестационарному режимі роботи інтерфейсів. Електронне наукове фахове видання–журнал «Проблеми телекомунікацій». №1(18). С. 3 – 21. УДК: 621.391

66. Кротов В.Д., Ткаченко А.Л., Науменко О.Г. (2019). Аналіз методів управління трафіком при забезпеченні QoS в мобільних радіомережах тактичної ланки управління. Збірник наукових праць ВІТІ. №1. С. 49–56.

67. Климаш М.М., Шпур О.М., Багрій В.О., Швець А.Л. (2014). МЕТОД ДИФЕРЕНЦІЙОВАНОГО МУЛЬТИПОТОКОВОГО КЕРУВАННЯ ТРАФІКОМ У ТРАНСПОРТНИХ ПРОГРАМНО-КЕРОВАНИХ МЕРЕЖАХ. Телекомунікаційні системи та мережі (Львівська політехніка). №796. С. 60 – 68. УДК 621-681

68. Кучернюк П.В. (2021). Технології моніторингу та трафік-інжинірингу в телекомунікаційних мережах. Навчальний посібник (КПІ). С. 257.
69. Беркман Л.Н., Климаш М.М., Костік Б.Я., Олійник В.Ф., Остапенко І.А., Поповський В.В., Проживальський О.П., Складенко С.М., Слюсар В.О. (2009). Система управління сучасними телекомунікаційними мережами. Монографія. К.: Зв'язок. (Бібліографічні дані наведені у профільній статті ДУТ).
70. Гніденко М.П., Вишнівський В.В., Ільїн О.О. (2019). Побудова SDN мереж: навчальний посібник. Київ: Державний університет телекомунікацій. 190 с. УДК 004.7(075)
71. Нікітюк Л.А. (2000). Телекомунікаційні технології цифрових мереж: навчальний посібник (за ред. М.В. Захарченка). Одеса: УДАЗ ім. О.С. Попова. 60 с.
72. Бірюков М.Л., Стеклов В.К., Костік Б.Я. (2005). Транспортні мережі телекомунікацій: системи мультиплексування (підручник). Київ: Техніка. 312 с.
73. Liyanage, M., Pham, Q.-V., Dev, K., Bhattacharya, S., Maddikunta, P.K.R., Gadekallu, T.R., Yenduri, G. (2022). A survey on Zero touch network and Service Management (ZSM) for 5G and beyond networks. *Journal of Network and Computer Applications*, 203, 103362. pp. 1 – 27. DOI: 10.1016/j.jnca.2022.103362.
74. Gallego-Madrid, J., Sanchez-Iborra, R., Ruiz, P.M., Skarmeta, A.F. (2022). Machine learning-based zero-touch network and service management: A survey. *Digital Communications and Networks*. Vol. 8(2). pp. 105–123. DOI: 10.1016/j.dcan.2021.09.001
75. El Rajab, M., Yang, L., Shami, A. (2024). Zero-touch networks: Towards next-generation network automation. *Computer Networks*. Vol. 243, 110294. pp. 1 – 15. DOI: 10.1016/j.comnet.2024.110294.
76. Mehmood, K., Kravlevska, K., Palma, D. (2023). Intent-driven autonomous network and service management in future cellular networks: A structured literature review. *Computer Networks*. Vol. 220, 109477. pp. 1 – 19. DOI: 10.1016/j.comnet.2022.109477

77. Zheng, X., Leivadeas, A., Falkner, M. (2022). Intent Based Networking management with conflict detection and policy resolution in an enterprise network. *Computer Networks*, 219, 109457. pp. 1 – 16. DOI: 10.1016/j.comnet.2022.109457.
78. Tanabe, K., (та ін.). (2023). Weaver Meets KANVAS: An Autonomous Closed-Loop Network Management System. *ACM (conference paper)*. pp. 28 – 36. DOI: 10.1145/3630590.3630594
79. Lysenko O.I., Valuiskyi S.V. Capacity Increasing of Sensor Telecommunication Networks. *Telecommunication Sciences*, 2012. Vol. 3, № 1. pp. 5–11. UDC 621.396.4
80. Wypiór, D., Klinkowski, M., Michalski, I. (2022). Open RAN–Radio Access Network Evolution, Benefits and Market Trends. *Applied Sciences*, Vol. 12(1), p. 408. DOI: 10.3390/app12010408
81. Attaran, M., Celik, B.G. (2023). Digital Twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal*. Vol. 6, 100165. pp. 1 – 10. DOI: 10.1016/j.dajour.2023.100165
82. Zhang, X., Li, L. (2024). Assessing the quality of experience in wireless networks for multimedia applications: A comprehensive analysis utilizing deep learning-based techniques. *Heliyon*. Vol. 10(9), e30351. pp. 1 – 11. DOI: 10.1016/j.heliyon.2024.e30351.
83. Efunogbon, A., Liu, E., Qiu, R., Efunogbon, T. (2025). Optimal 5G Network Sub-Slicing Orchestration in a Fully Virtualised Smart Company Using Machine Learning. *Future Internet*. Vol.17(2), 69. pp. 1 – 22. DOI: 10.3390/fi17020069
84. Mungari, F., Puligheddu, C., Garcia-Saavedra, A., Chiasserini, C.F. (2024). OREO: O-RAN intelligence Orchestration of xApp-based network services. *IEEE INFOCOM 2024 (accepted)*. pp. 1 – 10. DOI: 10.1109/INFOCOM52122.2024.10621166
85. Mungari, F., Puligheddu, C., Garcia-Saavedra, A., Chiasserini, C.F. (2024). OREO: O-RAN intelligence Orchestration of xApp-based network services. *arXiv preprint*. pp. 1 – 10. DOI: 10.48550/arXiv.2405.18198.

86. Rodrigo, M.S., Rivera D., Moreno J. I., Alvarez-Campana M. (2023). Digital Twins for 5G Networks: A Modeling and Deployment Methodology. IEEE Access (PDF). pp. 1 – 15. DOI: 10.1109/ACCESS.2023.3267548
87. ETSI (2020). ETSI GR ZSM 002: Zero-touch network and Service Management (ZSM); Reference Architecture. Standard. 80 p.
88. IETF (2024). Li, X., et al. A Language for Intent-based Networking. Internet-Draft.
89. ITU-T (2024). Focus Group on Autonomous Networks (FG-AN): Deliverables / reports (autonomous/intent-based management)
90. Lysenko O.I., Valuiskyi S.V. Capacity Increasing of Sensor Telecommunication Networks. Telecommunication Sciences, 2012, Vol. 3, № 1, pp. 5–11. ISSN 2219-9454
91. IEEE (2024). Kaur, M., et al. Survey: Large Language Models for Zero-Touch Network and Service Management
92. Gomes, P.H., Buhrgard M., Harmatos J., Mohalik S. K., Roeland D., Niemoller J. (2021). Intent-driven Closed Loops for Autonomous Networks. Journal of ICT Standardization. Vol. 9(2). pp. 257 – 290. DOI: 10.13052/jicts2245-800X.929
93. Zheng, Xiaoang, Leivadeas, Aris, Falkner, Matthias (2022). Intent Based Networking management with conflict detection and policy resolution in an enterprise network. Computer Networks. Vol. 219, 109457. Pp. 1 – 19. DOI: 10.1016/j.comnet.2022.109457
94. El Rajab, Mirna, Yang, Li, Shami, Abdallah (2024). Zero-touch networks: Towards next-generation network automation. Computer Networks, Vol. 243, 110294. DOI: 10.1016/j.comnet.2024.110294
95. K. Mehmood, K. Kravlevska, D. Palma (2023). Intent-driven autonomous network and service management in future cellular networks: A structured literature review. Computer Networks. Vol. 220, 109477. pp. 1 – 19. DOI: 10.1016/j.comnet.2022.109477
96. Mao V. Ngoa, Nguyen-Bao-Long Trana, Hyun-Min Yoob, Yong-Hao Puaa, Thanh-Long Lea, Xian-Loong Lianga, Binbin Chena, Een-Kee Hongb, Tony Q.S. Quek

(2024). RAN Intelligent Controller (RIC): From open-source implementation to real-world validation. *ICT Express*, Vol. 10(3). pp. 680–691. DOI: 10.1016/j.ict.2024.02.001

97. L. Cristobo, E. Ibarrola , I. Casado-O'Mara and L. Zabala (2024). Global Quality of Service (QoX). *Management for Wireless Networks. Electronics*, Vol. 13(16), 3113. pp. 1 – 30. DOI: 10.3390/electronics13163113.

