

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ**

**«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**Факультет менеджменту та маркетингу  
Кафедра математичного моделювання економічних систем**

«На правах рукопису»  
УДК 330.43

До захисту допущено:  
Завідувач кафедри  
\_\_\_\_\_ Володимир. КАПУСТЯН  
«\_\_» \_\_\_\_\_ 2020 р.

**Магістерська дисертація**  
**на здобуття ступеня магістра**  
**за освітньо-професійною програмою «Економічна кібернетика»**  
**зі спеціальності 051 «Економіка»**  
**на тему: « Економіко-математичне моделювання витрат компанії на**  
**інформаційну систему »**

Виконав:  
студент VI курсу, групи УК-91мп  
Шперлов Руслан Віталійович \_\_\_\_\_

Керівник:  
доц., к. ф.-м.н., доц.  
Жуковська О.А. \_\_\_\_\_

Рецензент:  
доц., к.е.н., доц.  
Черненко Н. О. \_\_\_\_\_

Засвідчую, що у цій магістерській  
дисертації немає запозичень з праць  
інших авторів без відповідних  
посилань.  
Студент \_\_\_\_\_

Київ – 2020 року

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**  
**Факультет менеджменту та маркетингу**  
**Кафедра математичного моделювання економічних систем**

Рівень вищої освіти – другий (магістерський)

Спеціальність – 051 «Економіка»

Освітньо-професійна програма «Економічна кібернетика»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Володимир КАПУСТЯН

«\_\_» \_\_\_\_\_ 2020 р.

**ЗАВДАННЯ**  
**на магістерську дисертацію студенту**  
**Шперлову Руслану Віталійовичу**

1. Тема дисертації «Економіко-математичне моделювання витрат компанії на інформаційну систему», науковий керівник дисертації Жуковська Ольга Анатоліївна, к.ф.-м.н., доц., затверджені наказом по університету від «02» листопада 2020 р. №3179-с
2. Термін подання студентом дисертації: 15.12.2020
3. Об'єкт дослідження – інформаційна система високого навантаження провайдера інтернет-реклами
4. Вихідні дані – наукова та навчально методична література, статистичні дані навантаження компанії за період карантинних обмежень та ціни оренди обчислювальних пристроїв
5. Перелік завдань, які потрібно розробити: 1) аналіз роботи інформаційної системи; 2) опрацювання наукових робіт з теми дослідження; 3) визначення задачі економічного змісту, яка має бути вирішена; 4) обґрунтування основних показників якості функціонування інформаційної системи; 5)

Визначення основних прогностичних моделей; 6) Створення динамічної імітаційної моделі інформаційної системи; 7) Формування стратегій масштабування інформаційної систем та оцінка заощаджень

6. Орієнтовний перелік ілюстративного матеріалу – презентація роботи

7. Перелік публікацій:

- 1) Жуковська О. А., Шперлов Р. В. Економіко-математичне моделювання приросту навантаження high-load систем / О.А. Жуковська, Р.В. Шперлов // Економічний вісник НТУУ "КПІ". – 2020. – №14;
- 2) Жуковська О. А., Шперлов Р. В. Прогнозування навантаження high-load систем / О.А. Жуковська, Р.В. Шперлов // Економічний вісник НТУУ "КПІ". – 2021. – №16;

8. Дата видачі завдання 02.11.2020

#### Календарний план

| № з/п | Назва етапів виконання магістерської дисертації                    | Термін виконання етапів магістерської дисертації | Примітка |
|-------|--------------------------------------------------------------------|--------------------------------------------------|----------|
| 1.    | Вступ. Огляд літератури і попередніх досліджень з обраної тематики | 02.11.2020 – 04.11.2020                          |          |
| 2     | Аналіз роботи інформаційної системи                                | 05.11.2020 – 09.11.2020                          |          |
| 3     | Вибір і постановка актуальної економічної задачі дослідження.      | 10.11.2020 – 14.11.2020                          |          |
| 4     | .                                                                  | 15.11.2020 – 19.11.2020                          |          |
| 5     | Розрахунок основних компонентів моделі.                            | 20.11.2020 – 25.11.2020                          |          |
| 6     | Визначення алгоритму розв'язання.                                  | 26.11.2020 – 30.11.2020                          |          |
| 7     | Симуляція моделі, програмна реалізація                             | 01.12.2020 – 04.12.2020                          |          |
| 8     | Оформлення дисертації та підготовка до захисту.                    | 05.12.2020 – 09.12.2020                          |          |
| 9     | Створення презентації.                                             | 10.12.2020 – 15.12.2020                          |          |

Студент

\_\_\_\_\_  
(підпис)

Руслан ШПЕРЛОВ

(ініціали, прізвище)

Науковий керівник

\_\_\_\_\_  
(підпис)

Ольга ЖУКОВСЬКА

(ініціали, прізвище)

## РЕФЕРАТ

У даній дисертації було розглянуто аналітико-прогностичне моделювання основного показника вимірювання навантаження інформаційної інфраструктури – кількість запитів до системи в секунду та побудовано імітаційну модель інформаційної системи – багатоканальну систему масового обслуговування. Досліджено статистичні дані кількості запитів з метою знаходження та виявлення тенденції руху його часового ряду та перевірка роботи імітаційної моделі для забезпечення надійності та доступності сервісів ІТ-компанії. Актуальність даного дослідження отримує певний резонанс у індустрії інформаційних технологій на даний час, адже розширення інформаційних систем до можливості передачі петабайтів даних в секунду потребує чіткого плану їх інфраструктуризації. Забезпечення мінімального прогнозу, оцінки даних та моделювання роботи інформаційної системи при прогнозованих навантаженнях може гарантувати не тільки стабільність програмного продукту чи технології але й неопосередковано вплинути на собівартість продукції та об'єм отриманого прибутку.

У роботі було проаналізовано особливості функціонування інформаційної системи компанії та об'єми витрат, що формуються на основі підтримки та моніторингу життєдіяльності системи. Обґрунтовано вибір прогностичних моделей для прогнозу навантаження. Відповідно теорії масового обслуговування було створено імітаційну модель інформаційної системи, що відображала реальний процес обробки запитів із заданими параметрами. Прораховано можливі варіанти масштабування інформаційної системи та об'єми заощаджень при дотриманні рекомендацій, оснований на даному дослідженні.

Загальний обсяг роботи – 109 сторінок, кількість ілюстрацій – 54, кількість таблиць – 8, кількість додатків – 5 , кількість бібліографічних найменувань за переліком посилань – 40.

Ключові слова: Інформаційна система, відмовостійкість, масштабованість, QPS, IT-інфраструктура, прогнозування, імітаційне моделювання, системи масового обслуговування.

## **ABSTRACT**

In current dissertation, an analytical and forecasting model of the main indicator of the change in the information infrastructure - the number of requests to the system per second was revealed. The imitational model of the information system is prompted - a multi-channel system of mass service. Previously, statistical data on the number of requests due to the fact that the tendencies to the decline of the data were shown, and the revision of the imitational model for securing the reliability and availability of IT-company services. The relevance of this prerequisite will negate the resonance of the information technology industry today, as well as the expansion of information systems to the possibility of transferring petabytes of data per second, requiring a clear plan of the infrastructure. Preservation of the minimum forecast, assessment of data and models of the system imitation, when forecasting the options, can guarantee the stability of the software product, technology and delivery.

The work has analyzed the peculiarities of the functionality of the company's information systems and costs, which are formed on the basis of adapting to the monitoring of the health of the system. Choice of predictive models for forecasting of load. Based on the theory of mass servicing, a model of the information system was introduced into an imitational model, which represented the real process of processing

power supplies from the given parameters. Possible options for scaling the information system and the amount of savings in compliance with the recommendations based on this study are calculated.

Total volume of work – 109 pages, number of illustrations – 54, number of tables – 8, number of appendices – 5, number of bibliographic names according to the list of references – 40.

Keywords: Information system, fault tolerance, scalability, QPS, IT-infrastructure, forecasting, simulation modeling, queuing systems.

## ЗМІСТ

|                                                                               |     |
|-------------------------------------------------------------------------------|-----|
| ВСТУП.....                                                                    | 8   |
| 1 КОМПЛЕКСНА ХАРАКТЕРИСТИКА ІНФОРМАЦІЙНИХ СИСТЕМ В ЕКОНОМІЦІ.....             | 11  |
| 1.1 Загальний опис інформаційної системи .....                                | 11  |
| 1.2 Формування витрат на впровадження та розвиток інформаційних систем .....  | 17  |
| 1.3 Тенденції розвитку ІС як хмарних технологій та їх вплив на економіку..... | 25  |
| 1.4 Система масового обслуговування – аналог процесу роботи ІС .....          | 32  |
| 2 МОДЕЛЮВАННЯ ВИТРАТ КОМПАНІЇ НА ІНФОРМАЦІЙНУ СИСТЕМУ.....                    | 40  |
| 2.1 Постановка економічної задачі .....                                       | 40  |
| 2.2 Загальний опис моделі.....                                                | 43  |
| 2.2.1 Прогностична модель часового ряду.....                                  | 43  |
| 2.2.2 Імітаційна модель інформаційної системи.....                            | 58  |
| 3 МОЖЛИВОСТІ ТА НАПРЯМИ ПРАКТИНОГО ЗАСТОСУВАННЯ РОЗРОБЛЕНОЇ<br>МОДЕЛІ .....   | 64  |
| 3.1 Реалізація моделі.....                                                    | 64  |
| 3.1.1 Побудова прогнозів .....                                                | 64  |
| 3.1.2 Реалізація імітаційної моделі ІС .....                                  | 82  |
| 3.2 Формування пропозицій щодо економії коштів компанії .....                 | 94  |
| ВИСНОВКИ.....                                                                 | 98  |
| ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....                                                 | 100 |
| ДОДАТОК А Прогнози кандидатних моделей.....                                   | 105 |
| ДОДАТОК В Обрані результати моделей .....                                     | 106 |
| ДОДАТОК С Модуль заявок і черги .....                                         | 107 |
| ДОДАТОК D Модуль обробки .....                                                | 108 |
| ДОДАТОК E Загальна імітаційна модель.....                                     | 109 |

## ВСТУП

*Актуальність теми.* Продуктивність інформаційних систем являє собою найбільш значимий параметр якості роботи ІТ-відділу компанії. Більшість найкрупніших ІТ-гігантів констатують нелінійне збільшення об'ємів оброблених запитів порівняно із ростом клієнтської бази. Домінуючу роль у формуванні такого швидкого приросту відіграє стрімка діджиталізація будь-яких комерційних або уже і державних процесів. Прискоренню цього процесу також посприяла світова коронокриза, що змушує підприємців розповсюджувати свої товари через інтернет-магазини. Пандемія дала різкий поштовх до освоєння нових методів продажу різним сферам господарства. Крива приросту навантажень на інформаційні системи все швидше набуває форми показникової функції, що ускладнює побудову високонавантажених інформаційних систем.

Звідси можна зробити логічний висновок, що для уникнення ситуацій, коли системи не зможуть справлятися із потоком даних, потрібно передбачити майбутнє навантаження, щоб наростити достатню кількість обчислювальних ресурсів і бути готовим для подальшого масштабування, що значно скоротить витрати на обслуговування та модернізацію апаратних компонентів інформаційних систем (далі ІС).

Таким чином, для вирішення задачі правильного управління обчислювальними ресурсами можуть сприяти додаткові відомості, які можна отримати з даних прогнозу про кількість запитів до системи на основі прогнозної моделі та симуляція роботи ІС, що відкине варіанти із можливими падіннями основної системи.

*Мета і завдання дослідження.* Метою даної роботи є прогнозування майбутнього навантаження на ІС, імітація її роботи для перевірки працездатності та складання об'ємів операційних витрат з урахуванням можливої економії.

Для досягнення поставленої мети сформульовано наступні завдання, що поділені на 3 етапи:

I Етап (Прогнозування):

- Проаналізувати особливості роботи ІС та визначити основні показники якості її роботи;
- Зібрати статистичні дані за період коронокризи;
- Підібрати прогностні методи та моделі для яких зібрана статистика буде відповідати їх умовам;
- Сформувати прогнози на виділений період та оцінка адекватності прогностичних моделей.

II Етап (Імітація):

- Побудувати імітаційну модель, що відтворює основні аспекти роботи реальної ІС;
- Ініціалізація вхідних даних, отриманих на минулому етапі дослідження;
- Симуляція процесу роботи ІС та оцінка її основних показників.

III Етап (Аналіз та стратегія):

- Проаналізувати роботу імітаційної моделі ;
- Побудувати інфографіку реальних витрат ;
- Сформувати стратегії масштабування ІС;
- Оцінити заощаджень.

*Об'єкт дослідження* - інформаційна система високого навантаження провайдера інтернет-реклами.

*Предмет дослідження* - методи прогнозування та імітаційні моделі функціонування інформаційних систем високого навантаження.

*Інформаційною базою* для дослідження слугує статистика роботи ІС ТОВ «Адміксер Технолоджис» за період початку карантинних обмежень в Україні, а також для роботи було застосовано дані із всесвітнього банку даних, звітів High-Load конференцій та робіт провідних спеціалістів в області високого навантаження.

*Методи дослідження.* Теоретичною та методологічною основою дослідження послужили прогностичні моделі Хольта, Хольта-Вінтерса та ARIMA-модель для першого етапу роботи, а також теорія масового обслуговування на якій базувався другий етап роботи - імітація.

За результатами дослідження було опубліковано 2 статті у науковому виданні «Економічний вісник НТУУ КПІ»:

- 1) Економіко-математичне моделювання приросту навантаження high-load систем(видання 2020 року);
- 2) Прогнозування навантаження high-load систем (видання 2021 року)

# 1 КОМПЛЕКСНА ХАРАКТЕРИСТИКА ІНФОРМАЦІЙНИХ СИСТЕМ В ЕКОНОМІЦІ

## 1.1 Загальний опис інформаційної системи

*Визначення інформаційної системи.* Розвиток технологій під час переходу до нового тисячоліття відкрив можливості для збору та обробки інформаційних потоків більш ефективно з меншими ресурсними затратами. Інформаційна система(ІС) - це інтегрований набір компонентів для збору, зберігання та обробки даних, а також для надання інформації, скомпільованих або цифрових продуктів. Бізнес-фірми та інші організації використовують ІС для аналізу та адаптації власних бізнес-процесів, взаємодії з клієнтами, постачальниками та контактними аудиторіями ринку. Наприклад, корпорації використовують інформаційні системи для обробки фінансових рахунків, управління своїми людськими ресурсами та охоплення потенційних клієнтів за допомогою онлайн-акцій. Сюди входять:

- eBay, аукціонний ринок;
- Amazon, електронний центр, постачальник послуг хмарних обчислень;
- Alibaba, електронний ринок для бізнесу;
- Google, компанія-пошуковик, яка отримує більшу частину доходу від реклами ключових слів при пошуку в Інтернеті.
- Microsoft Azure, хмарний провайдер з сервісами надання віддаленої хмарної платформи

Також ІС використовується і в державному секторі та є критичним не лише з точки зору ефективності. Уряди застосовують інформаційні системи для

економічного надання послуг громадянам. Цифрові товари - такі як електронні книги, відеопродукція та програмне забезпечення - та онлайн-послуги, такі як ігри та соціальні мережі, постачаються з інформаційними системами. Люди проводять значну частину свого особистого життя: для спілкування, навчання, покупок, банківської справи та розваг.

*Компоненти інформаційної системи.* Інформаційні системи складаються з різних компонентів, які працюють разом, щоб надати цінність для організації. Загальновизнаним є розподіл ІС на 5 компонентних складових (рис. 1.1)[1]:



Рисунок 1.1 – Діаграма компонентів ІС

1) Апаратне обладнання;

Так визначають фізичні компоненти технології. Сервери, персональні комп'ютери, мережеве обладнання та технічна периферія - це всі приклади апаратного забезпечення інформаційних систем.

2) Програмне забезпечення;

Програмне забезпечення - це набір інструкцій, які віддають сигнал апаратній частині щодо подальшої роботи. Програмне забезпечення є нематеріальною складовою ІС. Є декілька категорій програмного забезпечення,

основними є програмне забезпечення операційної системи, що робить апаратне забезпечення придатним для використання, та прикладне програмне забезпечення.

### 3) Дані;

Дані не можуть існувати самі по собі. Для інтеграції даних в інформаційну систему, необхідно надати їм контекстну складову. Після того, як дані вкладені в контекст, їх можна ефективно згрупувати і проаналізувати. Результатом цього повинні стати “інсайди”, які можна використовувати для прийняття управлінських рішень.

### 4) Людські ресурси(HR);

Починаючи від працівників довідкової служби до системних аналітиків, програмістів, аж до головного інформаційного директора, люди, які займаються інформаційними системами, є критичним елементом будь-якої ІС.

### 5) Процес.

Інформаційні системи все більше інтегруються в організаційні процеси, забезпечуючи більшу продуктивність та кращий контроль над цими процесами. Але просто автоматизувати діяльність за допомогою технологій недостатньо - підприємства, які прагнуть ефективно використовувати інформаційні системи, роблять більше. Кінцевою метою є використання технологій для управління та вдосконалення процесів як у межах компанії, так і в зовнішньому середовищі при роботі з клієнтами, конкурентами або постачальниками.

Перші три складові відносяться до технологічної частини ІС. Технологічну частину ІС визначають як застосування наукових знань у практичних цілях та розглядають як фундаментальну частину будь-якої інформаційної системи.

*Роль інформаційної системи.* Одна з ключових ролей інформаційних систем полягає у тому, щоб брати дані та перетворювати їх на інформацію, а потім трансформувати їх в організаційні знання. У міру розвитку технологій ця роль перетворилася на основу організації.

*Основні типи інформаційних систем.* Інформаційні системи бувають різних типів.

1) Корпоративні системи;

Основною функцією таких інформаційних систем призначені є забезпечення спільного доступу до інформації співробітниками. Використовуючи такі інструменти як дошки обговорень, обмін документами та відео, ці системи дозволяють фіксувати процеси мислення, що лежать в основі організаційних ідей. Прикладами спільних систем є: Google Drive, Microsoft SharePoint, Cisco WebEx, Atlassian Confluence, IBM Lotus Notes/Domino, OneDrive.

2) CRM-системи;

Визначення або аббревіатура CRM розшифровується як Customer Relationship Management, тобто «управління відносинами з клієнтами». CRM-система за допомогою автоматизації процесів допомагає ефективніше вибудовувати діалог з покупцем, не допускати помилок в роботі і в підсумку продавати йому більше.

3) Системи підтримки прийняття рішення(DSS).

Система підтримки прийняття рішень - це ІС, побудована для допомоги організації в прийнятті конкретного рішення або розробці конкретного процесу. DSS можуть існувати на різних рівнях прийняття рішень в організації. Ці системи призначені для отримання даних щодо відомого (або частково відомого) процесу прийняття рішень та надання інформації, необхідної для прийняття рішення. DSS, як правило, допомагають управлінському персоналу в процесі прийняття рішень, хоча деякі з них можуть бути призначені для автоматизації прийняття рішень.

*Класифікація інформаційних систем.*

Класифікація ІС доволі розширена та проникає в різні сфери життя людей (рис. 1.2):



Рисунок 1.2 – Діаграма класифікацій ІС

За ступенем автоматизації:

- ручні - опрацювання за допомогою людини;
- автоматизовані – деякі з функцій виконуються автоматично;
- автоматичні – всі функції відбуваються без участі людини.

За масштабом:

– одиначні, які реалізовано, переважно, на автономному персональному комп'ютері без необхідності під'єднання до комп'ютерної мережі будь-якої форми;

– групові – спроектовані для використання декількома суб'єктами для вирішення спільних задач;

– корпоративні – переважно орієнтовані на великі корпоративні групи та корпорації з підтримкою віддалених територіально інформаційних мереж або вузлових структур та мають ієрархічну клієнто-серверну форму та спеціалізацію серверів;

– глобальні - охоплюють територію держави або континенту (Інтернет).

За предметною галуззю: (вказано декілька):

- економічна;
- адміністративна;
- навчальна;
- криміналістична;
- географічна;
- військова;
- екологічна;
- медична;
- виробнича.

За місцем діяльності:

- моделювання об'єктів проектування;
- керування технологічними процесами, для автоматизації технологічних процесів.
- наукові – для автоматизації діяльності науковців, статистичного аналізу інформації, експериментів, тощо.
- автоматизованого проектування для автоматизації праці.
- розробку та проектування нових виробів та технологій;
- інженерні розрахунки.
- організаційного керування, для автоматизацій функцій управлінського персоналу промислових підприємств та непромислових об'єктів (бірж, готелів, страхових компаній, банків тощо) та окремими філіями;

Тепер, коли ми розуміємо конкурентну перевагу та деякі способи використання ІТ, щоб допомогти організаціям отримати її, ми звернемо свою увагу на деякі конкретні приклади. Стратегічна інформаційна система - це інформаційна система, розроблена спеціально для реалізації організаційної стратегії, призначеної для забезпечення конкурентних переваг. Такі види систем

почали з'являтися у 1980-х роках, як зазначається у статті Чарльза Уайзмана під назвою «Створення конкурентної зброї з інформаційних систем» [4].

Зокрема, стратегічною інформаційною системою є та, яка намагається зробити одне або кілька з наступного:

- доставити товар або послугу за нижчою вартістю;
- поставити товар або послугу, що відрізняються;
- допомогти організації зосередитися на певному сегменті ринку;
- активізувати інновації.

Отже, інформаційні системи сьогодні інтегровані у всі компоненти бізнесу та є критично важливою частиною бізнес-процесів будь-якої організації або підприємства. Дослідження показують, що вплив ІС на прибуток компанії є позитивним. Правильне налаштування та інтеграція технологій є конкурентною перевагою, яка є чинником екстенсивного розвитку.

## **1.2 Формування витрат на впровадження та розвиток інформаційних систем**

В умовах постійного вдосконалення процесів інформатизації в світі важливим чинником функціонування і розвитку підприємств є їх адаптація до цих динамічних змін. Ефективне використання інформаційних ресурсів – запорука підвищення рівня конкурентоспроможності підприємств у ринковому середовищі, тому важливим є вирішення питань вдосконалення інформаційної складової системи управління.

Теорія і практика доводить, що високий рівень інформаційного забезпечення виступає однією із умов успішної реалізації інноваційних та

інвестиційних проєктів. Водночас, впровадження нових інформаційних технологій, зокрема, інформаційних систем управління, є теж інвестиційним проєктом, успішна реалізація якого в значній мірі залежить від оцінки ефективності на початкових етапах.

В умовах сучасності для розвитку економіки характерний перехід різних підприємств на більш прогресивні умови господарювання, при залученні перспективних галузей науки та техніки та збільшенням коефіцієнта ефективності виробництва з метою отримання максимуму із наявних фінансових ресурсів. Враховуючи прагнення підприємств розвивати свої виробничі потужності для їх ефективного функціонування повинна бути створена певна ІС його внутрішнього середовища, оскільки саме в такий спосіб ймовірно зайняти конкурентні позиції в умовах ринкової економіки. Чому саме ІС? ІС створює додаткові уявні робочі руки, оскільки оптимізує виробничі процеси завдяки забезпеченню збору та обробки інформації, та застосуванню можливих алгоритмів для впровадження швидкого реагування на зміну ситуації в середовищі для якого вона створена. Сучасні ІС все далі масштабуються та створюють уже цілковиті технічні екосистеми, на яких базується бізнес. Тому кожний елемент ІС будь то операційна система, база даних, апаратна платформа, cloud - рішення для розгортання ІТ-інфраструктури реалізується з урахуванням витрат, окупності, рентабельності, ефективності, що дає можливість також в подальшому експлуатувати ІС із такими показниками як:

- надійність;
- безпечність;
- відмовостійкість;
- масштабованість;
- сумісність;
- мобільність програмного забезпечення ІС.

В першу чергу при проектуванні та розробці ІС доцільно вирішити ряд завдань пов'язаних із ефективною автоматизованою підтримкою тих бізнес процесів, що в першу чергу створюють дохідність на ринку відповідних товарів та послуг, для яких була створена.

На практиці часто до негативних наслідків розробки а також і впровадження ІС призводить недооцінка етапу аналізу ІС і її підсистем, що включає конфігурування та функціональність, принципи побудови та вибір, аналіз і оцінка оптимальних показників якості ІС. Ще одним важливим фактором в розробці ІС є сформована ціна, що повинна відповідати усім явним та неявним вимогам рівнів показника якості та сучасної технічної архітектури. Зважаючи на дослідження з Центру економічних і ділових досліджень (CEBR), а також дані дослідницьких компаній IDC та Gartner впровадження в високорозвинених країнах ІС склало доволі вагому часту ВВП, хоча створило переваги в конкуренції тих підприємств, де було застосовано ІС. При плановому та раціональному підході ці витрати оправдовували витрати.

Більшість керівників компанії не вважають, що повернення інвестицій у інформаційні системи є вагомим поштовхом до впровадження в дію реалізацію проектів. Ефективність системи в більшій мірі оцінюється підвищення продуктивності праці. Хоча є безліч розроблених методологічних підходів до оцінки ефективності від експлуатації інформаційних систем в міжнародній практиці.

Накопичений іноземний та вітчизняний досвід впровадження корпоративних інформаційних систем у компаніях, свідчить про те, що воно вимагає значних витрат і, водночас, сприяє переходу на якісно новий рівень управління організацією, результатом якого є отримання нових конкурентних переваг через покращення низки кількісних та якісних показників діяльності підприємства:

- підвищення продуктивності праці управлінського персоналу;

- організаційне впорядкування виробничих процесів;
- поліпшення якості продукції і зниження собівартості її виробництва.

Досить часто у фаховій літературі розглядається питання правильного визначення витрат на впровадження інформаційної системи.

До витрат включають налаштування техніки і програмного забезпечення, ремонт ПК, модернізацію і т. ін. З боку користувача до основних витрат належать підтримка, сервіс та супровід програмного забезпечення, яким він користується.

Найбільша питома вага в структурі витрат належить витратам на технічну підтримку і супровід системи (77 % ІТ-бюджету підприємства) [5].

На це впливають такі фактори :

1) Витрати на технічну підтримку і супровід мають місце протягом усього життєвого циклу інформаційної системи, а не лише на етапі її тестування і введення в експлуатацію.

2) Якщо до розробки та впровадження інформаційної системи застосовується галузевий підхід, тобто програмний продукт розробляється для підприємств певної галузі, з урахуванням специфіки їх діяльності, супровід такої системи здійснюють безпосередньо представники компанії-розробника.

У цій ситуації розробник є в деякій мірі "монополістом" і може встановлювати високі ціни на послуги технічної підтримки і супроводу свого програмного рішення.

Технічне оснащення та переведення штатних процесів в електронний формат забезпечують набагато більшу віддачу від капіталовкладень, ніж будь-які інші інвестиції, про що говорить економічної ефективності комп'ютеризації, що проводилось у 380 фірм Америки [5, с. 85].

Навіть високий моральний знос обчислювального обладнання не може конкурувати із іншим типом пристроїв, хоча створений чистий прибуток від їх роботи значною мірою перевищує об'єм прибутку від інших інвестицій.

Аналіз цих фактів свідчить про те, що на сучасному етапі розвитку, роль інформаційних систем управління в підвищенні ефективності діяльності компаній досить вагома. Звичайно, дані оцінки мають суб'єктивний характер, однак є свідченням позитивного ставлення представників компаній розвинених країн до впровадження і використання інформаційних систем на різних рівнях управління.

Водночас, актуальним питанням є підрахунок реального економічного ефекту, який отримує компанія завдяки впровадженню інформаційних систем.

Для здійснення аналізу витрат на інформаційне забезпечення принципове значення має вибір методів та показників ефективності, що характеризують його стан та специфіку.

Вибір методів оцінки ефективності впровадження інформаційних систем є вкрай важливим, оскільки за умови обмеженості ресурсів та значних витратах на інформаційні технології, необхідно забезпечити інформаційну оптимізацію процесів, що відбуваються на підприємстві. При цьому оцінка інформаційного забезпечення проводиться з позиції найбільш повного забезпечення керуючої системи та керованого об'єкта в інформації, що дає змогу розробити, прийняти та організувати виконання оптимальних рішень.

Зазначимо, що серед напрямів аналізу інформаційного забезпечення виділяють наступні:

- 1) Семантичний – передбачає оцінку повноти інформації, об'єктивності, достовірності, точності та надійності, тобто ступеня відображення останніх досягнень науково-технічного прогресу та передового досвіду, відповідності та взаємозв'язку даних різного рівня та призначення.
- 2) Прагматичний – це аналіз корисності інформації. Такий вид аналізу проводиться стосовно конкретного об'єкту та періоду часу. Він включає оцінку забезпечення суб'єкта необхідною інформацією що використовується в процесі

управління; достовірності та своєчасності надходження цієї інформації до суб'єкта та інтенсивності її використання в процесі прийняття рішень.

3) Структурний аналіз інформаційного забезпечення ґрунтується на виявленні закономірностей поступового агрегування інформації за різними напрямками, відповідності інформаційних блоків сутності та ступеня складності управлінських рішень, що приймаються, особливостей усереднення даних та формування документів.

4) Економічний аналіз базується на визначенні "економічності" побудови системи інформаційного забезпечення і передбачає її оцінку за наступними напрямками: комплексність інформаційного забезпечення, рівень витрат на інформаційне обслуговування, періодичність оновлення інформації, здатність взаємодіяти з іншими інформаційними системами.

Розглянувши сутність кожного виду аналізу, можна відзначити, що всі вони тісно пов'язані між собою і такий їх поділ можна вважати досить умовним. Це пояснюється тим, що в основу напрямів оцінки інформаційних потоків покладені техніко-економічні показники, значення яких залежить від специфіки діяльності підприємства, особливостей побудови організаційної структури управління, прийнятих способів реєстрації, обробки та передачі інформації та оперативності надання достовірних даних особі, що приймає рішення.

