Міністерство освіти і науки України

Системні технології

System technologies

2 (151) 2024

Регіональний міжвузівський збірник наукових праць

Засновано у січні 1997 року.

У випуску:

Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 2 (151). - Дніпро, 2024. – 198 с.
ISSN 1562-9945 (Print).
ISSN 2707-7977 (Online).

Редакційна колегія випуску:

Збірник друкується за рішенням Вченої Ради
Українського державного університету науки і технологій
від 31.01.2024 р., № 6

Адреса редакції: 49600, Дніпро, пр. Гагаріна, 4
Український державний університет науки і технологій,
ННІ «Інститут промислових та бізнес технологій»
кафедра Інформаційних технологій та систем.
Тел. +38(097)6854525
E-mail: st@nmetau.edu.ua
https://journals.nmetau.edu.ua/index.php/st

DOI 10.34185/1562-9945-2-151-2024-01
УДК 004.42:519.85

Б.І. Мороз, А.С. Круглик, Д.М. Мороз, А.А. Мартиненко

МАТЕМАТИЧНА МОДЕЛЬ РАЦІОНАЛЬНОЇ ОРГАНІЗАЦІЇ ОБРОБКИ ІНФОРМАЦІЙНИХ ПОТОКІВ В СИСТЕМІ ДОСТАВКИ ЛІТАЛЬНИМИ АПАРАТАМИ

Анотація. У роботі розглянуто математичну модель масового обслуговування, що дозволить розподіляти час між чергами таким чином, щоб гарантувати виконання(обробку) кожного повідомлення в чергах в межах часу, обумовленого для кожної з черг. Метою роботи є опис математичної моделі для квантування часу. Вектор управління повинен забезпечити системі гнучкість, що в свою чергу дозволить керувати системою обробки даних (СОД), в залежності від навантаження на певному проміжку часу чи кількісній зміні окремого типу повідомлень. Така модель може бути актуальною для систем доставки вантажу літальними апаратами, де навантаження на систему може змінюватись впродовж доби, а тримати парк літальних апаратів, для покриття пікових навантажень є економічно недоцільним.
Ключові слова: математична модель, квантування часу, вектор управління, пікові навантаження, літальні апарати.

Постановка проблеми. У сучасному світі, який переживає стрімке технологічне розширення, роль літальних апаратів у системах доставки стає надзвичайно важливою. З кожним роком кількість перевезень вантажу за допомогою літальних апаратів стрімко зростає. Зростання інтересу до безпілотних літаків, квадрокоптерів та інших літальних транспортних засобів визначає потребу у диспетчеризації вхідних інформаційних потоків (заявок) від множини користувачів на доставку вантажу. Питання диспетчеризації потоків повідомлень тісно пов’язане з вибором моделі раціональної організації обробки таких потоків. Тому що до такої системи пред’являються певні вимоги. Наприклад, вартість доставки може залежати від термінів, за які обумовлено таку доставку виконати. Можуть бути як термінові доставки, які потрібно виконувати якомога скоріше, або навпаки, заявка на доставку може виконуватись коли система найменш навантажена, при цьому вартість такої доставки буде нижча. Але при

цьому потрібно враховувати, незалежно від кількості заявок, їх терміновості і т. д., кожна заявка повинна бути виконана в межах її часових обмежень.

Отже розробка методів раціональної організації обробки інформаційних потоків в системах доставки літальними апаратами є актуальним завданням. Адже потрібно розробити таку систему, що буде відповідати пред'явленим до неї вимогам.

Аналіз останніх досліджень і публікацій. Вирішенням питання масового обслуговування інформаційних потоків займаються науковці і працівники починаючи з 1900-х років. Традиційно вважається, що вперше задачу масового обслуговування вирішував датський математик, статистик та інженер Копенгагенської телефонної компанії Агнер Ерланг, який вирішував завдання з упорядкування роботи телефонної станції.

Із останніх відомих робіт у напрямку розробки систем масового обслуговування в доставці товарів БПЛА є цікавою робота Б. П. Книш «Метод розподілу часу на доставку товарів за допомогою безпілотних літальних апаратів згідно пріоритету» [3]. В роботі розглянуто розподіл кванту часу на обслуговування згідно з пріоритетом. Але недолік системи масового обслуговування з пріоритетом може полягати в тому, що така система гарантує, що будуть доставлені товари в межах визначеного часу, лише з найвищим пріоритетом. Це може бути некоректно, якщо систему доставки літальними апаратами розглядати як таку, де вартість доставки змінюється в залежності від термінів доставки. Наприклад, закономірно, що вдень навантаження на систему обслуговування буде більше, ніж вночі. З цього можна зробити висновок, що доставка вночі може бути дешевшою, щоб мінімізувати простої літальних апаратів і отримати від цього економічний ефект. Але в той же час, наприклад, якщо з'являється економічно вигідніша доставка у часи простою, не можна нехтувати вже існуючою з більш низькою вартістю доставки. Система повинна обробляти заявки вчасно, або відмовити в обслуговуванні, якщо виконати таку заявку неможливо у вказаних часових обмеженнях.

Мета дослідження – описати концептуальну модель обробки інформаційних потоків в системах доставки літальними апаратами з використанням керованої дисципліни для раціональної організації обробки інформації з урахуванням її характеристик цінності і старіння, розроблену проф. Морозом Б.І. [6].

Виклад основного матеріалу дослідження. Концептуальна модель доставки вантажу за допомогою дронів представляє собою структурний та функці-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ональний зразок, що визначає основні етапи, аспекти та взаємодії у системі доставки. Цей опис включає ключові складові концепції, які спрямовані на оптимізацію процесу доставки вантажу та максимізацію ефективності використання безпілотних апаратів. Пропонується розглянути модель як систему, що має такі основні складові: мережа дронів, джерела даних (складів, аптек, точок поставки та ін.), множина користувачів, що є динамічною величиною і постійно змінюється в часі, а також диспетчерський пункт (див. рис. 1).

Рисунок 1 – Концептуальна модель організації обробки вхідних інформаційних потоків та доставки літальними апаратами

Замовлення та розміщення завдань. Користувач розміщує замовлення диспетчеру за допомогою додатку або веб інтерфейсу, вказуючи точний адресат та параметри доставки. Параметри доставки є характеристикою, що визначає час, необхідний для обробки заявки (пошук об’єкта доставки в динамічній системі магазинів, побудова оптимального маршруту доставки, погодження часу завантаження/відвантаження, терміновість доставки і т. ін.) а також час, необхідний безпосередньо для завантаження/відвантаження і транспортування літальним апаратом за адресатом. Тобто по суті, кожна заявка має різну цінність а також характеристики старіння.

Мережа дронів представляє собою цілісну систему – парк літальних апаратів з різними технічними характеристиками. Також включає в себе обслуговуючий персонал, необхідний для ремонту та обслуговування парку. Мережу парку літальних апаратів пропонується розглядати як економічно доцільну величину, де кількість дронів та їх характеристики визначаються мінімально достатньою величиною для виконання усіх заявок а також враховуючи амортизацію і поточні ремонти.

Джерела даних є динамічною величиною. Це можуть бути магазини, що пропонують свої товари та послуги, які може замовити користувач. Аптеки або склади, пункти перевантаження і т. д. А також є різновіддаленими від користувача. Цю складову системи слід розглядати як будь яке джерело, звідки можна отримати дані про вантаж, його наявність, достатню кількість, інші параметри, які можна отримати і проаналізувати для досягнення кінцевої мети – своєчасної доставки вантажу за вказаною замовником адресою.

Диспетчер приймає заявку від користувача, обробляє її і виконує. Під диспетчером слід розуміти автоматизоване програмне забезпечення (ПЗ) із мінімальним втручанням оператора або без втручання такого. Диспетчер, використовуючи алгоритми, займається пошуком товару, побудовою оптимального маршруту доставки, подає заявку до оператора з надання послуг використання літальних апаратів з відповідними характеристиками і т. ін.

Із вказаного вище стає зрозуміло, що саме диспетчерське ПЗ повинно раціонально організовувати обробку вхідних інформаційних потоків, щоб забезпечити виконання вхідної заявки згідно з накладеними на таке ПЗ вимогами.

Виконання функції диспетчеризації вимагає створення цілого ряду процедур, які забезпечать оптимальну (раціональну) організацію порядку їх обробки. При цьому також необхідно забезпечити реалізацію ряду проблем, пов’язаних з організацією етапів збору, очистки і зберігання даних від зовнішніх джерел. В якості джерел можуть використовуватися різні бази даних (БД), архівні справи і заявки (оперативні джерела).

При створенні інформаційно-аналітичних систем з використанням OLAP технології, існує ряд проблем, пов’язаних з організацією етапу збору, очистки, і погодження даних із зовнішніх джерел. В якості джерел використовуються різні бази даних різних інформаційних систем, що містять в собі оперативні, довідкові, архівні і т. ін. дані, рис. 2

З рис. 2 видно, що оперативні, довідкові, архівні і інші дані поступають у вигляді окремих повідомлень з БД1, БД2, БД3 і т. д. в систему попередньої обробки СОД (система обробки даних) чи СОД1, СОД2, СОД3 і т. д., обробляються цією системою (чи системами), накопичуються в базі даних попереднього зберігання (БДПЗ) або БДПЗ1, БДПЗ2, БДПЗ3 і т. д. Після цього завантажуються у сховище даних (СД) або вітрину даних (ВД).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

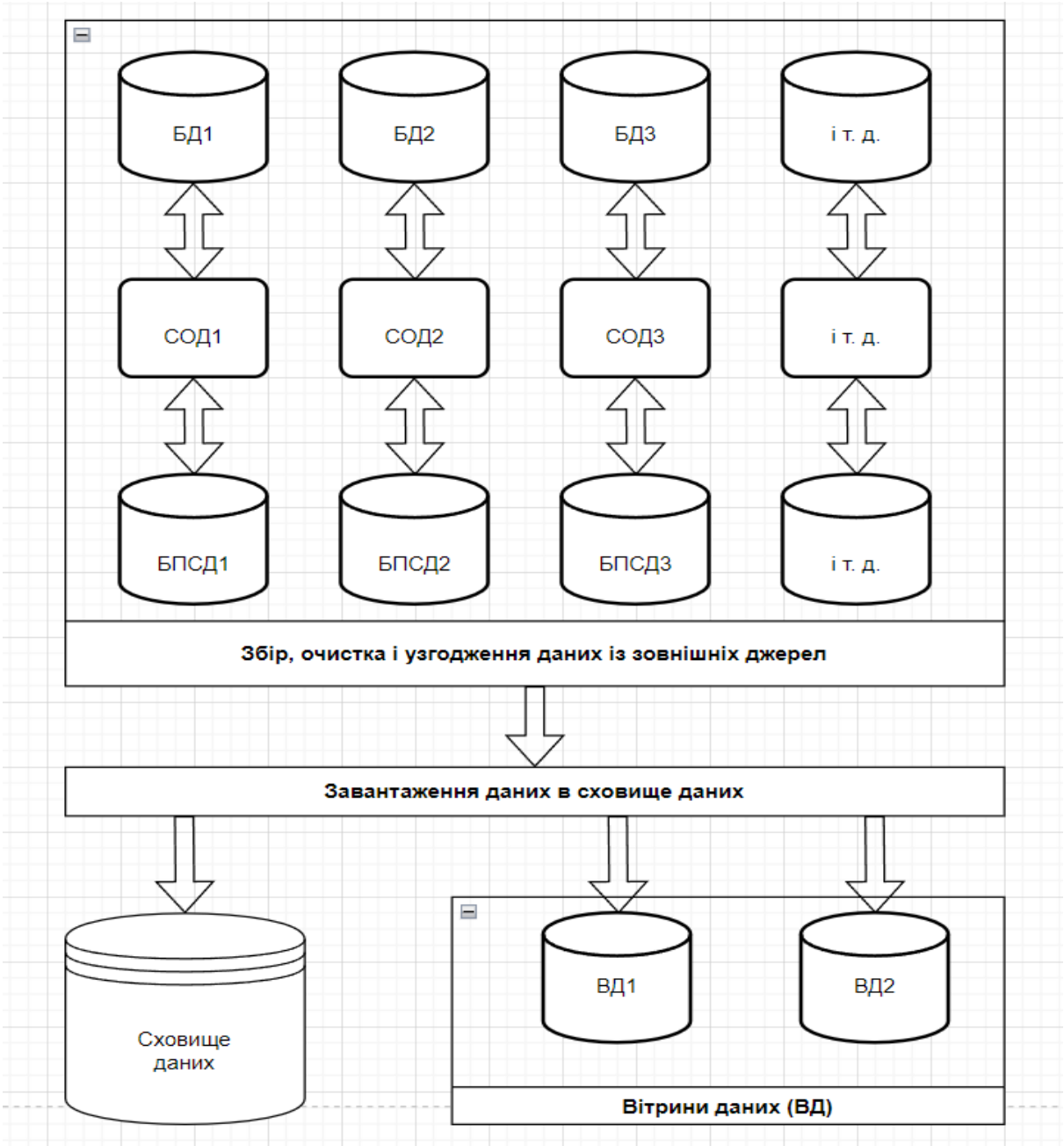


Рисунок 2 – Схема збору даних і їх завантаження в сховище даних

Усі оперативні, довідкові, архівні та інші дані відповідають певному моменту часу на часовій осі і врешті решт повинні бути сформовані у вигляді фактів пов'язаних з відповідним моментом часу.