98. X. Zhang, L. Li (2024). Assessing the quality of experience in wireless networks for multimedia applications: A comprehensive analysis utilizing deep learning-based techniques. *Heliyon*. Vol. 10(9), e30351. pp. 1 – 10. DOI: 10.1016/j.heliyon.2024.e30351

99. Hoßfeld, T., Pérez P. (2024). From theory to practice: System QoE assessment by user behaviour modelling. *ACM Conference Proceedings*. Vol. 16, № 3. pp. 1 – 5. DOI: 10.1145/3779008.3779011

100. Cristobo, L., Ibarrola, E., Casado-O'Mara, I., Zabala, L. (2024). Global Quality of Service (QoX) Management for Wireless Networks. *Electronics*. Vol. 13(16), 3113. pp. 1 – 30. DOI: 10.3390/electronics13163113

101. Omar, M.H.H., Kouraogo Justin P., Kabre Windmi Jonathan M., Tapsoba Stanislas David Wendkouni, Sie Oumarou Pr (2024). Using QoE Metric as a Decision Criterion in Multimedia Heterogeneous Network Optimization: Challenges and Research Perspectives. *Journal of Computer Networks and Communications*. Vol. 2024, Article ID 7864757. pp. 1 – 18. DOI: 10.1155/2024/7864757

102. T. Sadoun, S. Oubabas, S. Mokrani, R. Aoudjit, J. Lloret (2025). A Hybrid and Self-Adaptive QoS and QoE-Driven RAT selection strategy for IoT devices. *Ad Hoc Networks*. Vol 174 (2025), 103838. pp. 1 – 11. DOI: 10.1016/j.adhoc.2025.103838

103. Gilson Miranda Jr., Nina Slamnik-Kriještorac, Daniel Macedo, Johann Marquez-Barja (2025). A bottom-up approach for QoS-driven network slicing in TSN. *EURASIP Journal on Wireless Communications and Networking*. Vol. 2026:17. pp. 1 – 26. DOI: 10.1186/s13638-025-02474-9

104. Sadat, N., Rui Dai. (2025). In-network Caching and Flow Control for QoS/QoE Improvement in Future Networks – Survey. *MDPI Future Internet*. Vol. 5(2), 10. DOI: 10.3390/network5020010

105. Wypiór, D., Klinkowski, M., Michalski, I. (2022). Open RAN – Radio Access Network Evolution, Benefits and Market Trends. *Applied Sciences*. Vol. 12(1). pp. 1 – 18. DOI: 10.3390/app12010408
106. Efunogbon, A., Liu, E., Qiu, R., Efunogbon, T. (2025). Optimal 5G Network Sub-Slicing Orchestration in a Fully Virtualised Smart Company Using Machine Learning. *Future Internet*. Vol. 17(2), 69. pp. 1 – 22. DOI: 10.3390/fi17020069
107. Zhang, X., Li, L. (2024). Assessing the quality of experience in wireless networks for multimedia applications. *Heliyon*. Vol. 10(9), e30351. pp. 1 – 12. DOI: 10.1016/j.heliyon.2024.e30351
108. Perdigão, A., Quevedo, J., Aguiar R. L. (2025). Automating 5G Network Slice Management for Industrial IoT. *Computers Communications*. Vol. 229, 107991. pp. 1 – 19. DOI: 10.1016/j.comcom.2024.107991
109. El Rajab, M., Yang, L., Shami, A. (2024). Zero-Touch Networks: Towards Next-Generation Network Automation. *Computer Networks*. Vol. 243, 110294. pp 1 – 20. DOI: 10.1016/j.comnet.2024.110294
110. Liyanage, M., Pham, Q.V., Dev, K., et al. (2022). Zero-Touch Service Management in 5G/6G Networks: Survey. *Journal of Network and Computer Applications*. Vol. 203, 103362. pp. 1 – 27. DOI: 10.1016/j.jnca.2022.103362
111. Gallego-Madrid, J., Sanchez-Iborra, R., Ruiz, P.M., Skarmeta, A.F. (2022). Machine Learning-based Zero-Touch Network and Service Management: A survey. *Digital Communications and Networks*, Vol. 8(2). pp. 105 – 123. DOI: 10.1016/j.dcan.2021.09.001
112. Mehmood, K., Kralevska, K., Palma, D. (2023). Intent-Driven Autonomous Network and Service Management. *Computer Networks*. Vol. 220, 109477. pp. 1 – 13. DOI: 10.1016/j.comnet.2022.109477
113. Zheng, X., Leivadeas, A., Falkner, M. (2022). Intent Based Networking Management with Conflict Detection and Policy Resolution. *Computer Networks*. Vol. 219, 109457. pp. 1 – 23. DOI: 10.1016/j.comnet.2022.109457

114. Salhab, N., Langar, R., Rahim, R. (2021). 5G Network Slices Resource Orchestration using Machine Learning Techniques. *Computer Networks*. Vol. 188, 107829. pp. 1 – 24. DOI: 10.1016/j.comnet.2021.107829
115. Sulistyono, M. A., & Setiawan, D. (2025). Deep Reinforcement Learning-Based Algorithm for Dynamic Resource Allocation in Edge Computing. *ALCOM: Journal of Algorithm and Computing*. Vol. 1(1). pp. 13 - 22. DOI: 10.63846/fb7zns45
116. Ahmed, S., et al. (2024). AI-Enhanced 5G Networks: Machine Learning Approaches for Network Management. 4th Asian Conference on Innovation in Technology (ASIANCON). pp. 1–5. DOI: 10.1109/ASIANCON62057.2024.10838221.
117. Rodrigo, M.S., et al. (2023). Digital Twins for 5G Networks: Modeling and Deployment Methodology. *IEEE Access*. Vol. 11. pp. 1 – 15. DOI: 10.1109/ACCESS.2023.3267548
118. ITU-T (2024). Focus Group on Autonomous Networks (FG-AN) Report.
119. Gomes, P.H., et al. (2021). Intent-Driven Closed Loops for Autonomous Networks. *Journal of ICT Standardization*. Vol. 9(2). pp. 257 – 290. DOI: 10.13052/jicts2245-800X.929
120. ITU-T E.800 (2018). Terms and definitions related to quality of service and network performance including dependability. 64 p.
121. ITU-T Y.1541 (2011, kop.) Network performance objectives for IP-based services. 45 p.
122. ITU-T Y.3173 (2021). Framework for evaluating intelligence levels of future networks including ML/AI management automation. 30 p.
123. ITU-T FG-AN Deliverables (2023–2024). Focus Group on Autonomous Networks.
124. ISO/IEC 20000-1:2018. Information technology – Service management – Part 1: Service management system requirements. 30 p.
125. ISO/IEC 19770-1:2017. IT Asset Management – Overview and vocabulary. 25 p.
126. ISO/IEC 27004:2024. Information security management – Monitoring, measurement, analysis and evaluation. 50 p.



127. 3GPP TS 28.554 (2022). Network resource model for network slicing; Metrics for network slice serviceability. 80 p.
128. 3GPP TS 28.554 RED (2023). Network Slice management with operational metrics & service level monitoring. 95 p.
129. RFC 8325 (2020). Framework for Quality of Experience (QoE) Measurement. 35 p.
130. RFC 8326 (2020). QoE metrics for video streaming and real-time media. 40 p.
131. Халімов, Г.З. (2019). Аналіз мереж на основі моделей масового обслуговування. Навчально-методичний матеріал (НУРЕ). Електронне видання. 207 с.
132. Огірко О. І., Галайко Н. В. (2017). Теорія ймовірностей та математична статистика. Навчальний посібник. 292 с. УДК: 519.2(075.8)
133. Медведєв М.Г., Пащенко І.О. (2008). Теорія ймовірностей. Підручник. Електронний ресурс ВНТУ. УДК: 519.21
134. Jackson, J.R. (1957). Networks of Waiting Lines. Operations Research, Vol. 5(4) pp. 518–521. DOI: 10.1287/opre.5.4.518
135. Jackson, J.R. (1963). Jobshop-Like Queueing Systems. Management Science, Vol. 10(1). pp. 131–142. DOI: 10.1287/mnsc.10.1.131.
136. Chen, H., David D. Y. (2001). Jackson Networks. In: Fundamentals of Queueing Networks, pp. 15 – 35. DOI: 10.1007/978-1-4757-5301-1\_2
137. Weiss, G. (2021). Generalized Jackson Networks. Scheduling and Control of Queueing Networks (Cambridge). Chapter 9. pp. 143 – 156. DOI: 10.1017/9781108233217.014.
138. Whitt, W., You W. (2020). Heavy-traffic limits for stationary network flows. Queueing Systems. Vol. 95, Issue 1-2. pp. 53 – 68. DOI: 10.1007/s11134-019-09645-8.
139. Dai, J.G., Glynn, P., Xu, Y. (2025). Asymptotic product-form steady-state for generalized Jackson networks in multi-scale heavy traffic. arXiv. pp. 1 – 32. DOI: 10.48550/arXiv.2304.01499.