Таким чином, необхідно детально розглянути економічний аспект оцінки ефективності інформаційних систем управління. Слід зазначити, що принцип "економічності" є основоположним при побудові інформаційної системи управління. Спеціалісти в сфері менеджменту наводять наступне визначення терміну «економічність побудови системи управління», а саме: це ступінь використання системою необхідних ресурсів, що визначаються відношенням обсягу ресурсів, запланованого для використання до фактично витраченого обсягу ресурсів.[6]

Якщо ж розглядати безпосередньо впровадження інформаційної системи управління, дане визначення потребує уточнення: в даному випадку принцип економічності передбачає, що витрати на її створення, впровадження та розвиток не повинні перевищувати ефекту, що очікується від її експлуатації.

Всі методи оцінки ефективності впровадження та використання інформаційних систем, що на сьогоднішній день широко використовуються на практиці, реалізують даний принцип, оскільки базуються на твердженні, що ефективність визначається як відношення результату від експлуатації інформаційної системи управління до витрат на її придбання, впровадження і супровід протягом всього життєвого циклу, виражене в грошових одиницях.

Ефективного використання ресурсів інформаційної системи можна досягнути зменшивши, насамперед, витрати на її впровадження, експлуатацію і вдосконалення.

Для підвищення точності підрахунку загальних витрат, що мають місце на етапах впровадження і розвитку інформаційної системи управління доцільно використовувати метод загальної вартості володіння (ТСО), згідно якого до прямих відносяться витрати, що пов'язані з придбанням, впровадженням інформаційної системи управління, та введенням її в промислову експлуатацію, а саме:

- витрати на обладнання і програмне забезпечення;
- на впровадження, супровід і розвиток інформаційної системи, відділом інформаційних технологій.

Непрямі витрати можуть виникати непередбачено на різних етапах життєвого циклу інформаційної системи. До цієї групи витрат відносять витрати часу користувача на самостійну підготовку до роботи з новою системою, кількість годин простою на місяць у зв'язку із припиненням роботи інформаційної системи.

Якщо проаналізувати види витрат, що визначені згідно методу загальної вартості володіння системою, то можна зробити висновок, що значна їх частина тісно пов'язана зі зміною інформаційного обслуговування управлінських процесів. Це витрати на формування даних, а саме:

- вартість та трудомісткість організації запитів;
- вдосконалення форм відображення необхідної інформації, зокрема, документів;
- витрати на створення банків даних та забезпечення ефективних комунікацій як всередині підприємства, так і вихід в зовнішні мережі. Очевидно, що ці витрати досить важко калькулювати і тому вони вже включені у вартість програмного продукту та витрат на розвиток технічної бази.

Перспективним напрямком для подальших досліджень є встановлення взаємозв'язку між прямими та непрямими витратами на впровадження та розвиток інформаційних систем і розробка економіко-математичної моделі інвестування коштів у ці проекти.

### **1.3 Тенденції розвитку ІС як хмарних технологій та їх вплив на економіку**

Одним із найбільш широко розвиненим представленням інформаційної системи є «хмара». «Хмара» або ж хмарна технологія – інформаційне середовище, що включає в себе обчислювальне ядро, місце збереження великих об'ємів даних та являє собою технологію розподілених обчислень, в якій комп'ютерні ресурси і можливості, доступні користувачеві як інтернет-послуга. ІС, в цьому випадку представлена як комплексне рішення розміщення програмного забезпечення для різних сфер економіки. На сучасному етапі експоненціальний розвиток ІТ-технологій створюють і такий же швидкий темп розвитку хмарних технологій, оскільки вони є фізичним фундаментом, на якому формується прикладний рівень будь-якої інформаційної системи. На дослідженні з Центру економічних і ділових досліджень (CEBR), а також даних дослідницьких компаній IDC та Gartner можна провести оцінки економічної вигоди та побудувати прогноз, що стосується корисності хмарних технологій, як фундаментальної ІС, для економіки.

Дослідження CEBR, що було проведене в 2015 році стверджує, що економіка розвинених європейських країн додатково отримали 177,3 млрд. євро в рік. Якщо групувати цю суму по країнам, то отримаємо наступні результати: Іспанія - 25,2 млрд. євро, Великобританія - 30,0 млрд. євро, Італія - 35,1 млрд. євро, Франція - 37,4 млрд. євро, Німеччина - 49,6 млрд. євро (рис. 1.3).



Рисунок 1.3 – Діаграма прибутку від хмарних технологій

Для світу ІТ ІС представлені, як хмарні технології – це новий підхід, в якому масштабованість ІС є гнучкою та мобільною, тобто доступною в потрібних кількостях та в потрібний час. Результати досліджень показують, що це пришвидшує виведення програмних продуктів на ринок, та знімає певні обмеження входу та дозволяє компаніям використовувати нові можливості для бізнесу. Створюючи тиск на конкурентів, хмарні технології мають велечезний вплив на структуру ринку в багатьох секторах, що як наслідок і впливає на глобальні показники. СЕВР вважають, що хмарні технології стануть важливим каталізатором конкурентоспроможності, економічного зростання і створення нових підприємств по всьому світу.

Якщо брати до уваги світові показники загального обсягу ринку хмарних технологій, то його значення збільшилось в 10 раз порівняно із базовим роком (рис. 1.4):

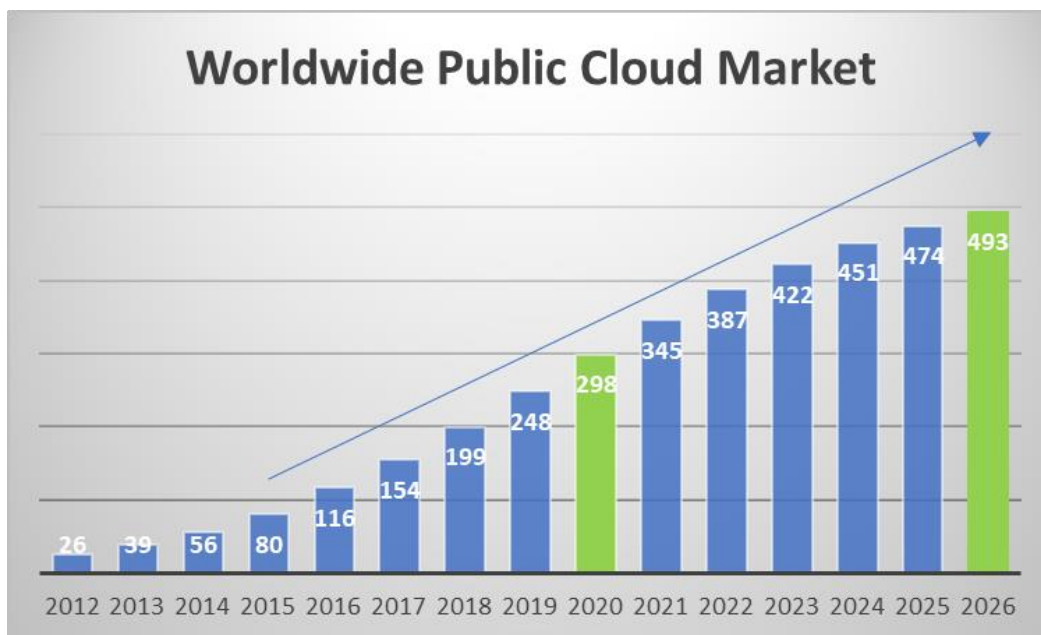


Рисунок 1.4 – Прогноз ринку хмарних обчислень компанією Wikibon

Провідний економіст CEBR Олівер Хоган сказав: "CEBR Дослідження показують, що хмарні обчислення - це не просто питання про короткострокову ефективність ІТ-інвестицій окремих компаній, а отже і їх продуктивність. Ця технологія може бути критичним макроекономічним фактором, який буде мати вирішальне значення для стимулювання економічного зростання, що особливо важливо в сьогоденній нестабільній економічній ситуації" [7].

В свою чергу президент ЕМС в регіоні ЕМЕА Райнер Ерлат прокоментував: "Рухливість і конкурентоспроможність, які забезпечують підприємству приватні та гібридні хмарні обчислення, створюють реальні можливості для глобального бізнесу - вони допомагають компаніям посилити свою перевагу, сприяючи росту економіки своїх країн" [7].

ІС в представленні хмарних технологій легко інтегровані, оскільки наразі на ринку є декілька основних гравців, що надають використання «хмари» як комплексного підходу формування та побудови своєї ІС: Microsoft Azure, Amazon, Google Cloud, Salesforce.com. Ці продукти допомагають різним підприємствам розширювати свої обчислювальні потужності при цім

економити значні суми фінансових ресурсів . Формат сплати послуг по факту їх використання призводить до зменшення капітальних витрат (CAPEX) і експлуатаційних витрат (OPEX), швидкої прибутковості інвестицій та реструктуризації ресурсів всередині компанії. Ці заощадження можуть бути реінвестовані, заохочуючи інновації, підвищуючи конкурентоспроможність та безпосередньо покращуючи рентабельність, що робить значний загальний позитивний ефект не тільки для певних компаній або підприємств, але і в цілому для економік цілих країн.

«Хмари» отримують широке застосування на основі наступних аргументів:

- скорочення капітальних витрат і витрат на обслуговування;
- підтримку і оновлення ІТ, зниження трудовитрат;
- можливість чіткого планування інвестицій в ІТ

Окрім цього «хмари» більш надійніші, безпечніші, легко масштабовані та прості у використанні. Вони сприяють розвитку інновацій та дозволяють на швидку зміну попиту на продукт чи диверсифікувати свої товари та послуги на ринку.

По всьому світі компанії витратять на хмарні технології порядку 204 мільярдів на 2021 рік, що на 16,5 % більше ніж в минулому 2020 році. Такі прогнози надає дослідницька компанія Gartner. Якщо порівняти, то , світовий ринок ІТ в цілому в 2021 р. повинен вирости всього на 0,6%. У свою чергу, в IDC очікують ще більшого зростання - на 19,4% до \$ 141 млрд проти \$ 70 млрд в минулому році. За Gartner, темп зростання ринку публічних хмарних сервісів в 27,5 разів перевищує темп зростання ІТ-ринку в цілому. У IDC назвали більш скромну цифру - темпи зростання відрізняються в шість раз [8].

Найбільший ріст отримує апаратний компонент ІС – інфраструктура (IaaS). Аналітики підрахували, що на 2021 рік приріст цього напрямку в грошовому еквіваленті складе 38,4%. Реклама в хмарі - найбільший сегмент світового ринку

хмарних технологій - продемонструє зростання на 13,6 \$ в 2021 р. і досягне \$ 90,3 млрд. Тоді як обсяг сегмента IaaS - \$ 22,4 млрд.

На противагу Gartner компанія IDC (International Data Corporation) виділяє сегмент ІС як програмного забезпечення як основний вектор зростання. Прогнозується, що саме на цей вид хмарних сервісів о в найближчі роки припаде більше ніж дві третіх всього ринку в грошовому вираженні. А ІС як інфраструктура відповідно їх же дослідженню складе в середньому 27% в період з 2020 по 2021 рр.

Світовий досвід роботи із «хмарою» спонукає підприємства України переглянути свої підходи до формування та побудови компонентів своїх ІС, та усвідомити зручність, надійність та безпечність переваги використання хмарних технологій. Діджиталізація державного сектору на території України та перехід до інформаційного суспільства, як ніколи близький до свого завершення, що робить такі технології лідерами на ІТ-ринку. Хоча зважаючи на всі аспекти «хмара» в Україні має свої позитивні та негативні сторони.

Для постачальника шляхом об'єднання ресурсів при нестійкому характеру споживання з боку клієнтів, хмарні обчислення дозволяють отримати економію за рахунок масштабу, оскільки використовується одні ті ж апаратні ресурси для декількох споживачів, та не створюється окремо для кожного із них за рахунок автоматизації процесу перерозподілу доступних ресурсів, що значно скорочує витрати на абонентське обслуговування.

Для споживача, хмарні технології надають високий рівень доступності і низький ризик відмови,, вони здатні забезпечити швидке масштабування завдяки гнучкості комп'ютерної системи, що не потребує підтримки та модернізації її апаратних компонентів. Простота доступу і універсальність забезпечує високу доступності послуг і підтримки різних класів термінальних пристроїв (ПК, мобільні телефони, інтернет-планшети). Експерти стверджують, що використання уже готових хмарних апаратних ІС в багатьох випадках знизить

витрати українських компаній в 2-3 рази порівняно із підтримкою власної інфраструктури. Не зважаючи на це, основна перевага використання даного типу ІС є здатність будь-якої установи швидко адаптуватися до змін навколишнього середовища, що в даний час, враховуючи бурхливий розвиток суспільства, дуже важливо [9].

Окрім переваг є і недоліки даного типу ІС. Для українського ринку тенденція до розвитку зросла відносно минулих років, але основною причиною такого повільного переходу на власні хмарні ІС є складність доступу до широкосмугового інтернету. Збільшення числа провайдерів інтернет-зв'язку та поліпшення ліній передачі даних поліпшить дану ситуацію. Окрім цього виникають проблеми обмеження в функціональних можливостях та сумісності програмного забезпечення із апаратною частиною ІС. На разі, найбільшнім недоліком «хмари» є її безпека, оскільки дані, що зберігаються можуть мати колосальне значення для бізнесу, а кібер-злочинці тільки і прагнуть проникнути та дістати важливу інформацію. Великі провайдери хмарних ІС впроваджують всі можливі засоби для збільшення рівня інформаційної безпеки і інвестують значні кошти в розробку нових, більш ефективних засобів захисту, але як і раніше, не слід зберігати або заливати особливо важливу інформацію в «хмару». Доволі дивним фактом є те, що на відміну від американських компаній, які, на подив, несерйозно сприймають інформаційну безпеку даних, українські підприємства, та підприємства Росії та ряду інших країн з інших країн з високим рівнем чорного ринку, з певною пересторогою відносяться до безпеки даних та наголошують, що це є головною проблемою хмарних технологій.

«Майкрософт Україна» в 2019 році провела дослідження, що визначило рівень використання «хмар» в Україні. Понад 65% українських компаній будуть використовувати хмарні технології вже в найближчі два роки. При дослідженні також було проведено опитування, що визначило 90% стурбованих опитуваних безпекою хмари.

Відповідно цьому ж опитуванню хмарний ринок України переходить на стадію становлення попиту та нагромадження певного практичного досвіду роботи із хмарними технологіями. Хоча зацікавленість поки низька, в найближчі 1-2 роки буде «хмарний вибух», оскільки такі альтернативні рішення вирішують просто величезний перелік проблем. Опитування показало мінімальне знання кінцевих користувачів про хмарні обчислення і низький рівень проникнення технологій (рис. 1.5):



Рисунок 1.5 – Рівень знання користувачів про хмарні обчислення

Підкреслимо, що 47% ІТ-компаній вважають, що їх розуміння хмарних рішень є поверхневим, а 88% менеджерів не знайомі з хмарними сервісами. Дослідження було проведено на підприємствах середнього та великого масштабу в сферах фінансів телекомунікацій, роздрібно́ї торгівлі, логістики та промислового секторів, оскільки ці сфери є основними споживачами ІТ-послуг в Україні.

Отже, підводячи підсумок визначень тенденцій хмарних ІС можна ствердити, що з кожним днем вони набирають обертів. У світі все більше і більше підприємств обирають процес переходу свого бізнесу на «хмару» оскільки це пропонує комплексний підхід до побудови та формуванню власної ІС на апаратних компонентах наданої ІС. Хмарні технології завойовують

передові позиції, а враховуючи набути світовий досвід компанії прямо зараз поліпшують своє становище на ІТ-ринку. Звісно існує певний ряд проблем, які на даний момент в силу можливостей не вирішений, але спроби до локалізації таких баг-задач наявні, що у світі, що і на території України. Слід відмітити, що вчасне впровадження хмарних ІС допомогло компанія закріпити свої позиції на ринку, перед тими підприємствами, що знехтували даною технологією.

#### **1.4 Система масового обслуговування – аналог процесу роботи ІС**

Складний характер ринкової економіки і сучасний рівень пропонованих до ній вимог стимулюють використання більш серйозних методів аналізу її теоретичних і практичних проблем. В останні десятиліття значну вагу в економічних дослідженнях придбали математичні методи. Математичне моделювання все більше і більше стає одним з основних і найбільш плідних методів вивчення економічних процесів і об'єктів. Математичний аналіз економічних задач органічно перетворюється в частину економіки. Позитивна оцінка цього підтверджується і тим, що починаючи з 1969 р Нобелівські премії в області економіки присуджуються, як правило, за економіко-математичні дослідження. Теорія систем масового обслуговування є одним із підрозділів імітаційного моделювання, що включає економіко-математичний аналіз та формує теоретичні основи ефективного конструювання СМО. Системи масового обслуговування – це математичні моделі таких реальних чи спроектованих систем, функціонування яких можна розглядати, як послідовність взаємодій деяких дискретних об'єктів з елементами системи. Ці об'єкти ,

зазвичай поступають в систему із зовнішнього джерела, деякий час взаємодіють із елементами системи по деяким правилам і згодом покидають її.

Теорія масового обслуговування почалася з роботи А. К. Ерланга 1909 р в якій він показав, що пуасонівський розподіл може бути застосовано до випадкового телефонному трафіку. Ерланг розробив теорію, яка багато десятиліть використовувалася для проектування телефонних мереж. Теорія дозволяє визначити ймовірність відмови в обслуговуванні. Для телефонних мереж з комутацією каналів відмова відбувалася, якщо всі канали зайняті і дзвінок не може бути здійснений. Імовірність цієї події потрібно було контролювати. Хотілося мати гарантію, що, наприклад, 95% всіх дзвінків будуть обслужені. Формули Ерланга дозволяють визначити, скільки потрібно серверів для виконання цієї гарантії, якщо відома тривалість і число дзвінків. По суті, це завдання як раз про гарантії якості, і сьогодні люди як і раніше вирішують схожі завдання. По суті уже на той момент люди розуміли, що очікування та відмова в обслуговуванні коштуватиме значно більше ніж просто втрата одного клієнта. На разі у боротьбу за клієнта в сучасній економіці вкладаються величезні кошти. За оцінками західних економістів, завоювання фірмою нового клієнта обходиться їй у 6 разів дорожче, ніж утримання існуючих покупців. А якщо клієнт пішов незадоволеним, то на його повернення доводиться витратити в 25 разів більше коштів. Для уникнення подібних ситуацій підприємства використовують теорію масового обслуговування як уже вирішений приклад задачі.

Системи масового обслуговування (СМО) містять один або кілька каналів які обробляють заявки, що надходять від декількох джерел (рис. 1.6).

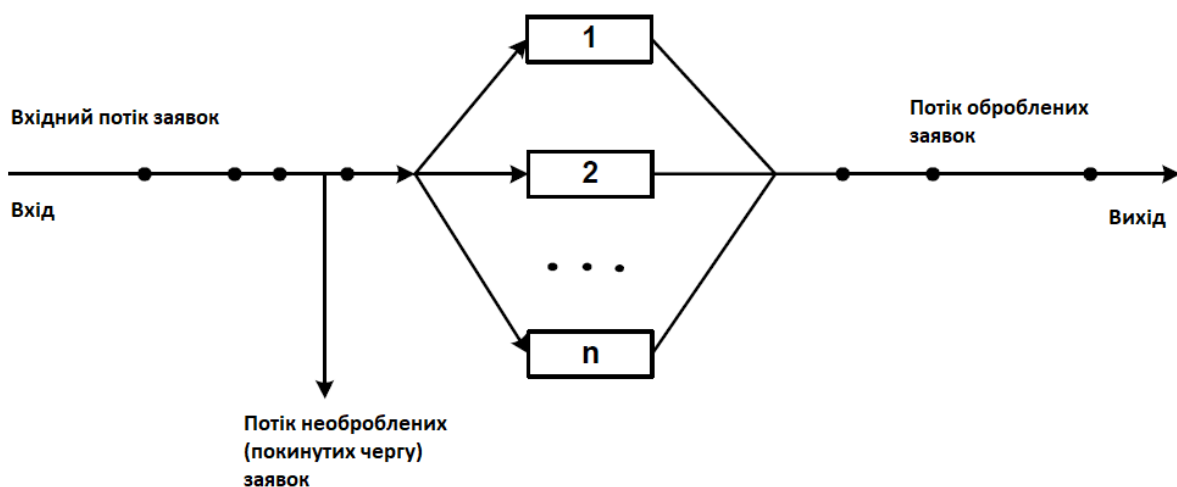


Рисунок 1.6 – Схема СМО

Обслуговуючий об'єкт в теорії систем масового обслуговування називається вимогою або заявкою. У загальному випадку під вимогою (заявкою) зазвичай розуміють запит на задоволення деякої потреби. Роль вимог або заявок виконують клієнти, відвідувачі, покупці, документи, товари, суду і так далі. Засоби, що обслуговують вимоги (заявки), називаються обслуговуючими пристроями або каналами обслуговування. Роль каналів обслуговування грають касири, фахівці, канали телефонного зв'язку, товарознавці, комп'ютери, майстри-ремонтники, вантажно-розвантажувальні точки на базах і складах і так далі. Вимоги (заявки) надходять в систему масового обслуговування (СМО) із стохастичним часом надходженням, утворюючи так випадковий потік вимог (заявок) із певною інтенсивністю. Обслуговування заявок також має випадковість розподілення. Оскільки потік заявок і часу мають випадковий характер, обслуговування призводить до того, що СМО в протягом часу роботи є завантаженою нерівномірно: в якісь періоди часу накопичується велика черга із відмовами (можливий варіант покинути СМО необробленою), в інші ж періоди СМО працює з недовантаженням, працюючи в «холосту».

Теорія масового обслуговування в деякій мірі має оптимізаційну задачу і в підсумку включає економічні питання за визначенням такого варіанту системи, при якому буде забезпечений мінімум сумарних витрат створених очікуванням обслуговування, використанням ресурсів на обслуговування, втрат часу на очікування і від простоїв каналів обслуговування. Найбільш часто використовуваними моделями потоків подій при аналізі СМО є регулярні і пуасонівські. Регулярні характеризуються однаковим часом між наступами подій, а пуасонівська - випадковим.

Таким чином, у будь-якій СМО можна виділити наступні основні елементи:

- 1) вхідний потік заявок;
- 2) чергу;
- 3) канали обслуговування;
- 4) виходить потік обслужених заявок.

Відповідно своїх параметрів, тобто кількості каналів обслуговування їх утилізації, інтенсивності потоку заявок, а також певних особливостей організації внутрішніх процесів кожна СМО володіє певною пропускнуою спроможністю, що дозволяє її оброблювати інтенсивні потоки заявок.

Як і будь-які системи відповідних теорії, СМО мають свої внутрішні класифікації.

СМО поділяється за кількість модулів обробки заявок (рис. 1.7):

- одно-канальні (коли є один модуль обслуговування);
- багатоканальні, точніше  $n$ -канальні (коли кількість модулів  $n \geq 2$ ).

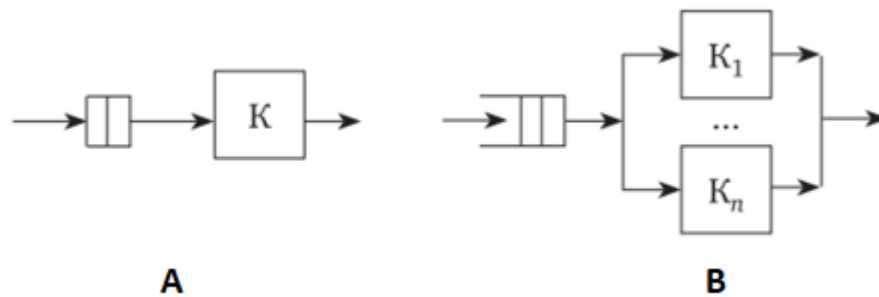


Рисунок 1.7 – А) одно-канальна СМО В) багатоканальна

В одно-канальних кожна заявка, яка обслуговується, знаходиться під обслуговування єдиного модуля (каналу). Багатоканальні СМО включають різні канали обслуговування, у яких час обслуговування  $T_{об}$  неоднаковий, де час обслуговування є випадковою неперервною величиною. Але зважаючи на те, що заявки та модулі обслуговування можуть бути однорідними, то  $T_{об} = const$ .

*СМО за дисциплінами обслуговування розділяється на три класи:*

1) СМО з відмовами - це СМО, в якій заявка, що надходить в момент, коли всі канали зайняті, отримує відмову, залишає СМО і в подальшому в процесі обслуговування не бере (наприклад, телефонна мережа). (рис. 1.8)

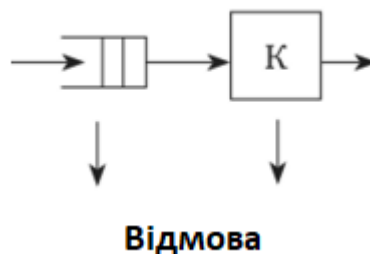


Рисунок 1.8 – СМО з відмовою

2) СМО чергою (з очікуванням) - це СМО, в якій запит, що прийшов в момент, коли всі канали зайняті, не йде, а стає в чергу на обслуговування. У свою чергу СМО з очікуванням (чергою) поділяються на (рис. 1.9):

- СМО з необмеженою чергою.
- СМО з обмеженою чергою.



Рисунок 1.9 – А) необмежена СМО В) обмежена СМО

3) СМО змішаного типу (з обмеженим очікуванням). Вхідна заявки, не прийнята оскільки всі пристрої зайняті, стають в чергу і чекає обслуговування протягом обмеженого часу. Якщо час очікування вийшов, то заявка залишає систему.

*По пріоритету обслуговування:*

1) СМО зі статистичним пріоритетом - СМО, в якій обслуговування проводиться в порядку надходження заявок (рис. 1.10.а).

2) СМО з відносним пріоритетом - СМО, в якій заявка високого пріоритету очікує закінчення обслуговування заявки з більш низьким пріоритетом (рис. 1.10.b) (СМО, де більш важлива заявка отримує лише «краще» місце в черзі).

3) СМО зі змішаним пріоритетом - СМО. в якій використовується абсолютний пріоритет, якщо заявка з нижчим пріоритетом обслуговувалася в плинні часу, менше критичного, і використовується відносний пріоритет в іншому випадку(рис. 1.10.c).

4) 4. СМО з абсолютним пріоритетом - СМО, в якій заявка високого пріоритету при надходженні витісняє заявку з більш низьким пріоритетом (рис. 1.10.d).

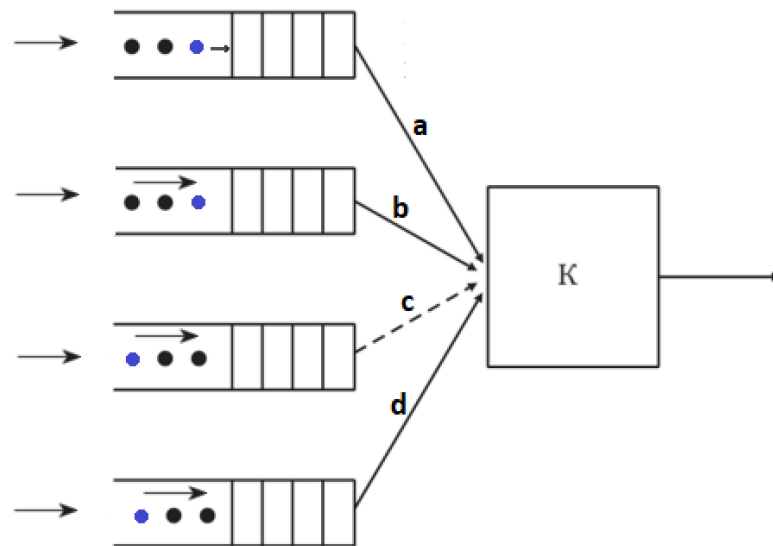


Рисунок 1.10 – А) статичний В) відносний С) змішаний D) абсолютний

*По принципу обслуговування:*

- 1) СМО з обслуговуванням за принципом "перший прийшов - останній обслужений" (наприклад, СМО з обслуговуванням за принципом "перший прийшов - останній обслужений")(рис. 1.11.a).
- 2) СМО з обслуговуванням за принципом "перший прийшов - останній обслужений"(рис. 1.11.b).

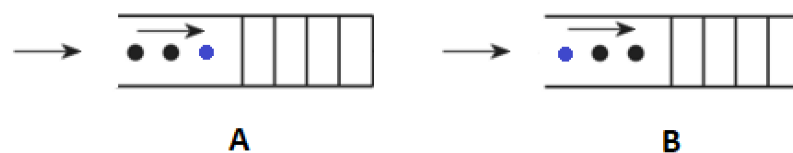


Рисунок 1.11 – А)FIFO В) LIFO

*В залежності від генерації заявок:*

- 1) Відкриті СМО - СМО, де циркулює кінцева, зазвичай постійна кількість вимог, які після завершення обслуговування повертаються в джерело (рис 1.12.a).
- 2) Замкнуті СМО - СМО, де джерело генерує нескінченну кількість вимог(рис. 1.12.b).

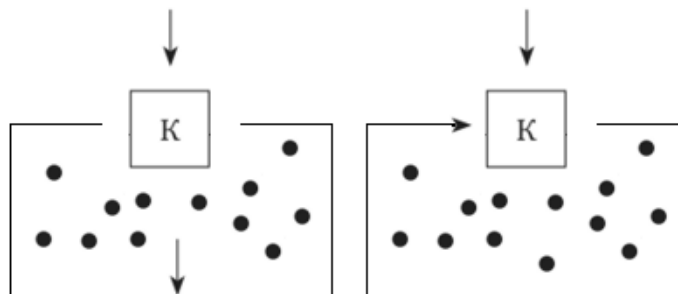


Рисунок 1.12 – А) Відкрита СМО В) Закрита СМО

Така класифікація є в більшості умовною, оскільки в реальності найчастіше СМО виступають в якості змішаних систем. Тобто в певні моменти часу, модель поведінки СМО змінюється.