Факти і їх агрегації в СД повинні з необхідною точністю і в необхідні моменти часу відображати стан предметної області, відносно якої приймаються рішення.

Необхідно відзначити, що дані з БД в СОД поступають у виді регулярних або випадкових потоків повідомлень з певною інтенсивністю, що змінюється в

часі. Повідомлення обробляються як в автоматичному, так і в автоматизованому режимах.

Головними і необхідними вимогами до обробки повідомлень є наступні:

1. Жодне повідомлення не повинно бути втрачено (тобто всі повідомлення повинні бути оброблені)

2. Усі повідомлення повинні бути оброблені в рамках певних часових обмежень. Ці часові обмеження визначаються виходячи із чітко визначеної необхідності завантаження кожного факту в чітко визначений час.

3. Обробка всіх повідомлень повинна бути виконана в умовах мінімальної фактичної продуктивності $P_{факт}$ СОД.

4. Організація обробки потоків повідомлень повинна виконуватися з урахуванням якісно-кількісних характеристик інформації (таких як цінність, старіння та ін.), якими характеризуються окремі повідомлення.

Розглянемо питання організації обробки повідомлень на деякому інтервалі часу Δt в СОД більш детально.

На вхід СОД поступає потік повідомлень, на інтервалі часу Δt характеризується вектором:

$$\Lambda\Delta t = \{\lambda_1\Delta t, \lambda_2\Delta t, \lambda_i\Delta t, \lambda_h\Delta t\}, \quad (1)$$

де, $\lambda_i\Delta t$ – інтенсивність потоку повідомлень і-го типу на проміжку часу Δt.

і – індекс повідомлення по його належності до деякого порогового часу обробки СОД.

Введемо вектор:

$$T_{порог.} = \{T_{порог.1}, T_{порог.2}, \ldots, T_{порог.i}, \ldots, T_{порог.h}\}, \quad (2)$$

де $T_{порог.i}$ – максимальний час перебування повідомлення *i*-го типу в СОД.

Час обробки повідомлення *i*-го типу визначимо як

$$b_i = \int_0^\infty t dBi(t), \quad (3)$$

де $B_i(t)$ функція розподілу часу обробки повідомлень *i*-го типу.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Повідомлення інтенсивностей $\lambda_1\Delta t$, $\lambda_2\Delta t$, $\lambda_i\Delta t$, $\lambda_h\Delta t$ з'являються в СОД на проміжку часу Δt через інтервали часу $t_{ц.г.i}$ і стають в черги кожного потоку в свою, рис. 3.

$t_{ц.г.i}$ – час циклу генерації повідомлення *i*-го типу.

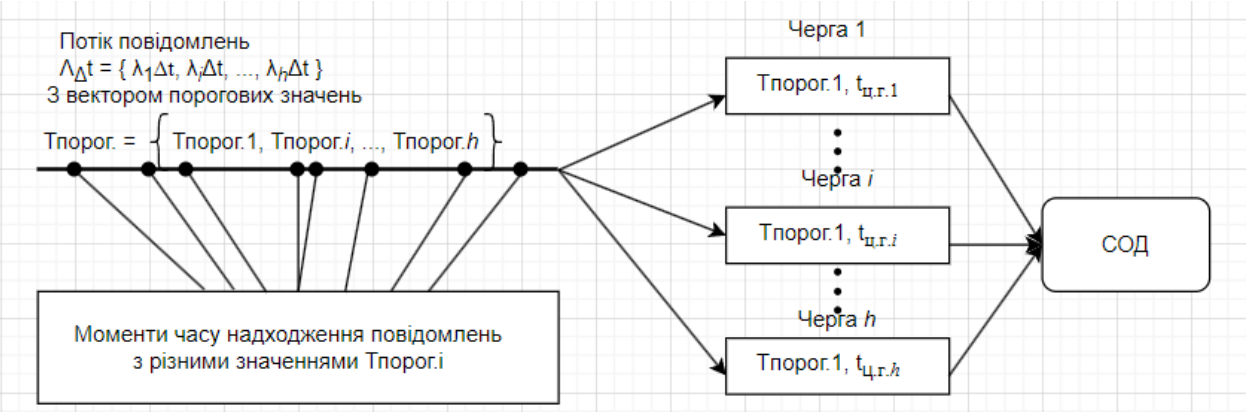


Рисунок 3 – Формування черг з різним пороговим часом обробки $T_{порог.i}$, що поступають через проміжки часу $t_{ц.г.i}$

Обробка повідомлень організується із наступною дисципліною: СОД віддає кожній черзі почергово квант часу T_{ki}, за який можна обробити 1, 2, 3…12 і т. д. повідомлень *i*-ї черги, тобто

$$t_{ki} / b_i = 1, 2, 3... \text{ і т. д.} \quad (4)$$

Така дисципліна дозволяє керувати обробкою потоків повідомлень при зміні вектору інтенсивності потоку повідомлень $\Lambda_\Delta t$.

Цикл обробки для *i*-ї черги дорівнює:

$$t_{ці} = t_{k1} + t_{k2} + ... t_{kh} = \sum_{i=1}^{h} tki \quad (5)$$

Кількість повних циклів на момент часу t – n_{ti} буде дорівнювати:

$$n_{ti} = \frac{t - t_{непр.0i}}{t_{ці}}, \; n_{ti} = 0, 1\ 2, 3 \text{ і т.д.} \quad (6)$$

Час закінчення n-го циклу для і-ї черги

$$t^{(n)}_{(зак.ці)} = t_{непр.0i} + t_{ці} n_i, \quad (7)$$

де n_i порядковий номер циклу і-ї черги.

Час початку n-го циклу для і-ї черги:

$$t_{(поч.ці)}^{(n)} = t_{(зак.ці)}^{(n)} + t_{ці}. \quad (8)$$

Введемо значення $c_i(t)$:

$$c_i(t) = r_i - \frac{t}{b_i} + \frac{t_{непр.i} n_{ti}}{b_i} + \frac{(t - t_{поч.непр.i})d}{b_i} + \frac{t_{непр.0i} d_1}{b_i} + \frac{\left(t - t_{поч.прод.i}^{(n)}\right) d_2}{b_i} + \left(\frac{t}{t_{ц.г.i}} - 1\right) - - \sum_{j_i^t=1}^{j_i^{t+\Delta t}} \left(C_i(t - t_{ц.г.i}) d_3\right)_{j_i^t} \quad (9)$$

де r_i – початковий стан черг;

$$d = \begin{cases} 0, t > t_{поч.ці}^{(1)}, \\ 1, t \le t_{поч.ці}^{(1)}; \end{cases}, \qquad d_1 = \begin{cases} 0, t \le t_{поч.ці}^{(1)}, \\ 1, t > t_{поч.ці}^{(1)}; \end{cases}$$

$$d_2 = \begin{cases} 0, t - t_{поч.непр.i}^{(n)} \le 0, \\ 1, t - t_{поч.непр.i}^{(n)} > 0; \end{cases}, \quad d_3 = \begin{cases} 0, Ci\left(t - t_{ц.г.i}\right) \ge 0, \\ 1, C_i\left(t - t_{ц.г.i}\right) < 0; \end{cases} \quad (10)$$

Величина $c_i(t)$ – це кількість повідомлень і-ї черги, яке повинно бути оброблене до того, як наступить черга обробки повідомлень і-ї черги в момент часу t.

Введемо вектор управління:

$$T_k = \{t_{k1}, t_{k2}, \ldots, t_{ki}, \ldots, t_{kh}\}, \quad (11)$$

де t_{ki} – квант часу, який СОД віддає на обробку повідомлень *i*-го типу.

Вибираючи вектор управління T_k можливо знайти таке управління, при якому виконуються умови:

$$\begin{cases} w_i(t) + b_1 \le T_{порог.1} \\ w_i(t) + b_i \le T_{порог.i} \\ w_i(t) + b_h \le T_{порог.h} \end{cases} \quad (12)$$

$W_i(t)$ – функція часу очікування обробки повідомленням, що надійшло в СОД у момент часу t.

Висновки. В роботі наведено аналіз публікацій, в яких проведені дослідження розглянутої вище дисципліни, а також наведена математична модель, що дозволяє визначити головну характеристику обслуговування – функцію часу очікування обробки повідомлення, що поступило в СОД в момент часу t.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЛІТЕРАТУРА

1. Agner Erlang, biography [Internet resource]. – Режим доступу: https://en.wikipedia.org/wiki/Agner_Krarup_Erlang.

2. Яганов П. О. Дослідження систем масового обслуговування: Тексти лекцій. Київ: Національний технічний університет України «КПІ», 2006. 40 с.

3. Книш Б. П., Кулик Я. А., Лісовенко А. І. Метод розподілу часу на доставку товарів за допомогою безпілотних літальних апаратів згідно з пріоритетом. *Вісник Хмельницького національного університету*. 2018. № 6. С. 232-240.

4. Книш Б.П., Кулик Я. А., Барабан М. В. Класифікація безпілотних літальних апаратів та їх використання для доставки товарів. *Вісник Хмельницького національного університету*. 2018. № 3. С. 246 – 252.

5. Shvachych G. G., Moroz B. I. Pobochii I. A., Sushco L. F., Busydin V. V. *Комп'ютерне моделювання та оптимізація складних систем*. Матеріали 4 Міжнародної науково – технічної конференції (Дніпро, 1-2.11.2018) / уклад. кафедра інформаційних систем ДВНЗ УДХТУ. Дніпро: Баланс-Клуб, 2018. С.196 – 199.

6. Moroz B., Pokotylenco O. Analysis of development on creation of delivery organization systems using drones. *Комп'ютерно-інтегровані технології: «освіта, наука, виробництво»*. 2019. № 34. С. 74 – 78.

REFERENCES

1. Agner Erlang, biography [Internet resource]. Available: https://en.wikipedia.org/wiki/Agner_Krarup_Erlang.

2. Yaganov P. O. Investigation of mass service systems: Texts of lectures. Kiev: National Technical University of Ukraine "KPI", 2006. 40 p.

3. Knysh B. P., Kulik J. A., Lisovenko A. I. The method of the time distribution for the goods shipment by the means of unpiloted aerial vehicles based on a priority. *Bulletin of Khmelnitsky National University*. 2018. № 6. P. 232-240.

4. Knysh B. P., Kulik J. A., Baraban M. V. Classification of unmanned aerial vehicles and their use for delivery of goods. *Bulletin of Khmelnitsky National University*. 2018. № 3. P. 246 – 252.

5. Shvachych G. G., Moroz B. I. Pobochii I. A., Sushco L. F., Busydin V. V. *Computer modeling and optimization of complex system*. 4-th International scientific-technical conference. (Dnipro, 1-2.11.2018) / prod. by the the Department of Information Technology SHEI USCaTU. Dnipro : Balance-Club, 2018. P. 196 – 199.

Moroz B., Pokotylenco O. Analysis of development on creation of delivery organization systems using drones. *Computer-integrated technologies: "education, science, production"*. 2019. № 34. P. 74 – 78.

Received 19.02.2024.
Accepted 21.02.2024.

Mathematical model of the rational organization of the information flows processing in aircraft delivery system

The growing interest in unmanned aircrafts, quadcopters and other flying vehicles determines the need for dispatching incoming information flows from multiple users for the delivery of cargo by these flying vehicles. The issue of dispatching message flows is closely related to the choice of a model for the rational organization of processing such flows.

Analyzing the latest research and publications in this area, I would like to point the work of B. P. Knysh " The method of the time distribution for the goods shipment by the means of unpiloted aerial vehicles based on a priority." But the such priority mass service system can guarantees that the only highest priority goods will be delivered in time.