140. Guang, J., et al. (2025). Uniform Moment Bounds for Generalized Jackson Networks in Multiscale Heavy Traffic. *Mathematics of Operations Research*. Vol. 51, №1. Pp. 668 – 685. DOI: 10.1287/moor.2024.0408.

141. Lindley, D.V. (1952). The theory of queues with a single server. *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 48(2). pp. 277–289. DOI: 10.1017/S0305004100027638.

142. Boxma, O., Löpker, A., Mandjes, M., Palmowski, Z. (2021). A multiplicative version of the Lindley recursion. *Queueing Systems*. Vol. 98(3). pp. 225–245. DOI: 10.1007/s11134-021-09698-8

143. Karpelevich, F.I. (1994). Higher-order Lindley equations. *Stochastic Processes and their Applications*. Vol. 53, Issue 1. pp. 65 – 96. DOI: 10.1016/0304-4149(94)90058-2

144. Lui, J.C.S. (2008). G/G/1 Queueing Systems (конспект): Lindley's equation,

145. Schrage, L.E. (1966). The Queue M/G/1 with the Shortest Remaining Processing Time Discipline. *Operations Research*. Vol. 14 №4. pp. 670 – 684. DOI: 10.1287/opre.14.4.670

146. Wu, D.A., Takagi, H. (2006). M/G/1 queue with multiple working vacations. *Performance Evaluation*. Vol. 63, Issue 7. Pp. 654 – 681. DOI: 10.1016/j.peva.2005.05.005

147. Eckberg Jr., A.E. (1977). Sharp Bounds on Laplace–Stieltjes Transforms, with Applications to Various Queueing Problems. *Mathematics of Operations Research*. Vol. 2, №2. pp. 135 – 142. DOI: 10.1287/moor.2.2.135

148. Kella, O. (2020). Synchronized Lévy queues. *Journal of Applied Probability*. Vol. 57, Issue 4. pp. 1222 – 1233 DOI: <https://doi.org/10.1017/jpr.2020.75>

149. Dimitriou, I. (2025). On Markov-dependent reflected autoregressive processes and related models. *Queueing Systems*. Vol. 109, 20. pp. 1 – 20. DOI: 10.1007/s11134-025-09948-z

150. Thomasian, A. (2014). Analysis of Fork/Join and Related Queueing Systems. *ACM Computing Surveys*. Vol. 47, Issue 2. pp. 1 - 71. DOI: 10.1145/2628913

151. Baccelli, F. (1989). The fork-join queue and related systems with synchronization constraints: stochastic ordering and computable bounds. *Advances in Applied Probability*. Vol. 21, Issue 3. pp. 629 – 660. DOI: 10.2307/1427640

152. A. Rizk, F. Poloczek, F. Ciucu (2016) Stochastic bounds in Fork–Join queueing systems under full and partial mapping *Queueing Systems: Theory and Applications*. Springer. Vol. 83(3). pp. 261-291. DOI: 10.1007/s11134-016-9486-x

153. Kim, J.H., Tcha D.-W. (1997). Approximate analysis of finite fork/join queueing networks. *Computers & Industrial Engineering*. Vol. 32, Issue 4. pp. 799 -810.

154. Zeng, Y., et al. (2018). Throughput Scalability Analysis of Fork-Join Queueing Networks. *Operations Research*. Vol. 66, №6. pp. 1728-1743. DOI: 10.1287/opre.2018.1748

155. Zeng, Y., et al. (2018). Fork and Join Queueing Networks with Heavy Tails: Scaling Dimension and Throughput Limit. *SIGMETRICS '18: Abstracts of the 2018 ACM International Conference on Measurement and Modeling of Computer Systems*. pp. 122 – 124. DOI: 10.1145/3219617.3219668

156. Y. Narahari and P. Sundarrajan (1994). Performability Analysis of Fork-Join Queueing Systems. *The Journal of the Operational Research Society*. Vol. 46, № 10. pp. 1237 – 1249. DOI: 10.2307/2584619

157. M. Bolla (2014). SVD, discrepancy, and regular structure of contingency tables. *Discrete Applied Mathematics*. Vol. 176. pp. 3 – 11. DOI: 10.1016/j.dam.2014.03.022

158. T. Rogers, A.J. McKane (2011). A unified framework for Schelling's model of segregation. *Journal of Statistical Mechanics Theory and Experiment*. Vol. 7. pp. 1 – 22. DOI: 10.1088/1742-5468/2011/07/P07006

159. B. Berg, M. Harchol-Balter, B. Moseley, W. Wang, J. Whitehouse (2020). Optimal Resource Allocation for Elastic and Inelastic Jobs. *SPAA '20: Proceedings of the 32nd ACM Symposium on Parallelism in Algorithms and Architectures*. pp. 75 – 87. DOI:10.48550/arXiv.2005.09745

160. B. Kemper, M. Mandjes (2012). Mean sojourn times in two-queue fork-join systems: bounds and approximations. *OR Spectrum*. Vol. 34. pp 723–742 DOI 10.1007/s00291-010-0235-y
161. Palomo, S., Pender, J. (2020). Learning Lindley's Recursion. *Proceedings of the 2020 Winter Simulation Conference*. pp. 644–655. DOI: 10.1109/WSC48552.2020.9384121
162. Palomo, S., Pender, J. (2021). Learning the Tandem Network Lindley Recursion. *WSC 2021 proceedings*. pp. 1-12. DOI: 10.1109/WSC52266.2021.9715530
163. Grosz, I. (2023). Optimal Scheduling in Multiserver Queues. PhD Thesis (PDF). Copyright. CMU-CS-23-128
164. Пепа Ю.В., Поночовний П.М. (2024) Інтегральне оцінювання стійкості систем управління. *Наук.-практ. конф. «Перспективи та проблематика інтелектуальних систем: збірник тез науково-практичної конференції»*, Київ, 31 травня 2024 року. К.: ДУІКТ, 2024. С. 3 - 4. УДК 004.89:621.39(043.2)
165. Хоменко А.О., Пепа Ю.В., Шавловський Я.С. (2024) Ефективність застосування привілейованого доступу. *Всеукр. наук.-практ. конф. «Актуальні проблеми безпеки інформаційно-телекомунікаційних систем»*, Київ, 3 - 5 листопада 2024 р. К: РВЦ ДУІКТ, 2024. С.10-11. УДК 004.89:621.39(043.2)
166. В. Хорошко, О. Лаптев, Ю. Хохлачова, Абдуллах Аль-Далваш, Ю. Пепа. (2024) Особливості проектування захищених інформаційних мереж. *Наукоємні технології*. Том 62. №2. С. 154-163. DOI: <https://doi.org/10.18372/2310-5461.62.18709>
167. P. Holme, J. Saramäki (2019) *Temporal Network Theory*. Copyright. 375 p. DOI: 10.1007/978-3-030-23495-9
168. V. Akhramovych, Y. Pepa, A. Zahynei, T. Dzyuba, I. Danylov. (2024) Calculating the Information Security Indicator in Social Media with Consideration of the Path Duration Between Clients. *Informatyka, Automatyka, Pomiar w Gospodarce i Ochronie Srodowiska*. Vol. 14(1). pp. 71-77. DOI: 10.35784/iapgos.5720
169. Mohammed, M. M., Viktor, Z., Yurii, K., Lyubov, B., Konstantyn, P., & Turovsky, O. L. (2022). Methods for assessing the impact of bandwidth of control channels on the quality of telecommunications networks in the transmission of data

packets of different types. *International Journal of Communication Networks and Information Security*. Vol. 13, №2. pp. 220-229. DOI: 10.35784/iarpgos.5720

170. Биченков В.В., Заїка В.Ф., Заїка Л.А., Власов О.М. (2025) Удосконалення методики побудови моделей складних систем з використанням нового алгоритму селекції критеріїв. *Телекомунікаційні та інформаційні технології*. № 3 (88). С. 88 – 94. DOI: 10.31673/2412-4338.2025.038708

171. Заїка В.Ф., Ружинський В.Г., Брезіцький С.М., Аношков Г.В. (2025) Удосконалення алгоритмів розподілу навантаження в мережі спільноканальної сигналізації № 7. *Зв'язок* №1 (173). С. 36 – 42. DOI:10.31673/2412-9070.2025.017521

172. Беркман Л.Н., Каток В.Б., Скрипнік В.В., Зіненко Ю.М. (2025) Аналіз методів визначення затримки інформації в інфокомунікаційних мережах. *Наукові записки ДУІКТ*. №2(8). DOI: 10.31673/2786-8362.2025.022538