Відповідно описаного базису теорії масового обслуговування ІС є одним із багатьох різновидів систем, що може бути описані за допомогою СМО, оскільки основним її функціоналом є збір, обробка або ж зберігання даних, тобто виконання певних операцій на поставлені задачі. ІС відтворює формування черг, обробку запитів та потоку відхилень. Усе це базується на природі обраної ІС. Ми зустрічаємося із теорією масового обслуговування в ІС повсякчасно, тільки не прокинувшись та оновивши повідомлення в месенджері (запит на нові повідомлення на сервер – оброблення сервером запиту на оновлення), розраховуючись в супермаркеті безготівковим шляхом (запит в бази банку). Система масового обслуговування – це переведена тривіальна версія на теорію будь-якої ІС.

## 2 МОДЕЛЮВАННЯ ВИТРАТ КОМПАНІЇ НА ІНФОРМАЦІЙНУ СИСТЕМУ

### 2.1 Постановка економічної задачі

Швидке зростання цифрових технологій та віртуалізація ведення бізнесу призвели до збільшення масштабованості інфраструктурної архітектури інформаційних систем, як горизонтальної так і вертикальної. Слід одразу зазначити, що поняття інформаційної системи є досить широким і застосовується як в різних сферах так і в різних напрямках економіки, техніки, суспільства. Стійкі високонавантажені системи докорінно змінили вектор ведення бізнесу, що переросло у своєрідну гонку озброєності компаній конкурентів. Основними якісними показниками якої є не точність та масштабність, а надійність та продуктивність. Ефективність функціонування бізнесу побудованого на інформаційній системі суттєво залежить від потужності ІТ-інфраструктури та продуктивності використання її ресурсів. Будь-яке довготривале падіння таких систем тягне за собою ряд значних проблем: великі фінансові втрати, зниження репутації компанії та втрату своїх позицій на ринку, як надійного партнера, що є одним із найголовніших каталізаторів розвитку та руху діяльності компанії. Під терміном «високонавантажена інформаційна система» ховається складна, взаємопов'язана структура серверів або ж сервісів, що характеризуються великою кількістю запитів за певну одиницю часу, що повинні бути оброблені і передані. Перенавантаження таких систем, тобто неспроможність обробити масивний потік даних, спричиняє колапс, падіння продуктивності та переводить весь бізнес-процес у небоєздатний стан, що спричиняє ряд економічних проблем зазначених вище.

При таких ризиках, будь-яка інформаційна система, при збільшенні своїх масштабів та нарощенні обчислювальних потужностей контролює дві основних критерії: відмовостійкість та доступність. Одним із основних показників навантаження системи є так званий QPS (query per second) – кількість запитів на систему в секунду. Значення даного показника є пульсометром для високонавантаженої системи типу «Запит-Відповідь». Якість моніторингу та контроль якого є обернено-пропорційними відмовостійкості системи.

Даний показник в собі несе також і економічний зміст, адже обслуговування його значення закладено в собівартість випущеної компанією технології чи то продукту. При динамічних змінах його від сподіваного можна або ж скоротити витрати на обслуговування ІТ-інфраструктури (як варіант, банально зменшивши кількість серверів) або ж відповідно їх збільшити задля розширення об'ємів технології для споживача, що потягне за собою збільшення прибутку. Оптимізація гри із цим показником має дві сторони медалі:

- Позитивно-якісний вплив на собівартість випущеної технології та закріплення конкурентних позицій компанії на ринку;
- Неочікувані фінансові втрати та ризик банкрутства.

Задля стабілізації зміни цих можливих варіантів розвитку подій під час оптимізації даного показника на доброякісному стані - невід'ємною частиною для компанії буде прогноз майбутніх значень показника, що обумовлює велику практичну потребу в проведенні дослідження над рядом даних і знаходженням оптимальних, адекватних та теоретично обґрунтованих методів прогнозу. Тобто основні задачі дослідження полягають у прогнозуванні майбутнього навантаження на інформаційну систему для імітації поведінки ІС за для експериментальних перевірок стресостійкості системи, що в свою чергу, безпосередньо впливає на основні витрати компанії. Додатковими, паралельно

досліджуваними питаннями є можливі методи прогнозування та стратегії планування витрат на підтримку інформаційної системи.

Основа алгоритму дослідження в даній роботі базується на комбінованому підході, що виражається у аналізі часового ряду, застосуванні відповідних прогностичних моделей та формуванні оцінок адекватності побудованих прогнозів, які для наступного етапу надають верхню межу значень показника завантаженості системи для імітаційної моделі. Оскільки основна задача інформаційної системи – це збір, обробка або ж зберігання даних [10], тобто виконання певних операцій на поставлені задачі, ця система є одним із екземплярів систем масового обслуговування [12,13]. Процес системи масового обслуговування базується на декількох основних стовпах: розподілу часу між надходженнями, інтенсивності надходження заявок, числу обслуговуючих пристроїв (що в нашому випадку є модульними частинами системи, які відповідають за певну долю потоку) та обмеження на одночасне число запитів на систему. За нотацією Кендалла [14], яка використовується для опису різних систем, система набуває виду  $A/S/k/K$ , де:

$A$  – розподіл часу між запитами;

$S$  – розподіл розмірів запитів;

$k$  – кількість пристроїв обслуговування;

$K$  – обмеження довжини черги.

Моделювання поведінки системи із обмеженістю у верхньому піці навантаження дає змогу оцінити утилізацію модулів системи. Якщо певний період часу модулі не є завантаженими, оскільки навантаження що в цілому виробляється на систему не досягає тієї потужності при якій масштабується інфраструктура, вони відповідно відключаються із загальної системи та не створюють фінансові експлуатаційні та амортизаційні витрати.

## 2.2 Загальний опис моделі

### 2.2.1 Прогностична модель часового ряду

В основі будь-якого дослідження закладений збір та первинний аналіз даних об'єкта дослідження. Відштовхуючись від природи об'єкта та його властивостей, дані можуть бути представлені в різних формах: випадкові вибірки, часові ряди. На відміну від аналізу випадкових вибірок, аналіз часових рядів ґрунтується на припущенні, що послідовні значення даних спостерігаються через рівні проміжки часу[15]. Існують дві основні цілі аналізу часових рядів:

- Визначення природи ряду (наявність тренду, сезонності, стаціонарності);
- Прогнозування (передбачення майбутніх значень часового ряду по теперішнім і минулим значенням).

Зв'язок між цими цілями несе в собі причинно-наслідковий характер, адже без визначення природи ряду важко підібрати відповідні методи прогнозування. По оцінці Е.С. Тихонова в його роботі «Прогнозування у умовах ринку» [16] кількість методів та моделей прогнозування становить більше сотні модифікованих варіантів. Основу цих моделей та методів будують по наступним групам:

- 1) Регресійні моделі прогнозування;
- 2) Авторегресійні моделі прогнозування (ARIMA, GARCH, ARDLN);
- 3) Моделі експоненціального згладжування (ES);
- 4) Модель за вибіркою максимального подібності (MMSP);
- 5) Модель на нейронних мережах (ANN);

- 6) Модель на ланцюгах Маркова (Markov chains);
- 7) Модель на класифікаційно-регресійних деревах (CART);
- 8) Модель на основі генетичного алгоритму (GA);
- 9) Модель на опорних векторах (SVM);
- 10) Модель на основі передавальних функцій (TF);
- 11) Модель на нечіткій логіці (FL).

Кожна із цих моделей базує свої прогнози відповідно природи часового ряду, що досліджує. Одними із найширше використовуваними класами моделей є авторегресійні моделі прогнозування та моделі експоненціального згладжування. Їх широкому застосуванню можна завдячувати простоті побудови, ефективності прогнозу та легкості у реалізації. Окрім цих факторів, дані моделі враховують та опираються у своїх прогнозах на основні складові ряду – наявність тренду, сезонність, стаціонарність, кривизна.

Сформований ряд даних нашого об'єкту дослідження, який базується на метриці показника навантаження на систему QPS, адаптований до того, щоб мати кожен із цих ознак. Діяльність підприємства, що надає дані, націлена на надання технології показу інтернет-реклами. Кожна одиниця показника QPS по своїй суті є запитом на показ реклами для відповідного клієнта на стороні партнерського сайту. Відповідно цього кількість одиниць QPS плаває протягом доби, що надає ряду сезонної складової. Оскільки компанія є одним із провідних провайдерів інтернет-реклами на просторах України та відкриває свої горизонти на міжнародній арені, кількість споживачів даної технології росте, що говорить про наявність зростаючого тренду в ряді.

Тому для чистоти експерименту в дослідженні використовуються декілька методів із даних груп, а саме:

- Авторегресійна модель ARIMA;
- Експоненціальне згладжування;

- Метод Хольта або ж подвійне експоненціальне згладжування;
- Модель Хольта-Вінтерса або потрійне експоненціальне згладжування.

Оцінка точності прогнозу і визначення найоптимальнішої моделі буде базуватись на абсолютній процентній помилці MAPE прогнозованих значень вище згаданих моделей та методів.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - \hat{Y}_t|}{Y_t}, \quad (2.1)$$

де  $n$  – кількість спостережень;

$Y_t$  – фактичне значення;

$\hat{Y}_t$  – прогнозоване значення.

На результаті найбільш достовірних прогнозів буде створено обмеження в навантаженні при симуляції роботи інформаційної системи на наступному етапі роботи - імітаційному моделюванні.

*Модель ARIMA.* Вперше систематичний підхід до побудови моделі ARIMA був викладений Боксом і Дженкінсом в 1976 році [17]. Методологія побудови ARIMA-моделі (рис. 2.1) для досліджуваного часового ряду включає наступні основні етапи:

- Ідентифікацію пробної моделі;
- Оцінка параметрів моделі;
- Діагностична перевірка адекватності моделі.



Рисунок 2.1 – Структурна схема підбору моделі ARIMA

Сама модель має наступний характерний запис:

$$(\Delta^d X_t) = \sum_{t=1}^p \varphi(\Delta^d X_{t-1}) + \varepsilon_t + \sum_{j=1}^q \theta_j(\Delta^d \varepsilon_{t-1}), \varepsilon_t \sim N(0, \sigma_t^2). \quad (2.2)$$

Можемо використати і більш короткий запис:

$$\varphi(B)(1 - B)^d X_t = \theta(B) \varepsilon_t, \quad (2.3)$$

де  $\varphi(\cdot), \theta(\cdot)$  – поліноми степені  $p$  і  $q$ ;

$B$  – лаговий оператор ( $B^j X_t = X_{t-j}, B^j \varepsilon_{t-j}, j = 0, \pm 1, \dots$ );

$d$  – порядок взяття послідовної різниці (2.4)

$$(\Delta X_t = X_{t-1} - X_t = (1 - B)X_t, \Delta^2 X_t = \Delta^2 X_{t+1} - \Delta X_t = (1 - B)^2 X_t, \dots). \quad (2.4)$$

Таким чином, спочатку (в блоці 1-3) необхідно отримати стаціонарний ряд. На цьому етапі рекомендується проводити аналіз автокореляційної функції (АКФ) та часткової автокореляційної функції (ЧАКФ).

Швидке загасання значень АКФ - простий тест на стаціонарність. На цьому етапі використовуються також статистичні тести на наявність одиничного кореня (розширений тест Дікі-Фуллера або ADF-тест)[18].

Часовий ряд має одиничний корінь, або порядок інтеграції один, якщо його перші різниці утворюють стаціонарний ряд. Ця умова записується як:

$$y_t \sim I(1). \quad (2.5)$$

Якщо ряд перших різниць  $\Delta y_t = y_t - y_{t-1}$  є стаціонарним, то  $\Delta y_t \sim I(0)$ . За допомогою цього теста перевіряються значення коефіцієнта  $a$  в авторегресійному рівнянні першого порядку AR(1):

$$y_t = a y_{t-1} + \varepsilon_t, \quad (2.6)$$

де  $y_t$  – часовий ряд;

$\varepsilon$  – помилка.

Якщо  $a = 1$ , то процес має одиничний корінь, в цьому випадку ряд  $y_t$  не стаціонарний, є інтегрованим часовим рядом першого порядку –  $I(1)$ . Якщо  $|a| < 1$ , то ряд стаціонарний –  $I(0)$ .

Наведене авторегресійне рівняння AR (1) можна переписати у вигляді:

$$y_t = b y_{t-1} + \varepsilon_t, \quad (2.7)$$

де  $b = a - 1$ ;

$\Delta$  – оператор різниці першого порядку  $\Delta y_t = y_t - y_{t-1}$ .

Тому перевірка гіпотези про одиничний корень в даному поданні означає перевірку нульової гіпотези про рівність нулю коефіцієнта  $b$ .

Статистика тесту ( $DF$ -статистика) - це звичайна  $t$ -статистика для перевірки значущості коефіцієнтів лінійної регресії. Однак, розподіл даної статистики відрізняється від класичного розподілу  $t$ -статистики (розподіл Стьюдента або асимптотичний нормальний розподіл). Розподіл  $DF$ -статистики виражається через вінерівський процес і називається розподілом Дікі - Фуллера. Існує три версії тесту (тестових регресій):

1. Без константи:

$$y_t = by_{t-1} + \varepsilon_t. \quad (2.8)$$

2. З константою, але без тренду:

$$y_t = b_0 + by_{t-1} + \varepsilon_t. \quad (2.9)$$

3. З константою і лінійним трендом:

$$y_t = b_0 + b_1t + by_{t-1} + \varepsilon_t. \quad (2.10)$$

Для кожної з трьох тестових регресій існують свої критичні значення  $DF$ -статистики, які беруться зі спеціальної таблиці Дікі - Фуллера (МакКіннона). Якщо значення статистики лежить лівіше критичного значення (критичні значення - негативні) при даному рівні значущості, то нульова гіпотеза про одиничні корені відхиляється і процес визнається стаціонарним (в сенсі даного тесту). В іншому випадку гіпотеза не відкидається і процес може містити одиничні корені, тобто бути нестаціонарним (інтегрованим) часовим рядом.

Якщо відповідно до статистики Дікі-Фуллера або оцінок АКФ ряд є нестаціонарним, то для переходу до стаціонарного ряду традиційно застосовують

оператор взяття послідовних різниць, тим самим визначається значення параметра  $d$ .

У блоці 4 після отримання стаціонарного ряду досліджується характер поведінки вибірових АКФ і ЧАКФ і висуваються гіпотези про значення параметрів  $p$  (порядок авторегресії) і  $q$  (Порядок ковзного середнього). На вході блоку 4 може формуватися базовий набір, що включає одну, дві або навіть більше число моделей, іншими словами, портфель моделей.

У блоці 5 після здійснення ідентифікації моделі необхідно оцінити їх параметри. Для цих цілей використовується метод максимальної правдоподібності (ММП).

У блоці 6 для перевірки кожної пробної моделі на адекватність аналізується її ряд залишків. У адекватної моделі ряд залишків повинен бути схожий на білий шум, тобто їх часткові АКФ не повинні відрізнятися від нуля. Для перевірки гіпотези про те, що спостережувані дані є реалізацією «білого шуму», використовується також  $Q$ -статистика.

$$Q^* = n(n+2) \sum_{k=1}^m \frac{r_k^2}{n-k}, \quad (2.11)$$

де  $n$  – кількість спостережень;

$m$  – максимальна кількість лагів;

$r_k$  – коефіцієнти автокореляційної функції.

Якщо в результаті перевірки кілька моделей виявляються адекватні вихідними даними, то при остаточному виборі слід врахувати два фактори:

- підвищення точності (якість підгонки моделі);
- зменшення числа параметрів моделі.

За допомогою моделі в блоці 7 можна будувати точний і інтервальний прогноз на  $L$  кроків вперед. Для оцінки точності прогнозу використовується ряд

стандартних показників (1). Відповідність якості прогнозу та значення похибки представлено (табл. 2.1 ).

Таблиця 2.1– Проміжки оцінок прогнозу

|       |             |              |                          |
|-------|-------------|--------------|--------------------------|
| 0% <  | <i>MAPE</i> | $\leq 10\%$  | Висока точність          |
| 10% < | <i>MAPE</i> | $\leq 20\%$  | Вища середнього точність |
| 20% < | <i>MAPE</i> | $\leq 50\%$  | Задовільна точність      |
| 50% < | <i>MAPE</i> | $\leq 100\%$ | Слабка точність          |

Відповідно такого алгоритму побудовано портфель моделей-кандидатів та оцінено результати прогнозування в підрозділі 2.3.

*Експоненціальне згладжування.* Експоненціальне згладжування - це дуже популярний метод прогнозування багатьох часових рядів. Історично метод був незалежно відкритий Броуном і Хольтом. Броун служив на флоті США під час Другої світової війни, де займався виявленням підводних човнів і системами наведення. Пізніше він застосував відкритий ним метод для прогнозування попиту на запасні частини. Свої ідеї він описав у книзі, що вийшла в світ в 1959 році. Дослідження Хольта були підтримані Департаментом військово-морського флоту США. Незалежно один від одного, Броун і Холт відкрили експоненціальне згладжування для процесів з постійним трендом, з лінійним трендом і для рядів з сезонної складової.

Проста і прагматично ясна модель часового ряду має такий вигляд:

$$X_t = b + \varepsilon_t, \quad (2.12)$$

де  $b$  – константа;

$\varepsilon$  – помилка.

Константа  $b$  відносно стабільна на кожному часовому інтервалі, але може також повільно змінюватися з часом. Один з інтуїтивно ясних способів виділення

$b$  полягає в тому, щоб використовувати згладжування ковзним середнім, в якому останнім спостереженням приписуються більші ваги, аніж передостаннім, передостаннім більші ваги, аніж перед-передостаннім. Просте експоненціальне саме так і влаштовано. Тут більш старим спостереженням приписуються експоненціально спадні ваги, при цьому, на відміну від змінного середнього, враховуються всі попередні спостереження ряду, а не ті, що потрапили в певне вікно. Точна формула простого експоненціального згладжування має наступний вигляд:

$$S_t = \alpha X_t + (1 - \alpha)S_{t-1}. \quad (2.13)$$

Коли ця формула застосовується рекурсивно, то кожне нове згладжене значення (яке є також прогнозом) обчислюється як зважене середнє поточного спостереження і згладженого ряду. Очевидно, результат згладжування залежить від параметра  $\alpha$  (альфа). Якщо рівно 1, то попередні спостереження повністю ігноруються. Якщо рівно 0, то ігноруються поточні спостереження. Значення між 0, 1 дають проміжні результати.

Розрахунок найкращого значення параметра  $\alpha$  (альфа) досліджувалось в роботах [20][21][22]. В роботі [20] обговорюються різні теоретичні та емпіричні аргументи на користь вибору певного параметра згладжування. Очевидно, з формули, наведеної вище, слідує, що  $\alpha$  повинна потрапляти в інтервал між 0 (нулем) і 1 (хоча дослідження [21] для подальшого застосування аналізу АРПСС вважають, що  $0 < \alpha < 2$ ). В цій же роботі наголошують, що практиці зазвичай рекомендується брати  $\alpha$  менше 0.3. Однак в дослідженні [22], рекомендації до вибору  $\alpha$  завжди приймають значення більше 0.3, оскільки саме воно найчастіше дає кращий прогноз. Загалом ще однією рекомендацією, якою пронизані дослідження є той факт, що краще оцінювати оптимально  $\alpha$  по даним, ніж просто "гадати" або використовувати штучні рекомендації.

На практиці параметр згладжування часто шукається з пошуком на сітці. Можливі значення параметра розбиваються сіткою з певним кроком. Наприклад, розглядається сітка значень від  $\alpha = 0.1$  до  $\alpha = 0.9$ , з кроком 0.1. Потім вибирається, для якого сума квадратів (або середніх квадратів) залишків (спостережувані значення мінус прогнози на крок вперед) є мінімальними.

*Метод Хольта або ж подвійне експоненціальне згладжування.* В простому методі експоненціального згладжування припускається, що рівень значень часових рядів міняється рідко, тому необхідна оцінка лише поточного рівня. В деяких випадках дані спостережень можуть мати помітні відхилення. В подібній ситуації необхідна функція лінійного тренду. Оскільки часові ряди рідко характеризуються фіксованим лінійним трендом, слідує розглянути можливість локального лінійного тренду, що змінюється в часі. В 1957 р. Ч. Хольт розібрав метод експоненціального згладжування[29], який отримав назву двухпараметричного метода Хольта. В цьому методі враховується локальний лінійний тренд, присутній в часовому ряді.

Якщо в часових рядах є тенденція до росту, то окрім оцінки поточного рівня необхідна і оцінка нахилу. В методі Хольта значення рівня і нахилу згладжуються неопосередковано, при цьому для них використовуються різні постійні згладжування, які дозволяють оцінити поточний рівень та нахил, уточнюючи їх кожен раз як тільки з'являються нові спостереження. Одна із переваг методу Хольта – його гнучкість, що дозволяє вибрати співвідношення, в якому відслідковуються рівень і нахил.

Основою для застосування методу прогнозування Хольта є наявність тренду. Для виявлення його наявності найкраще застосувати метод Фостера-Стюарта. Він не тільки дозволить знайти тренд в середньому, а й в дисперсії часового ряду, що може сказати про «розхитування» часового ряду у випадку

збільшенні дисперсії та постійного розкиду рівнів ряду у випадку коли її немає. Статистика цього методу має вигляд :

$$S = \sum_{t=2}^n s_t; \quad D = \sum_{t=2}^n d_t, \quad (2.14)$$

$$s_t = u_t + l_t; \quad d_t = u_t - l_t, \quad (2.15)$$

$$u_t = \begin{cases} 1, & \text{якщо } x_t > x_{t-1}, x_{t-2} \dots x_1, \\ 0, & \text{у всіх інших випадках} \end{cases}, \quad (2.16)$$

$$l_t = \begin{cases} 1, & \text{якщо } x_t < x_{t-1}, x_{t-2} \dots x_1. \\ 0, & \text{у всіх інших випадках} \end{cases}. \quad (2.17)$$

Значення величин  $u_t$  та  $l_t$  знаходяться шляхом послідовного порівняння рівнів ряду. Якщо будь-який рівень перевищує по своїй величині кожний із попередніх то величині  $u_t$  прирівнюється значення 1, в інших випадках – 0. І таким же чином тільки навпаки – якщо рівень ряду менше всіх попередніх, то  $l_t$  прирівнюється значення 1, а в наступних випадках – 0. Статистика  $S$  використовується для перевірки наявності тренду в дисперсіях, а статистика  $D$  – для знаходження тренду в середніх, межі яких очевидно належать таким проміжкам:

$$0 \leq S \leq n - 1; \quad -(n - 1) \leq D \leq n - 1. \quad (2.18)$$

Наступним кроком методу є перевірка гіпотези про те, чи можна вважати різниці  $D - 0$  та  $S - \mu$ , де  $\mu$  - математичне сподівання величини  $S$  – випадковими. Для цього використаємо t-критерій Стюдента.

$$t_s = \frac{S - \mu}{\partial_s}; \quad t_d = \frac{D - 0}{\partial_d}, \quad (2.19)$$

де  $\partial_s$  та  $\partial_d$  – середньоквадратичні помилки величин  $S$  та  $D$  та рівні:

$$\partial_s = \sqrt{2 \ln(n) - 3,4253}; \quad \partial_d = \sqrt{2 \ln(n) - 0,8456}, \quad (2.20)$$

де  $n$  – кількість ступенів свободи.

Фінальним кроком є порівняння табличного значення  $t$ -критерія Стьюдента при  $n$ -ступенів свободи та рівні значимості  $\alpha = 0,05$  з отриманими в результаті розрахунків. В табл. 2.2 наведені можливі варіанти порівняння значень.

Таблиця 2.2 – критерії визначеності наявності тренду

| Нерівність     |                | Наявність тренда                  |
|----------------|----------------|-----------------------------------|
| $t_d > t_{кр}$ | $t_s > t_{кр}$ | Тренд у середньому                |
| $t_d < t_{кр}$ | $t_s > t_{кр}$ | Тренд у середньому та в дисперсії |
| $t_d > t_{кр}$ | $t_s < t_{кр}$ | Невизначеність                    |
| $t_d < t_{кр}$ | $t_s < t_{кр}$ | Тренд відсутній                   |

Підтвердивши що тренд присутній у середньому, можна застосовувати метод Хольта. Метод Хольта використовується для прогнозування часових рядів, коли є тенденція до зростання або падіння значень часового ряду, як було зазначено вище. А також для рядів, коли дані є не за повний цикл, і сезонність ще не виділити (наприклад, за неповний рік для прогнозу по місяцях).

У запропонованому алгоритмі значення рівня та тренду згладжуються за допомогою експоненціального згладжування. Для управління рівнем і нахилом в моделі Хольта вводиться 2 коефіцієнта згладжування – коефіцієнт згладжування ряду і тренду. Причому параметри згладжування у них різні:

$$\hat{y}_{0k} = \alpha y_k + (1 - \alpha)(\hat{y}_{0k-1} - y_{k-1}), \quad (2.21)$$

$$t_k = \beta(\hat{y}_{0k} - \hat{y}_{0k-1}) + (1 - \beta)t_{k-1}, \quad (2.22)$$

$$\hat{y}_{k+p} = \hat{y}_{0k} + pt_k, \quad (2.23)$$

де  $\hat{y}_{0k}$  – згладжувана величина на даний період;

$\alpha$  – коефіцієнт згладжування ряду;

$y_k$  – поточне значення ряду;

$\hat{y}_{0k-1}$  – згладжена величина на попередній період;

$t_{k-1}$  – значення тренду за попередній період;

$\beta$  – коефіцієнт згладжування тренду;

$\hat{y}_{k+p}$  – прогноз на  $p$ -періодів;

$t_k$  – тренд за останній період.

Перше рівняння описує згладжений ряд загального рівня. Друге рівняння служить для оцінки тренду. Третє рівняння визначає прогноз на  $p$ -періодів вперед. Підбір коефіцієнтів сформований на мінімізації суми середньоквадратичних відхилень на основі лінійного програмування.

*Модель Хольта-Вінтерса або потрійне експоненціальне згладжування.* Пітер Вінтерс в своїй роботі [23] розвив модель експоненціального згладжування з трендом Хольта та додав в неї сезонність. Даний метод отримав свою назву на основі імен своїх авторів Чарльза Хольта та Пітера Вінтерса. Метод Хольта-Вінтерса - це удосконалення методу експоненціального згладжування часового ряду. Експоненціальне згладжування забезпечує загальне представлення про тренд і дозволяє робити короткострокові прогнози. Відмінність від експоненціального згладжування полягає в здатності методу виявляти тренди, які стосуються коротких періодів в моменти часу, що безпосередньо передують прогнозованим, і екстраполювати ці тренди на майбутнє. Метод Хольта-Вінтерса заснований на тому, що досліджуваний часовий ряд може бути представлений у вигляді трьох компонент: коефіцієнту згладжування, лінії тренду і сезонного ефекту. Алгоритм передбачає, що кожна з цих компонентів змінюється в часі. До умов, що змінюються значенням кожної з компонент застосовується експоненціальне згладжування.

Як і в методі експоненціального згладжування, прогноз на наступний період обчислюється застосуванням до поточного значення прогнозу коефіцієнтів:  $\alpha$ ,  $\beta$ ,  $\gamma$ . Дані коефіцієнти являються параметрами моделі і можуть

приймати значення в межах від 0 до 1. При більш високих значеннях коефіцієнтів минулі значення враховуються в більшій мірі, ніж поточні, а при більш низьких найбільший вплив на прогноз надають поточні значення.

Прогнозом на  $p$ -кроків вперед:

$$\hat{y}_{k+p} = (\hat{y}_k + p t_k) s_{k-s+p}, \quad (2.24)$$

де  $\hat{y}_{k+p}$  – прогноз на  $p$ -кроків;

$p$  – порядковий номер кроку, на який робимо прогноз;

$t_k$  – значення тренду;

$\hat{y}_k$  – експоненціально згладжена величина;

$s_{k-s}$  – коефіцієнт сезонності попереднього періоду.

Оцінка сезонної компоненти (Сезонність в першому і другому періоді дорівнює 1):

$$s_k = \gamma \left( y_k / \hat{y}_k \right) + (1 - \gamma) s_{k-s}, \quad (2.25)$$

де  $\gamma$  – константа згладжування циклів  $s$  періодом;

$y_k$  – поточне значення ряду.

Оцінка тренду:

$$t_k = \beta (\hat{y}_k - \hat{y}_{k-1}) + (1 - \beta) t_{k-1}, \quad (2.26)$$

де  $\beta$  – коефіцієнт згладжування тренду.

Експоненціально згладжений ряд:

$$y_k = \alpha (y_k / s_{k-s}) + (1 - \alpha) (\hat{y}_{k-1} + t_{k-1}), \quad (2.27)$$

де  $\alpha$  – коефіцієнт згладжування ряду.

У першій точці ряду значення експоненціального згладжування і тренду не розраховуються, так як для їх розрахунку не існує попередніх експериментальних значень. У другій точці ряду значення коефіцієнту згладжування приймається

рівним її спостереженому значенню, а мікро-тренд за цей період вважається лінійним і розраховується як різниця між поточним і минулим значенням.

Починаючи з третьої точки, використовуються формули (2.24) - (2.27):

1. Будуємо експоненційно-згладжений ряд, формула (2.27);
2. Визначаємо значення тренду за формулою (2.26);
3. Оцінюємо сезонність. Відповідно розрахунку в початковій точці третього спостереження застосовуємо формулу (2.25);
4. Прогнозуємо  $p$ -періодів за формулою (2.24);

При появі нових даних прогноз за методом Хольта -Вінтерса бажано перерахувати для уточнення ряду, тренду і сезонності.