The purpose of this work is to describe a conceptual model of information flow processing in aircraft delivery systems using the discipline for the rational organization of information processing based on several characteristics. There are provides mathematical model that allows determining the main characteristic of the service - function of the waiting time for processing of a message received in the data processing system(DPS) at the time t.

Мороз Борис Іванович – д.т.н., професор, професор кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет «Дніпровська політехніка».

Круглик Андрій Сергійович – здобувач вищої освіти рівня доктор філософії, Національний технічний університет «Дніпровська політехніка».

Мороз Дмитро Максимович – доктор філософії, доцент кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет «Дніпровська політехніка».

Мартиненко Андрій Анатолійович – доктор філософії, доцент кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет «Дніпровська політехніка».

Boris Moroz - Doctor of Technical Sciences, Professor, Professor at the Department of Software Engineering, Dnipro University of Technology.

Andrii Kruhlyk – PhD student, Dnipro University of Technology.

Dmytro Moroz -Doctor of Philosophy (PhD), Associate Professor of the Software Engineering Department, Dnipro University of Technology.

Andrii Martynenko - Doctor of Philosophy (PhD), Associate Professor of the Software Engineering Department, Dnipro University of Technology.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-2-151-2024-02
УДК 004.67,004.042, 004.622

К. Руденко, Л. Божуха

ПРО ІНТЕГРАЦІЮ РЕКОМЕНДАЦІЙНИХ АЛГОРИТМІВ В СИСТЕМУ МОБІЛЬНОЇ ТОРГІВЛІ

Анотація. Розглянуто процес розроблення, впровадження та оптимізації рекомендаційних систем у сфері мобільної торгівлі з застосуванням хмарних сервісів. У роботі використано методи машинного навчання та рекомендаційні алгоритми з метою підвищення персоналізації та точності пропозицій. Результати роботи рекомендовані для підвищення персоналізації послуг у бізнесі (особливо в роздрібній торгівлі та онлайн платформах) для зростання задоволеності та лояльності клієнтів, а також для поліпшення маркетингових стратегій та аналізу трендів ринку.
Ключові слова: amazon personalize, amazon sagemaker, amazon services, рекомендаційні системи, персоналізація, мобільна торговля.

Постановка проблеми. Активне використання рекомендаційних систем в мобільних торгових додатках відкриває нові можливості для підвищення продажів та забезпечення високого рівня задоволення клієнтів. Ці системи, адаптовані до індивідуальних потреб та вподобань споживачів, дозволяють бізнесам ефективно реагувати на зміни у споживацьких трендах, підвищуючи лояльність та залученість клієнтів.

Важливою задачею стає інтеграція та оптимізація рекомендаційних систем у сфері мобільної торгівлі з застосуванням хмарних сервісів, що вимагає глибокого аналізу історії запитів і поведінкових патернів користувачів та вибору існуючих інформаційних технологій.

Аналіз останніх досліджень і публікацій. Рекомендаційні системи мають великий вплив у розвиток мобільної торгівлі, на залучення та утримання клієнтів, на оптимізацію маркетингових стратегій [1]. Використання машинного навчання забезпечує рекомендаційним системам можливість покращувати користувацький досвід та збільшувати продажі, створюючи особистісні та влучні пропозиції для кожного користувача [2]. Технологічний аналіз рекомендаційних систем включає детальне вивчення та оцінку різних алгоритмів та підходів, які є фундаментальними для їх функціонування.

Ключові алгоритми рекомендаційних систем охоплюють широкий спектр технік та підходів. У підході, заснованому на користувачах, система шукає користувачів із схожими інтересами або вподобаннями до заданого користувача та рекомендує продукти, які сподобалися цим схожим користувачам [4]. Контент-орієнтовані фільтри дозволяють системі визначати рекомендації на основі конкретних атрибутів продукту, таких як жанри, теми, стилі, автори та інші відомості [5]. Гібридні моделі використовується для вирішення обмежень, іноді пов'язаних з одним методом, забезпечуючи більшу гнучкість та точність у рекомендаціях [6].

Актуальним питанням є задача адаптації рекомендаційних алгоритмів до мобільних платформ при врахуванні обмежених обчислювальних ресурсів та специфіки мобільних інтерфейсів [7]. Використані алгоритми повинні бути достатньо ефективними для оброблення даних без значного навантаження на апаратне забезпечення пристроїв. Вирішення цієї задачі включає оптимізацію обчислень, мінімізацію використання пам'яті та забезпечення швидкої відповіді системи на запити користувачів. До важливих питань можна віднести задачу адаптації інтерфейсів рекомендаційних систем до мобільних екранів. Ефективне використання кешування та локального зберігання даних може допомогти зменшити навантаження на мережу та підвищити швидкість роботи додатків. Крім того, важливим є забезпечення безпеки та приватності даних користувачів, особливо в умовах зростання уваги до цих питань у цифровому світі.

Аналіз проблем у рекомендаційних системах набуває особливого значення, оскільки вони впливають на ефективність та точність цих систем. Дві ключові проблеми - це "холодний старт" та упередженість даних. Проблема холодного старту в рекомендаційних системах виникає, коли необхідно зробити рекомендації для нових користувачів або продуктів, про які система має мало або зовсім не має інформації. Гібридні системи, які комбінують колаборативну фільтрацію та контент-орієнтовані методи, можуть ефективно вирішувати проблему холодного старту, використовуючи переваги обох підходів [8].

Упередженість даних в рекомендаційних системах може мати різні причини від спотвореної або неповної інформації у зборі даних до навіть алгоритмічної упередженості [9].

Мета дослідження. Метою роботи є порівняльний аналіз різних підходів до інтеграції рекомендаційних алгоритмів у мобільну торгівлю, з акцентом на оптимізацію користувацького досвіду та підвищення ефективності продажів.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Значна увага приділяється також визначенню ключових викликів та можливостей, що стоять перед рекомендаційними системами в сучасному динамічному середовищі.

Викладення основного матеріалу дослідження.

Вибір архітектурного рішення. Для архітектурної організації проєкту обрана багаторівнева N-tier архітектура, що є класичним рішенням у сфері розробки клієнт-серверних додатків. Цей підхід передбачає розділення на різні рівні функцій представлення, бізнес-логіки та управління даними, кожен з яких може розміщуватися на окремих серверах або кластерах для оптимізації роботи та ресурсів. Запропонована архітектура спрощує масштабування, підтримку та ізоляцію потенційних проблем на окремих рівнях, покращуючи загальну надійність системи.

У програмного продукті приділена увага проєктуванню бази даних для категоризації товарів, системі ціноутворення та інформації по торговим точкам, задачі зберігання комерційних документів та ін. Інформація про торгові точки включає деталі про місцезнаходження, демографічні характеристики району, патерни трафіку клієнтів та інші ключові метрики, що сприяє у визначенні потенціалу ринку для впровадження нових продуктів або послуг та в управлінні ризиками, пов'язаними з географічним розміщенням. У мобільному додатку інформація про торгові точки та історію продажів може використовуватися для оптимізації робочих процесів.

Оцінка хмарних платформ для мобільної торгівлі. Хмарні платформи є новим інструментом для мобільної торгівлі, пропонуючи гнучкість, масштабованість та доступність інфраструктури, необхідної для сучасних бізнес-додатків. Вибір хмарних сервісів впливає на здатність їх інтегруватися з мобільними додатками, забезпечувати безперервну роботу, високий рівень безпеки та відповідність стандартам індустрії. Для вирішення поставленої задачі обраний Amazon Web Services для динамічних бізнес-вимог за рахунок його гнучкості, масштабованості та всебічним послугам. Amazon повністю автономно вирішує, чи є проблема класифікацією або регресією, а потім створює і тренує кілька конвеєрів з різними моделями. Отже, найкращий конвеєр можна легко розгорнути, тоді як прогнози потрібно генерувати, надсилаючи запити до змонтованої кінцевої точки. Це відрізняється від Amazon Personalize, де рекомендації можна отримати безпосередньо з графічного інтерфейсу, але Autopilot створює більше конвеєрів. Дані клієнтів зберігаються на льоту в DynamoDB і експортуються в S3 за допомогою AWS Data Pipeline. Потім викликається Athena

для виконання необхідного ETL і об'єднання інформації про сесії з даними клієнта. Нові дані зберігаються в S3-бакеті разом з історичними даними тренувань, а Sage Maker запускається для повторного запуску моделі з оновленими даними. Sage Maker може бути налаштований так, щоб він слугував кінцевою точкою для передачі даних назад клієнту. Частота оновлення моделі залежить від бізнес-використання. Якщо модель потрібно оновлювати часто, тоді виникає необхідність адаптувати частину оброблення даних у вище описаній архітектурі. У цьому випадку Amazon Kinesis можна використовувати для передачі нових даних до кластера EMR, який також зчитуватиме дані клієнтів, доступні в S3 bucket, для необхідного ETL. Також має сенс інтегрувати оновлення рекомендаційної моделі у фреймворк обробки даних, що працює на кластері EMR, оскільки це дозволить скористатися перевагами великої кількості вузлів, що використовуються в кластері, а також уникнути затримок між надходженням нових даних про сесію та оновленням моделі. Потім результати можна було б знову вставити в таблицю DynamoDB, з якої була б створена кінцева точка для обслуговування клієнтів з відповідними рекомендаціями.

Для зберігання даних використовується база даних ключ-значення або документів DynamoDB, яка масштабується до петабайтів даних. У DynamoDB за базову інфраструктуру відповідає AWS, за винятком пропускної здатності таблиці. У випадку з гарячими сесіями можна активувати прискорювач DAX, який може як зменшити латентність запитів.

AWS DataPipeline використовується для вилучення записів з DynamoDB і зберігання їх у S3. Об'єднання двох наборів даних (клієнтських і сеансових) можна здійснити за допомогою AWS Athena або Glue, які є безсерверними сервісами обробки даних в екосистемі AWS. Athena має бути сервісом вибору, якщо для обробки даних використовується в основному SQL. Glue краще підходить для патернів ETL, які значно відрізняються від простого SQL, або задіяно багато аналітики.

Для створення завдань з алгоритмами машинного навчання використаний сервіс AWS Sage Maker, який може автоматично створювати кінцеву точку та робити результати моделі машинного навчання легкодоступними для очикуємого на рекомендації клієнта.

Для захищення додатку від атак AWS впроваджує практики безпеки за допомогою додаткових налаштувань.

Методи збору та аналізу даних для рекомендацій. AWS, Azure та Google Cloud пропонують різноманітні інструменти для роботи з великими да-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ними, які можуть бути застосовані для розроблення рекомендаційних систем. Зокрема, AWS Personalize та Google Cloud AI використовують передові алгоритми машинного навчання для персоналізації пропозицій [10], тоді як Azure ML забезпечує широкий спектр сервісів для аналізу споживацької поведінки. Вибір конкретної платформи залежить від специфіки бізнес-вимог, масштабів даних та бажаної інтеграції з існуючими системами.

При проєктуванні рекомендаційної системи застовані алгоритми класифікації для точного сегментування клієнтів, що дозволяє персоналізувати пропозиції та покращити взаємодію з користувачами, забезпечуючи їм релевантність і цінність. Аналіз часових рядів є ключовим для розуміння сезонних тенденцій у продажах, що сприяє прогнозуванню майбутніх попитів та запасів. Методи зниження розмірності даних, як-от головні компоненти аналізу (PCA), можуть використовуватися для ефективної візуалізації та витягу інсайтів з великих наборів даних, полегшуючи інтерпретацію та прийняття рішень. Пакетний аналіз передбачає періодичну обробку даних. Важливим є вибір оптимального методу фільтрації та ефективного алгоритму навчання моделі при проєктуванні рекомендаційної системи.

Етапи розроблення рекомендаційної системи. У запропонованому рішенні рекомендаційна система є інтеграцією обраних сервісів (Amazon S3, Amazon RDS, Amazon SageMaker, та Amazon Personalize) у вже існуючу інфраструктуру мобільного додатку.

На основі навчальної вибірки за допомогою Amazon Personalize створена модель рекомендацій, яка використовує старіші 90% даних кожного користувача з тестового набору, як вхідні дані для новоствореного рекомендатора. Оцінка ефективності включає порівняння рекомендацій, сформованих моделлю, із реальними взаємодіями користувачів, використовуючи останні 10% даних тестового набору. Змінюючи запропоновані метрики можна сприяти подальшому удосконаленню системи рекомендацій, дозволяючи вносити необхідні корективи для підвищення задоволеності користувачів [11].

Система створена на Python 3.8 з Flask для бекенду та React з C# для клієнтського додатка. Дані обробляються та зберігаються через Amazon Web Services, включаючи Amazon S3 та RDS. Моделі розроблено та розгорнуто через Amazon SageMaker та Amazon Personalize.