173. Berkman, L., Zakharzhevskiy, A., & Lavrinets, K. (2023). Improving the technology for processing the aggregated data flow of a secure corporate multiservice communication network. *Eastern-European Journal of Enterprise Technologies*, Vol. 4(9 (124)). pp. 14 – 23. DOI: 10.15587/1729-4061.2023.285414

174. Bertsekas, D.P. (2006) Optimal Routing and Flow Control Methods for Communication Networks. Analysis and Optimization of Systems. Proceedings of the Fifth International Conference on Analysis and Optimization of Systems Versailles, December 14–17, 1982. pp. 613 – 643. DOI:10.1007/BFb0044418

175. Uehara, K., Hirota, K. (1997) Fuzzy Inference Based Connection Admission Control in ATM Networks. *Journal of Advanced Computational Intelligence and Intelligent Informatics*. Vol. 1, No. 1. pp. 14 – 22. DOI: 10.20965/jaciii.1997.p0014

176. Uehara, K., Hirota, K. (1997) Fuzzy Connection Admission Control for ATM Networks Based on Possibility Distribution of Cell Loss Ratio. *IEEE Journal on Selected Areas in Communications*. Vol. 15, No. 2. pp. 179–190. DOI: 10.1109/49.552068

177. ISO/IEC 20000-1:2018. Information technology – Service management – Part 1. 34 p.

## **ДОДАТКИ**

## ДОДАТОК А

# Акти впровадження результатів дисертаційних досліджень



## А К Т

Комісія у складі:

голови:

директор ТОВ «ТЕСЕМ», Невгод Максим Юрійович;

членів комісії:

менеджер у будівництві, Михайленко Юлія Іванівна;

Інженер електрозв'язку та інженерних мереж, Шафранович

В'ячеслав Валентинович

цим Актом підтверджує, що результати дисертаційного дослідження аспіранта кафедри Телекомунікаційних систем та мереж Державного університету інформаційно-комунікаційних технологій **ГЕРАСИМЧУКА Владислава Сергійовича** на здобуття наукового ступеня доктора філософії за спеціальністю 172 - «Телекомунікації та радіотехніка» впроваджені в діяльність ТОВ «ТЕСЕМ».

Аспірант Герасимчук В.С. запропонував імітаційні моделі для оцінювання стану телекомунікаційної мережі при сеансовому способі надання телекомунікаційних послуг, що дозволяє оцінити поведінку самої системи за різних умов і оперативно здійснити коригування параметрів мережі для досягнення оптимального рівня надання послуг користувачам.

Також розроблено програмний комплекс оцінки станів телекомунікаційних мереж, який дозволяє провести оцінку ефективності роботи телекомунікаційної мережі під час її завантаження та провести прогнозування її пікового навантаження та можливих станів відмови.

ТОВ «ТЕСЕМ» використало програмне рішення для оцінки кількісних параметрів стану телекомунікаційної мережі, що поліпшує якість обслуговування для отримання оптимального рівня надання послуг користувачами.

Запропоновані імітаційні моделі дозволяють на стадії проектування телекомунікаційних мереж оцінити пікове навантаження та виконати адаптацію параметрів мережі під потреби кінцевого користувача, що підвищує якість надання телекомунікаційних послуг і збільшує пропускну здатність мережі приблизно на 27%.

Голова комісії

PhD

Максим НЕВГОД

Члени комісії

Юлія МИХАЙЛЕНКО

В'ячеслав ШАФРАНОВИЧ





ЗАТВЕРДЖУЮ

Директор

Інституту прикладних систем

управління Національної академії

наук України



Олег КОПІЙКА

03 2026 року

## АКТ

про реалізацію результатів дисертаційної роботи здобувача вищої освіти ступеня  
доктора філософії ГЕРАСИМЧУКА Владислава Сергійовича на тему:  
«Методи підвищення ефективності управління в телекомунікаційних мережах  
при наданні послуг кінцевому користувачеві»

Комісія у складі: голови – заступника директора з наукової роботи, доктора  
технічних наук Єгорова Віктора Богдановича, та членів – провідного наукового  
співробітника, доктора технічних наук, доцента  
Семка Віктора Володимировича і завідувача відділу наукової освіти та і портальних  
рішень, кандидата технічних наук, професора Дробика Олександра Васильовича,  
склала цей акт про те, що наукові результати, які отримані Герасимчуком Владиславом  
Сергійовичем у дисертаційному дослідженні, а саме:

- метод адаптивного управління мережевими ресурсами на основі замкненого контуру моніторингу та динамічної зміни пропускної здатності каналів;
- метод багатокритеріальної маршрутизації та пріоритезації критичних даних на основі комбінованої вагової функції та багатошляхового розподілу потоків;

використані при виконанні науково-дослідної роботи «Розроблення інформаційної платформи для інтеграції інформаційно-комунікаційних систем Збройних сил України».

Голова комісії  
д.т.н.

Віктор ЄГОРОВ

Члени комісії:  
д.т.н., доцент

Віктор СЕМКО

к.т.н., професор

Олександр ДРОБИК



ЗАТВЕРДЖУЮ

Перший проректор  
Державного університету інформаційно-  
комунікаційних технологій,  
член-кореспондент НАН України,  
доктор технічних наук, професор  
Олександр КОРЧЕНКО

“ 20 ”

2025 року

### АКТ

**впровадження результатів наукових досліджень  
Герасимчука Владислава Сергійовича  
у наукову та науково-технічну діяльність  
Державного університету інформаційно-комунікаційних технологій**

Комісія в складі:

голови комісії – завідувача кафедри Робототехніки та технічних систем  
Навчально-наукового інституту телекомунікацій, доктора технічних наук,  
професора Мельника Ю.В.;

членів комісії:

завідувача кафедри Телекомунікаційних систем та мереж Навчально-  
наукового інституту телекомунікацій, доктора технічних наук, професора  
Заїки В.Ф.;

професора кафедри Мобільних та відеоінформаційних технологій  
Навчально-наукового інституту телекомунікацій, доктора технічних наук,  
професора Беркман Л.Н.;

заступника директора Інституту проблем штучного інтелекту з наукової  
роботи та міжнародної діяльності Державного університету інформаційно-  
комунікаційних технологій, кандидата технічних наук, професора Дробика О.В.;

завідувача кафедри Мобільних та відеоінформаційних технологій  
Навчально-наукового інституту телекомунікацій, кандидата технічних наук,  
доцента Галаган Н.В.

в період з 18 по 20 листопада 2025 року розглянула надані матеріали та  
встановила, що нові наукові положення, які розроблені особисто аспірантом  
Герасимчуком В.С. в межах дисертаційного дослідження, а саме:

- математична модель функціонування телекомунікаційної мережі, яка  
ґрунтується на використанні математичного апарату мереж Джексона для аналізу  
характеристик багатокористувацьких систем, що дозволяє врахувати закони  
розподілу імовірностей завантаженості маршрутів;



- метод багатокритеріальної оптимізації побудови маршруту передачі інформації з пріоритезацією критичних даних, який ґрунтується на використанні комбінованої вагової функції каналів та механізму багатошляхового розподілу потоків;

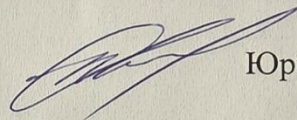
- метод оцінювання часових характеристик комбінованих телекомунікаційних послуг який, дозволяє враховувати складні сценарії взаємодії між компонентами мікросервісної архітектури

були використані під час проведення досліджень та підготовки матеріалів науково-дослідної роботи, що виконувалася у Державному університеті інформаційно-комунікаційних технологій за темою “Методи моніторингу продуктивності телекомунікаційних мереж на основі аналізу показників їх якості” (договір з ПНВП «Сардер Телеком» № 11/24 від 03.06.2024 р.),

а також використовуються в навчальному процесі Навчально-наукового інституту телекомунікацій Державного університету інформаційно-комунікаційних технологій при виконанні кваліфікаційних робіт і викладанні навчальних дисциплін «Телекомунікаційні системи та мережі наступного покоління», «Проектування телекомунікаційних систем та мереж» та «Теоретичні та прикладні основи проектування сучасних радіоелектронних та інфокомунікаційних систем».