### 2.2.2 Імітаційна модель інформаційної системи

Інформаційна система по своїй природі має декілька концептуальних інтерпретацій, оскільки вона є складною динамічною системою, в своїй структурі містить декілька підсистем різного роду: фізичну систему на якій розміщується, бізнес-систему, що впроваджує бізнес-логіку ІС, абстрактну систему зв'язків, що будують відносини між компонентами. Складність функціонування таких ІС формує не тільки строкатість рівнів підсистем, але й кількість множин змінних параметрів, що впливають на систему. Імітаційна модель, в свою чергу, відтворює поведінку цієї системи взаємодіючих підсистем та вхідних параметрів, базуючись на наслідуванні реального процесу (імітації). Для опису такої імітаційної моделі в роботі застосовується математична модель динамічної системи, що визначається сукупністю множин і відображень:

$$MM_{dc} = \langle T, X, \Omega, H, Z, Y, \Phi_1, \Phi_2 \rangle, \quad (2.28)$$

де  $T$  – універсум відношень між елементами системи;

$X = X_1 \times \dots \times X_{NX}$  – простір вихідних впливів;

$\Omega = \Omega_1 \times \dots \times \Omega_{N\Omega}$  – простір впливів зовнішнього середовища;

$Z = Z_1 \times \dots \times Z_{NZ}$  – простір внутрішніх станів системи;

$Y = Y_1 \times \dots \times Y_{NY}$  – простір вихідних і залежних змінних

$\Phi_1(z_0, x(t), h(t), w(t), t) = z_1(t), \dots, z_{NZ}(t))^T \in Z$  – оператор зміни внутрішніх станів системи в залежності від початкового внутрішнього стану та зміни з часом вхідних впливів, внутрішніх параметрів, зовнішнього середовища.

$\Phi_2(z(t), t) = y(t) = (y_1(t), \dots, y_{NY}(t)))^T \in Y$  – оператор зміни виходів системи в залежності від зміни внутрішніх станів системи.

Математична модель динамічної системи включає основні елементи реальної інформаційної системи, що можуть бути використанні в симуляції її роботи.

Дана імітаційна модель ІС являє собою систему масового обслуговування, оскільки її основна задача – це обробка запитів користувачів на показ інтернет-реклами, де формується черга та виникають стохастичні процеси. Системи масового обслуговування (СМО), до яких відносяться клієнт-серверні системи, є класом математичних схем, розроблених в теорії масового обслуговування і різних застосуваннях для формалізації процесів функціонування систем, які за своєю суттю є процесами обслуговування.

Дана система масового обслуговування є багатоканальною (рис 2.2) з необмеженою чергою, що містить нескінченну кількість станів ( $D_0, D_1, \dots, D_L, D_{L+1}, \dots$ ), де перші  $L$  станів відповідають фізичній підсистемі інформаційної системи у якій є незадіяні процесорні ядра та черга запитів пуста.

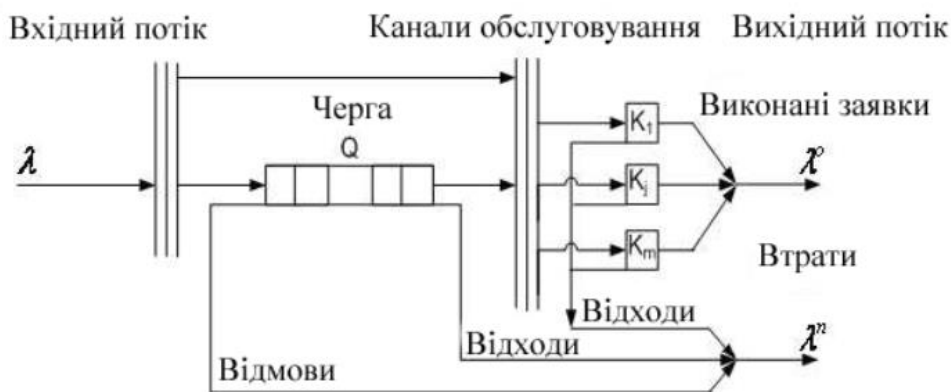


Рисунок 2.2 – Модель обробки запитів

Граф станів інформаційної системи як системи масового обслуговування зображений на рис.2.3:

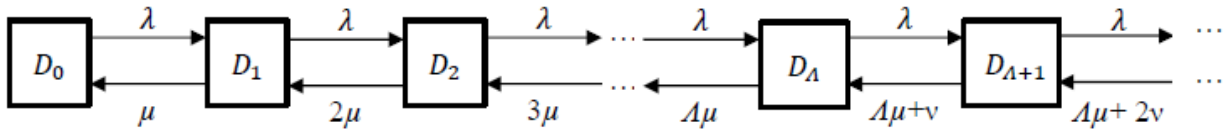


Рисунок 2.3 – Граф станів інформаційної системи як системи масового обслуговування

Потік запитів послідовно переводить систему з будь якого стану  $D_0$  у стан  $D_{i-1}$ , що знаходиться правіше на рисунку з постійною інтенсивністю  $\lambda$ .

$$\lambda = \frac{1}{\bar{\tau}_{\text{над}}}, \quad (2.29)$$

де  $\bar{\tau}_{\text{над}}$  – середній час надходження заявки.

Інтенсивність запитів обмежена максимальними піками навантаження в прогнозованому ряді. Інтенсивність потоку обслуговування зростає зі збільшенням кількості задіяних каналів від  $\mu$  до  $\Lambda\mu$ , коли задіяні всі канали і рівна:

$$\mu = \frac{1}{\bar{\tau}_{\text{об}}}, \quad (2.30)$$

де  $\bar{\tau}_{\text{об}}$  – середнє значення часу обслуговування.

Також на систему діє потік відхилених запитів, який переводить систему зі стану  $D_i$  у стан  $D_{i-1}$ . Інтенсивність потоку відхилених запитів зростає при збільшенні черги на  $\nu$  для кожного наступного стану СМО.

Запити мають право знаходитись в СМО не більше  $\tau_{\text{доп}}$  одиниць часу. Цей час є максимальним для обслуговування запиту та встановлюється при заданні параметрів обробки інформаційної системи. Якщо час в системі перевищує  $\tau_{\text{доп}}$ , то запит покидає систему та вважається відхиленим. Величина  $\tau_{\text{доп}}$  є сталою та фіксованою. По протоколу торгу інтернет реклами IAB OpenRTB, час відповіді на запит реклами не повинен перевищувати  $\tau_{\text{доп}} \leq 200\text{мс}$ .

Зручним для подальшого розгляду є представлення про найпростіший потік відходів із СМО з інтенсивністю

$$\eta = \frac{1}{\tau_{\text{доп}}}, \quad (2.31)$$

Відходи запитів можливі або з черги, якщо  $t_{\text{оч}} > \tau_{\text{доп}}$ , або ж з каналу обслуговування якщо  $t_{\text{оч}} \leq \tau_{\text{доп}} \leq t_c$ . Методично зручно розглядати два потоки відходів із інтенсивністю відповідно (відходи черги і обробки):

$$\eta_{\text{оч}} \text{ та } \eta_{\text{об}} : \eta_{\text{об}} = \eta_{\text{оч}} = \eta = \frac{1}{\tau_{\text{доп}}}, \quad (2.32)$$

Дана модель обробки запитів є одним із класичних варіантів СМО, за нотацією Кендалла має вигляд: M/M/5/FIFO [4]. Тобто:

/M – на вході подається Пуасонівський потік запитів (заявки приходять в випадкові моменти часу із заданою середньою швидкістю запитів в секунду або, більш формально: час між двома сусідніми подіями має експоненціальний розподіл).

/M – Обслуговування цих заявок теж має експоненціальний розподіл і в середньому обробляється  $\mu$  запитів в секунду.

/5 – Канали обслуговування відображають фізичні пристрої обслуговування, що в нашому випадку є модулями ІТ-інфраструктури (фізичної підсистеми).

/FIFO – дисципліна обслуговування (*first in - first out*)

Основні показники такої системи визначаються теорією Марківських ланцюгів. При зведенні інтенсивності всіх потоків до інтенсивності потоку обслуговування матимемо:

- Зведена інтенсивність вхідного потоку, яка відображає середнє число заявок, що поступають на вхід СМО за середній час обслуговування однієї заявки.

$$\rho = \frac{\bar{\lambda}}{\bar{\mu}}. \quad (2.33)$$

- Зведена інтенсивність потоку відходів з черги:

$$\alpha_{оч} = \frac{\eta_{оч}}{\mu}. \quad (2.34)$$

- Зведена інтенсивність потоку відходів з каналу обслуговування.:

$$\alpha_{об} = \frac{\eta_{об}}{\mu}. \quad (2.35)$$

- Ймовірність  $P_0$  (немає заявок в системі)

$$P_0 = \left[ 1 + \sum_{i=1}^m \frac{\rho^i}{i!(1+\alpha_{об})^i} + \frac{\rho^m}{m!(1+\alpha_{об})^m} \sum_{l=1}^n \prod_{j=1}^l \frac{\rho}{m(1+\alpha_{об})+j\alpha_{оч}} \right]^{-1}. \quad (2.36)$$

- Ймовірність знаходження системи в одному зі станів  $P_i, i = \overline{1, m}$ , тобто при нулі заявок та частковому завантаженню каналів рівна:

$$P_i = \frac{\rho^i}{i!(1+\alpha_{об})^i} P_0. \quad (2.37)$$

При повному завантаженні каналів та наявності заявок в черзі очікування матимемо:

$$P_{m+l} = \frac{\rho^m}{m!(1+\alpha_{об})^m} \prod_{j=1}^l \frac{\rho}{m(1+\alpha_{об})+j\alpha_{оч}} P_0, l = \overline{1, n}. \quad (2.38)$$

Одним із важливих показників ефективності є середнє число каналів, що зайняті обслуговуванням:

$$\bar{K} = \sum_{i=1}^m i P_i + m(1 - \sum_{i=0}^m P_i). \quad (2.39)$$

При цьому середня довжина черги:

$$\bar{l} = \sum_{i=1}^m l P_{m+l}. \quad (2.40)$$

У розглядуваній СМО втрати заявок можливі або у формі відмови внаслідок переповнення системи, або у формі відходу «нетерплячих» заявок із системи. Справедливі наступні формули:

$$P_{\text{відм.}} = \frac{\rho^{m+n}}{m!(1+\alpha_{\text{об}})^m \prod_{j=1}^l (m(1+\alpha_{\text{об}}) + j\alpha_{\text{оч}})} P_0 . \quad (2.41)$$

де  $P_{\text{відм.}}$  – ймовірність відмови;

$P_{\text{в}}^{\text{оч}} = \frac{\bar{l}\eta_{\text{оч}}}{\lambda}$  – ймовірність відходу заявки під час очікування;

$P_{\text{в}}^{\text{об}} = \frac{\bar{K}\eta_{\text{об}}}{\lambda}$  – ймовірність відходу заявки під час обробки;

Моделювання процесу обробки запитів викладений в розділі 3.1

## З МОЖЛИВОСТІ ТА НАПРЯМИ ПРАКТИНОГО ЗАСТОСУВАННЯ РОЗРОБЛЕНОЇ МОДЕЛІ

### 3.1 Реалізація моделі

#### 3.1.1 Побудова прогнозів

*ARIMA model.* Для висвітлення практичного значення і корисності застосовуваних методів та моделей потрібно використати не менш практичні дані високонавантаженої системи. В дослідженні були використанні дані однієї із провідних компаній-провайдерів інтернет-реклами на ринку України. На рис 3.1 побудовано динаміку зміни значення агрегованого середнього показника QPS на серверну інфраструктуру за 6 місяців 2020 року.

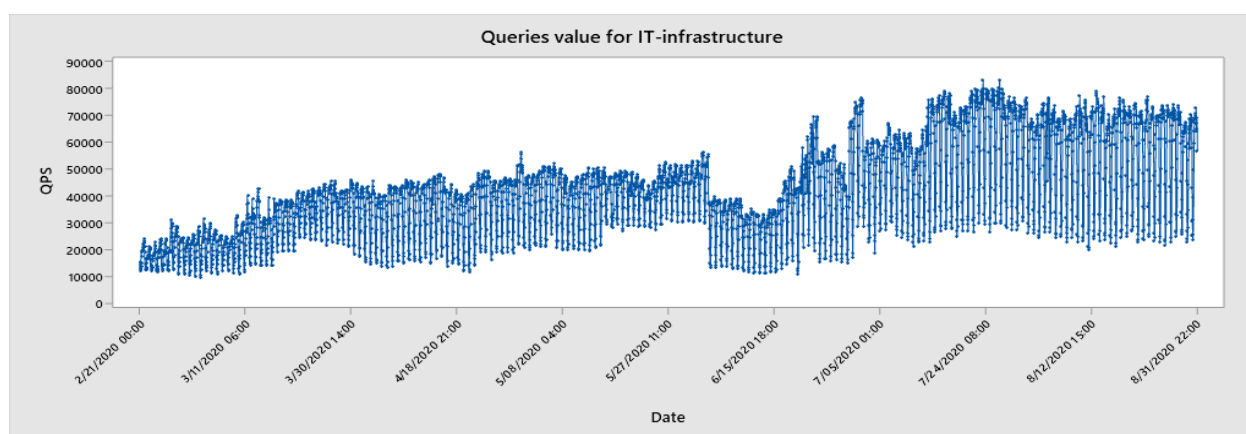


Рисунок 3.1 – Графік динаміки зміни показника QPS

Часовий ряд сформований із цілодобового погодинного середнього значення навантаження на систему, що унеможлиблює довгостроковий прогноз, оскільки періодичність спостереження є година. Збільшити період прогнозу в такому випадку можливо за рахунок розбиття часового ряду. Дані подані щодобовим погодинним середнім значенням, тому для збільшення проміжку перегрупуємо дані відносно години доби (рис. 3.2 – 3.3)

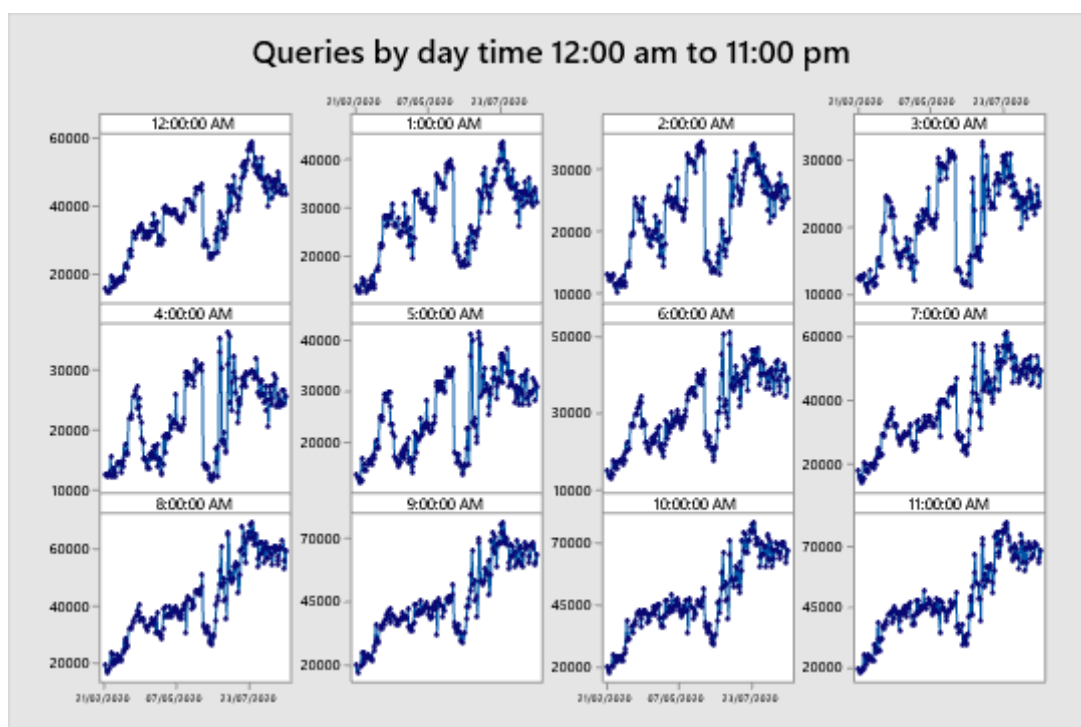


Рисунок 3.2 – Груповані дані по годинах з 00:00 до 11:00

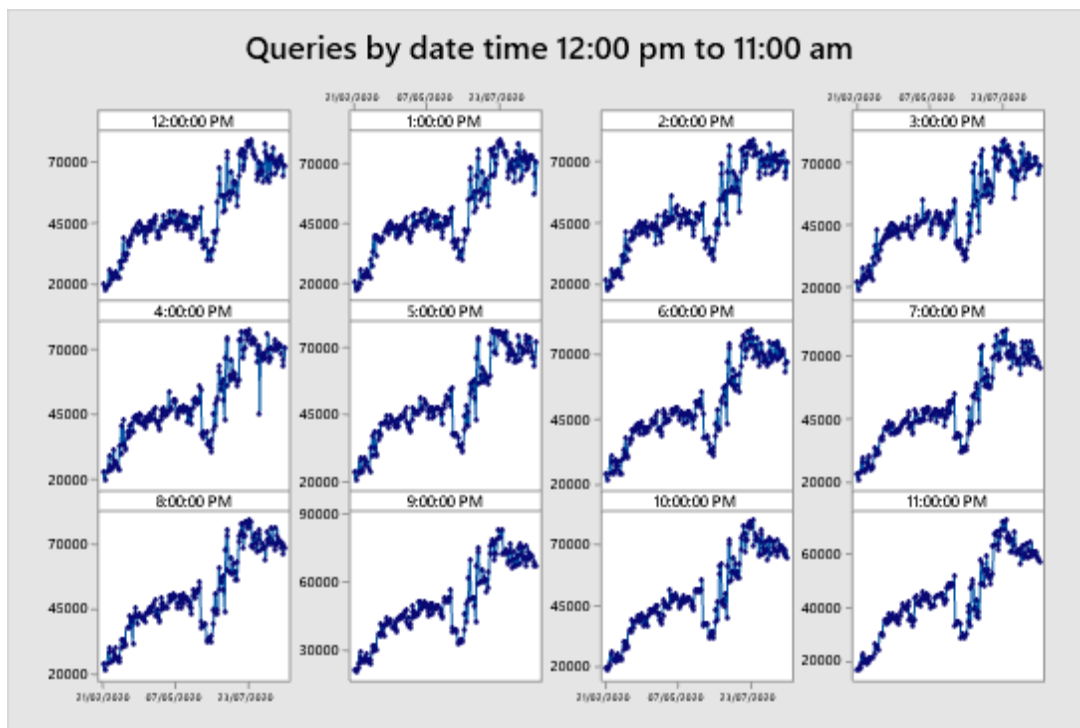


Рисунок 3.3 – Груповані дані по годинах з 12:00 до 23:00

Відповідно структурній схемі підбору моделі ARIMA визначимо стаціонарність розподілених часових рядів шляхом проведення загального тесту Дікі-Фуллера на наявність одиничних коренів.

Згідно значенням критерію ADF-тесту ряди є нестационарними, оскільки нульова гіпотеза про наявність одиничного кореня підтверджується на 5% рівні значущості, а фактичні значення t-критерія Стюдента знаходиться вище табличного\* ( -1,94248) (рис. 3.4), при цьому кількість лагів, яке потрібно включати в модель при застосуванні ADF-тесту, дорівнювало 0. Для дослідження було використано версію тесту без константи та тренду. Для переходу до стаціонарності проведемо послідовне взяття різниці ряду, та проведемо повторний тест (рис.3.5).



Рисунок 3.4 – Порівняльний графік табличного значення Дікі-Фуллера та t-критерія



Рисунок 3.5 – Порівняльний графік табличного значення Дікі-Фуллера та t-критерія Стюдента послідовних різниць

Після взяття першої послідовної різниці ADF-тест показав наяву стаціонарність кожного групованого часового ряду. Оскільки операція послідовної різниці була застосована тільки один раз, то значення коефіцієнту в моделі  $ARIMA(p, d, q)$   $d$  рівний 1.

Оскільки ряди є стаціонарними, проведемо оцінку порядку параметрів  $p$  і  $q$  моделі  $ARMA(p, q)$ , яка, в свою чергу, являє собою модель  $AR(p)$  і  $MA(q)$ . Для цього скористаємося АКФ (рис. 3.6-3.7): і ЧАКФ відповідно (рис. 3.8-3.9):

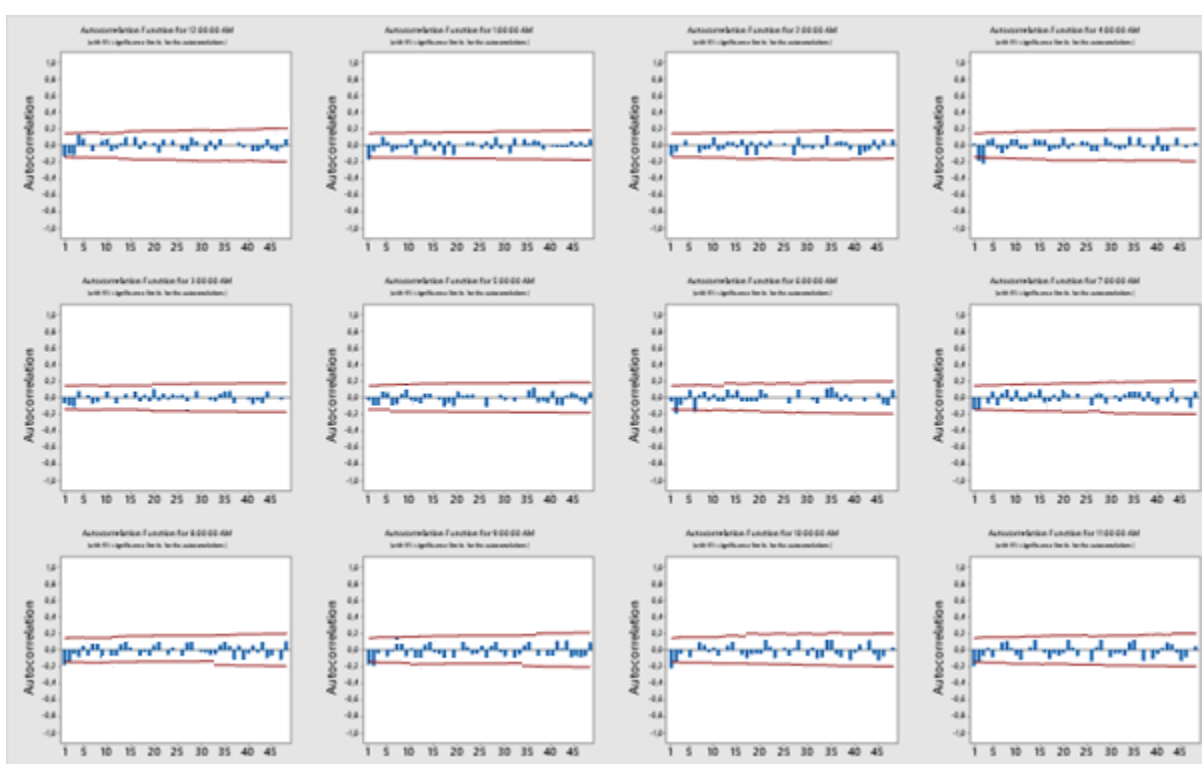


Рисунок 3.6 – Автокореляція часових рядів 00:00 - 11:00

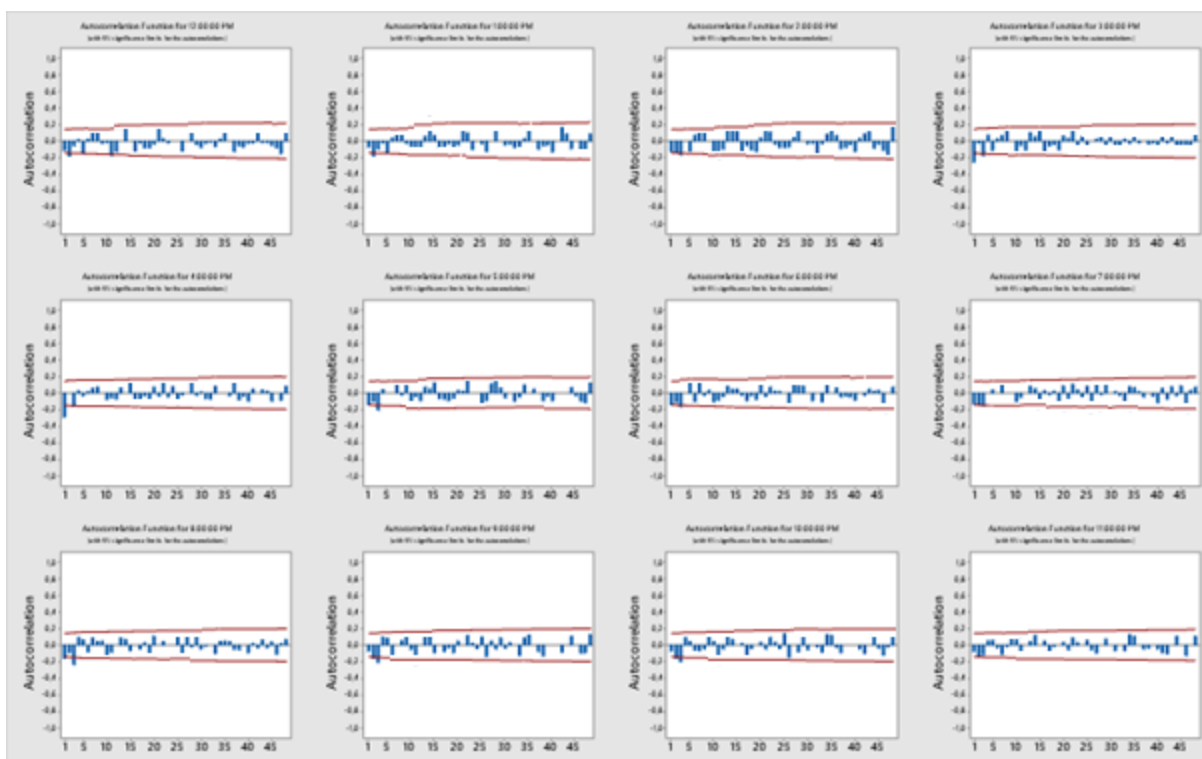


Рисунок 3.7 – Автокореляція часових рядів 12:00 - 23:00

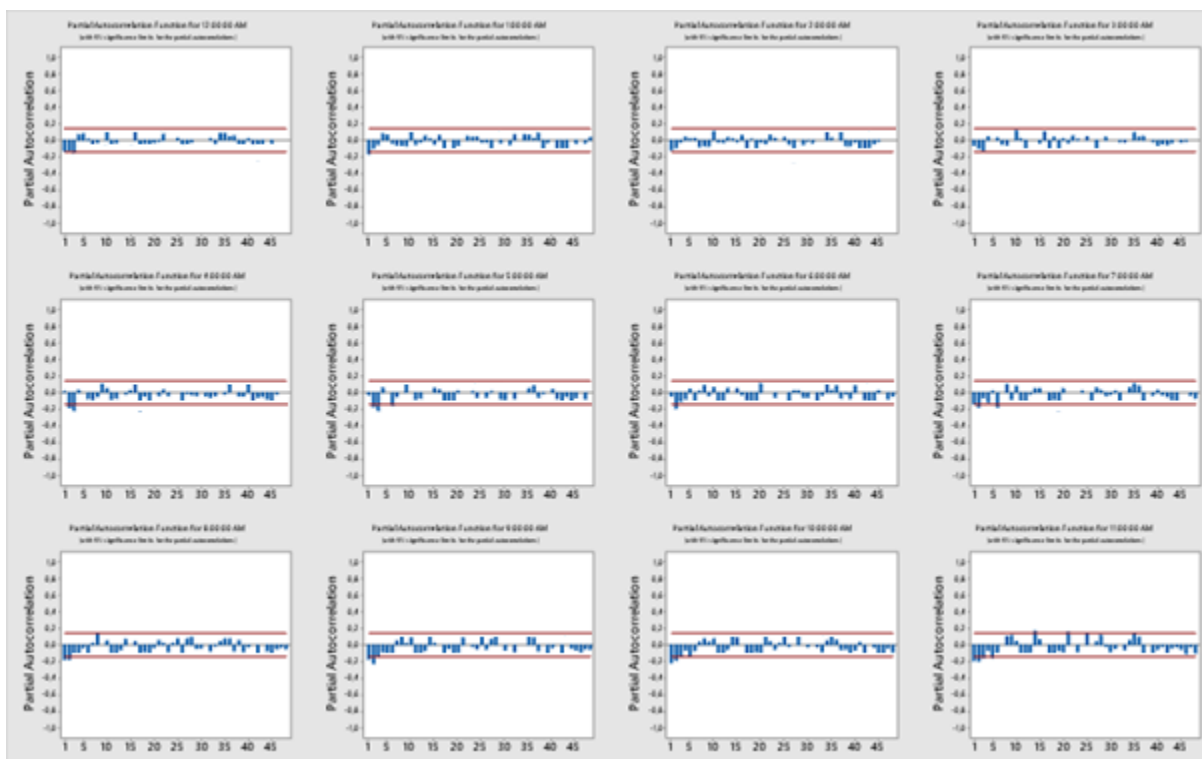


Рисунок 3.8 – Часткова Автокореляція часових рядів 00:00 - 11:00

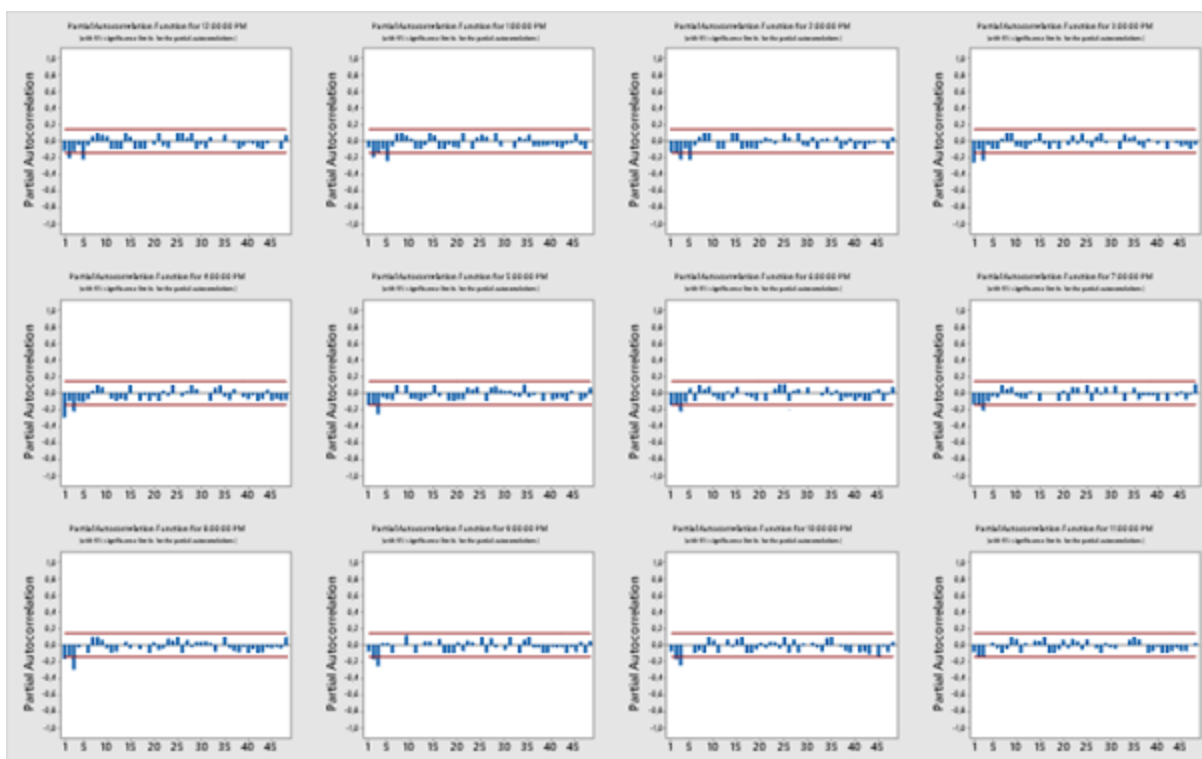


Рисунок 3.9 – Часткова Автокореляція часових рядів 12:00 - 23:00

Порядок моделі AR ( $p$ ) вибирається з ЧАКФ і відповідає останньому ненульовому коефіцієнту ЧАКФ (Рис. 3.8-3.9). Відповідна процедура застосовується для знаходження порядку моделі MA ( $q$ ), при цьому використовується АКФ (Рис. 3.6-3.7).