Висновки. Використання технологій Amazon Personalize та SageMaker Data Wrangler дозволило оптимізувати оброблення великих об'ємів даних та надати гнучкість у розробці рекомендаційних систем. Розглянуто технологія

практичного впровадження алгоритмів машинного навчання у мобільний додаток для ефективної взаємодії з користувачами та підвищення їхньої лояльності. Інтеграція рекомендаційних алгоритмів в мобільні додатки при використанні сервісів хмарних послуг демонструє потенціал подальшого розвитку в динамічних галузях.

ЛІТЕРАТУРА

1. The rise of “big data” on cloud computing: Review and open research issues / I. A. T. Hashem та ін. Information Systems. 2015. Т. 47. С.98–115. URL: https://doi.org/10.1016/j.is.2014.07.006.

2. Ricci F. Mobile Recommender Systems. Information Technology & Tourism. 2010. Т. 12, № 3. С.205–231. URL: https://doi.org/10.3727/109830511x12978702284390.

3. AutoML and AutoAI - IBM Watson Studio. IBM in Deutschland, Österreich und der Schweiz | IBM. URL: https://www.ibm.com/products/watson-studio/autoai.

4. Content-based group recommender systems: A general taxonomy and further improvements / Y. Pérez-Almaguer та ін. Expert Systems with Applications. 2021. Т. 184. С. 115444. URL: https://doi.org/10.1016/j.eswa.2021.115444.

5. Dixit V. S., Gupta S., Jain P. A Propound Hybrid Approach for Personalized Online Product Recommendations. Applied Artificial Intelligence. 2018. Т. 32, № 9-10. С.785–801. URL: https://doi.org/10.1080/08839514.2018.1508773.

6. OmniSuggest: A Ubiquitous Cloud-Based Context-Aware Recommendation System for Mobile Social Networks / O. Khalid та ін. IEEE Transactions on Services Computing. 2014. Т. 7, № 3. С 401–414. URL: https://doi.org/10.1109/tsc.2013.53.

7. Linden G., Smith B., York J. Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Computing. 2003. Т. 7, № 1. С.76–80. URL: https://doi.org/10.1109/mic.2003.1167344.

8. Biased Recommender Systems And Supplier Competition / A. Fletcher та ін. SSRN Electronic Journal. 2023. URL: https://doi.org/10.2139/ssrn.4319311.

9. Autonomic Resource Management in a Cloud-Based Infrastructure Environment / B. K. Singh та ін. Autonomic Computing in Cloud Resource Management in Industry 4.0. Cham, 2021. С.325–345. URL: https://doi.org/10.1007/978-3-030-71756-8_18.

10. Dzulhikam D., Rana M. E. A Critical Review of Cloud Computing Environment for Big Data Analytics. 2022 International Conference on Decision Aid Sciences and Applications (DASA), м. Chiangrai, Thailand, 23–25 берез. 2022 р. 2022. URL: https://doi.org/10.1109/dasa54658.2022.9765168.

11. Zangerle E., Bauer C. Evaluating Recommender Systems: Survey and Framework. ACM Computing Surveys. 2022. URL: https://doi.org/10.1145/3556536.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

REFERENCES

1. The rise of “big data” on cloud computing: Review and open research issues / I. A. T. Hashem та ін. Information Systems. 2015. Т. 47. p.98–115. URL: https://doi.org/10.1016/j.is.2014.07.006.
2. Ricci F. Mobile Recommender Systems. Information Technology & Tourism. 2010. Т. 12, № 3. p.205–231. URL: https://doi.org/10.3727/109830511x12978702284390.
3. AutoML and AutoAI - IBM Watson Studio. IBM in Deutschland, Österreich und der Schweiz | IBM. URL: https://www.ibm.com/products/watson-studio/autoai.
4. Content-based group recommender systems: A general taxonomy and further improvements / Y. Pérez-Almaguer та ін. Expert Systems with Applications. 2021. Т. 184. С. 115444. URL: https://doi.org/10.1016/j.eswa.2021.115444.
5. Dixit V. S., Gupta S., Jain P. A Propound Hybrid Approach for Personalized Online Product Recommendations. Applied Artificial Intelligence. 2018. Т. 32, № 9-10. С.785–801. URL: https://doi.org/10.1080/08839514.2018.1508773.
6. OmniSuggest: A Ubiquitous Cloud-Based Context-Aware Recommendation System for Mobile Social Networks / O. Khalid та ін. IEEE Transactions on Services Computing. 2014. Т. 7, № 3. P.401–414. URL: https://doi.org/10.1109/tsc.2013.53.
7. Linden G., Smith B., York J. Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Computing. 2003. Т. 7, № 1. p.76–80. URL: https://doi.org/10.1109/mic.2003.1167344.
8. Biased Recommender Systems And Supplier Competition / A. Fletcher та ін. SSRN Electronic Journal. 2023. URL: https://doi.org/10.2139/ssrn.4319311.
9. Autonomic Resource Management in a Cloud-Based Infrastructure Environment / B. K. Singh та ін. Autonomic Computing in Cloud Resource Management in Industry 4.0. Cham, 2021. p.325–345. URL: https://doi.org/10.1007/978-3-030-71756-8_18.
10. Dzulhikam D., Rana M. E. A Critical Review of Cloud Computing Environment for Big Data Analytics. 2022 International Conference on Decision Aid Sciences and Applications (DASA), м. Chiangrai, Thailand, 23–25 берез. 2022 р. 2022. URL: https://doi.org/10.1109/dasa54658.2022.9765168.
11. Zangerle E., Bauer C. Evaluating Recommender Systems: Survey and Framework. ACM Computing Surveys. 2022. URL: https://doi.org/10.1145/3556536.

Received 20.02.2024.
Accepted 24.02.2024.

Integration algorithms of recommendation in the mobile trade system

The algorithms of recommender systems must be efficient enough to process data without a significant load on the hardware of the devices. The solution to this problem includes optimization of calculations, minimization of memory usage and provision of quick response of the system to user requests.

An important task is the integration and optimization of recommender systems in the field of mobile commerce using cloud services.

The purpose of the work is a comparative analysis of different approaches to the integration of recommendation algorithms in mobile commerce, with an emphasis on optimizing the user experience and increasing sales efficiency.

In the software product, attention are paid to the design of a database for product categorization, a pricing system and information on retail outlets, the task of storing commercial documents, etc.

Amazon Web Services was chose for dynamic business requirements due to its flexibility, scalability and comprehensive services. To create tasks with machine learning algorithms, the AWS Sage Maker service was used. When designing a recommender system, classification algorithms are using for accurate segmentation of customers. Time series analysis is key to understanding seasonal trends in sales, which helps predict future demand and inventory. Data dimensionality reduction techniques such as principal component analysis (PCA) can be used to efficiently visualize and extract insights from large data sets, facilitating interpretation and decision making. Batch analysis involves periodic processing of data.

In the proposed solution, the recommendation system is the integration of selected services (Amazon S3, Amazon RDS, Amazon SageMaker, and Amazon Personalize) into the already existing mobile application infrastructure.

The technology of practical implementation of machine learning algorithms in a mobile application for effective interaction with users and increasing their loyalty is considered. The integration of recommendation algorithms into mobile applications when using cloud services demonstrates the potential for further development in dynamic industries.

Keywords: amazon personalize, amazon sagemaker, amazon services, recommendation systems, personalization, mobile commerce.

Божуха Лілія Миколаївна - к.ф.-м.н., доцент, доцент Дніпровського національного університету ім..Олеся Гончара.

Руденко Кирило Олегович – магістр Дніпровського національного університету ім..Олеся Гончара.

Bozhukha Liliia - Ph.D., associate professor, associate professor of Dnipro National University named after Oles Honchar.

Rudenko Kyrylo - Master of Dnipro National University named after Oles Honchar.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-2-151-2024-03
УДК 621.385.6

А.Ю. Зимогляд, А.І. Гуда, С.М. Кліщ

АПАРАТНИЙ КОМПЛЕКС ДЛЯ ВИМІРЮВАННЯ ПОТУЖНОСТІ UHF СИГНАЛІВ

Анотація. У статті описана схема та конструкція апаратного пристрою, для вимірювання потужності радіосигналів діапазону UHF. Розроблений апаратний пристрій може вимірювати потужність сигналу в діапазоні частот 0.8 – 6 GHz.
Ключові слова: UHF, вимірювання потужності UHF сигналів, радіо керовані пристрої.

Вступ

В сучасному світі все більше пристроїв переходять на діапазон UHF (Ultra High Frequency) сигналів. Сучасні телефони з підтримкою стандарту 5G отримають і передають данні на частотах 4.8 – 4.9 GHz, частоти близько 5.8 GHz використовуються в радіокерованому авіамоделізмі для пілотування за зображенням відеокамери. Авіамоделі для керування використовують частоти 0.9 – 2.4 GHz. Зв'язок із космічними апаратами на орбіті Землі (зокрема і супутниковому телебаченні) використовує переважно діапазони C (3.4 – 7 GHz) і Ku (12 – 18 GHz). Це лиш мала частина сучасного світу де використовують діапазон UHF [1,2].

Для досліджень приймально - передавальних пристроїв, а також антен в діапазонах UHF потрібно обладнання великої вартості. Часто саме це і зупиняє дослідження з цим діапазоном хвиль. Проте якщо виділити область 0.8 – 6 GHz, то можливо зробити деякі пристрої для оцінки потужності сигналу на основі сучасних мікросхемотехнічних рішень.

Звичайно комплекс, що буде розглянуто у цієї роботі, не дозволить замінити повноцінний спектральний аналізатор або інший високочастотний вимірюючий пристрій. Проте стане можлива якісна оцінка, стане можливо оцінити змінився сигнал або яка антенна більше передає сигнал. Це вказую на те, що розробка такого комплексу є актуальною задачею.

Підсилювач UHF сигналів на базі QPL9547

Для підсилення UHF (Ultra High Frequency) сигналів застосовувався підсилювач на базі мікросхеми QPL9547. Дана мікросхема QPL9547 – це високо

лінійний, надмало шумний підсилювач у невеликому корпусі розміром 2x2 mm для поверхневого монтажу. На частоті 1.9 GHz підсилювач зазвичай забезпечує максимальне посилення +19.5 dB або +39 dBm на навантаженні 50 Ω. При установці струму зміщення в 65 mA коефіцієнт шуму – 0 3 dB. Мікросхема працює в широкому діапазоні частот від 0.1 GHz до 6 GHz. На максимальній частоті підсилення складає +12 dB.