Голова комісії:

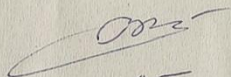
доктор технічних наук, професор



Юрій МЕЛЬНИК

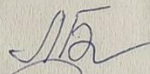
Члени комісії:

доктор технічних наук, професор



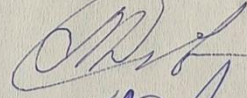
Віктор ЗАЙКА

доктор технічних наук, професор



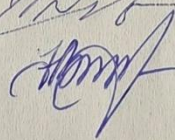
Любов БЕРКМАН

кандидат технічних наук, професор



Олександр ДРОБИК

кандидат технічних наук, доцент



Наталія ГАЛАГАН

## ДОДАТОК Б

Таблиця Б.1

Характеристики складної системи сучасних телекомунікаційних послуг

Номер виклику	Тип	Підхід 1			Підхід 2		
		Прийнято (A)	Відхилено (R)	Режим / Кількість	Прийнято (A)	Відхилено (R)	Режим / Кількість
1	EX-RA	A/BL		Простий	A		Простий
2	LX-RA	A			A		
3	MX-RA	A		6-A	A		
4	CX-RA	A		2-A/BL	A		
5	EX-RB	A/BL		1-R	A		9-A
6	LX-RB	A			A		
7	MX-RB	A			A		
8	CX-RB	A			A		
9	EX-RC		R		A		
10	LX-RC	A/BL		Нечіткий	A/BL		Нечіткий
11	MX-RC	A				R	2-A
12	CX-RC	A		4-A		R/BL	2-R
13	EX-RD		A/BL	1-A/BL	A		1-A/BL
14	LX-RD	A		1-R/BL	A		1-R/BL
15	MX-RD	A				R	
16	CX-RD	A		Повний		R	Повний
17	EX-RE		R			R	
18	LX-RE		R			R	
19	MX-RE		R	1-A		R	
20	CX-RE		R	8-R		R	9-R
21	EX-RF		R			R	
22	LX-RF		R			R	
23	MX-RF		R			R	
24	CX-RF		R			R	

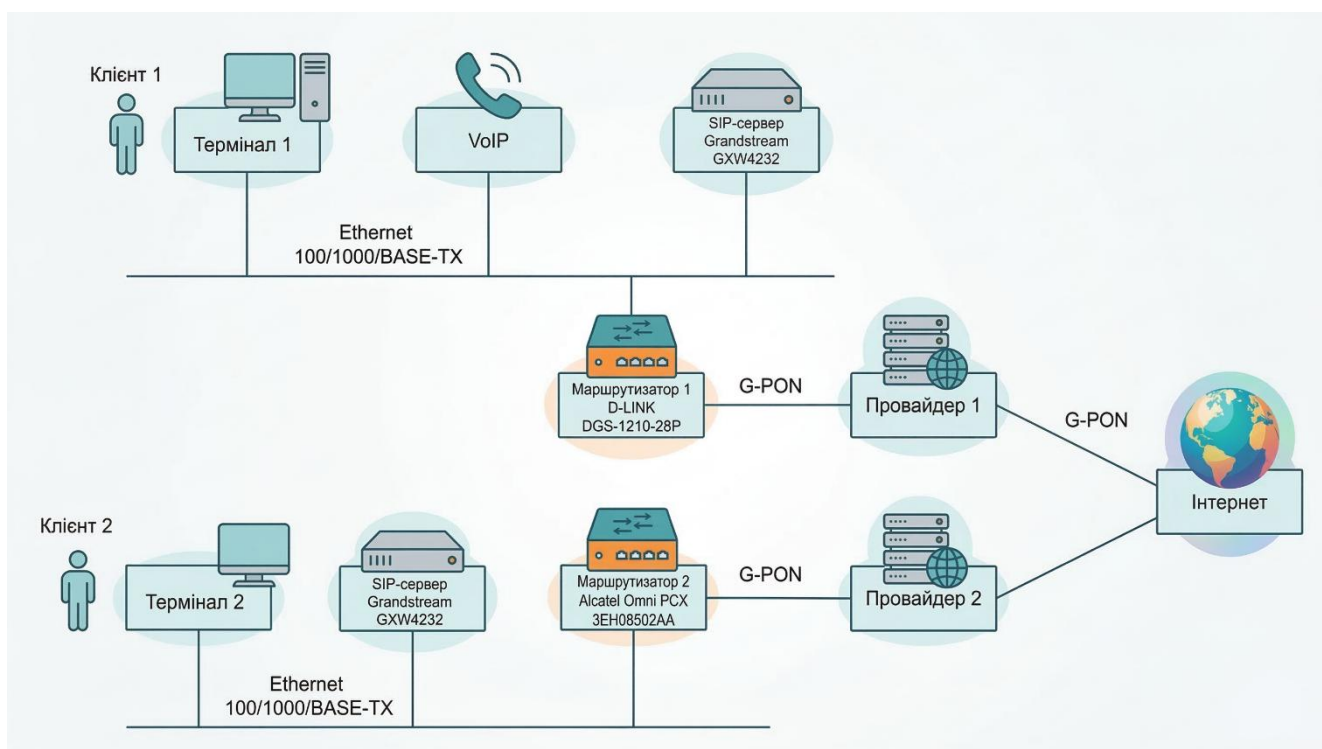
## Рішення щодо прийняття/відхилення викликів VoIP (підходи 1 та 2)

Номер виклику	Тип	Підхід 1			Підхід 2		
		Прийня- то (A)	Відхиле- но (R)	Режим / Кількість	Прийня- то (A)	Відхиле- но (R)	Режим / Кількість
1	EX-RA	A/BL		Простий	A		Простий
2	LX-RA	A			A		
3	MX-RA	A		7-A	A		8-A
4	CX-RA	A		1-A/BL	A		1-A/BL
5	EX-RB	A/BL		1-R/BL	A		
6	LX-RB	A			A		
7	MX-RB	A			A		
8	CX-RB	A			A		
9	EX-RC		R/BL		A		
10	LX-RC	A/BL		Нечіткий	A/BL		Нечіткий
11	MX-RC	A				R	1-A
12	CX-RC	A		4-A		R/BL	1-A/BL
13	EX-RD		R/BL	1-A/BL	A		2-R
14	LX-RD	A		1-R/BL	A		2-R/BL
15	MX-RD	A				R	
16	CX-RD	A		Повний		R	Повний
17	EX-RE		R			R	
18	LX-RE		R			R	
19	MX-RE		R	1-A		R	
20	CX-RE		R	8-R		R	9-R
21	EX-RF		R			R	
22	LX-RF		R			R	
23	MX-RF		R			R	
24	CX-RF		R			R	

Підсумкові рішення щодо прийняття/відхилення викликів  
(включає граничні випадки рішень)

Набір даних	Режим	Рішення	Підхід 1 (% викликів)	Підхід 2 (% викликів)
I	Простий	Прийнято	34	38
		Відхилено	4	–
	Нечіткий	Прийнято	18	12
		Відхилено	6	12
	Повний	Прийнято	4	–
		Відхилено	34	38
II	Простий	Прийнято	34	37
		Відхилено	5	–
	Нечіткий	Прийнято	17	8
		Відхилено	5	17
	Повний	Прийнято	5	–
		Відхилено	34	38

## ДОДАТОК В



Вихідні дані:

1. Network/Name: EVANGELION-LAB-NET
2. Hosts (PCs): LAB-HOST-1/HOST-2
3. CPU: Intel Core i7-12700H (14C/20T)
4. RAM: 32 GB DDR5
5. Storage: 1 TB NVMe SSD
6. OS: Windows 11 Pro x64
7. Datasets: I, II
8. Regimes: VoIP (Simple / Gray / Full)
9. Tested: 2026-01-07 08:08:49