Візуальний аналіз ACF/PACF дає змогу зробити висновок, чи можна вважати часовий ряд чистим AR- або MA-процесом, чи він є змішаним ARMA процесом. Зважаючи на поведінку автокореляційної та часткової автокореляційної функції, які нескінченно зменшуються до нуля, можна зробити висновок про змішаний характер процесу кожного окремого ряду[22]. Для кожного часового ряду виділим значущі лаги ACF/PACF (табл.3.1 ).

Таблиця 3.1 – Таблиця значущих лагів для кожного ряду

|          | ACF (lag) | PACF(lag) |          | ACF (lag) | PACF(lag) |
|----------|-----------|-----------|----------|-----------|-----------|
| 00:00:00 | 1         | 1,2       | 12:00:00 | 2         | 2         |
| 01:00:00 | 1         | 1         | 13:00:00 | 2         | 2         |
| 02:00:00 | 1         | 1         | 14:00:00 | 3         | 3         |
| 03:00:00 | 2,3       | 3         | 15:00:00 | 1,3       | 1,3       |
| 04:00:00 | 1         | 2,3       | 16:00:00 | 1,3       | 1,3       |
| 05:00:00 | 2,3       | 2,3       | 17:00:00 | 3         | 3         |
| 06:00:00 | 2         | 2         | 18:00:00 | 3         | 3         |
| 07:00:00 | 2         | 1,2       | 19:00:00 | 3         | 2,3       |
| 08:00:00 | 1,2       | 1,2       | 20:00:00 | 1,3       | 1,3       |
| 09:00:00 | 1,2       | 1,2       | 21:00:00 | 2,3       | 2,3       |
| 10:00:00 | 1         | 1,2       | 22:00:00 | 2,3       | 2,3       |
| 11:00:00 | 1         | 2,3       | 23:00:00 | 2         | 2,3       |

Методом перебору для кожної пари коефіцієнтів будуємо відповідну ARIMA-модель, перевіряємо на адекватність за допомогою Q-статистики Льюнга – Бокса та аналізу отриманого ряду залишків на подібність до "білого" шуму. При нульовій гіпотезі про відсутність автокореляції статистика Q має  $\chi^2$  -розподіл. Якщо  $Q > \chi^2$ , то група перших коефіцієнтів автокореляції значущі, так як не всі коефіцієнти рівні нулю.

Для кожної пари коефіцієнтів було перевірено адекватність моделі та подібність отриманого ряду залишків на подібність до "білого" шуму. В Додатку А наведені прогнози відповідних моделей.

Для оцінки точності прогнозу і визначення найоптимальнішої моделі порівнюємо прогнозовані значення використовуючи середню абсолютну процентну помилку MAPE (Додаток В) із реальними та визначимо траєкторію руху прогнозованої кривої(рис. 3.10).

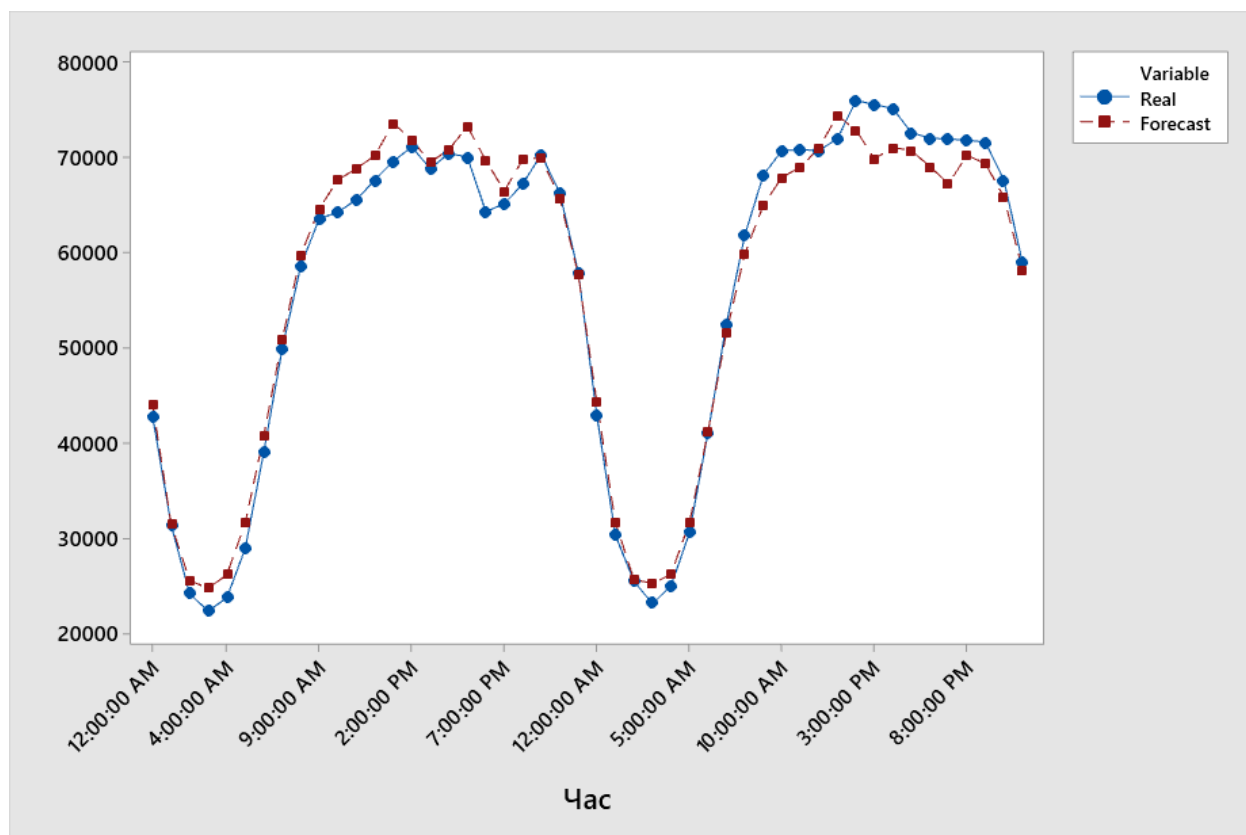


Рисунок 3.10 – Траєкторія прогнозованої кривої за дві доби

*Метод Хольта або двохпараметричний метод згладжування.* Подальші дії з цим часовим рядом вимагають певних маніпуляцій, які допоможуть виявити та усунути аномалії, не характерні для часових рядів, що спотворюють загальну картину. Одним з таких методів очищення ряду від викидів є метод ковзного середнього. Алгоритм методу ковзного середнього виконуються на основі з усереднення існуючих даних протягом певного періоду часу, який є так званим вікном згладжування.

У випадку цієї серії вікно з  $n = 3$  є найбільш оптимальним згладжувальним вікном, оскільки воно мінімально спотворює ряд, одночасно очищаючи викиди (рис. 3.11). Оскільки серія готова до подальшої роботи з прогнозом, ми визначимо наявність тенденції, яка є базовою умовою при застосуванні методу прогнозування Хольта. Тенденцію в цьому часовому ряді

можна виявити без використання відповідних методів - просто поглянувши на сформовану лінію тренду, яка поступово збільшується (рис. 3.12).

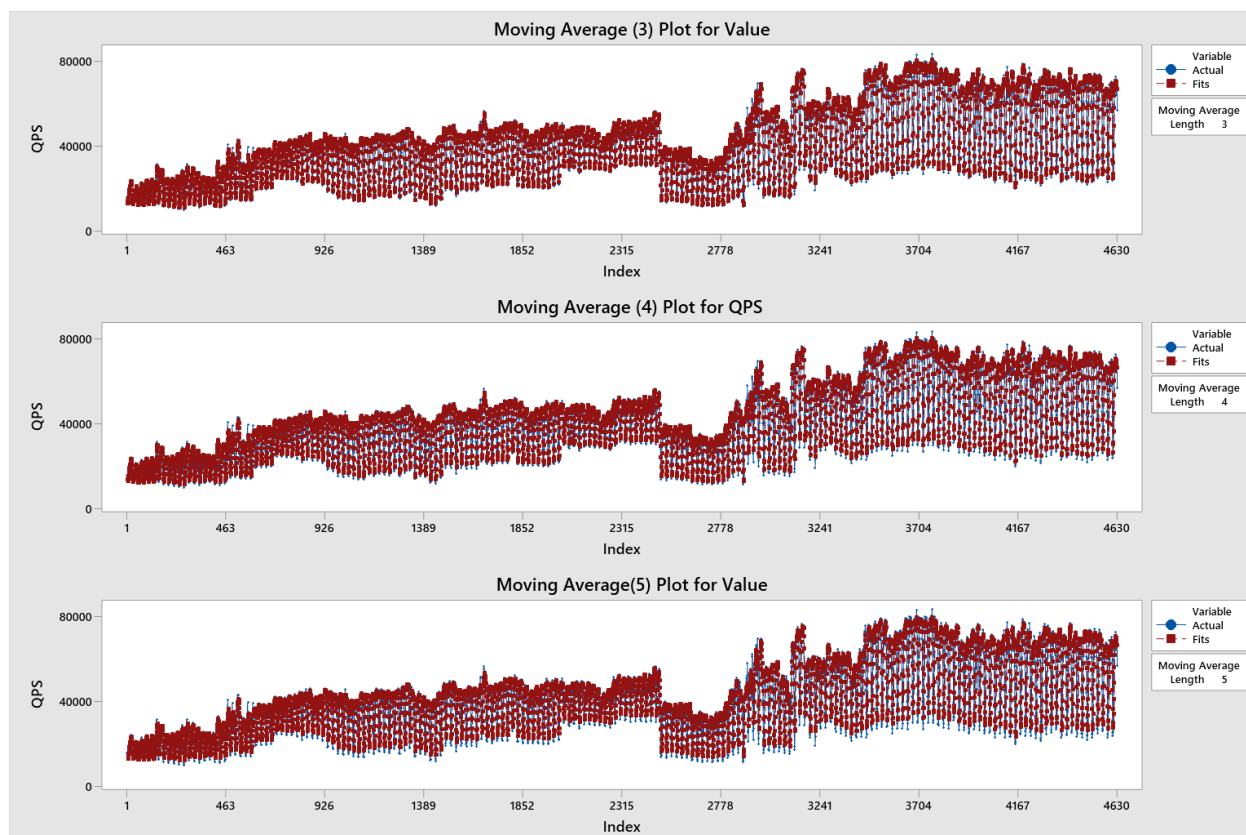


Рисунок 3.11 – Середні ковзні різного вікна

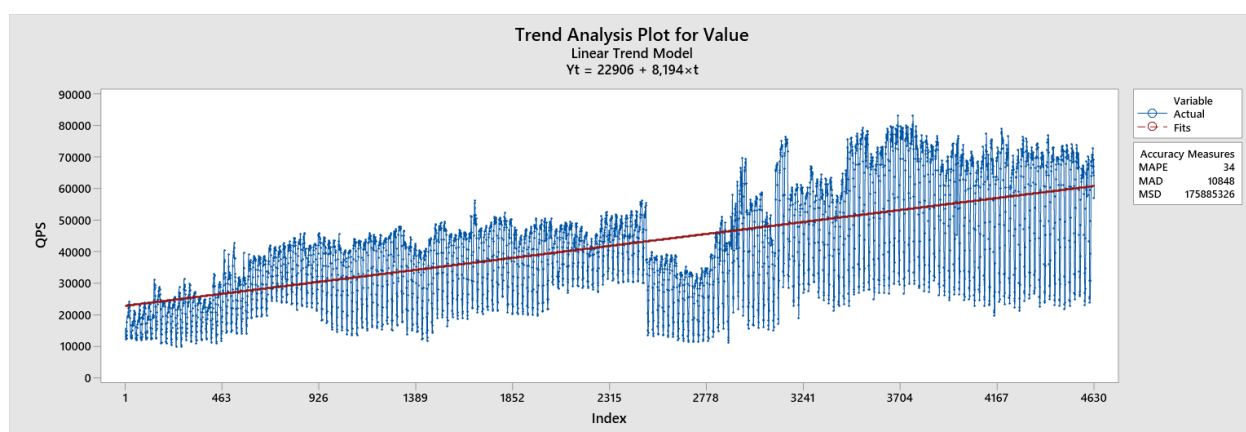


Рисунок 3.12 – Тренд ряду

Підтвердивши, що тенденція присутня в середньому, переходимо до реалізації методу Хольта. Метод Хольта використовується для прогнозування часових рядів, коли існує тенденція до зростання чи спадання часових рядів. А також для рядків, де дані не для повного циклу, а сезонність ще не виділена (наприклад, для неповного року для щомісячного прогнозу).

Одним з основних недоліків методу Хольта є прогноз на один або два майбутні періоди [24] . За таких масивів даних важко передбачити значення значень, які б фактично допомогли спланувати кількість серверів, які будуть використовуватися для обробки запитів. У цьому випадку можна збільшити прогнозований період діленням часових рядів. Повний часовий ряд, у нашому випадку, є відображенням стану системи протягом дня з годинним діапазоном. Тобто, щодня протягом кожної години значення QPS усереднюються в одне. Групування часових рядів дозволить прогнозувати майбутні показники за той самий період на наступні кілька днів. Алгоритм згладжування та виявлення присутності тренду буде повторений для кожної групи даних часових рядів (рис. 3.13-3.14):

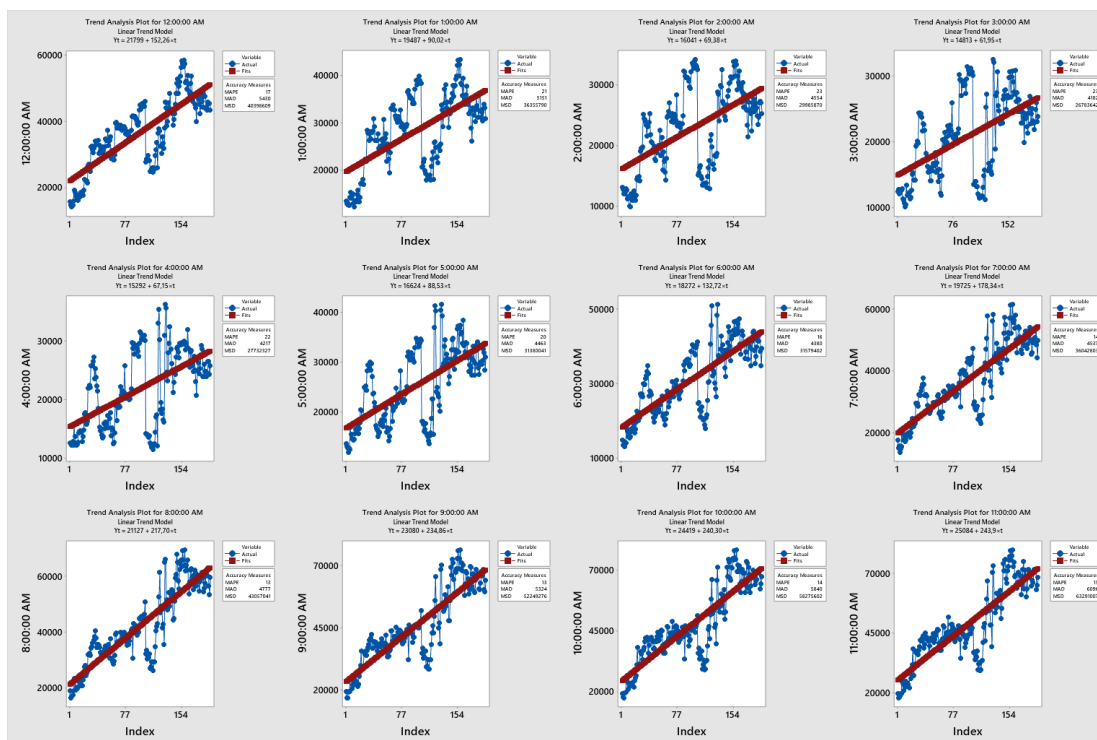


Рисунок 3.13 – Тренд рядів АМ групи

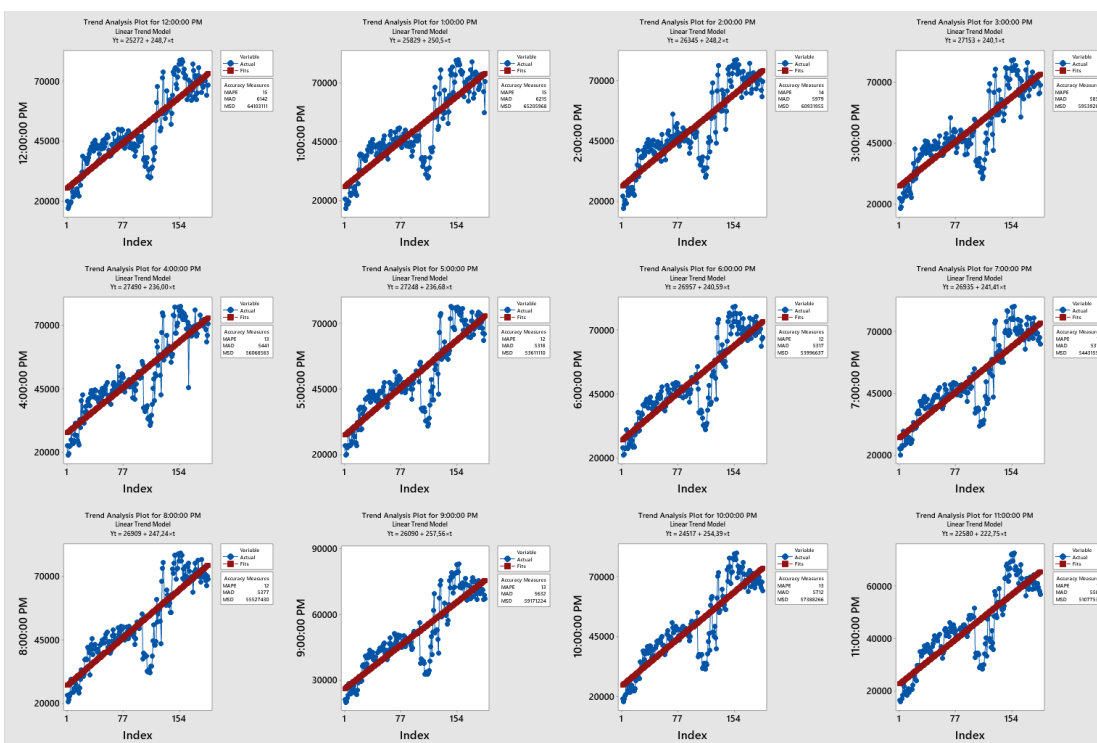


Рисунок 3.14 – Тренд рядів РМ групи

Відповідно до груп даних прогнози на наступні два значення вперед побудовані із застосуванням математичного пакету для аналізу MiniTab (рис. 3.15-3.16).

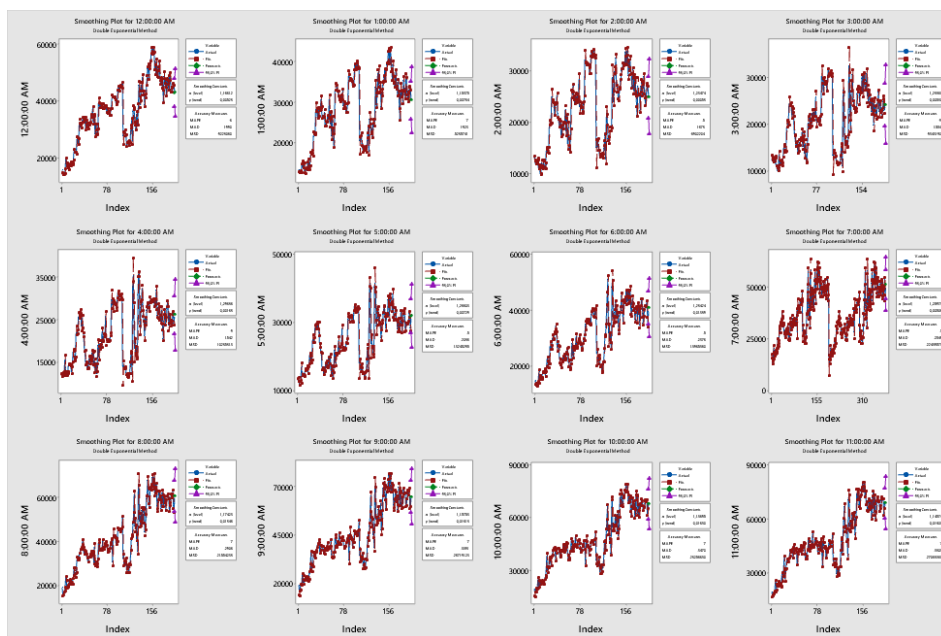


Рисунок 3.15 – Прогноз значень АМ групи

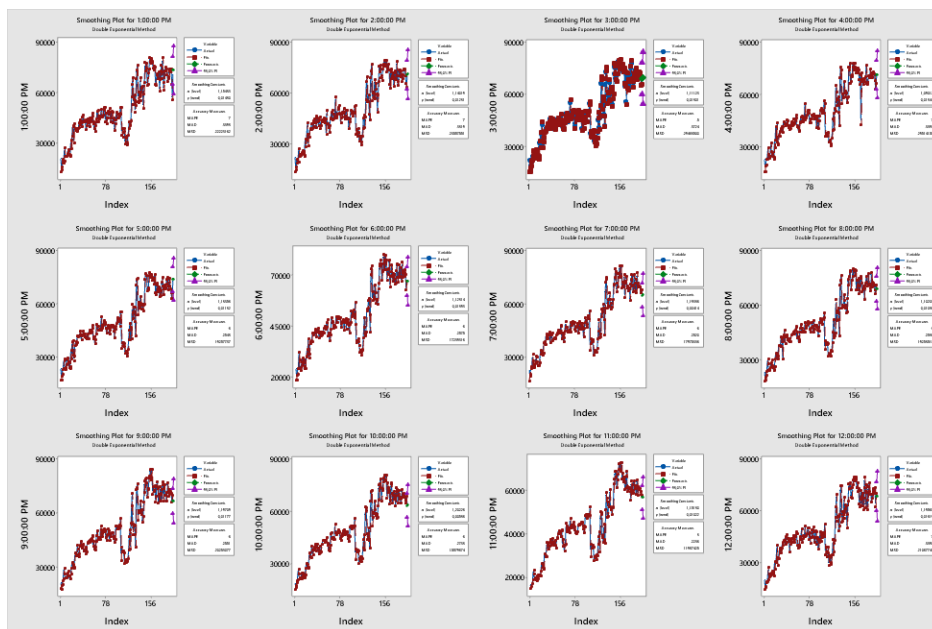


Рисунок 3.16 – Прогноз значень РМ групи

Підсумковий прогноз на наступні два дні представлений на рис. 3.17

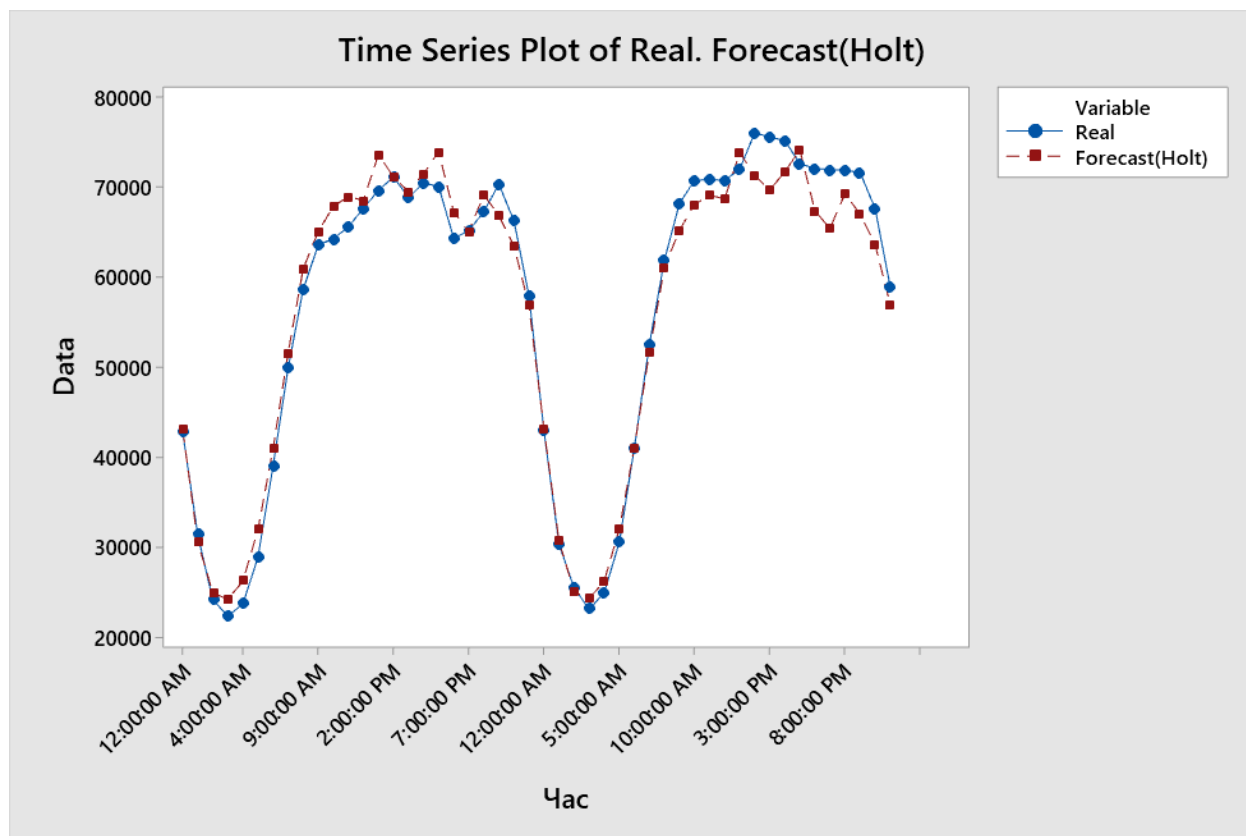


Рисунок 3.17 – Прогноз Хольт

*Модель Хольта-Вінтерса або потрібне експоненціальне згладжування.*

Метод Хольта-Вінтерса в першу чергу передбачає умову наявності сезонної складової. Часовий ряд повинен відображати реальний процес роботи, що зв'язана із сезонними коливаннями: збільшення витрат на комунальні послуги із зміною пори року, добове навантаження на електромережу, попит на літній одяг тощо. Навантаження на ІС має сезонну складову, адже її економічний зміст несе в собі попит на рекламу в певний період доби. Оскільки часовий ряд має «ступінчастий» тренд, тобто відхилення коливання значень відносно певного рівня відбувається із змінною місяця. Тобто для побудови більш точного прогнозу, базовим часовим рядом є дані за останній місяць(рис. 3.18).