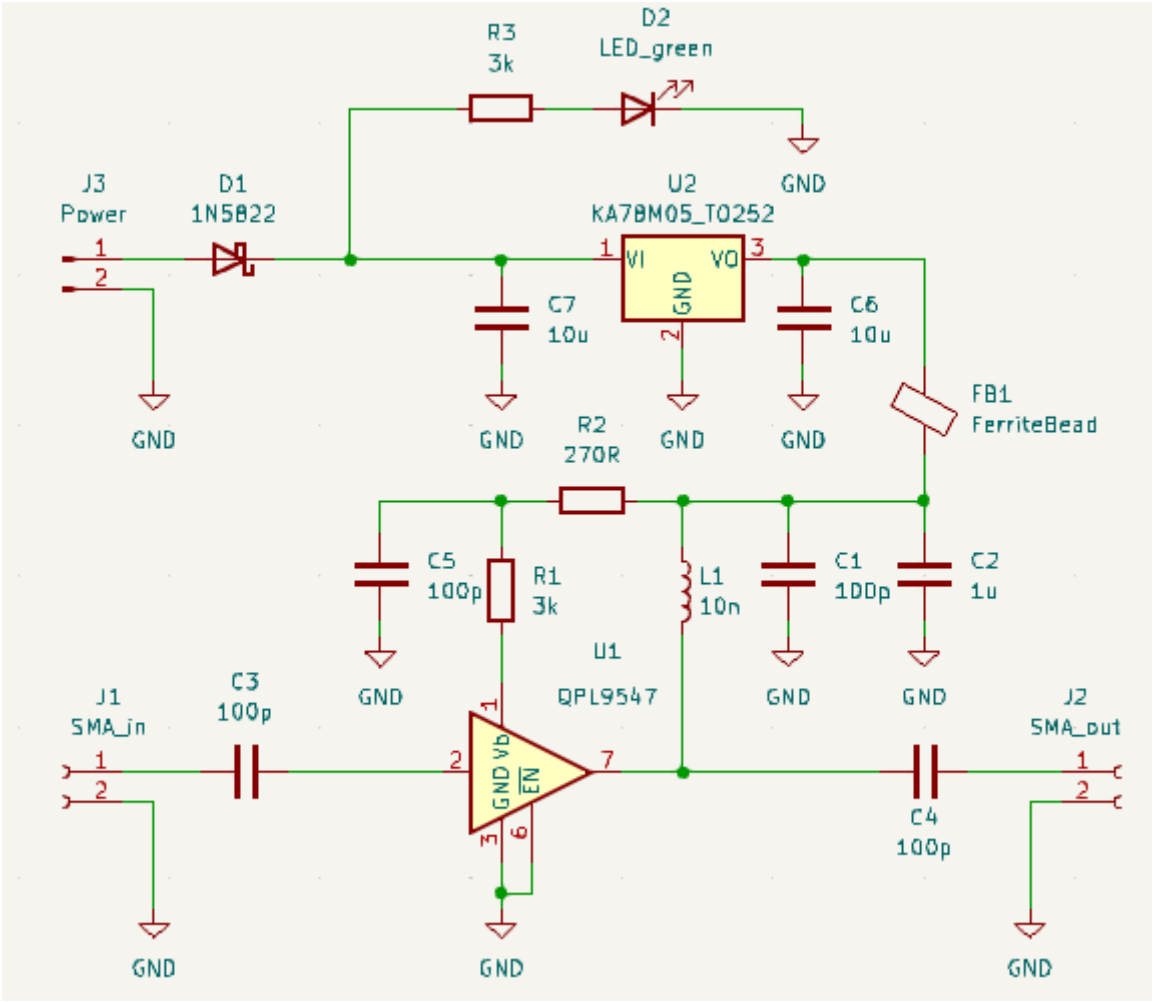


Рисунок 1 – Схема UHF підсилювача на основі QPL9547

На рис. 1 представлено розроблену схему підсилювача. Основним елементом схеми є мікросхема U1 – QPL9547. Для відокремлення постійного струму, а також частот нижчих за 0.1 GHz на вході і виході схеми стоять розділяючи конденсатори C3 та C4. Індуктивність L1 (10 nH) було використано в корпусі 0805 фірми Coilcraft з спеціальної серії для UHF сигналів. Одразу біля неї стоять конденсатори C1 та C2 для блокування проходження над високочастотного сигналу в шину живлення. Також для стабільної роботи QPL9547 живиться від стабілізатору напруги 78M05, який забезпечує стабілізоване живлення +5 V. Для

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

захисту від збудження мікросхеми лінійного стабілізатору живлення 78M05 над високочастотними сигналами стоїть феритова бусина FB1. Без неї збуджувався лінійний стабілізатор при підсиленні сигналів після 500 MHz. Хоча в документації на мікросхему підсилення вказується, що можна для встановлення струму зміщення в 65 mA використовувати один резистор в 3.2 kΩ, в реальності для запобігання збудження підсилювача треба використовувати послідовне поєднання 2 резисторів, а саме R1 та R2 з точки поєднання яких стоїть конденсатор C5. Ця схема є простішим ФНЧ (Фільтром Низької Частоти) і захищає ланцюг зміщення від попадання високочастотних сигналів. Діод Шоттки D1 захищає від неправильної полярності живлення, а світлодіод D2 – індикатор живлення.

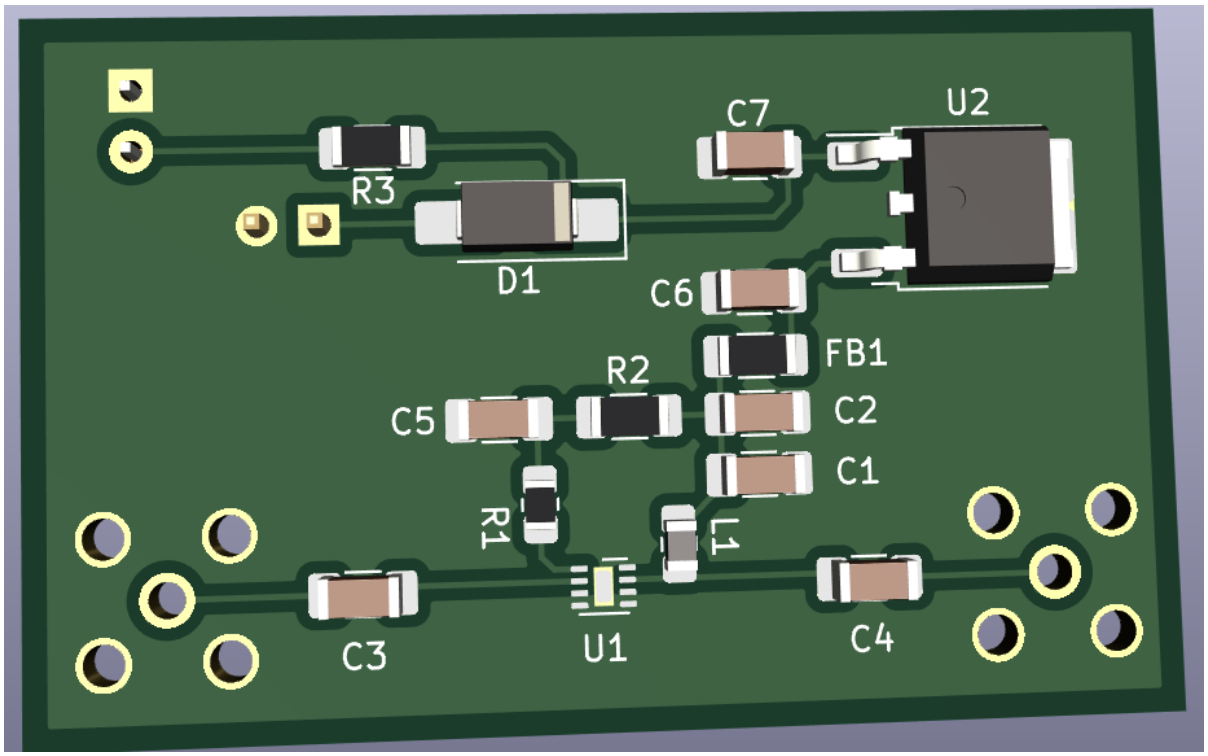


Рисунок 2 – 3Д модель друкованої плати в KiCAD

Оскільки підсилювач UHF сигналів працює в діапазоні частот від 0.1 GHz – 6 GHz, до плати пред'являються особливі вимоги. По перше – це суцільне заливання земляним полігоном без термобар'єрів у місцях пайки, по друге – це досить шильне встановлення компонентів один до одного (особливо компонентів, що блокують над високочастотний сигнал). По третє – потрібно враховувати певні параметри плати, такі як довжина, товщина доріжок та відстань їх до земляного полігону. Параметри плати можна взяти з документації на мікросхему, в який приводяться зразки правильної розводки

плати та вимоги до матеріалу текстоліту. На рис. 2 зображено 3D модель друкованої плати в середовищі KiCAD. Параметри плати задовольняють вказаним вимогам.

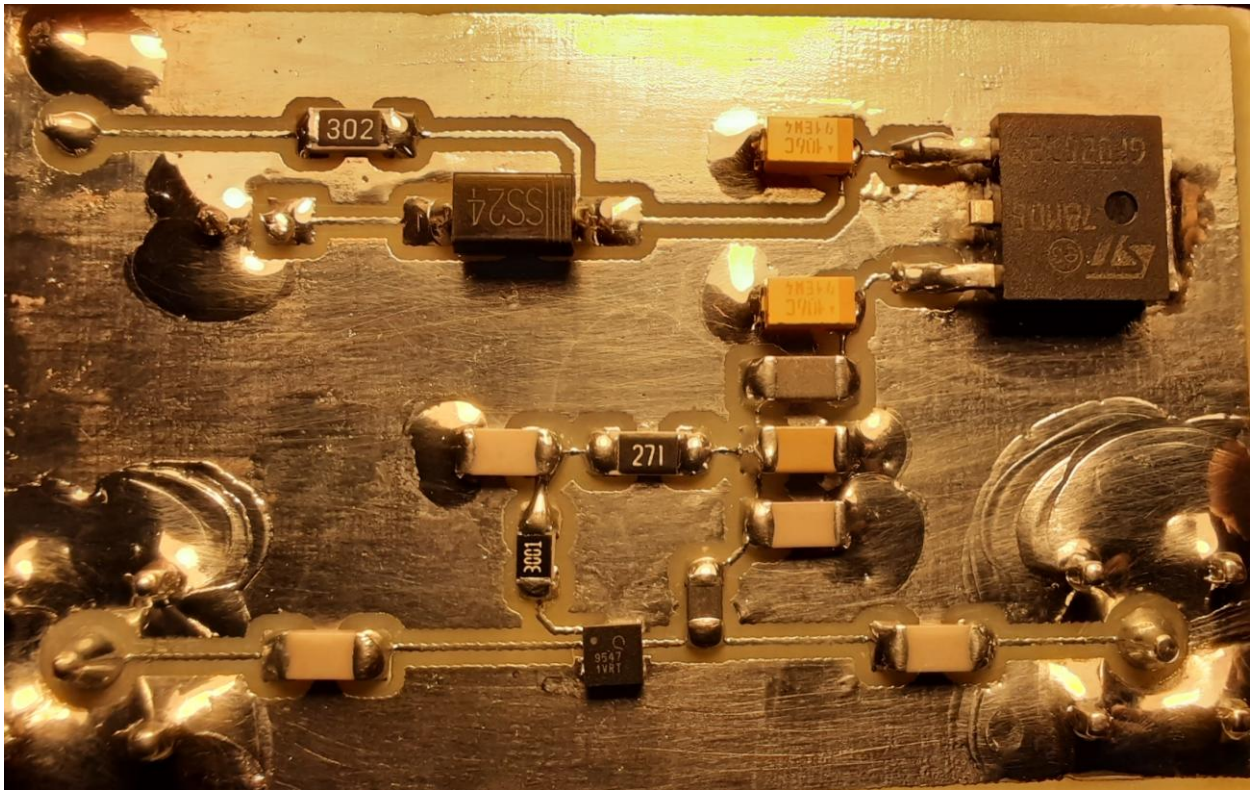


Рисунок 3 – Розроблена плата UHF підсилювача

На рис. 3 представлено фотографію розробленого підсилювача. Плата після витравлювання покривалась припоєм з вмістом срібла 5% для кращої роботи в UHF діапазоні частот.

Для перевірки працездатності UHF підсилювача було зроблено тестовий стенд (рис. 4). В якості джерела вхідного сигналу використовувався векторний аналізатор спектру nanoVNA. В ньому є режим CW, при якому можливо його використовувати, як генератор ВЧ сигналу близького до синусу. При цьому сигнал береться з виходу S11, потужність вихідного сигналу з виходу S11 – +10 dBm.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