Рис. В.1 – Топологія VoIP мережі для практичного тестування

## ДОДАТОК Г

# Алгоритми експериментальних тестувань мережі та практичні результати

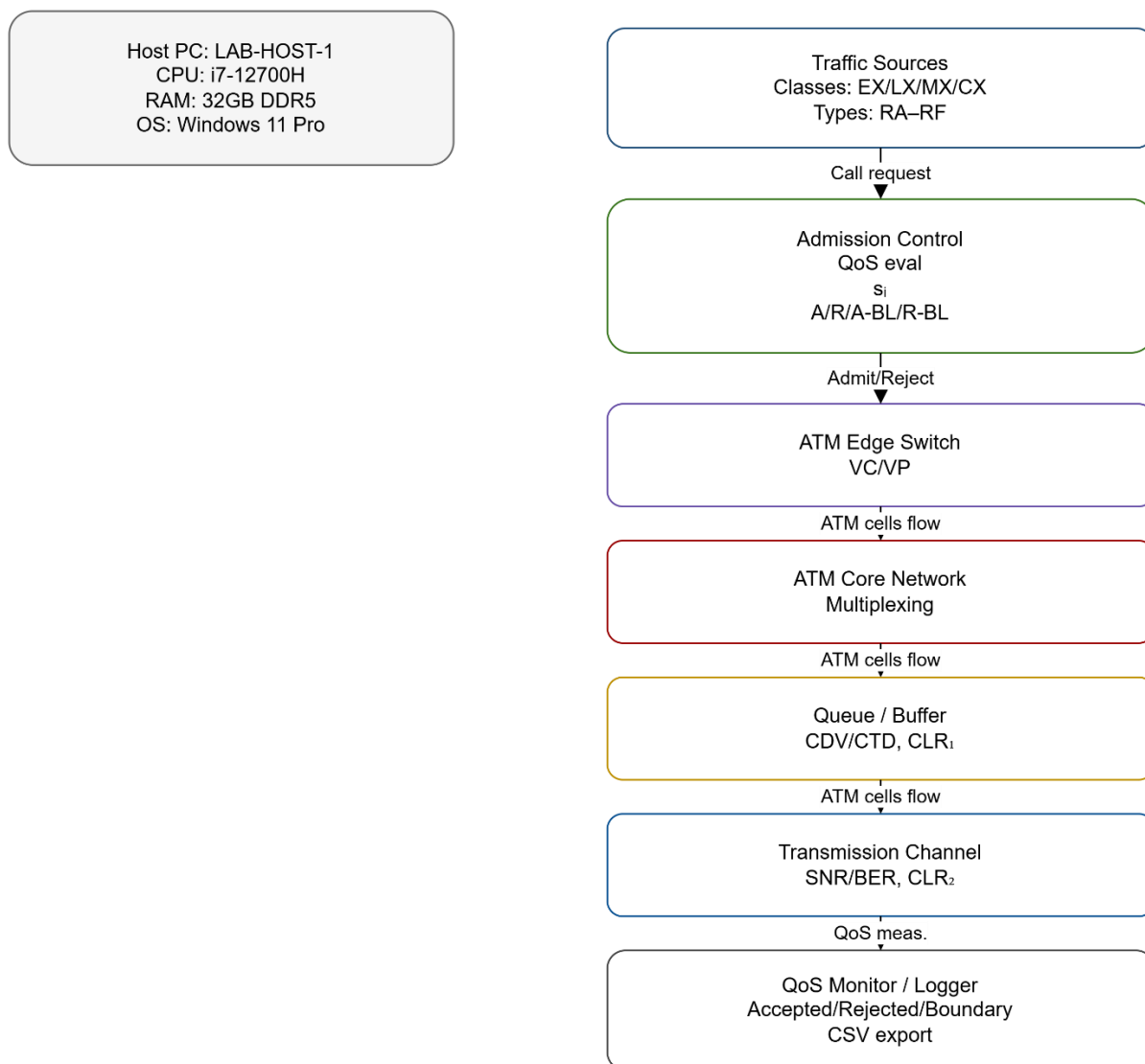


Рис. Г.1 – Алгоритм тестування мережі та відкриття шлюзу  
для надання послуги VoIP



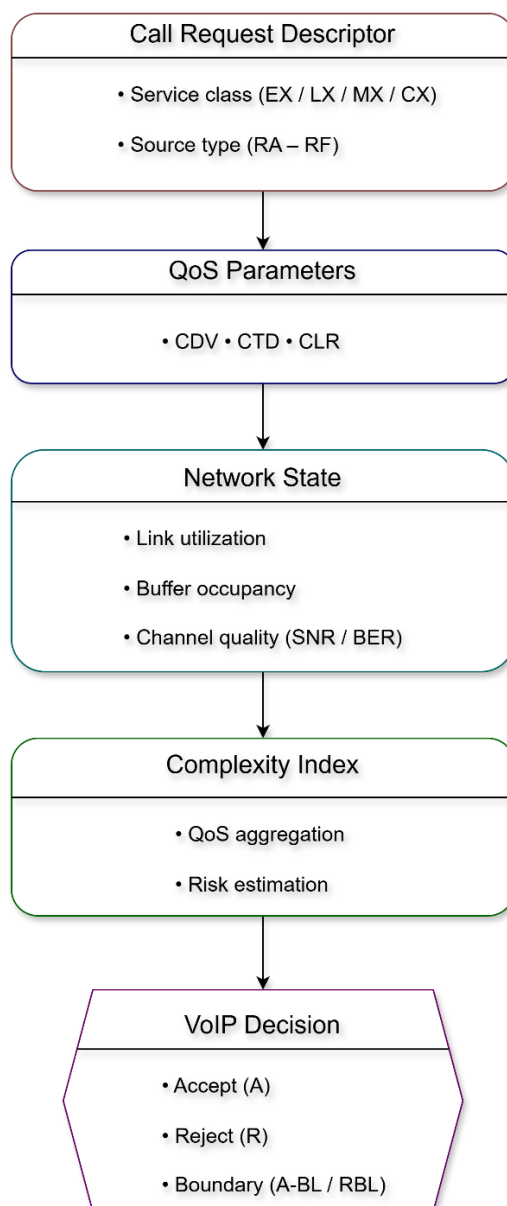


Рис. Г.2 – Алгоритм контролю допуску викликів VoIP  
на основі нечіткого оцінювання параметрів QoS