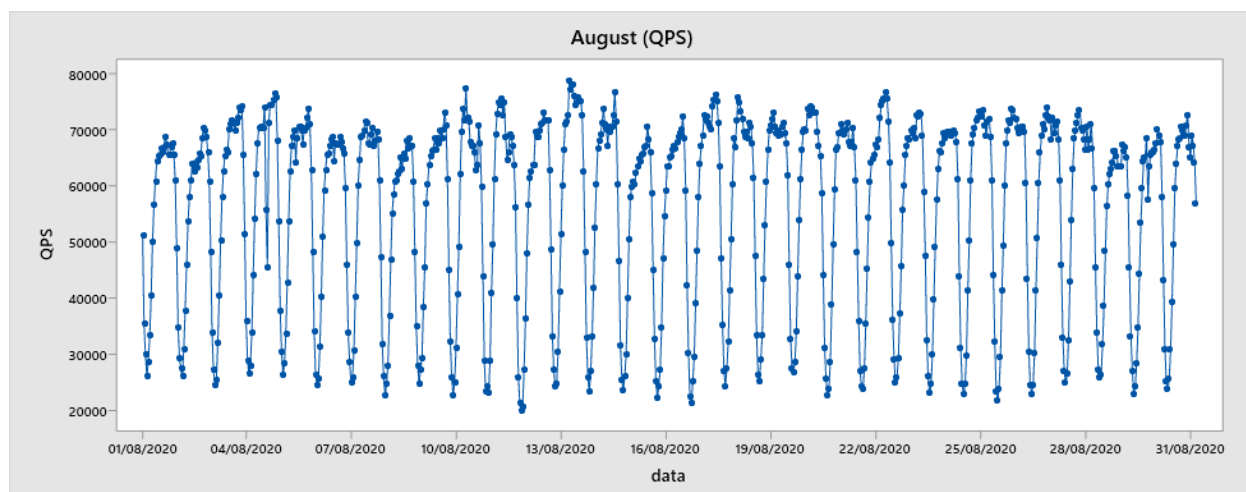


Рисунок 3.18 – Навантаження в серпні місяці

Визначити сезонність та її період можливо за рахунок автокореляційного аналізу(рис. 3.19)

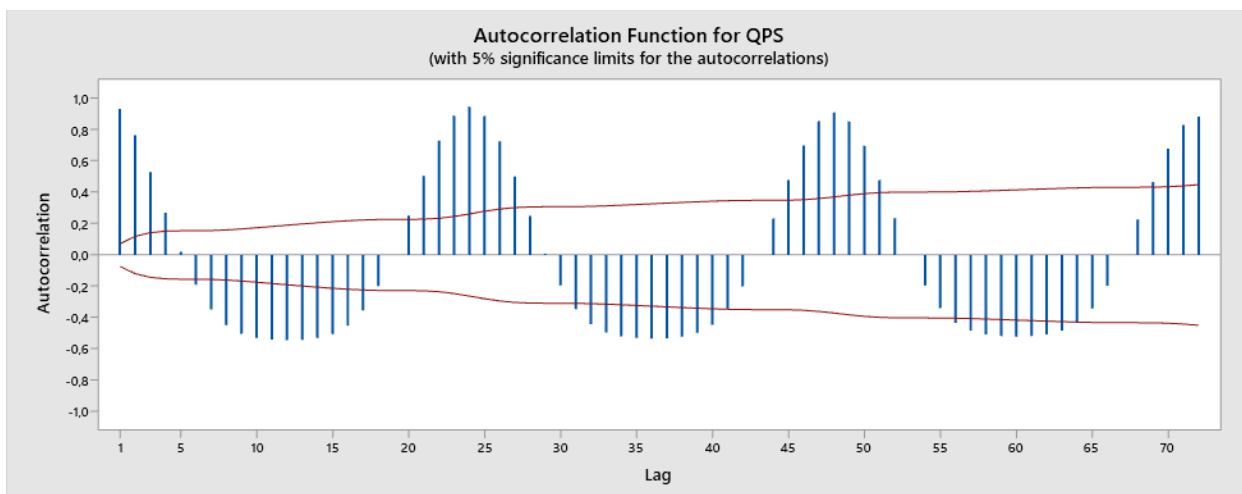


Рисунок 3.19 – Автокореляція ряду

Як можна замітити із графіка значення автокореляцій між лагами 1 та 24 повторюють глобальний цикл, що і є періодом циклічності часового ряду. Для знаходження оптимальних значень компонентів моделі Хольта-Вінтерса використано пакет «Пошук рішення», де цільовою функцією задано:

$$\sum_{k=0}^n (p_k - \hat{p}_k)^2 \rightarrow \min, \quad (3.1)$$

з дотриманням обмежень (3.2):

$$\begin{cases} 0 \leq \alpha \leq 1 \\ 0 \leq \beta \leq 1, \\ 0 \leq \gamma \leq 1 \end{cases} \quad (3.2)$$

де  $\alpha$  – коефіцієнт згладжування ряду;

$\beta$  – коефіцієнт згладжування тренду;

$\gamma$  – константа згладжування циклів с періодом.

$p_k$  – реальні значення;

$\hat{p}_k$  – прогноз.

Результат розрахованого прогнозу відображений на рис. 3.20:

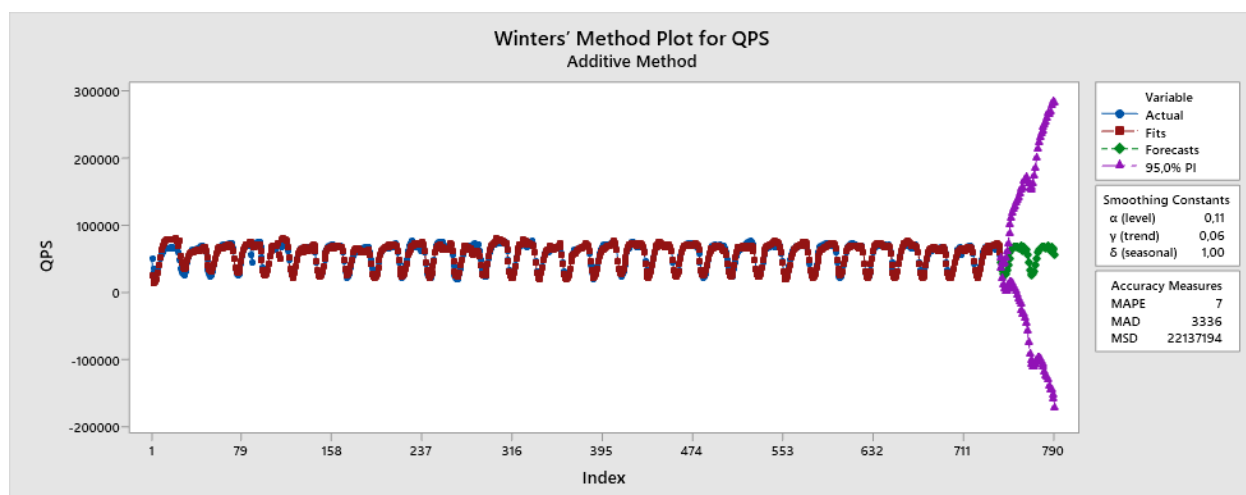


Рисунок 3.20 – Прогноз Хольта-Вінтерса

Як бачимо оптимальні значення компонентів логічно вибудовують структуру часового ряду: сезонна компонента  $\gamma=1$ , що вказує на повне врахування сезонності часового ряду. Даний часовий ряд стаціонарний (значення коливаються в околі одного рівня), що говорить нам про відсутність тренду: тому і значення компоненти відповідне  $\beta=0,06$ . Згладжування ряду із збільшенням компоненти  $\alpha$  призводить до росту різниці між прогнозом та

реальними даними у екстремумах циклу. Дане значення  $\alpha=0,11$  максимально приближує траєкторії часових рядів (рис. 3.21)

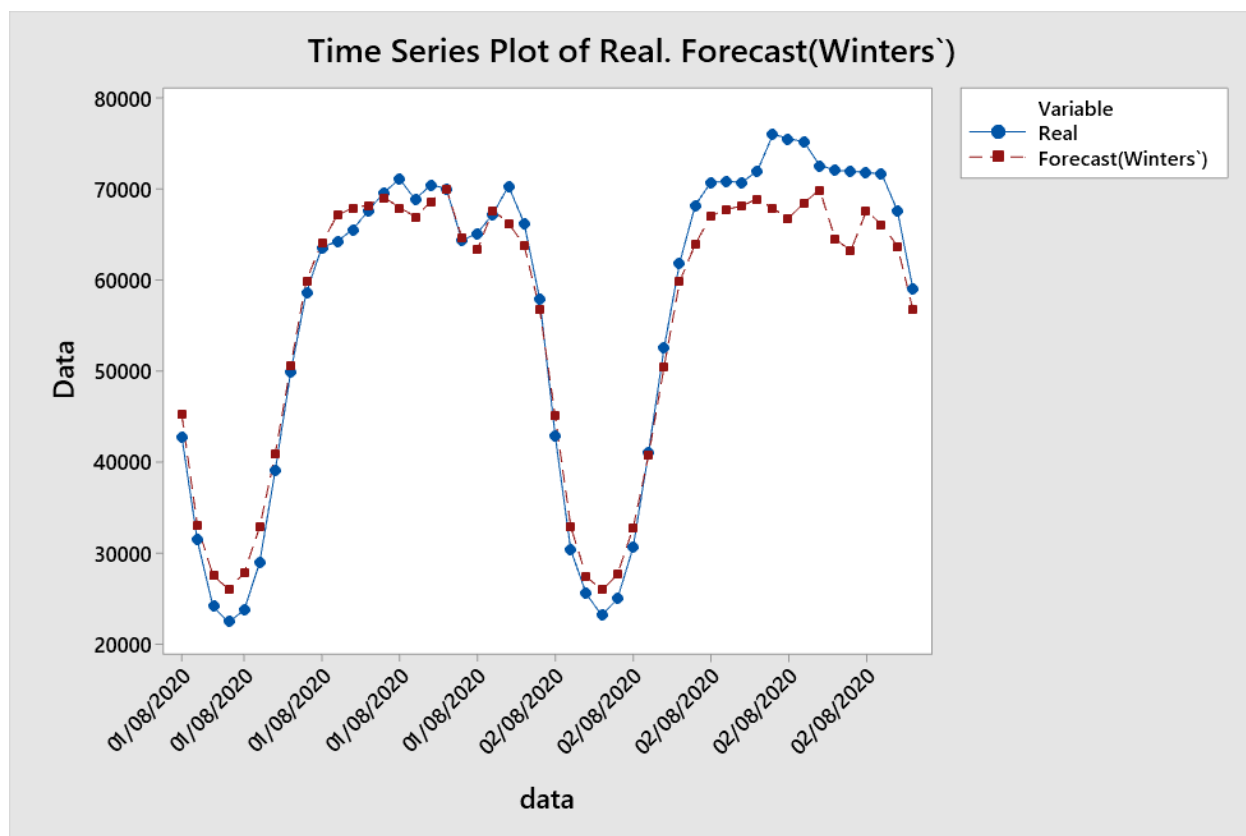


Рисунок 3.21 – Прогноз Хольта-Вінтерса vs Реальні дані

*Порівняння прогнозів.* Після проведення аналізу часового ряду та застосування різних методів прогнозу для наступного етапу дослідження потрібно виділити найкращий варіант на основі якого будуть базуватися інтенсивності надходження заявок для імітації роботи ІС. Для кожного із трьох методів визначимо MAPE (1) та порівняємо отримані значення (табл. 3.2):

Таблиця 3.2 – MAPE для різних методів

|      | ARIMA модель | Метод Хольта | Метод Вінтерса |
|------|--------------|--------------|----------------|
| MAPE | 3,63%        | 3,9 %        | 3,43 %         |

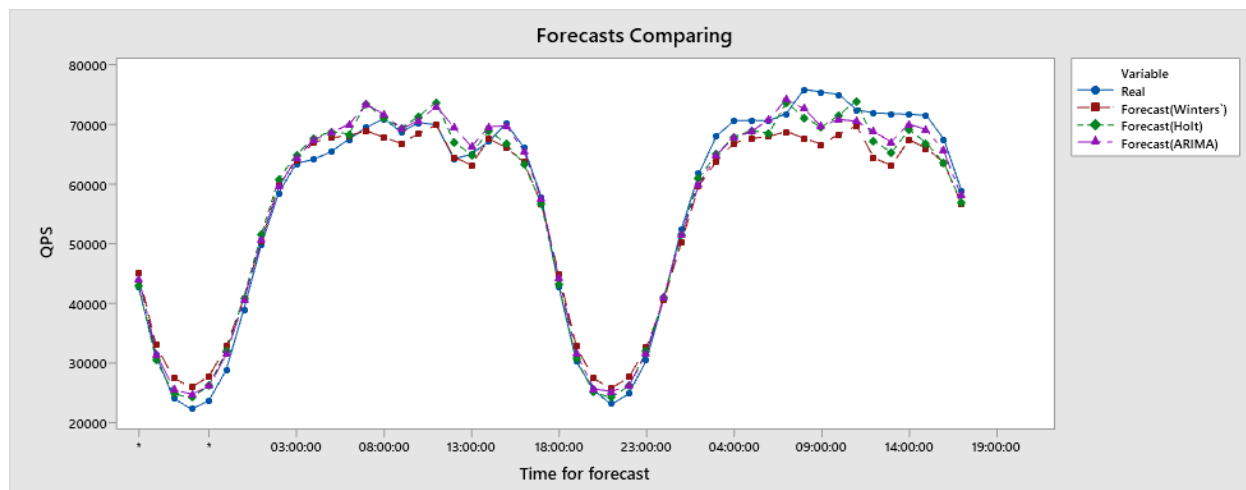


Рисунок 3.22 – Порівняння прогнозів

Результати прогнозу кожного методу є доволі високими, оскільки процентна помилка не перевищує 10% табл. 2.1. Після порівняння прогнозів, найменша абсолютна процентна помилка належить методу Хольта-Вінтерса, що враховує сезонність часового ряду. Метод Хольта-Вінтерса є найбільш вдалим для прогнозування такого показника ще по низці причин:

- Врахування тренду ряду.
- Потреби в статистичних даних за невеликий часовий проміжок.
- Легкості в програмній реалізації.

В порівнянні із цим методом, метод Хольта та ARIMA-модель для такого виду економічних задач мають менші переваги, оскільки, вони є ресурсно затратними в реалізації, потребують великого об'єму даних та мають більшу помилку прогнозу. Тому прогноз для імітації роботи ІС буде базуватися на методі Хольта-Вінтерса.

### 3.1.2 Реалізація імітаційної моделі ІС

Імітаційне моделювання широко використовується при проектуванні бізнес-процесів та систем. Найбільш широко розповсюджені системи моделювання Arena [34], AnyLogic [35] та Matlab дозволяють вводити структури модельованих систем та процесів у графічному вигляді.

Matlab є пакетом прикладних програм (ППП) для рішень обчислювальних задач, у яких використовується мова програмування Matlab [36]. У склад ППП Matlab входить система структурного та імітаційного моделювання (CCIM) Simulink [37]. Моделі в CCIM Simulink складаються з окремих блоків, які можуть бути об'єднані в підсистеми. Підсистеми дозволяють представляти складні моделі в ієрархічному вигляді, де на кожному рівні декомпозиції визначається склад функціональних підсистем і зв'язок між ними.

Simulink реалізує ідею аналогових обчислювальних машин. Модельована система представляє сукупність функціональних блоків, за зв'язком між якими передаються аналогові сигнали. На відміну від Arena та AnyLogic, в структурних моделях Matlab Simulink окрім самих об'єктів-заявок між блоками можуть передаватися і інші сигнали, що дає можливість представити всю модель цілком в графічному вигляді.

Для побудови різних моделей в CCIM Simulink включено велику кількість бібліотек, що дозволяють вирішувати різні завдання, включаючи математичні обчислення і перетворення, побудова нейронних мереж, кореляційний аналіз, імітаційне моделювання систем масового обслуговування (СМО), моделювання механічних взаємодій об'єктів, електричних схем, гідравліки.

У Simulink є безліч бібліотек, що містять графічні блоки для моделювання системи однією із яких є Simulink-StateEvents. Її компоненти дозволяють проектувати моделі і оптимізувати їх параметри. Основними компонентами бібліотеки є блоки: генерації заявок, черг, серверів, маршрутизації руху заявок і виведення заявок з системи. Опис компонент SimEvent приведено в [38]. З їх допомогою легко побудувати структурну модель СМО для планування часу завантаження ресурсів в розподілених системах управління, програмних і апаратних інфраструктур [39].

Як описано в розділі 2.2 система масового обслуговування за нотацією Кендалла має вигляд : M/M/5/FIFO. Для кожного із елементів бібліотека Simulink має відповідний модуль, що відтворює певний динамічний процес роботи СМО - генерація заявок - черга – модуль обробки запитів. Час моделювання навантаження на ІС рівний часу спостереження між двома значеннями прогнозу отриманого на попередньому етапі роботи, тобто 1 годині. Інтенсивність надходження заявок представлена піковим навантаженням даного періоду доби для якого відбувається симуляція. Генерація заявок реалізована за рахунок модуля «*Time-Based Entity Generator*», де формування заявок є пуасонівським потоком із експоненціальним розподілом (рис. 3.23).

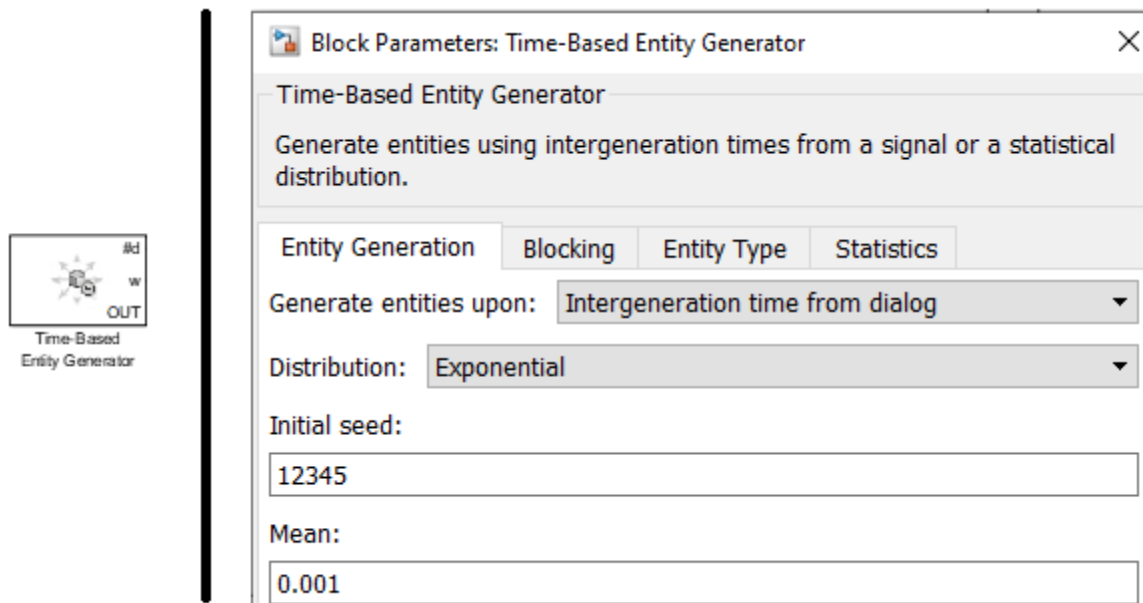


Рисунок 3.23 – Модуль генерації заявок

Параметр “Mean” є середнім часом між формуванням заявок за секунду, тобто є інтенсивністю, але представленою часом. Із формули (2.29) ми можемо задати інтенсивність через цей середній час, розділивши секунду на пік навантаження. Для кожної години вона подана наступним чином (табл. 3.3):

Таблиця 3.3 – Інтенсивності надходжень заявок для кожної симуляції

| $\lambda$ (сек.) $\times 10^4(-4)$ | 00:00:00 | 01:00:00 | 02:00:00 | 03:00:00 | 04:00:00 | 05:00:00 | 06:00:00 | 07:00:00 | 08:00:00 | 09:00:00 | 10:00:00 |
|------------------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 День:                            | 0,221469 | 0,302871 | 0,362874 | 0,384241 | 0,359783 | 0,304772 | 0,245015 | 0,197848 | 0,166929 | 0,156226 | 0,149026 |
| 2 День:                            | 0,222096 | 0,304044 | 0,364559 | 0,386132 | 0,36144  | 0,30596  | 0,245782 | 0,198348 | 0,167285 | 0,156537 | 0,149309 |
| $\lambda$ (сек.) $\times 10^4(-4)$ | 13:00:00 | 14:00:00 | 15:00:00 | 16:00:00 | 17:00:00 | 18:00:00 | 19:00:00 | 20:00:00 | 21:00:00 | 22:00:00 | 23:00:00 |
| 1 День:                            | 0,144881 | 0,147231 | 0,149491 | 0,14587  | 0,142843 | 0,154788 | 0,15796  | 0,147847 | 0,151067 | 0,156771 | 0,176039 |
| 2 День:                            | 0,145149 | 0,147507 | 0,149776 | 0,146141 | 0,143104 | 0,155094 | 0,158278 | 0,148126 | 0,151359 | 0,157085 | 0,176435 |

Після формування заявки прямує в чергу при цій проходить через модуль «*Time-Out Schedule*», що фіксує для кожної заявки час в 200мс для можливості переходити у відмову у разі недотримання часових періодів (рис. 3.24).

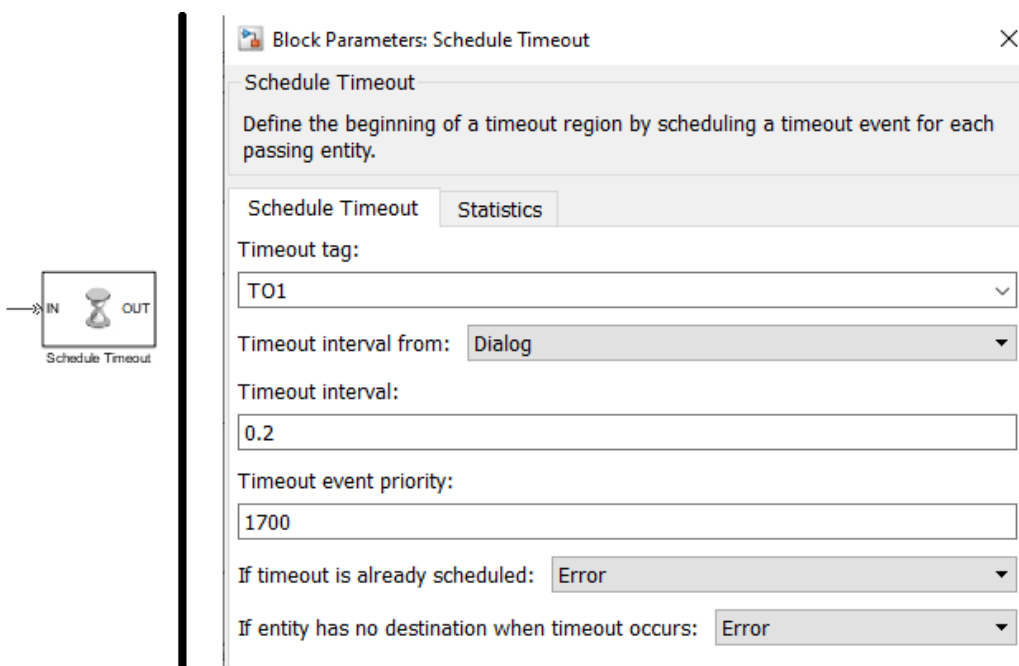


Рисунок 3.24 – Модуль Таймінгу

Черга, представлена модулем «*FIFO*» (рис. 3.25), являє собою дисципліну обслуговування, коли перша заявка, що надійшла, буде оброблена також першою.

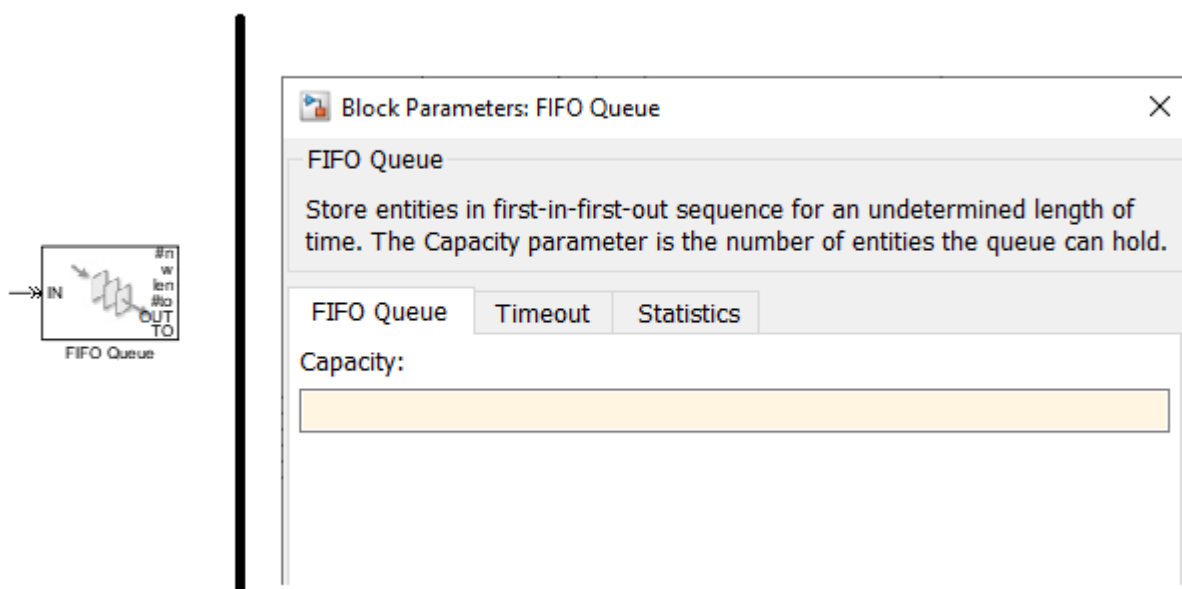


Рисунок 3.25 – Модуль формування черги

Відповідно специфіки роботи ІС довжина черги не є лімітованою, оскільки основним лімітом на перебуванні в черзі є обмеження в часі по протоколу ІАВ OpenRTB. Після виходу із черги заявка потрапляє в розподілювач «*Output Switch*» (рис. 3.26).

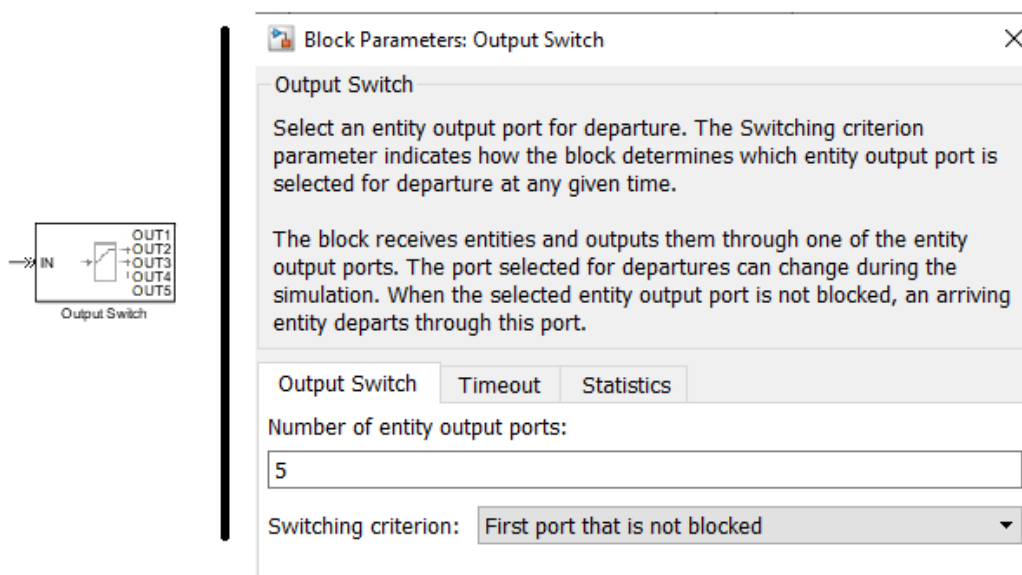


Рисунок 3.26 – Модуль розподілювача

Даний модуль розподіляє заявки між 5-ма обслуговуючими модулями. Якщо перший обслуговуючий модуль повністю завантажений, розподілювач надсилає заявку на другий і по такому ж алгоритму перевіряє навантаження на нім. У випадку завантаженості усіх модулів заявка формує чергу в модулі «*FIFO*». Система формування заявок та формування черги представлена в додатку С. Наступною подією після розподілювача для заявки є оброблювальний модуль «*N-Server*» (рис. 3.27).

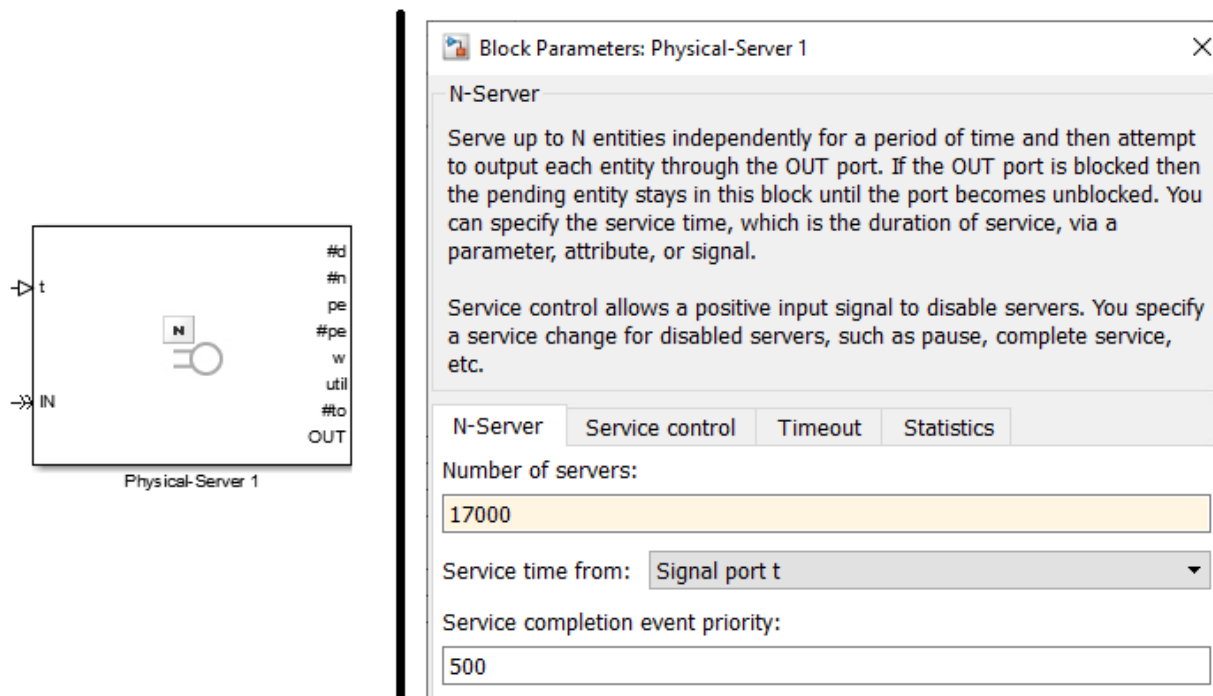


Рисунок 3.27 – Модуль обробки запитів

Даний модуль являє собою скомпоновану модель декількох реальних обчислювальних пристроїв, а саме: 1 фізичний веб-сервер, 2 фізичних сервери База Даних, 3 сервери ядра обробки (Додаток D). Такий модуль на практиці обслуговує 17 тисяч заявок в секунду за  $\bar{\tau}_{об} \sim 100 - 200$  мс. враховуючи технічні особливості отриманого запиту. Для моделювання стохастичного часу обробки запиту в заданих обмеженнях  $100 \text{ мс} \leq \bar{\tau}_{об} \leq 200 \text{ мс}$  застосовано модуль «*Event-Based Random Number*» (рис. 3.28).