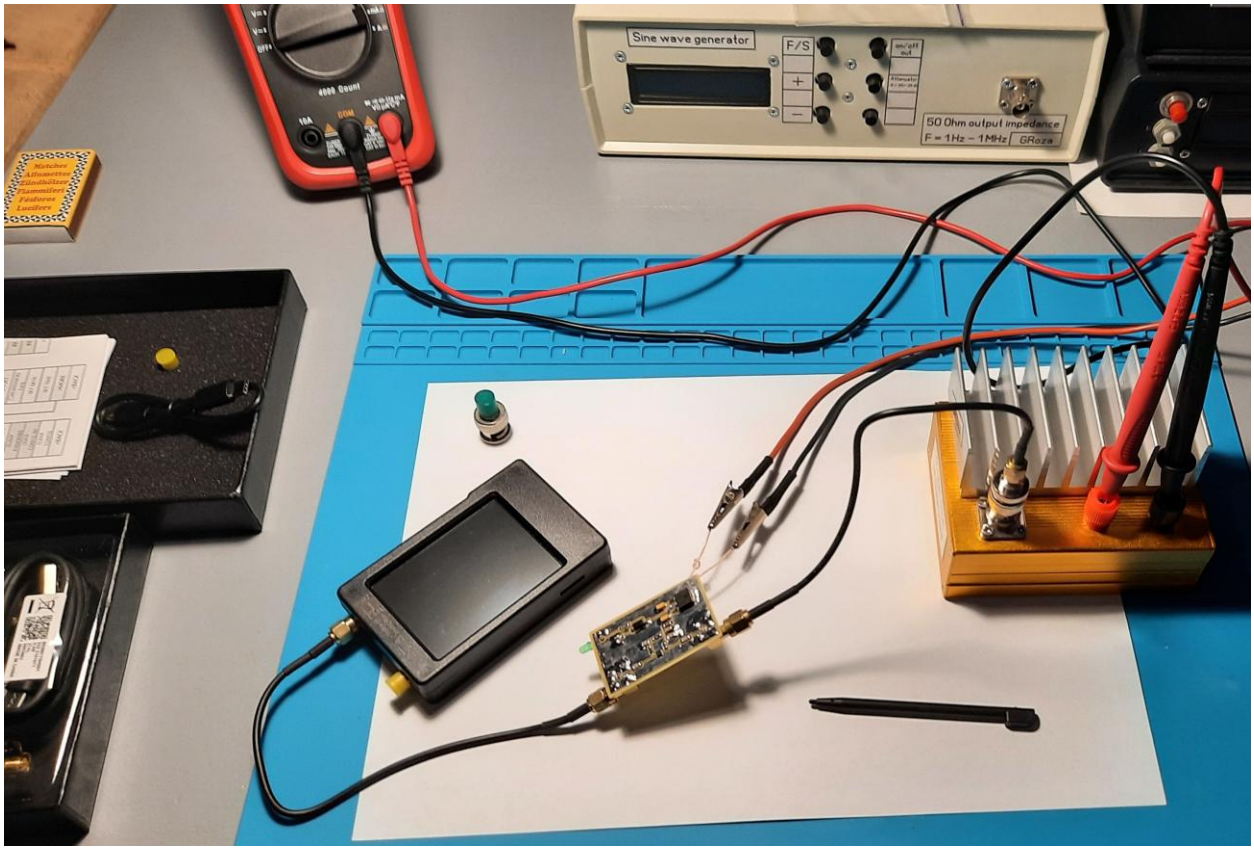


Рисунок 4 – Тестовий стенд для дослідження UHF підсилювача

В якості генератора nanoVNA може генерувати сигнали від 50 kHz до 1.5 GHz. Це було використано для дослідження властивостей підсилення, а також перевірки працездатності. Для цього подавався сигнал від 100 MHz до 1.5 GHz на вхід створеного підсилювача на базі QPL9547. В якості навантаження підсилювача використовувався спеціальний UHF резистор 50 Ω, який був поміщений у металевий корпус. В цьому корпусі окрім резистора також знаходиться найпростіший піковий детектор з діода 1N5711 та конденсатора 100 nF. В якості джерела живлення використовувався лабораторний блок живлення, на якому було встановлено напругу 12 V і струм обмеження 500 mA. Вимірювалось постійна напруга на піковому детекторі, що, у відповідності до документації, дозволяє оцінити потужність сигналу:

$$P = \frac{((U + 0.35) \cdot 0.707)^2}{50} \quad (1)$$

Для перетворювання напруги U на піковом детекторі у потужність P, що виділяється на резисторі 50 Ω було використано формулу 1. Формула також враховує падіння на діоді в 0.35 V.

Спочатку вимірювалась потужність насамперед від джерела вхідного сигналу nanoVNA на навантаження 50 Ω, потім між навантаженням в 50 Ω і джерелом сигналу встановлювався підсилювач на QPL9547. За допомогою вище описаного стенду було перевірено підсилювач у діапазоні частот від 100 MHz до 1.5 GHz при цьому підсилювач видав підсилення +23 dB. На жаль, за відсутністю джерела на більш високі частоти вимірювання підсилювання не відбувалось більше 1.5 GHz, проте можна сказати про працездатність схеми.

Вимірювач потужності UHF сигналу на базі AD8319

Для вимірювання потужності UHF сигналу було прийнято рішення використовувати сучасну спеціалізовану мікросхему – AD8319. AD8319 – це демодулюючий логарифмічний підсилювач, здатний виконувати точне перетворення вхідного сигналу UHF на відповідну вихідну напругу в логарифмічному масштабі. Він заснований на методі прогресивної компресії і складається з послідовного з'єднання підсилювальних каскадів, кожен з яких має власний елемент, що детектує. Компонент може працювати в режимі вимірювання або контролера. AD8319 підтримує високу точність логарифмічної характеристики під час роботи з сигналами в смузі частот від 1 MHz до 8 GHz та забезпечує прийнятну якість при частотах до 10 GHz. Типовий діапазон вхідних сигналів, у якому підтримується похибка менше ±1 dB, становить 40 dB (при навантаженні 50 Ω). AD8319 має час відгуку 8 ns/10 ns (час наростання/час спаду), що дозволяє детектувати за його допомогою імпульсні радіосигнали з частотою повторення понад 50 MHz. У документації заявлено, що компонент має безпрецедентну стабільність точки перетину логарифмічної характеристики при змінах температури навколишнього середовища. Для роботи AD8319 необхідне одне джерело напруги живлення в діапазоні від 3.0 V до 5.5 V. Типовий струм споживання дорівнює 22 mA, і зменшується до 200 mkA при перемиканні компонента в неактивний стан. AD8319 можна налаштувати для формування керуючої напруги, що подається на підсилювач зі змінним коефіцієнтом посилення, або видачі результату вимірювань на виводі VOUT. Оскільки вихідний сигнал може бути використаний у завданнях управління, при проектуванні компонента особливу увагу було приділено мінімізації широкосмугового шуму.

У режимі контролера керуюча напруга, що задає робочу точку підсилювача, прикладається до виводу VSET. Контур керування підсилювачем ВЧ замикається на вивод VOUT, вихідний сигнал, на якому керує вихідним сигналом підсилювача. Шляхом зміни напруги VOUT рівень вихідного сигналу підсилювача встановлюється рівним величині, що задається за допомогою

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

VSET. Вихідна напруга на виводі VOUT AD8319 має діапазон від 0 до (VPOS – 0.1 V), достатній для завдань управління підсилювачами. Для роботи в якості вимірювального пристрою вивід VOUT з'єднується зовнішнім ланцюгом з виводом VSET, і компонент формує вихідну напругу, VOUT, що є спадною, лінійною в логарифмічному масштабі функцією амплітуди вхідного радіосигналу. Номінальний нахил логарифмічної характеристики дорівнює −22 mV/dBm і визначається ланцюгом VSET. Точка перетину становить 15 dB (відносно 50 Ω, для не модульованого сигналу) і задається з допомогою входу INHI. Ці параметри зберігають високу стабільність при змінах напруги живлення та температури. AD8319 виробляється за кремній-германієвою (SiGe) технологією виготовлення IMC на біполярних транзисторах, випускається в 8-вивідний корпус LFCSP_VD, що має габарити 2 мм × 3 мм, і працює в температурному діапазоні від -40°C до +85°C.

Все вище зазначене дозволяє використовувати AD8319 в нашій задачі в режимі вимірювального пристрою. Було спроектовано схему (рис. 5) з урахуванням усіх особливостей описаних в документації на AD8319.

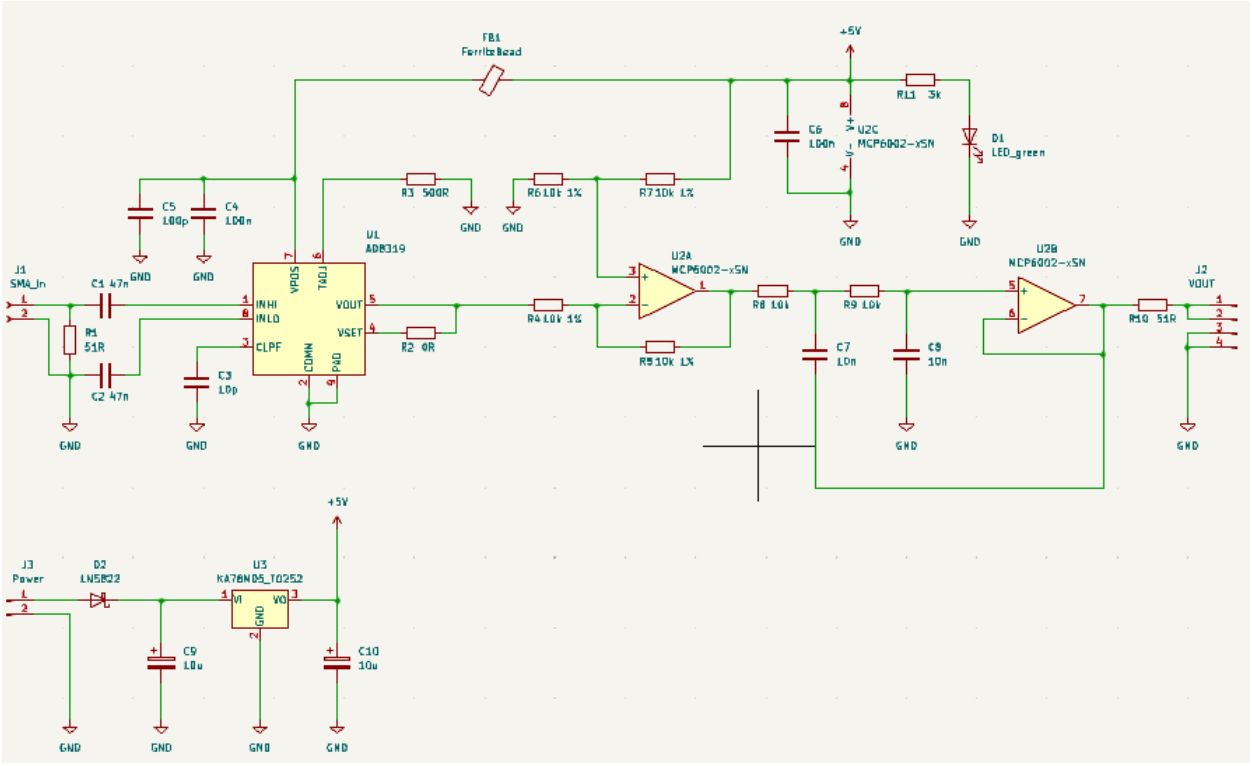


Рисунок 5 – Схема UHF вимірювача потужності на базі демодулюючого логарифмічного підсилювача AD8319

На вході схеми UHF вимірювача потужності стоїть резистор R1 номіналом 51 Ω, в якості навантаження. З нього диференційно знімається сигнал через 2

конденсатори C1 та C2 номіналом 47 nF, після них сигнал поступає на входи мікросхеми AD8319. Конденсатор C3 обирається для фільтрації, для діапазону частот 1–6 GHz його номінал становить 10 pF. Резистор R3 номіналом в 510 Ω задає температурну компенсацію AD8319. Конденсатори C5 та C4 стоять по живленню мікросхеми і також блокують подальше просування UHF сигналу в шину живлення. Для стабільного живлення схеми стоїть лінійний стабілізатор 78M05, який в свою чергу подає напругу +5 V через феритову бусину FB1 на AD8319, а на операційний підсилювач MCP6002 напряму. Відокремлення живлення MCP6002 від AD8319 феритовою бусиною є в край важливим, щоб не мати похибок на виході пов'язаних з протіканням UHF сигналу через шину живлення.

Операційний підсилювач MCP6002 використовується для двох задач: перша це – інвертувати вихідний сигнал, друга – це фільтрація вихідного сигналу.