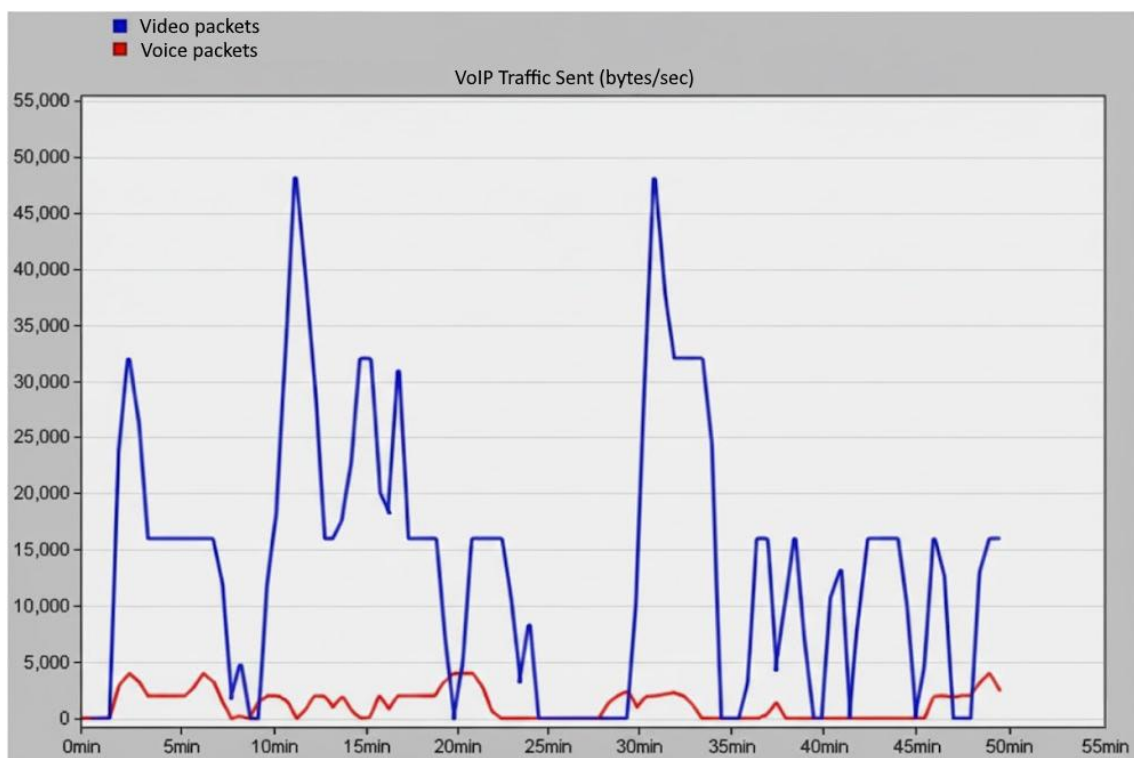


Рис. Г.3 – Переданий трафік VoIP

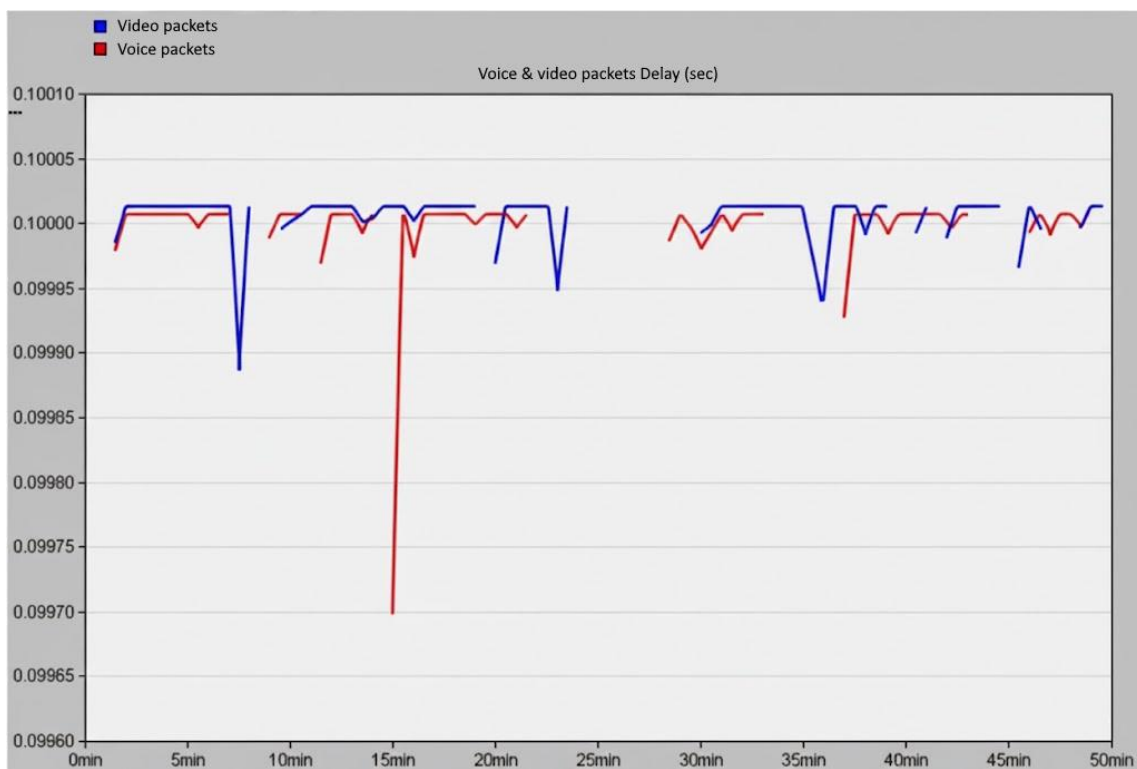


Рис. Г.4 – Затримка/втрата пакетів VoIP

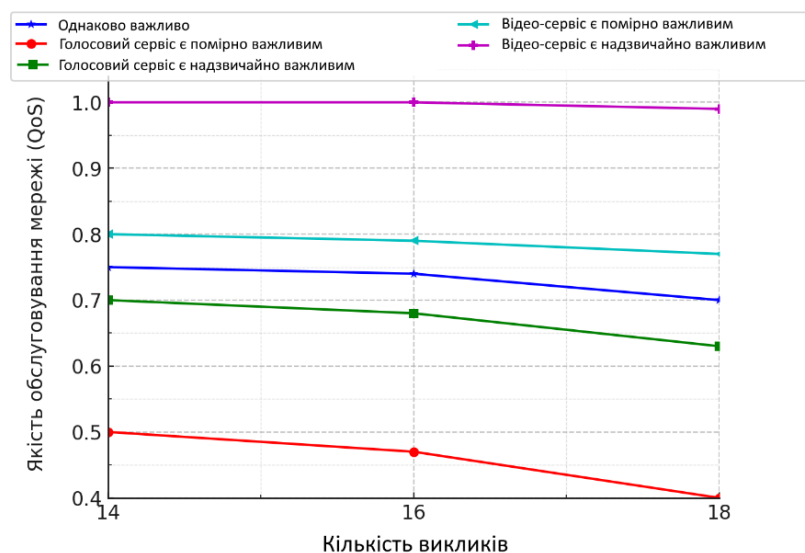


Рис. Г.5 – Показник QoS при різних пріоритетах під час дзвінків

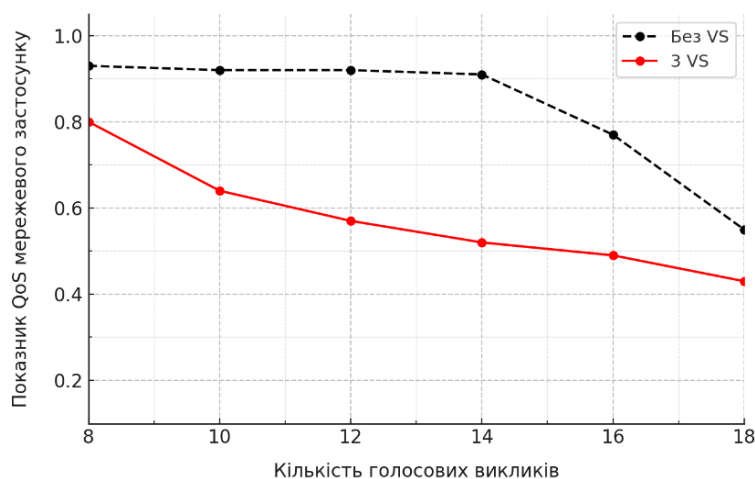


Рис. Г.6 – Показник QoS з/без передачі відео сигналу

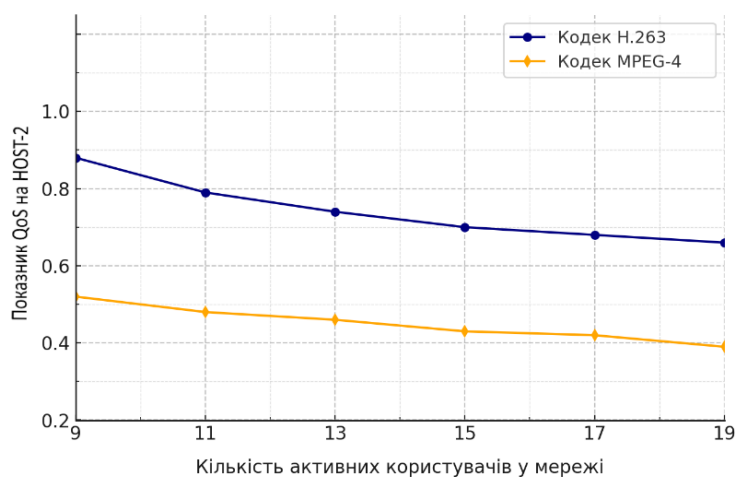


Рис. Г.7 – Показник QoS на приймальному HOST-2 при різних відеокodeках

## ДОДАТОК Д

### Код реалізації алгоритму на мові Python

```

from pathlib import Path
import pandas as pd
import numpy as np
from openpyxl import Workbook
from openpyxl.chart import LineChart, BarChart, AreaChart, Reference
from openpyxl.chart.label import DataLabelList
from zipfile import ZipFile, ZIP_DEFLATED
import textwrap, json, datetime

base = Path("/mnt/data/cac_variant3_pack")
(base/"csv").mkdir(parents=True, exist_ok=True)
(base/"excel").mkdir(parents=True, exist_ok=True)
(base/"scripts").mkdir(parents=True, exist_ok=True)
(base/"spec").mkdir(parents=True, exist_ok=True)
(base/"drawio").mkdir(parents=True, exist_ok=True)
# -----
# Base results
# -----
summary = pd.DataFrame([
    ["I", "Simple", 1, 34, 4],
    ["I", "Simple", 2, 38, 0],
    ["I", "Gray", 1, 18, 6],
    ["I", "Gray", 2, 12, 12],
    ["I", "Full", 1, 4, 34],
    ["I", "Full", 2, 0, 38],
    ["II", "Simple", 1, 34, 5],
    ["II", "Simple", 2, 37, 0],
    ["II", "Gray", 1, 17, 5],
    ["II", "Gray", 2, 8, 17],
    ["II", "Full", 1, 5, 34],
    ["II", "Full", 2, 0, 38],

```

```

],
columns=["dataset", "regime", "approach", "accepted_percent", "rejected_
percent"])
summary["boundary_percent"] = (100 - summary["accepted_percent"] -
summary["rejected_percent"]).clip(lower=0)
summary.to_csv(base/"csv"/"results_summary.csv", index=False)
# -----
# Tables for graphics
# -----
order = ["Simple", "Gray", "Full"]
wide_acc = summary.pivot_table(index="regime",
columns=["dataset", "approach"], values="accepted_percent",
aggfunc="sum").reindex(order).fillna(0)
wide_rej = summary.pivot_table(index="regime",
columns=["dataset", "approach"], values="rejected_percent",
aggfunc="sum").reindex(order).fillna(0)
wide_bnd = summary.pivot_table(index="regime",
columns=["dataset", "approach"], values="boundary_percent",
aggfunc="sum").reindex(order).fillna(0)

def save_wide(df, fname):
    out = df.copy()
    out.insert(0, "Regime", out.index)
    out.columns = ["Regime"] + [f"{d}-A{a}" for d,a in df.columns]
    out.to_csv(base/"csv"/fname, index=False)
    return out

acc_w = save_wide(wide_acc, "accepted_by_regime.csv")
rej_w = save_wide(wide_rej, "rejected_by_regime.csv")
bnd_w = save_wide(wide_bnd, "boundary_by_regime.csv")
# -----
# Descriptor host PC/network EVANGELION)
# -----
descriptor = pd.DataFrame([
    ["Network/Testbed name", "EVANGELION-LAB-NET"],
    ["Host (PC)", "LAB-HOST-1"],