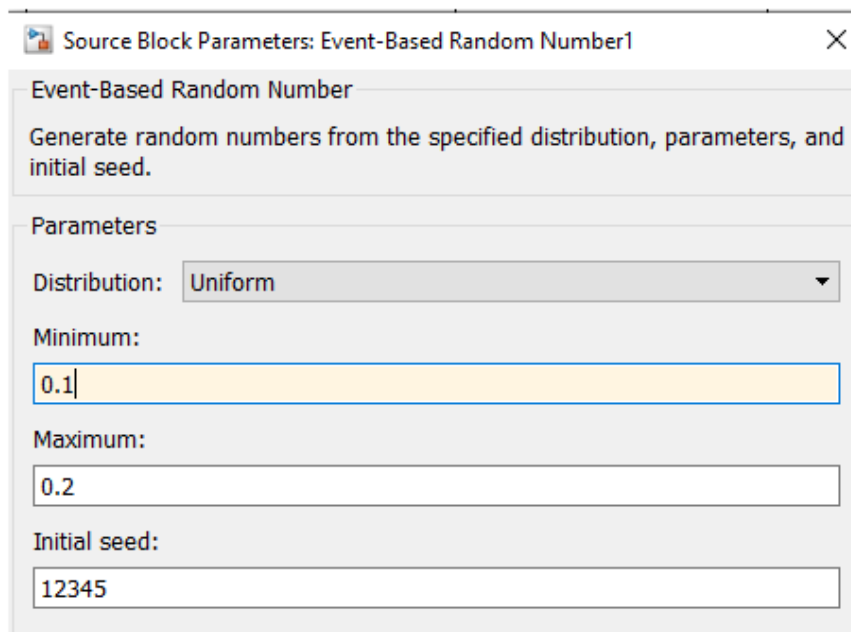


Рисунок 3.28 – Модуль часу обробки

Даний модуль формує для кожної заявки експоненціально розподілений час обробки в межах 100 мс та 200 мс та передає дане значення на обробляючий модуль. Модуль обробки представлений в додатку В. Загальна сформована імітаційна модель системи масового обслуговування в ППП Matlab бібліотеки Simulink з використанням модулів SimEvents представлена в додатку Е.

Для кожної із 24-х пікових точок для кожної з двох діб було проведено симуляцію її роботи. Для періоду з 00:00 AM до 01:00 AM представлені основні показники роботи моделі.

Надходження заявок для даного періоду  $\bar{t}_{\text{над}} = 0,221 \times 10^{-4}$  формує потік запитів за годину моделювання (рис. 3.29):

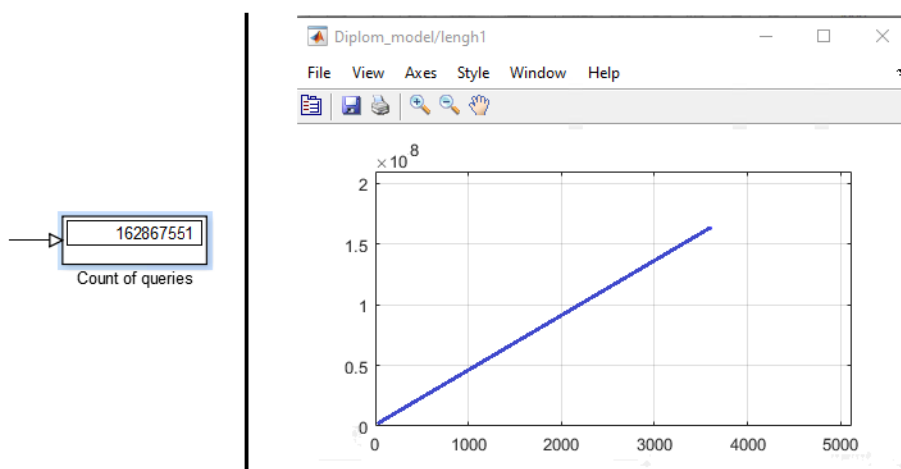


Рисунок 3.29 – Графік приросту кількості запитів на систему

Згенерована кількість заявок на систему  $1,623 \times 10^8$ , що відображає реальну поведінку навантаження системи в даний період, оскільки об'єм запитів коливається в межах від  $1,60 \times 10^8$  до  $1,62 \times 10^8$ . Середній час між надходженнями заявок (графік представлений на останньому мільйоні запитів, через технічні обмеження обчислювального пристрою на якому проводились симуляції) (рис. 3.30):

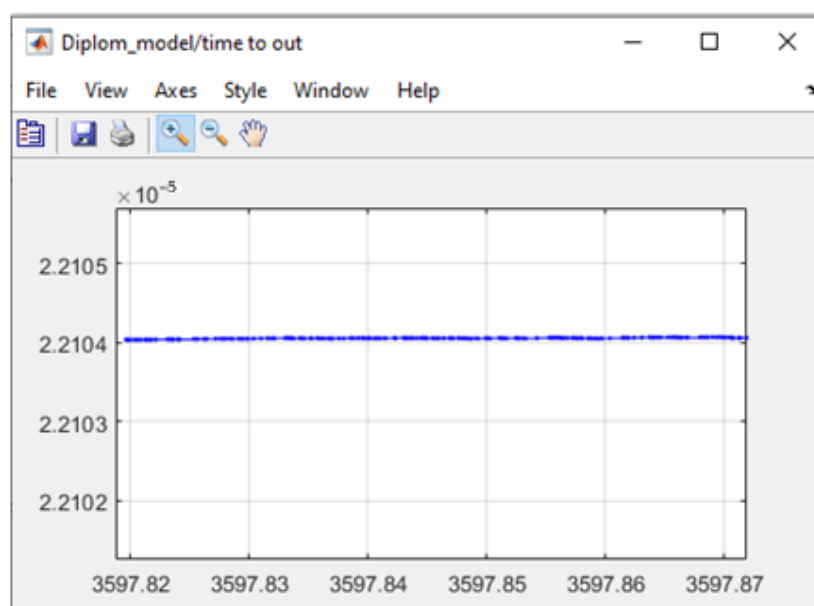


Рисунок 3.30 – Час надходження запитів

Для обслуговуючого модулю основним показником навантаження є його утилізація, якщо модуль завантажений повністю - значення його утилізації рівне 1 (рис. 3.31). Розподільувач перемикає потік заявок на інші модулі обслуговування, якщо утилізація модуля більша 90%. Для даної симуляція навантаження модулів розподілилось наступним чином (рис. 3.31):

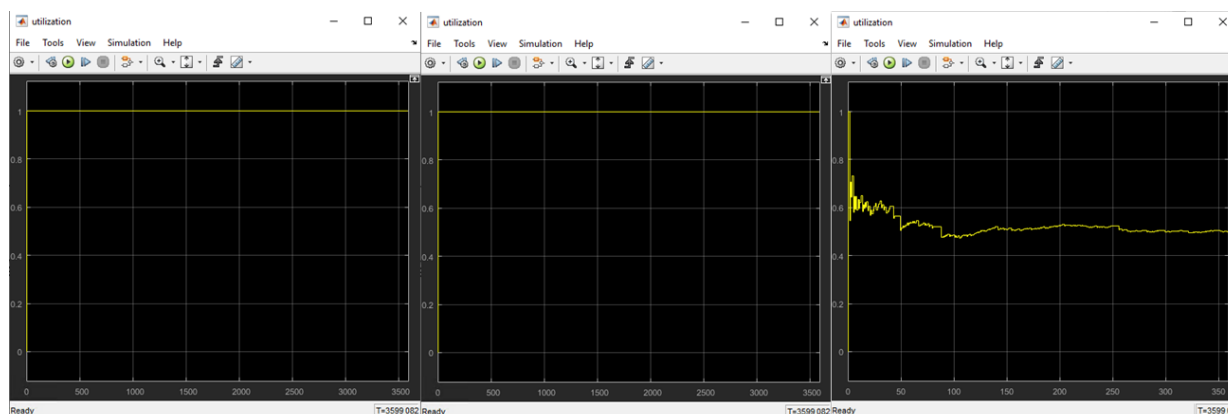


Рисунок 3.31 – Утилізація I,II,III обслуговуючого модуля

Для наступних двох модулів обслуговування утилізація рівна нулю (рис. 3.32), адже для поточного навантаження достатньо 3-х при тому третій модуль завантажений на 60-70%.

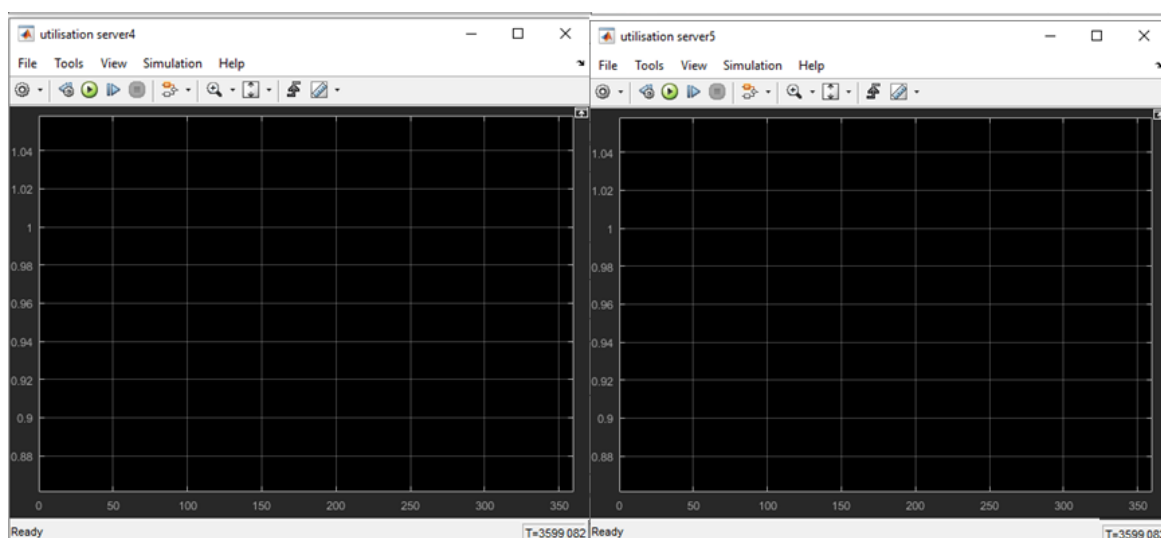


Рисунок 3.32 – Утилізація IV, V обслуговуючого модуля

Час обслуговування кожної заявки коливається в межах 100-200 мс, що відповідає реальній роботі ІС, на рис. 3.33 представлено графік часу обслуговування кожної заявки для першого модуля.

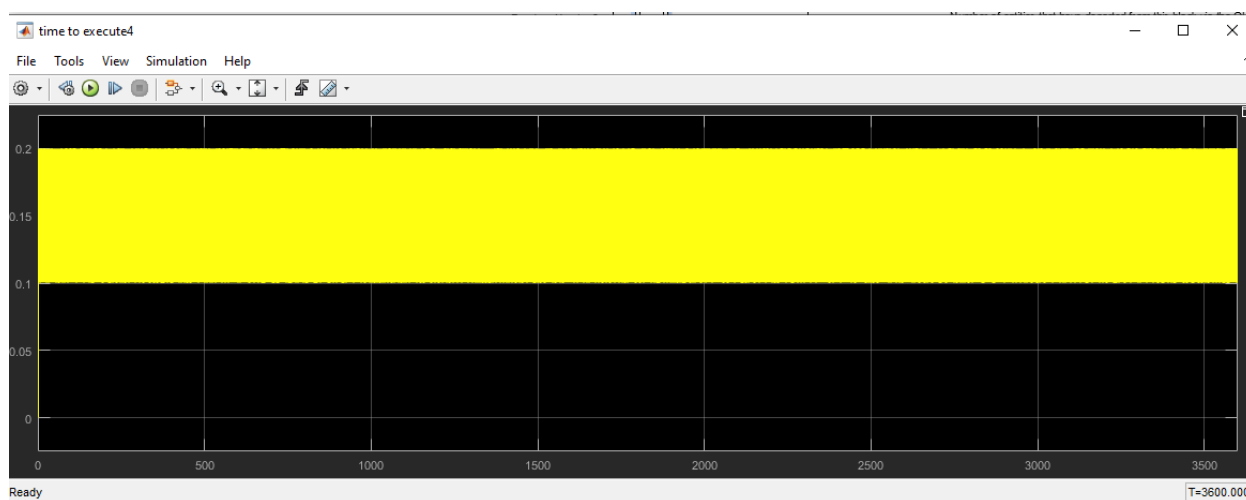


Рисунок 3.33 – Час обслуговування заявок І модуля

На рис. 3.34 представлено кількість одночасно оброблюваних заявок на І модулі.

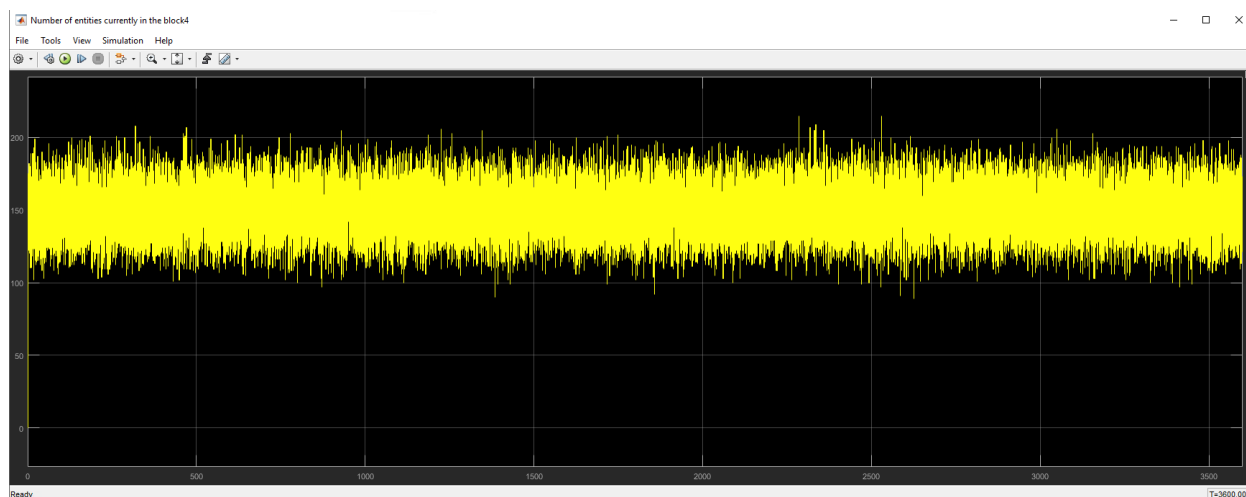


Рисунок 3.34 – Кількість одночасних заявок на І модулі

Кількість оброблених заявок розподілено між модулями наступним чином (табл. 3.4):

Таблиця 3.4 – Кількість заявок оброблених кожним модулем

| I модуль | II модуль | III модуль |
|----------|-----------|------------|
| 61233455 | 61123221  | 40510875   |

Прослідковуючи динаміку обробки заявок можна виділити швидкий і рівний один одному приріст заявок на I та II модулях та плавний на III модулі (рис. 3.35).

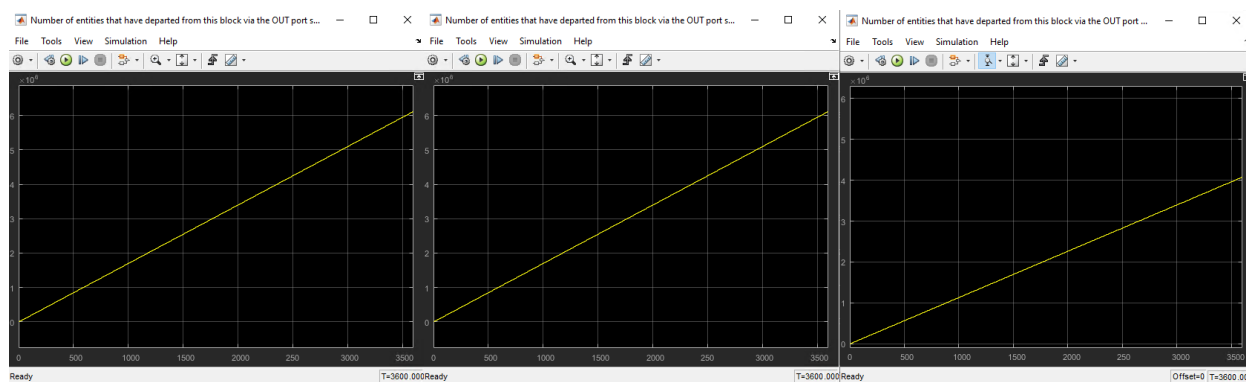


Рисунок 3.35 – Динаміка росту оброблених заявок

Оскільки завантаження модулів не повне, то черга даної симуляції нульова, оскільки кожна заявка, що зайняла чергу одразу потрапляє на обробку (рис. 3.36). Час перебування в черзі пропорційний довжині черзі (рис. 3.36):

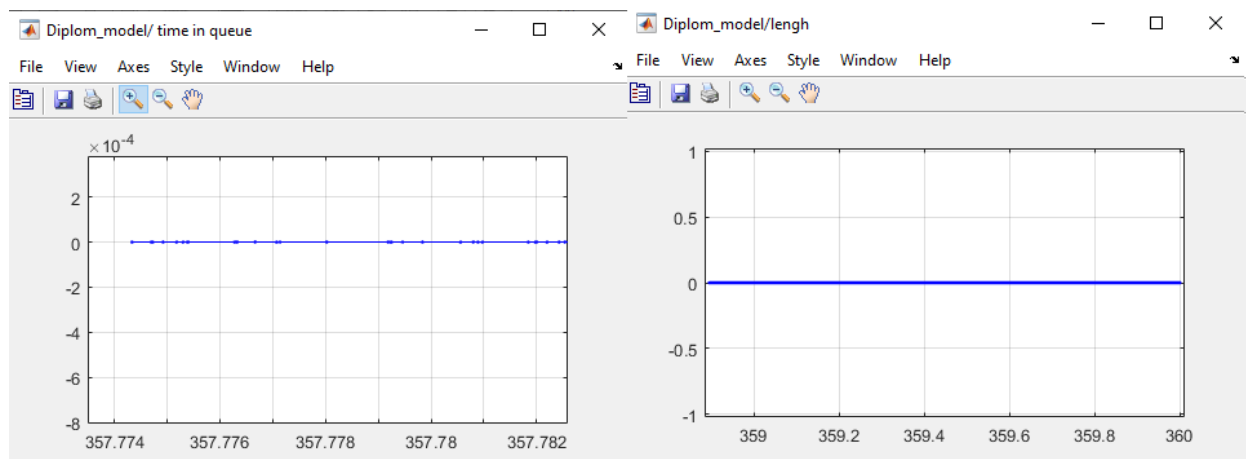


Рисунок 3.36 – Час та довжина черги

Логічним висновком для даної довжини черги та часу перебування заявки в черзі буде відсутність потоку відмов (рис. 3.37 ):

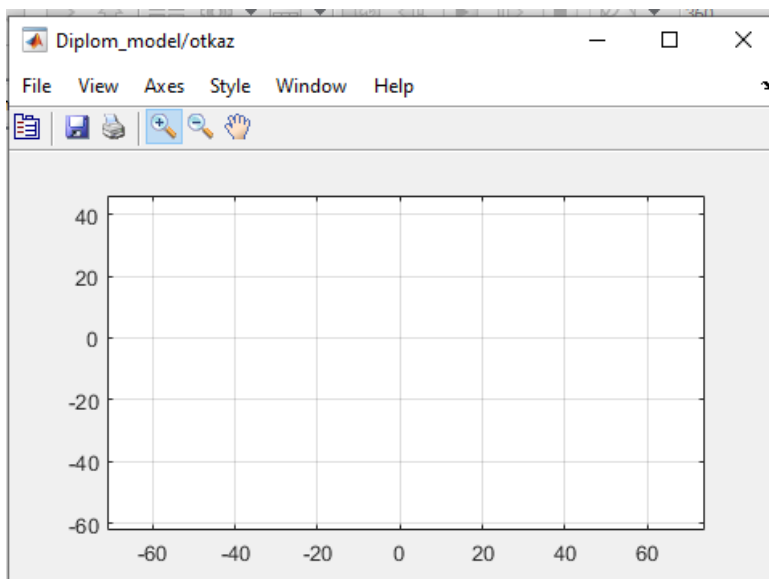


Рисунок 3.37 – Відмови черги

Однак при процесі обробки потік заявок що отримали відмову зростає (рис. 3.38):

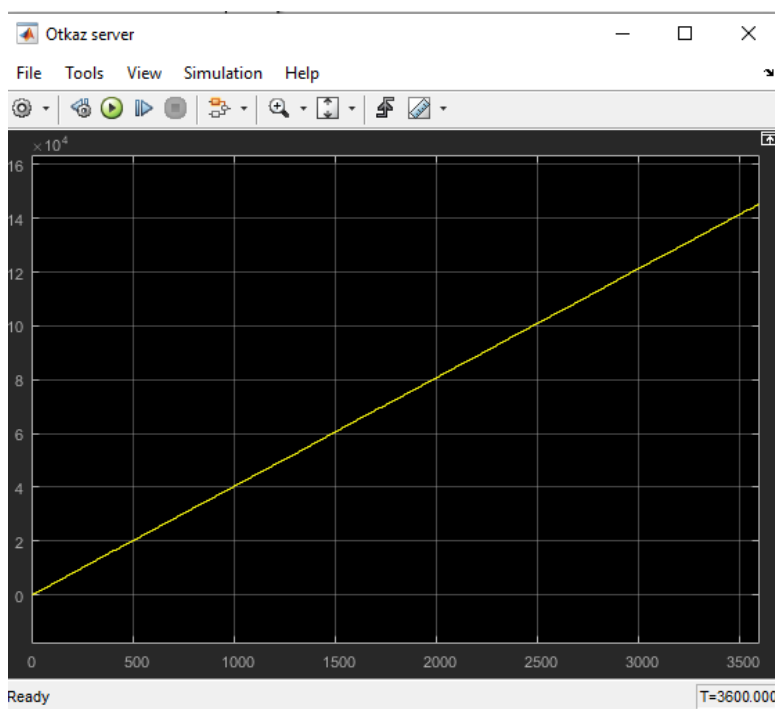


Рисунок 3.38 – Відмови обслуговування

### 3.2 Формування пропозицій щодо економії коштів компанії

Симуляції були здійснені для кожного часового періоду, та прораховані кількості модулів, що були задіяні для даного навантаження. На рис. 3.39 зображено їх відношення.

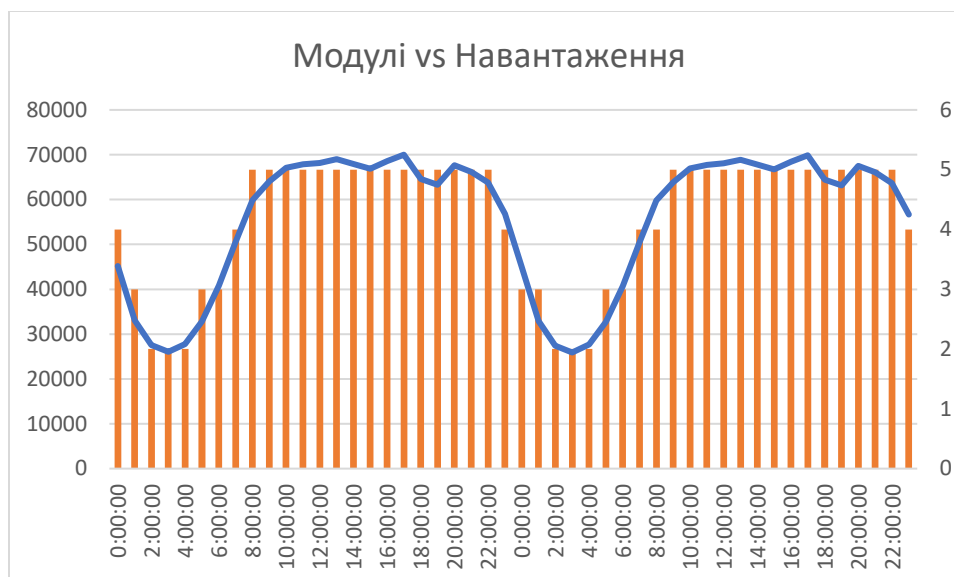


Рисунок 3.39 – Відношення модулів до навантаження

Для даної конфігурації одного модуля використовується 1 фізичний веб-сервер, 2 фізичних сервери Баз Даних, 3 сервери ядра обробки. Оренда обчислювальних пристроїв у дата-провайдерів вартує значних коштів, що створює сильний фінансовий тиск на бюджет компанії. В табл. 3.5 наведені вартості оренди обчислювальних пристроїв модуля.

Таблиця 3.5 – Вартість оренди обчислювальних пристроїв

| К-сть | Сервер(характеристики)                                                                              | Помісячна оренда/ | Денна оренда   | Погодинна оренда |
|-------|-----------------------------------------------------------------------------------------------------|-------------------|----------------|------------------|
| 1     | 2                                                                                                   | 3                 | 4              | 5                |
| 1     | Supermicro SYS-5038MR (10 ядер, ОЗУ 32 ГБ, 3 петабайт інтернет-трафіку )                            | 170 €./ 1 у.о     | 5,50 €./ 1 у.о | 0,24 €./ 1 у.о   |
| 2     | HP ProLiantDL120Gen9 (10 ядер, ОЗУ 128 ГБ, 1 петабайт інтернет-трафіку, 2.4 Тб дискового простору ) | 150 €./ 1 у.о.    | 5 €./ 1 у.о    | 0,20 €./ 1 у.о   |
|       |                                                                                                     |                   |                |                  |

Кінець Таблиці 3.5

| 1 | 2                                                                 | 3              | 4            | 5                   |
|---|-------------------------------------------------------------------|----------------|--------------|---------------------|
| 3 | HGST HDN72 (20 ядер,<br>ОЗУ 64 ГБ, 1 петабайт<br>інтернет-трафіку | 90 €. / 1 у.о. | 3 €. / 1 у.о | 0, 125 €. /<br>1у.о |

Для одного модуля місячна вартість оренди складе:

$$1 * 170 + 2 * 150 + 3 * 90 = 740 \text{ €} . ;$$

Відповідно за 5 модулів компанія сплачує 3700 €. Компанія не оптимізує використання модулів, сплачуючи за їх «холосту» роботу. Якщо спрогнозувавши майбутнє навантаження та оцінивши можливості роботи інформаційної системи із використанням не повного набору модулів можна значно скоротити витрати. Базуючись на двох добовому прогнозі та змодельованій роботі ІС визначимо вартість оренди при можливих відключеннях модулів:

$$\text{Годинна оренди модуля} - 1 * 0,24 + 2 * 0,20 + 3 * 0,125 = 1,01 \text{ €}.$$

$$\text{Дві доби компанії коштуватиме } 5 * 48 * 1,01 = 242,4 \text{ €} . \text{ відповідно;}$$

При вимкненні модулів отримаємо наступні значення (табл. 3.6) :

Таблиця 3.6 – Вартість оренди обчислювальних пристроїв

| 1 Доба   |       |          |       | 2 Доба   |       |          |       |
|----------|-------|----------|-------|----------|-------|----------|-------|
| 00:00:00 | 4     | 12:00:00 | 5     | 00:00:00 | 3     | 12:00:00 | 5     |
| 01:00:00 | 3     | 13:00:00 | 5     | 01:00:00 | 3     | 13:00:00 | 5     |
| 02:00:00 | 2     | 14:00:00 | 5     | 02:00:00 | 2     | 14:00:00 | 5     |
| 03:00:00 | 2     | 15:00:00 | 5     | 03:00:00 | 2     | 15:00:00 | 5     |
| 04:00:00 | 2     | 16:00:00 | 5     | 04:00:00 | 2     | 16:00:00 | 5     |
| 05:00:00 | 3     | 17:00:00 | 5     | 05:00:00 | 3     | 17:00:00 | 5     |
| 06:00:00 | 3     | 18:00:00 | 5     | 06:00:00 | 3     | 18:00:00 | 5     |
| 07:00:00 | 4     | 19:00:00 | 5     | 07:00:00 | 4     | 19:00:00 | 5     |
| 08:00:00 | 5     | 20:00:00 | 5     | 08:00:00 | 4     | 20:00:00 | 5     |
| 09:00:00 | 5     | 21:00:00 | 5     | 09:00:00 | 5     | 21:00:00 | 5     |
| 10:00:00 | 5     | 22:00:00 | 5     | 10:00:00 | 5     | 22:00:00 | 5     |
| 11:00:00 | 5     | 23:00:00 | 4     | 11:00:00 | 5     | 23:00:00 | 4     |
| Сума     | ×1,01 |          | ×1,01 |          | ×1,01 |          | ×1,01 |
|          | 43,43 |          | 59,59 |          | 41,41 |          | 59,59 |

Сума складених витрат на оренду після вимкнень модулів 204,2 €, що на 38,2 € менше від поточних витрат. Не складними обчисленнями можна знайти приблизну суму заощаджень на місяць:  $\sim 3700 - 15 \cdot 204,2 = 637$  €. Така сума складає 84% вартості одного модуля за місяць, що дозволяє компанії розширювати ІС із меншими витратами.