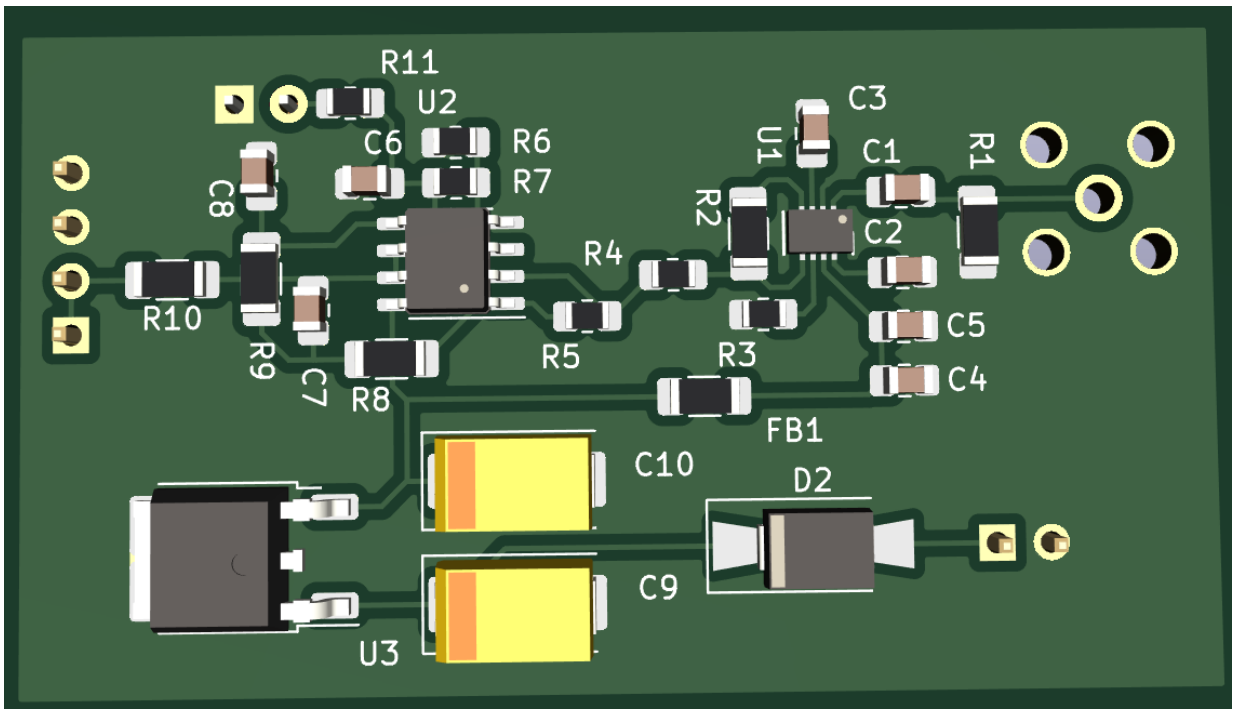


Рисунок 6 – 3Д модель UHF вимірювача потужності на базі демодулюючого логарифмічного підсилювача AD8319 в KiCAD

AD8319 формує вихідну напругу VOUT, що є спадною, лінійною в логарифмічному масштабі функцією амплітуди вхідного радіосигналу з номінальним нахилом логарифмічної характеристики −22 mV/dBm, тому при максимальну сигналі на виході буде приблизно 0.3 V, а при мінімальному приблизно 1.5 V.

Щоб полегшити зчитування сигналу була зроблена схема, що перегортає шкалу на диференційному підсилювачі. Один з входів диференційного підси-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

лювача поєднано до шини +5 V інший на вихід AD8319. За рахунок цього тепер при найменшому сигналі −60 dBm буде 3.5 V, а при сигналі в 0 dBm (1 mW при 50 Ω) буде 4.7 V. Також операційний підсилювач виконує ще функцію узгодження вихідного опору AD8319. Нахил в 22 mV/dBm при цьому зберігається (рис. 6).

Фільтр після диференційного підсилювача побудований по схемі Саллена Кі. Це фільтр другого порядку низької частоти, він необхідний для того, щоб на виході UHF вимірювача потужності не було високочастотного шуму, який може бути на виході VOUT AD8319.

Створена плата UHF вимірювача потужності на базі AD8319 (рис. 7).

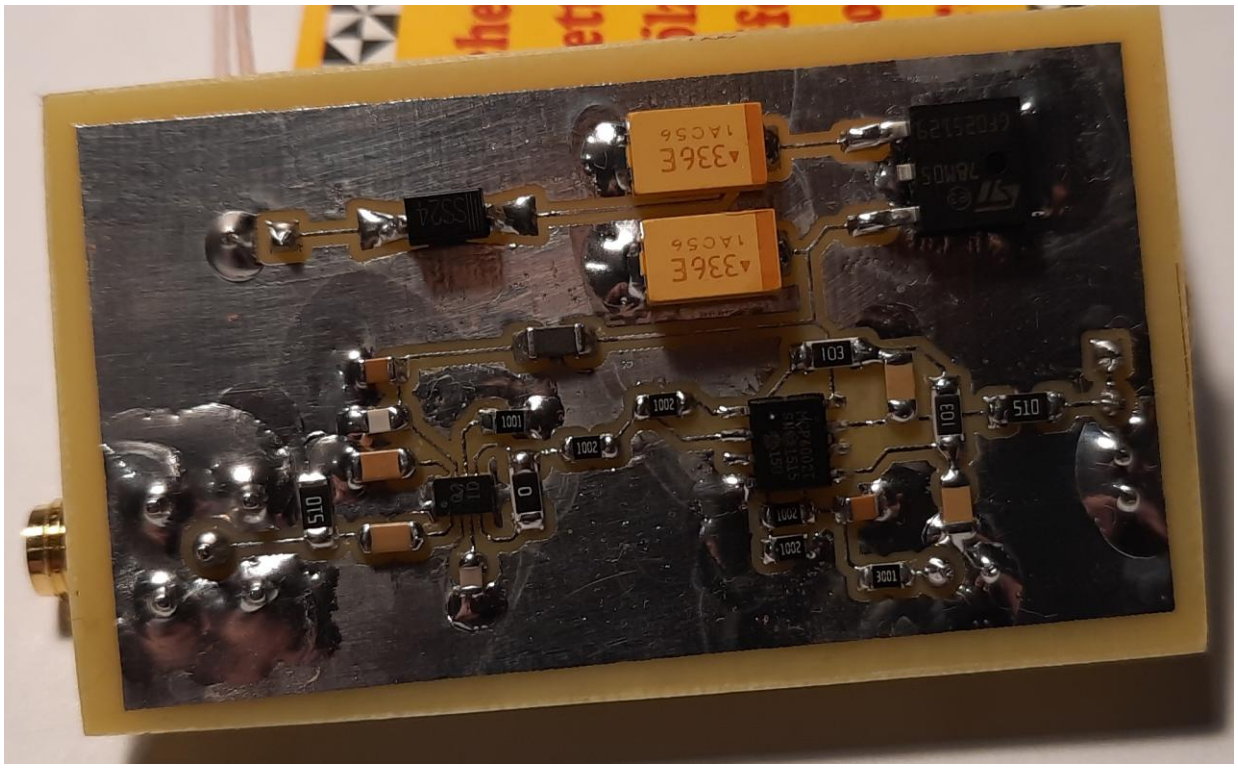


Рисунок 7 – Створена плата UHF вимірювача потужності на базі AD8319

Для перевірки працездатності UHF вимірювача потужності був проведено експеримент, при якому з джерела вхідного сигналу nanoVNA 1 GHz (+10 dBm) через атенюатори подавався сигнал на вхід розробленої схеми з AD8319 та MCP6002. Вихідний сигнал вимірювався мультиметром UT136C+.

Таблиця 1

Вихідна напруга на UHF вимірювачі в залежності від потужності вхідного сигналу

Вхідна потужність (dBm)	Вихідна напруга (V)
0	4.80
-10	4.52
-20	4.27
-30	4.05
-40	3.77
-50	3.70
-60	3.60

Представлені данні в таблиці вказують на працездатність створеного UHF вимірювача. Графік цієї залежності представлено на рис. 8.

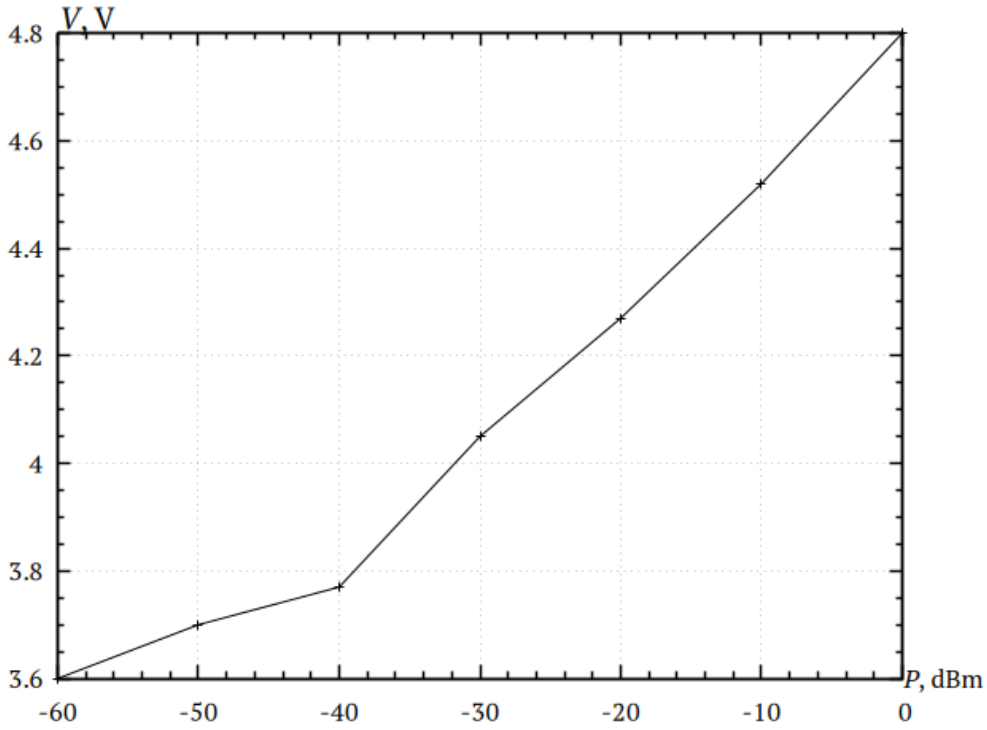


Рисунок 8 – Залежність вихідної напруги V від потужності вхідного сигналу

Залежність проявляє лінійність на діапазоні –40 – 0 dBm. Втрата лінійності на більш слабких сигналах пов'язана, скоріш усього, з суттєвим рівнем зовнішніх завад.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Висновки. Описаний апаратний комплекс, який складається з 2 блоків, а саме з підсилювача UHF сигналів на мікросхемі QPL9547 та демодулюючого логарифмічного підсилювача на AD8319 показали свою працездатність.

Результати перевірки, що були представленні у табл. 1 показують, як реагував UHF вимірювач потужності сигналів на низьку вхідних сигналів різної потужності.

Цей комплекс може бути використано у лабораторних дослідженнях вихідної потужності передаючих пристроїв, для якісної оцінки антен або порівнянь антен. За рахунок підсилювача на вході можливо дослідження потужності сигналу близько до –80 dBm.

Описаний комплекс також має доволі помірну ціну, в порівнянні з промисловими аналогами.

ЛІТЕРАТУРА / REFERENCE

1. Radio Frequency Transistors: Principles and Practical Applications. Second Edition (EDN Series for Design Engineers) / *Norman Dye, Helge Granberg* /2001/ ISBN 978-0-7506-7281-8 / DOI https://doi.org/10.1016/B978-0-7506-7281-8.X5017-6

2. The VHF/UHF DX Book *Radio Society of Great Britain / Roger Blackwell, David Butler, Geoff Grayer, Günter Hoch, Sam Jewell, John Nelson, Dave Powis, Dave Robinson, IIan White, John Wilkinson / 1995/ ISBN 0 9520468 0 6*

Received 23.02.2024.
Accepted 27.02.2024.

Hardware complex for measuring the power of UHF signals

The article describes the scheme and construction of a hardware device for measuring the power of UHF radio signals. The developed hardware device can measure the signal strength in the frequency range of 0.8 – 6 GHz.

For the research of receiving and transmitting devices, as well as antennas in the UHF bands, expensive equipment is required. This is often what stops research with this wave range. However, if we single out the 0.8-6 GHz range, it is possible to make some devices for evaluating the signal power based on modern microcircuit engineering solutions. Of course, the complex that will be considered in this work will not replace a full-fledged spectrum analyzer or other high-frequency measuring device. However, a qualitative assessment will be possible, it will be possible to assess the changed signal or which antenna transmits the signal more. This indicates that the development of such a complex is an urgent task.

A hardware complex has been developed, which consists of 2 units, the UHF signal amplifier on the QPL9547 microcircuit and the demodulating logarithmic amplifier on

the AD8319. The results of the research, which were presented in the table. 1 show how the UHF signal strength meter responded to a series of input signals of varying strength.

This complex can be used in laboratory studies of the output power of transmitting devices, for qualitative evaluation of antennas or antenna comparisons. Due to the amplifier at the input, it is possible to study the signal power up to –80 dBm. The described complex also has a fairly moderate price, compared to industrial analogues.

Key words: UHF, measuring the power of UHF signals, radio-controlled devices.

Зимогляд Андрій Юрійович - к.т.н., доцент кафедри ІТС УДУНТ.
Гуда Антон Ігорович - д.т.н., професор кафедри ІТС УДУНТ.
Кліщ Сергій Михайлович – старший викладач кафедри ІТС УДУНТ.