```

```

["CPU", "Intel Core i7-12700H (14C/20T)"],
["RAM", "32 GB DDR5"],
["Storage", "1 TB NVMe SSD"],
["GPU (optional)", "NVIDIA RTX 4070 (32 GB)"],
["OS", "Windows 11 Pro x64"],
["Simulation mode", "Event-driven (call-level) + monitoring
DUICTs (practically)"],
["Datasets", "I, II"],
["CAC approaches", "Approach 1 (FIS rule-based), Approach 2
(conservative threshold)"],
["Regimes", "Simple, Gray, Full"],
["Generated", datetime.datetime.now().strftime("%Y-%m-%d
%H:%M:%S")]
], columns=["Field", "Value"])
ws_desc = add_sheet_from_df("Descriptor", descriptor)

```

## ДОДАТОК Е

## Log-файл тестування мережі EVANGELION в режимі VoIP

```
# fields: ts_ms,run_id,dataset,approach,regime,call_id,combo,service_class
0,II-A2-Full-run0,II,2,Full,00000,EX-RA,EX,RA,0.8637,0.6404,0.044002,0.9590,R/BL
20,II-A2-Full-run0,II,2,Full,00001,LX-RA,LX,RA,0.8736,0.7541,0.051987,0.8866,R
40,II-A2-Full-run0,II,2,Full,00002,MX-RA,MX,RA,0.8722,0.6767,0.045150,0.7349,R/BL
60,II-A2-Full-run0,II,2,Full,00003,CX-RA,CX,RA,0.8607,0.7121,0.046054,0.6581,R/BL
80,II-A2-Full-run0,II,2,Full,00004,EX-RB,EX,RB,0.9050,0.7788,0.048883,1.0444,R/BL
100,II-A2-Full-run0,II,2,Full,00005,LX-RB,LX,RB,0.9348,0.8487,0.057061,0.9567,R/BL
120,II-A2-Full-run0,II,2,Full,00006,MX-RB,MX,RB,0.8635,0.6746,0.044460,0.7299,R/BL
140,II-A2-Full-run0,II,2,Full,00007,CX-RB,CX,RB,0.8641,0.7611,0.046422,0.6733,R/BL
160,II-A2-Full-run0,II,2,Full,00008,EX-RC,EX,RC,0.8459,0.6068,0.044237,0.9331,R
180,II-A2-Full-run0,II,2,Full,00009,LX-RC,LX,RC,0.9554,0.7917,0.055298,0.9477,R/BL
200,II-A2-Full-run0,II,2,Full,00010,MX-RC,MX,RC,0.9401,0.8331,0.051314,0.8178,R/BL
220,II-A2-Full-run0,II,2,Full,00011,CX-RC,CX,RC,0.8726,0.6911,0.049230,0.6575,A/BL
240,II-A2-Full-run0,II,2,Full,00012,EX-RD,EX,RD,0.8858,0.8220,0.054669,1.0499,R/BL
260,II-A2-Full-run0,II,2,Full,00013,LX-RD,LX,RD,0.8805,0.6786,0.050424,0.8629,R/BL
280,II-A2-Full-run0,II,2,Full,00014,MX-RD,MX,RD,0.9009,0.7308,0.050108,0.7661,R/BL
300,II-A2-Full-run0,II,2,Full,00015,CX-RD,CX,RD,0.9145,0.7822,0.055584,0.7014,R
320,II-A2-Full-run0,II,2,Full,00016,EX-RE,EX,RE,0.9115,0.8020,0.049699,1.0584,R/BL
340,II-A2-Full-run0,II,2,Full,00017,LX-RE,LX,RE,0.9215,0.7409,0.052278,0.9094,R/BL
360,II-A2-Full-run0,II,2,Full,00018,MX-RE,MX,RE,0.8469,0.6744,0.047103,0.7217,R
380,II-A2-Full-run0,II,2,Full,00019,CX-RE,CX,RE,0.8925,0.8183,0.051237,0.7018,R
400,II-A2-Full-run0,II,2,Full,00020,EX-RF,EX,RF,0.8917,0.6507,0.044695,0.9818,R/BL
420,II-A2-Full-run0,II,2,Full,00021,LX-RF,LX,RF,0.9069,0.8369,0.053428,0.9363,R/BL
440,II-A2-Full-run0,II,2,Full,00022,MX-RF,MX,RF,0.8144,0.5869,0.050577,0.6780,R/BL
460,II-A2-Full-run0,II,2,Full,00023,CX-RF,CX,RF,0.8805,0.7605,0.047606,0.6804,R/BL
480,II-A2-Full-run0,II,2,Full,00024,EX-RA,EX,RA,0.9351,0.8061,0.046685,1.0757,R/BL
500,II-A2-Full-run0,II,2,Full,00025,LX-RA,LX,RA,0.9117,0.7757,0.051816,0.9166,R/BL
520,II-A2-Full-run0,II,2,Full,00026,MX-RA,MX,RA,0.9042,0.7761,0.050646,0.7821,R/BL
540,II-A2-Full-run0,II,2,Full,00027,CX-RA,CX,RA,0.8611,0.7205,0.051065,0.6606,R/BL
560,II-A2-Full-run0,II,2,Full,00028,EX-RB,EX,RB,0.8217,0.6985,0.041374,0.9557,R/BL
580,II-A2-Full-run0,II,2,Full,00029,LX-RB,LX,RB,0.9017,0.7844,0.057522,0.9140,R/BL
600,II-A2-Full-run0,II,2,Full,00030,MX-RB,MX,RB,0.9047,0.7892,0.051332,0.7864,R/BL
620,II-A2-Full-run0,II,2,Full,00031,CX-RB,CX,RB,0.9069,0.8369,0.053428,0.7134,R/BL
640,II-A2-Full-run0,II,2,Full,00032,EX-RC,EX,RC,0.8595,0.7157,0.052277,0.9879,R/BL
660,II-A2-Full-run0,II,2,Full,00033,LX-RC,LX,RC,0.9240,0.8924,0.055144,0.9666,R/BL
680,II-A2-Full-run0,II,2,Full,00034,MX-RC,MX,RC,0.8684,0.6989,0.041919,0.7400,R/BL
700,II-A2-Full-run0,II,2,Full,00035,CX-RC,CX,RC,0.9895,0.8520,0.056108,0.7539,R
720,II-A2-Full-run0,II,2,Full,00036,EX-RD,EX,RD,0.9398,0.7833,0.048324,1.0693,R/BL
740,II-A2-Full-run0,II,2,Full,00037,LX-RD,LX,RD,0.9768,0.9182,0.054881,1.0065,R/BL
760,II-A2-Full-run0,II,2,Full,00038,MX-RD,MX,RD,0.8712,0.6457,0.048491,0.7246,R
780,II-A2-Full-run0,II,2,Full,00039,CX-RD,CX,RD,0.9215,0.8325,0.053678,0.7186,R/BL
800,II-A2-Full-run0,II,2,Full,00040,EX-RE,EX,RE,0.9380,0.8329,0.055360,1.0889,R
820,II-A2-Full-run0,II,2,Full,00041,LX-RE,LX,RE,0.8576,0.6769,0.043567,0.8490,R
840,II-A2-Full-run0,II,2,Full,00042,MX-RE,MX,RE,0.9293,0.7272,0.046584,0.7791,R/BL
860,II-A2-Full-run0,II,2,Full,00043,CX-RE,CX,RE,0.8681,0.7132,0.043871,0.6617,R/BL
880,II-A2-Full-run0,II,2,Full,00044,EX-RF,EX,RF,0.9431,0.9119,0.063940,1.1254,R
900,II-A2-Full-run0,II,2,Full,00045,LX-RF,LX,RF,0.8452,0.7481,0.050982,0.8680,R
920,II-A2-Full-run0,II,2,Full,00046,MX-RF,MX,RF,0.8716,0.7187,0.051537,0.7478,R/BL
940,II-A2-Full-run0,II,2,Full,00047,CX-RF,CX,RF,0.9631,0.9099,0.056766,0.7585,R
960,II-A2-Full-run0,II,2,Full,00048,EX-RA,EX,RA,0.8180,0.6451,0.045949,0.9308,R/BL
980,II-A2-Full-run0,II,2,Full,00049,LX-RA,LX,RA,0.8963,0.7440,0.058597,0.8960,R
1000,II-A2-Full-run0,II,2,Full,00050,MX-RA,MX,RA,0.8963,0.7440,0.058597,0.7680,R/BL
```