## ВИСНОВКИ

У роботі було проаналізовано особливості функціонування інформаційної системи компанії та об'єми витрат, що формуються на основі підтримки та моніторингу життєдіяльності системи. Обґрунтовано вибір прогностичних моделей для прогнозу навантаження. Відповідно теорії масового обслуговування було створено імітаційну модель інформаційної системи, що відображала реальний процес обробки запитів із заданими параметрами. Прораховано можливі варіанти масштабування інформаційної системи та об'єми заощаджень при дотриманні рекомендацій, оснований на даному дослідженні.

Дане дослідження окреслює основні межі роботи технічної інфраструктури, що дає змогу налаштувати роботу компонентів із зменшенням обчислювальних потужностей, що в свою чергу впливає на кінцеві витрати на підтримку або ж оренду обчислювального обладнання для інформаційної системи. Правильний підбір моделі під певну інформаційну систему дасть змогу не тільки заощадити, але й унеможливити варіанти падіння систем, що може призвести до збитків компанії.

Наукова новизна дослідження полягає у застосуванні методів прогнозу та імітаційних моделей у ІТ-галузі, що дає змогу правильно оцінити варіанти масштабування для обчислювальних потужностей.

Головним результатом дослідження є проведення економіко-математичного моделювання роботи ІС та, як наслідок, розробка імітаційної моделі ІС. Моделювання даного об'єкту дає змогу створити позитивно-якісний вплив на собівартість випущеної технології та закріплення

конкурентних позицій компанії на ринку, а також унеможливити варіанти неочікуваних фінансових втрат та ризику банкрутства.

Теоретична значимість дослідження полягає у тому, що воно показало необхідність створення комплексних методів оцінки моделей прогнозу навантажень із урахуванням економії витрат на підтримку інфраструктури.

Практична значимість роботи визначається тим, що використання розроблених в ній алгоритмів, методів та моделей може слугувати як для малого так і великого підприємства інструментом заощадження на ІС.

Отже, все вище викладене дозволяє стверджувати, що прогнозні моделі та створена імітаційна модель дає можливість відтворити роботу ІС на майбутні періоди. Опираючись на проведений аналіз результатів обчислення моделі та співставлення їх з динамікою реальних значень навантажень на систему, можна зробити висновок про адекватність прогнозних моделей та правильну побудову імітаційної моделі, що говорить про можливість застосування основи даного дослідження для прогнозування схожих показників у майбутньому та симуляцій роботи систем базованих на теорії масового обслуговування.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Алексеев Е.Р., Чеснокова О.В., MATLAB 7. Самоучитель: [підручник] /Алексеев Е.Р., Чеснокова О.В М: НТ Пресс. 2006.
2. Бокс Дж., Аналіз часових рядів. Прогноз і Управління. [підручник]/ Дж. Бокс, Г. Дженкинс: Мир, 1974. – С. 406.
3. Глівенко С.В. Економічне прогнозування: навч. посібник / Глівенко С.В., Соколов М.О., Таліженко О.М. – 2-ге вид., перероб. та доп. – Суми: Видавництво «Університетська книга», 2001. – 207 с.
4. Дуброва Т.А., Статистичні метод прогнозування/ Т.А. Дуброва // UNYTY-DANA.2013. – С. 206.
5. Дьяконов В.П., Matlab 6.5 SP1/7 + Simulink 5/6 в математиці і моделюванні: [підручник]/ В.П. Дьяконов М.: СОЛОН-Пресс. 2005. – С.576.
6. Загородна, О. Оперативне управління грошовими потоками підприємства / О. Загородна, В. Серединська // Економічний аналіз : збірник наукових праць. – Випуск 10, Ч. 4. – Тернопіль : Економічна думка, 2012. – С. 100-104
7. Івницький В.А, Теорія мереж масового обслуговування: [підручник] [електронний ресурс]/ В.А Івницький: Видавництво фізико-математичної літератури. 2004. Режим доступу: <http://fizmatlit.narod.ru/Ivnickii.htm>
8. Канторович Г.Г., Анализ временных рядов [Текст] / Г. Г. Канторович. М.: Экономический журнал ВШЭ. 2002. – №3.
9. Клейнрок Л. Теорія масового обслуговування [підручник]/ Л. Клейнрок: Машиностроение. 1979.

10. Коваленко А.А., Кучук Г.А., Методи синтезу інформаційної та технічної структур системи управління об'єктом критичного застосування //Advanced Information Systems. 2018. – №18.

11. Крюков, Ю. А., Исследование самоподобия трафика высоко скоростного канала передачи пакетных данных [электронный ресурс] / Ю. А. Крюков, Д. В. Чернягин // Режим доступа: <http://sanse.ru/download/33> (датаобращения09.10.2010).

12. Тихонов Е.Є., Методи прогнозування в умовах ринку: [підручник] / Е. Є. Тихонов: Невинномиськ. 2006. – С.21.

13. Турчаева И.Н., Финансовая среда предпринимательства и предпринимательские риски [підручник] [электронный ресурс]/ И.Н.Турчаева, В.А.Матчинов: Саратов: Вузовское образование, 2018. – С. 248 -278. Режим доступа: <http://www.iprbookshop.ru/77575.html>

14. Якимов І. М., А.П. Кирпичников, Комплексный подход к моделированию сложных систем в системе BPwin-Arena/ І. М. Якимов , А.П. Кирпичников // Вестник Казанского технологического университета 2013. – № 12 – С. 21

15. Якимов І. М., Кирпичников А.П., Зайнуллина Г., Оценка достоверности результатов имитационного моделирования по результатам аналитического моделирования./ Якимов І. М., Кирпичников А.П., Зайнуллина Г // Вестник технологического университета. 2015. – № 6 – С. 173-178

16. Якимов І. М., Кирпичников А.П., Моделирование сложных систем в имитационной среде ANYLOGIC/ І. М. Якимов , А.П. Кирпич // Вестник Казанского технологического университета. 2014. № 13 – С. 352-357.

17. Якимов І. М., Кирпичников А.П.,Мокшин В. В.. Обучение имитационному моделированию в пакете SIMULINK системы MATLAB./

Якимов И. М., Кирпичников А.П., Мокшин В. В. // Вестник технологического университета. 2015. – № 22. – С. 184-188.

18. Anderson T. W., Testing Multiple Equation System for Common Nonlinear Components / T. W. Anderson, F. Vahid // Journal of Econometrics. 1998. – №84. – P. 1–36.

19. Baillie R. T., Long Memory Processes and Fractional Integration in Econometrics / R. T. Baillie // Journal of Econometrics. 1996. – No 73. – P. 5–60.

20. Blanchard Oliver J., What is Left of the Multiplier Accelerator / Oliver J. Blanchard // The American Economic Review. 1981. – №71. – P. 150–154

21. Davidson R. Several Tests for Model specification in the Presence of Alternative Hypothesis / R. Davidson, J. MacKinnon // Econometrica. 1981. – №49. – P. 781–793.

22. Dickey D. A., Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root / D. A. Dickey // Econometrica. 1981. – № 49. – P. 1057–1072.

23. Dimas Setyo., Time Series Analysis for Predicting the Total Patient Treated in Blora Health Center Using Minitab Software 14/ Dimas Setyo// FMIPA UNNES Publisher. 2011.

24. Doronina A. I., Time-series models the example of oil products consumption in France/ Doronina A. I.// Financial University under the Government of the Russian Federation. 2012.

25. Enders W. Applied Econometrics Time Series : 3rd ed. / W. Enders. – Jon Wiley and Sons, Inc., – 2010. – 517 p.

26. Frances P. H. Spurious Deterministic Seasonality / P. Frances, S. Hylleberg, H. Lee // Economics Letters. 1995. – № 48. – P. 249–256.

27. Graham E., Economic Forecasting / E. Graham, A. Timmermann // Journal of Economic Literature. 2008. – № 46. – P. 504–508.36.
28. Granger C. W. J., Spurious Regressions in Econometrics / C. W. J. Granger, P. Newbold // Journal of Econometrics. 1974. – № 2. – P. 111–120.
29. Hadijah., Forecasting of Operational Reserve with Program Approach Using Minitab Arima / Hadijah.// Journal THE WINNERS.2013. – P. 13-19.
30. Hamilton J. D. Time Series Analysis / James D. Hamilton. – New Jersey : Princeton University Press, 1994. – 799 p.
31. Hannan E. J. The Estimation of the Order of an ARMA Process / E. J. Hannan // Annals of Statistics. 1980. – № 8. – P. 1071–1081.
32. Hansen L. P., Seasonality and Approximation Errors in Rational Expectations Models / L. P. Hansen, T. J. Sargent // Journal of Econometrics. 1993. – № 55. – P. 21–56.
33. Holt C.C., Forecasting seasonal and trends by exponentially weighted moving averages / Holt C.C.// Journal of forecasting. 2004 – №20 – P. 5-10.
34. Jingfei Yang M., Sc. Power System Short-term Load Forecasting: Thesis for Ph.d degree. / Jingfei Yang // Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universität, 2006. – P.139.
35. Kendall D.G., Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain./ Kendall D.G// The Annals of Mathematical Statistics. 1953. – №3. – P. 338–254.
36. Lyuys K.D., Methods of forecast of economics values/ Lyuys K.D // Finance ta statistics. 1986. – P. 133.
37. Pagan A. R. The Econometrics of Financial Markets / A. R. Pagan // Journal of Empirical Finance. 1996. – № 3. – P. 15–102.
38. Pattabi Seshadri and Lizy K. John, University of Texas at Austin. Characterization of Web Server Workloads on Three Generations of IBM

PowerPC Microarchitectures [электронный ресурс] – Режим доступа:  
<http://www.research.ibm.com/acas/projects/00seshadri.pdf>

39. Rutka G., Network Traffic Prediction using ARIMA and Neural Networks Models / Rutka G.// Electronics And Electrical Engineering. 2008 – №4

40. Salah K., Analysis of a two-stage network server/Salah K., // Applied Mathematics and Computation. 2011 – No 217. – P. 9635–9645.

## ДОДАТОК А

### Прогнози кандидатних моделей

|                  |              |              |                           |              |              |                           |
|------------------|--------------|--------------|---------------------------|--------------|--------------|---------------------------|
|                  | 00:00:00     |              |                           | 12:00:00     |              |                           |
|                  | Arima(1,1,1) |              | Arima(1,1,2)              | Arima(2,1,2) |              |                           |
| forecast (lag 1) | 439289       |              | 43984                     | 70170        |              |                           |
| forecast (lag 2) | 44261        |              | 44305                     | 70965        |              |                           |
|                  | 01:00:00     |              |                           | 13:00:00     |              |                           |
|                  | Arima(1,1,1) |              |                           | Arima(2,1,2) |              |                           |
| forecast (lag 1) | 31482        |              |                           | 73542        |              |                           |
| forecast (lag 2) | 31678        |              |                           | 74360        |              |                           |
|                  | 02:00:00     |              |                           | 14:00:00     |              |                           |
|                  | Arima(1,1,1) |              |                           | Arima(3,1,3) |              |                           |
| forecast (lag 1) | 25548        |              |                           | 71814        |              |                           |
| forecast (lag 2) | 25694        |              |                           | 72852        |              |                           |
|                  | 03:00:00     |              |                           | 15:00:00     |              |                           |
|                  | Arima(2,1,3) |              | Arima(3,1,3)              | Arima(1,1,1) | Arima(1,1,3) | Arima(3,1,1) Arima(3,1,3) |
| forecast (lag 1) | 24768        |              | 24801                     | 69471        | 69755        | 69961 69779               |
| forecast (lag 2) | 25255        |              | 25348                     | 69825        | 69782        | 70277 69898               |
|                  | 04:00:00     |              |                           | 16:00:00     |              |                           |
|                  | Arima(2,1,2) |              | Arima(2,1,3)              | Arima(1,1,1) | Arima(1,1,3) | Arima(3,1,1) Arima(3,1,3) |
| forecast (lag 1) | 26149        |              | 26367                     | 70293        | 70782        | 70900 70780               |
| forecast (lag 2) | 26207        |              | 26827                     | 70279        | 70592        | 70646 71010               |
|                  | 05:00:00     |              |                           | 17:00:00     |              |                           |
|                  | Arima(2,1,2) | Arima(2,1,3) | Arima(3,1,2) Arima(3,1,3) | Arima(3,1,3) |              |                           |
| forecast (lag 1) | 31618        | 31788        | 31946 31780               | 73208        |              |                           |
| forecast (lag 2) | 31596        | 32226        | 32578 32244               | 70692        |              |                           |
|                  | 06:00:00     |              |                           | 18:00:00     |              |                           |
|                  | Arima(2,1,2) |              |                           | Arima(3,1,3) |              |                           |
| forecast (lag 1) | 40720        |              |                           | 69670        |              |                           |
| forecast (lag 2) | 41128        |              |                           | 69000        |              |                           |
|                  | 07:00:00     |              |                           | 19:00:00     |              |                           |
|                  | Arima(2,1,1) |              | Arima(2,1,2)              | Arima(3,1,2) |              | Arima(3,1,3)              |
| forecast (lag 1) | 49149        |              | 50843                     | 66448        |              | 65757                     |
| forecast (lag 2) | 48681        |              | 51579                     | 67199        |              | 66253                     |
|                  | 08:00:00     |              |                           | 20:00:00     |              |                           |
|                  | Arima(1,1,1) | Arima(1,1,2) | Arima(2,1,1) Arima(2,1,2) | Arima(1,1,1) | Arima(1,1,3) | Arima(3,1,1) Arima(3,1,3) |
| forecast (lag 1) | 59739        | 59948        | 59995 59814               | 69767        | 70376        | 69737 69993               |
| forecast (lag 2) | 59903        | 59810        | 59850 59703               | 70206        | 69746        | 69130 69467               |
|                  | 09:00:00     |              |                           | 21:00:00     |              |                           |
|                  | Arima(1,1,1) | Arima(1,1,2) | Arima(2,1,1) Arima(2,1,2) | Arima(2,1,2) | Arima(2,1,3) | Arima(3,1,2) Arima(3,1,3) |
| forecast (lag 1) | 64568        | 64801        | 64960 65331               | 68225        | 69319        | 67573 69928               |
| forecast (lag 2) | 64978        | 64834        | 64930 65565               | 69121        | 69012        | 67554 69316               |
|                  | 10:00:00     |              |                           | 22:00:00     |              |                           |
|                  | Arima(1,1,1) |              | Arima(1,1,2)              | Arima(2,1,2) | Arima(2,1,3) | Arima(3,1,2) Arima(3,1,3) |
| forecast (lag 1) | 67618        |              | 67615                     | 65093        | 65233        | 64912 65662               |
| forecast (lag 2) | 67882        |              | 67767                     | 66186        | 65495        | 65244 65782               |
|                  | 11:00:00     |              |                           | 23:00:00     |              |                           |
|                  | Arima(1,1,2) |              | Arima(1,1,3)              | Arima(2,1,2) |              | Arima(2,1,3)              |
| forecast (lag 1) | 68775        |              | 69122                     | 57626        |              | 57134                     |
| forecast (lag 2) | 68935        |              | 68825                     | 58172        |              | 57434                     |

Рисунок А.1 – Прогноз кандидатних моделей

## ДОДАТОК В

### Обрані результати моделей

|              |              |              |              |              |              |              |              |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 00:00:00     |              |              |              | 12:00:00     |              |              |              |
| Arima(1,1,1) |              | Arima(1,1,2) |              | Arima(2,1,2) |              |              |              |
| 0,03090182   |              | 0,030856183  |              | 0,020957458  |              |              |              |
| 01:00:00     |              |              |              | 13:00:00     |              |              |              |
| Arima(1,1,1) |              |              |              | Arima(2,1,2) |              |              |              |
| 0,023768844  |              |              |              | 0,045376211  |              |              |              |
| 02:00:00     |              |              |              | 14:00:00     |              |              |              |
| Arima(1,1,1) |              |              |              | Arima(3,1,3) |              |              |              |
| 0,032232906  |              |              |              | 0,025721706  |              |              |              |
| 03:00:00     |              |              |              | 15:00:00     |              |              |              |
| Arima(2,1,3) |              | Arima(3,1,3) |              | Arima(1,1,1) | Arima(1,1,3) | Arima(3,1,1) | Arima(3,1,3) |
| 0,098946793  |              | 0,101693704  |              | 0,042614705  | 0,044963573  | 0,043183411  | 0,044369977  |
| 04:00:00     |              |              |              | 16:00:00     |              |              |              |
| Arima(2,1,2) |              | Arima(2,1,3) |              | Arima(1,1,1) | Arima(1,1,3) | Arima(3,1,1) | Arima(3,1,3) |
| 0,075778799  |              | 0,092788337  |              | 0,032782579  | 0,03275134   | 0,033229952  | 0,029954037  |
| 05:00:00     |              |              |              | 17:00:00     |              |              |              |
| Arima(2,1,2) | Arima(2,1,3) | Arima(3,1,2) | Arima(3,1,3) | Arima(3,1,3) |              |              |              |
| 0,063795311  | 0,077046567  | 0,085539853  | 0,077202901  | 0,03558928   |              |              |              |
| 06:00:00     |              |              |              | 18:00:00     |              |              |              |
| Arima(2,1,2) |              |              |              | Arima(3,1,3) |              |              |              |
| 0,023254868  |              |              |              | 0,062704714  |              |              |              |
| 07:00:00     |              |              |              | 19:00:00     |              |              |              |
| Arima(2,1,1) |              | Arima(2,1,2) |              | Arima(3,1,2) |              | Arima(3,1,3) |              |
| 0,043673061  |              | 0,018210987  |              | 0,042950856  |              | 0,044223479  |              |
| 08:00:00     |              |              |              | 20:00:00     |              |              |              |
| Arima(1,1,1) | Arima(1,1,2) | Arima(2,1,1) | Arima(2,1,2) | Arima(1,1,1) | Arima(1,1,3) | Arima(3,1,1) | Arima(3,1,3) |
| 0,025485954  | 0,028021622  | 0,028099316  | 0,027743168  | 0,02998788   | 0,037721974  | 0,037259112  | 0,036816199  |
| 09:00:00     |              |              |              | 21:00:00     |              |              |              |
| Arima(1,1,1) | Arima(1,1,2) | Arima(2,1,1) | Arima(2,1,2) | Arima(2,1,2) | Arima(2,1,3) | Arima(3,1,2) | Arima(3,1,3) |
| 0,031227642  | 0,034117876  | 0,034664522  | 0,032923825  | 0,031789284  | 0,024767042  | 0,047373473  | 0,018310717  |
| 10:00:00     |              |              |              | 22:00:00     |              |              |              |
| Arima(1,1,1) |              | Arima(1,1,2) |              | Arima(2,1,2) | Arima(2,1,3) | Arima(3,1,2) | Arima(3,1,3) |
| 0,046342132  |              | 0,047132252  |              | 0,018634061  | 0,022691039  | 0,026972588  | 0,01732754   |
| 11:00:00     |              |              |              | 23:00:00     |              |              |              |
| Arima(1,1,2) |              | Arima(1,1,3) |              | Arima(2,1,2) |              | Arima(2,1,3) |              |
| 0,037776979  |              | 0,041201597  |              | 0,008891163  |              | 0,019400596  |              |

Рисунок В.1 – Вибір кращих моделей прогнозу

## ДОДАТОК С

### Модуль заявок і черги

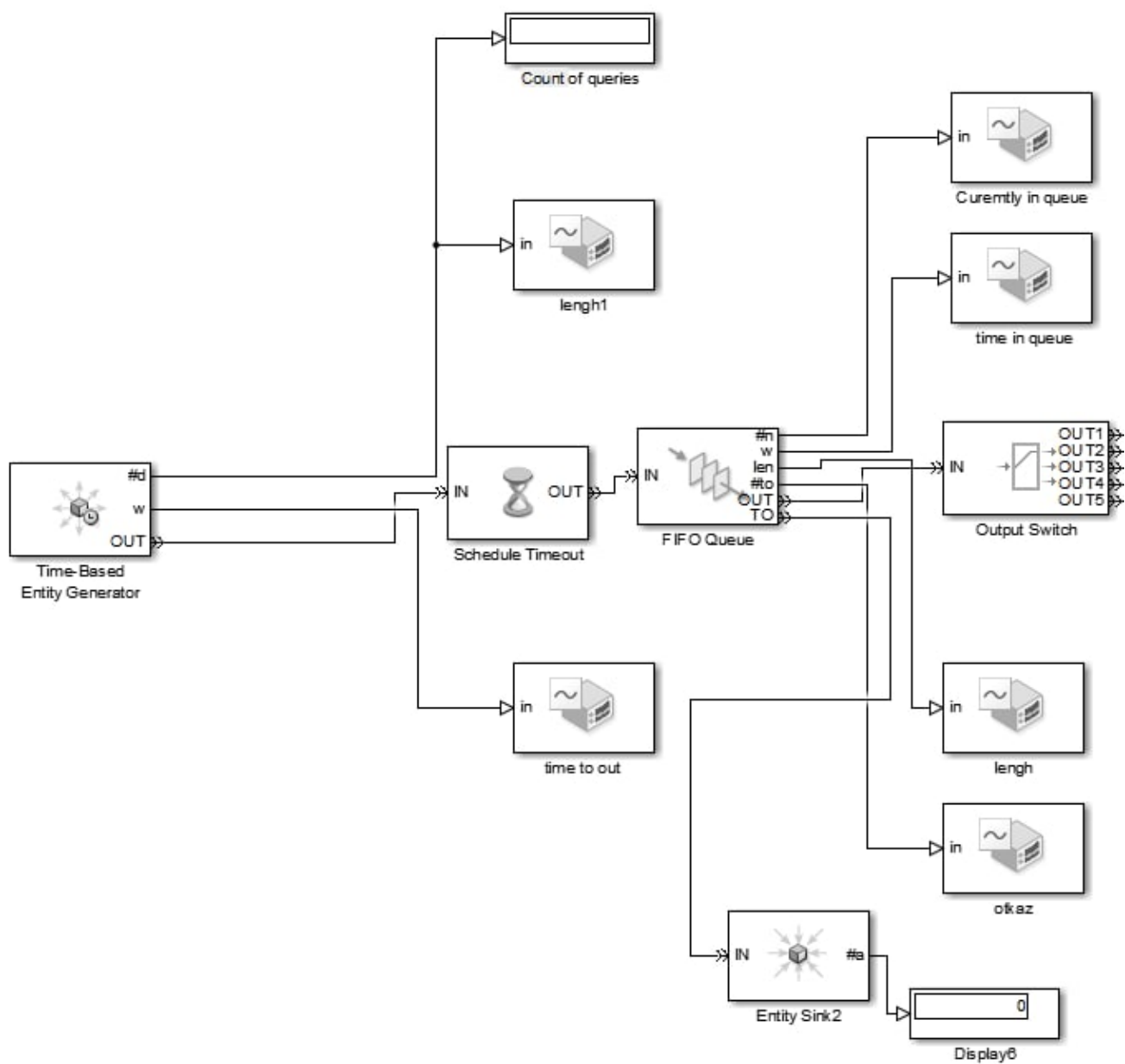


Рисунок С.1 – Модулі формування заявок та черги в Matlab Simulink

## ДОДАТОК D

### Модуль обробки

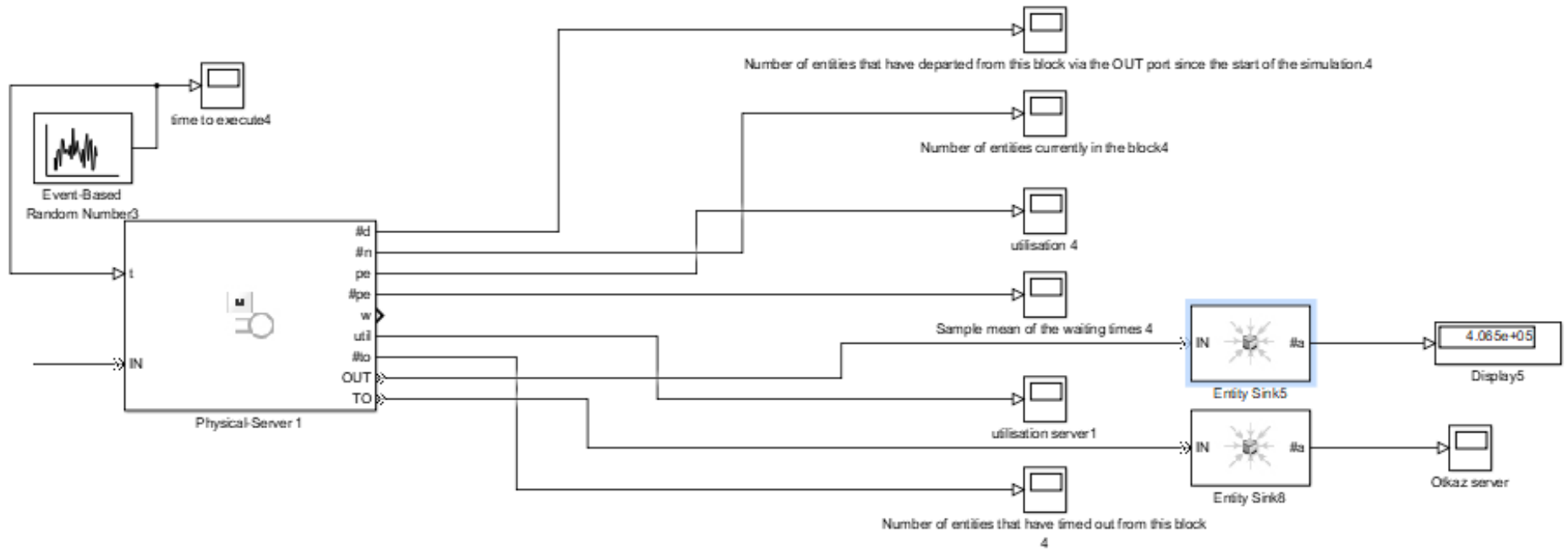


Рисунок D.1 – Модуль обробки заявок в Matlab Simulink

## ДОДАТОК Е

### Загальна імітаційна модель

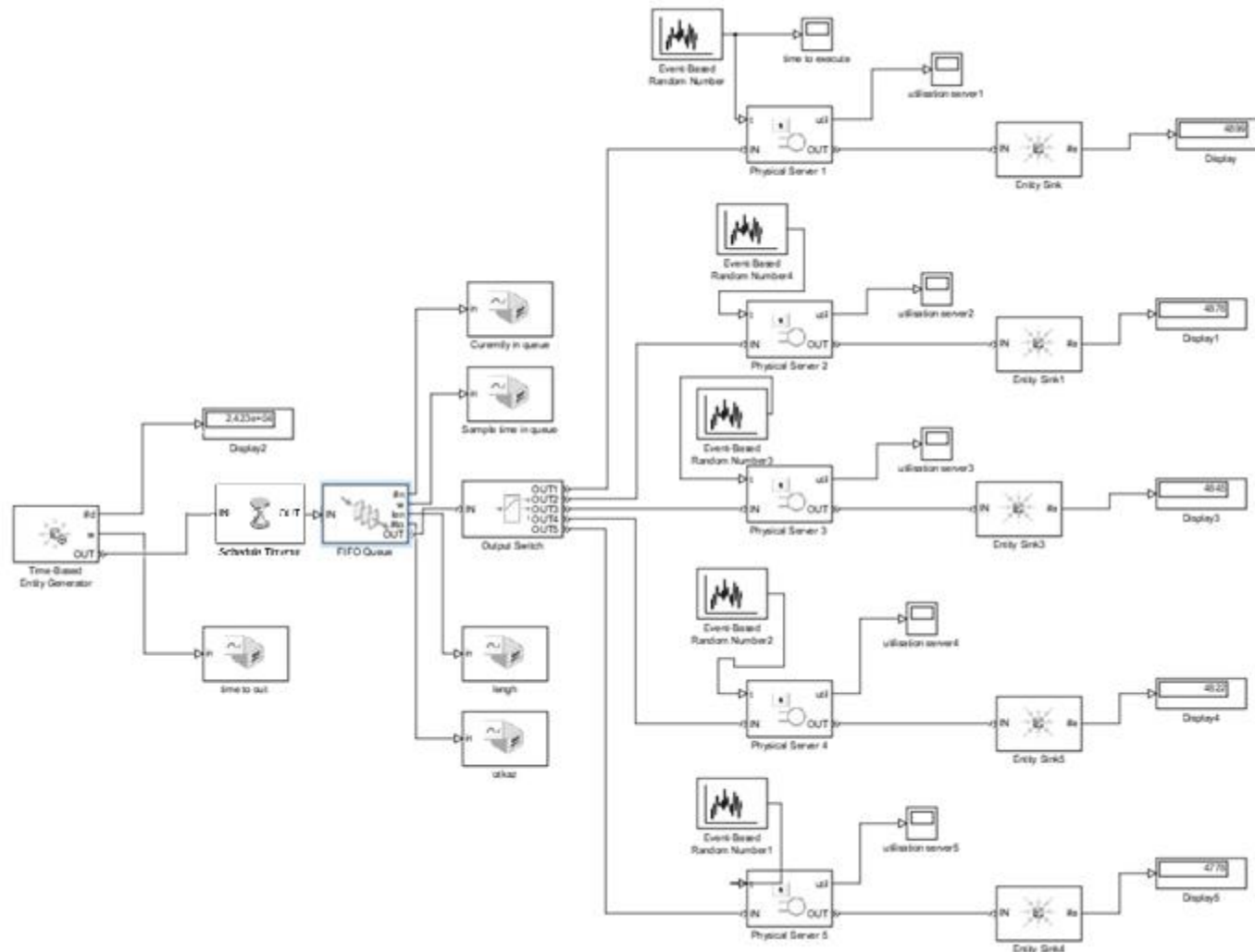


Рисунок Е.1 – Загальна імітаційна модель навантаження ІС в Matlab Simulink