Zimoglyad Andrew Yuriyovych - Ph.D., associate professor of the ITS department of UDUNT.
Guda Anton Igorovich - doctor of technical sciences, professor of the ITS department of UDUNT.
Klishch Serhii - senior teacher of the ITS department of UDUNT.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-04
УДК 004.42:519.25

О.Д. Земляний, О.Г. Байбуз

МЕТОДИ ІМПУТУВАННЯ ПРОПУСКІВ У ДАНИХ ПРО ІШЕМІЧНУ ХВОРОБУ СЕРЦЯ

Анотація. Пропонуються декілька модифікацій алгоритмів ітеративного множинного імпутування пропусків у змішаних даних, що представлені кількісними і якісними ознаками. Серед кількісних є і неперервні, і дискретні. Серед якісних є порядкові та бінарні. Для аналізу підходів використовується два типи тестів. В першому тесті датасет з відомими даними штучно заповнюється пропусками у випадкових позиціях, проводиться імпутування різними методами, оцінюється середньо квадратична похибка та час виконання алгоритмів. В другому тесті навчають моделі бінарної класифікації на датасетах, з імпутацією пропусків різними методами, та порівнюють точність класифікації на тестовій вибірці. Для перетворення якісних ознак на кількісні запропоновано власні алгоритми, які працюють з пропущеними даними та дозволяють виконувати зворотне перетворення знов до якісних ознак. Розглядаються два відомих датасети про спостереження стосовно ішемічної хвороби серця.

Ключові слова: імпутування даних, ітеративне множинне імпутування даних, обробка змішаних даних, регресія, бінарна класифікація, перетворення якісних ознак на кількісні, мова програмування python.

Вступ. У сфері науки про дані дослідницький аналіз є важливим етапом. Цей аналіз дозволяє дослідити певні характеристики даних, виявити закономірності, аномалії, провести очищення від викидів та побудувати початкові моделі для подальшої роботи. На цьому етапі можна встановити вид розподілу, оцінити його основні параметри, виявити викиди, побудувати матрицю кореляції ознак і таке інше.

У попередньому аналізі досить серйозною проблемою є виявлення пропущених значень, і найскладнішим є те, що немає жодного універсального алгоритму. Для кожної конкретної задачі доводиться підбирати відомі методи, їх комбінації, модифікації, або зовсім нові підходи.

Більшість моделей машинного навчання не можуть обробляти пропущені значення, тому ми не можемо просто проігнорувати пропуски в даних. Проблему пропущених даних необхідно вирішувати під час попередньої обробки.

Найпростішим рішенням є видалення кожного спостереження, що містить пропущені значення. Таке рішення реалізовано у відомих бібліотеках мови програмування Python, таких як Numpy або Pandas. Але цей підхід є крайнім випадком, бо ми втрачаємо всю корисну інформацію, яка може бути важливою для аналізу даних.

Існують наступні основні стратегії імпутування пропущених даних:

1. Замінити пропущені значення середнім/медіаною чи модою.
2. Замінити найчастішим значенням, що зустрічається, або константою.
3. Імпутування даних за допомогою методу k найближчих сусідів (алгоритм kNN).
4. Множинне імпутування пропущених даних (алгоритм MICE).
5. Імпутування даних за допомогою глибокого навчання.

Ці методи застосовуються для вирішення проблеми відсутніх даних у наборах даних.

Постановка проблеми в загальному вигляді. Розглядаємо два популярних набори даних, що є у відкритому доступі на платформі Kaggle.com. Це збірний датасет UCI Heart Disease Data [1, 2], створений чотирма кардіологічними центрами Угорщини, Швейцарії та США, та набір даних Framingham Heart Study [3, 4], який отримано з активного кардіологічного дослідження мешканців Массачусетсу та Фремінгему.

В обох наборах даних наявні пропуски. Дані містять порядкові якісні та кількісні показники, причому процент якісних показників вище кількісних. В наборі даних UCI Heart Disease Data частина пропусків спричинена об'єднанням кількох датасетів, в яких не всі з представлених показників були в наявності. Цей датасет містить цільовий стовпець зі значеннями, що відповідають статусу захворювання на ішемічну хворобу серця, від 0, що відповідає відсутності хвороби, до 4 – критична стадія. В цьому дослідженні ми будемо розглядати набір даних лише з точки зору бінарної класифікації: є хвороба чи ні. Причина такого вибору в тому, що вихідний набір даних є незбалансованим, а при розгляді його лише як задачу бінарної класифікації ми отримуємо вже збалансований датасет. Для якісного аналізу кардіологічних даних потрібно виконати імпутування пропущених даних. Дана робота присвячена пошуку можливих підходів до імпутування даних в рамках заданих умов.

Аналіз останніх досліджень. Р.Дж.Літтл та Д.Б.Рубін в своїй роботі надали класифікацію методів обробки даних з пропусками [5]. Для імпутування

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

пропусків даних популярними є такі підходи: аналіз повних спостережень (listwise deletion); методи, які використовують доступну інформацію (pairwise deletion); підстановка середнього по вибірці (mean substitution); метод хот-дек (hot deck); регресійний аналіз (regression) [6, 7]; оцінка за допомогою максимізації правдоподібності (maximum likelihood estimation); підстановка за допомогою факторного аналізу (factor analysis substitution) [8]; модель множинного відновлення даних (MICE) [9 – 13]. В нашому попередньому дослідженні [14] запропоновано ідеї імпутування даних гідрологічного моніторингу в межах знайдених груп постів спостереження, де підтверджено збіг емпіричних функцій розподілу, або в межах груп постів, де підтверджено гіпотезу про однорідність, а також за рахунок побудови регресійних залежностей між ознаками. Такий підхід був доречним, коли в межах екземплярів є ряди спостережень, по яким можна шукати групи схожих об'єктів – постів, а також, коли наявні кореляційні зв'язки між ознаками.

Мета роботи. Метою роботи є розроблення методів імпутування пропусків у кардіологічних даних з переважною більшістю якісних ознак на основі методів класифікації та регресії.

Основна частина. Почнемо розгляд з опису датасетів.

Датасет UCI Heart Disease Data містить 920 рядків даних, 15 ознак та один цільовий стовпець. Стовпці id та назва датасету ми виключаємо як неінформативні. Серед 13 ознак, що залишаються, 5 ознак – кількісні, 8 ознак – якісні порядкові, серед яких 2 – бінарні.

Первинний аналіз даних показав наявність пропусків. Серед 920 екземплярів лише 299 є повними.

Датасет Framingham Heart Study містить 4240 рядків даних, 15 ознак та один цільовий стовпець. Стовпець з ознакою куріння виявився неінформативним, тому ми його виключаємо. Серед 14 ознак, що залишаються, 8 ознак – кількісні, 6 ознак – якісні порядкові, серед яких 5 – бінарні.

Первинний аналіз даних показав наявність пропусків. Серед 4240 екземплярів 3658 є повними. Датасет є незбалансованим, кількість екземплярів одного класу в 5 разів перевищує кількість екземплярів другого класу.

Перетворення якісних ознак у кількісні. Популярним підходом для кодування якісних порядкових ознак є Label Encoding [15]. Цей метод виконує кодування унікальних якісних ознак у числа 0, 1, 2, …, k – 1, де k – кількість унікальних значень ознаки. Він реалізований в модулі

sklearn.preprocessing бібліотеки scikit-learn на мові програмування Python [16]. У цього методу є дві головні проблеми: він кодує пропущене значення як окреме; він не враховує частоту конкретного значення. В результаті кодування пропущених значень втрачається інформація про пропуски. Тому прийнято рішення запропонувати два власних алгоритми, які не кодують пропущені дані. Перший le_non_missing – це модифікація Label Encoding, який лише пропускає пропущені дані, а для інших застосовує Label Encoding (лістинг 1), другий le_non_missing_frequent – враховує частоту значення. Цей алгоритм дозволяє надавати максимальний номер значенню з найменшою частотою або з найбільшою в залежності від параметра ascending (лістинг 2). Обидва рішення мають декодер для зворотного перетворення.

Лістинг 1. Алгоритм le_non_missing

```
def label_encoding_non_missing_values(df):
    df1 = df.copy()
    encoders = dict()
    for col_name in df1.select_dtypes(
                include=['object', 'category']).columns:
        series = df1[col_name]
        label_encoder = LabelEncoder()
        df1[col_name] = pd.Series(
                        label_encoder.fit_transform(
                         series[series.notnull()]),
                   index=series[series.notnull()].index
        )
        encoders[col_name] = label_encoder
    return df1, encoders
def label_decoding_non_missing_values(df, encoders):
    df1 = df.copy()
    for col_name in encoders:
        enc = encoders[col_name]
        nmap = dict(zip(enc.transform(enc.classes_),
                        enc.classes_))
        df1[col_name] = df1[col_name]
        .apply(lambda x: x if np.isnan(x) else nmap[x])
    return df1
```

Лістинг 2. Алгоритм le_non_missing_frequent

```
def label_encoding_non_missing_values_frequent(df,
                                               ascending=False):
    df1 = df.copy()
```


ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

```
    encoders = dict()
    for col_name in df1.select_dtypes(
                include=['object', 'category']).columns:
        d = df[col_name].value_counts().sort_values(
                            ascending=ascending).to_dict()
        k = 0
        for key in d:
            d[key] = k
            k = k + 1
        encoders[col_name] = d
        df1[col_name] = df1[col_name]
            .apply(lambda x: x if x != x else d[x])
    return df1, encoders
def label_decoding_non_missing_values_frequent(df,
                                                encoders):
    df1 = df.copy()
    for col_name in encoders:
        d = encoders[col_name]
        nd = dict()
        for key, value in d.items():
            nd[value] = key
                  df1[col_name] = df1[col_name]
                       .apply(lambda  x:  x  if  x  !=  x  else
nd[x])
    return df1
```

Особливістю запропонованих алгоритмів є формування словнику для кожної якісної ознаки для подальшого зворотного перетворення та пропуск пропущеного значення. В подальшому ці методи будуть переписані на класи в архітектурі scikit-learn для універсального застосування. Слід врахувати те, що алгоритми спрацюють тільки для стовпців, типи яких відносяться до якісних (object та category). Наприклад, датасет Framingham Heart Study містить вже закодовані ознаки і це не те, що нам потрібно. Тому для аналізу спочатку фактично якісні ознаки датасету було перекодовано на їх первинний стан, що містить дійсно якісні ознаки згідно словнику, наданому в описі даних.

Запропоновані алгоритми дають можливість одразу перетворити всі якісні ознаки на кількісні без втрати інформації про пропуски.

Імпутування даних. Розглянемо ідею імпутування даних як задачу прогнозу значення в стовпці, що містить пропуск, на основі даних зі всіх інших

стовпців, враховуючи цільовий стовпець. Якщо цільова ознака є функцією від інших ознак:

$$y_i = f\left(x_{i,1}, x_{i,2}, \dots, x_{i,k}, \dots, x_{i,p}\right) + \varepsilon ,$$

то можна очікувати, що існує також стохастична залежність *k*-ї ознаки від інших ознак, враховуючи цільову:

$$x_i = \varphi\left(x_{i,1}, x_{i,2}, \dots, x_{i,k-1}, x_{i,k+1}, \dots, x_{i,p}, y_i\right) + \varepsilon .$$

Ця ідея певною мірою лежить в основі методів MICE [10, 11] та NoNa [13].

Запропоновано 4 власні реалізації такої ідеї з кількома модифікаціями. Ми будемо виходити з можливої наявності таких патернів у пропусках даних:

a) в даних немає жодного рядка чи стовпця з повними даними;

b) в даних є *k* стовпців з повністю наявними даними у всіх екземплярах;

c) в даних є *n* рядків з повністю наявними даними за всіма ознаками.

Ми також будемо враховувати, що серед ознак можуть бути присутні як кількісні, так і якісні ознаки (вже закодовані).

Запропоновані такі алгоритми:

Алгоритм 1. Метод *fillna_k_columns* – враховує патерни a), b), застосовує регресор чи класифікатор в залежності від типу ознаки.

Алгоритм 2. Метод *fillna_k_sorted_columns* – враховує патерни a), b), застосовує регресор чи класифікатор в залежності від типу ознаки, обробляє стовпці в залежності від кількості пропусків у них.

Алгоритм 3. Метод *fillna_2steps_rg_class* – враховує патерни a), b), c), застосовує регресор чи класифікатор в залежності від типу ознаки, обробляє стовпці в залежності від кількості пропусків у них.

Алгоритм 4. Метод *fillna_2steps_rg* – враховує патерни a), b), c), застосовує тільки регресор з коригуванням значення для якісних ознак, обробляє стовпці в залежності від кількості пропусків у них.

Далі наведемо ці алгоритми.

Алгоритм 1. Метод *fillna_k_columns*

1. Цикл по стовпцям датасету *for j in columns*:

2. Якщо в *j*-му стовпці є пропущені дані, переходимо на крок 3, інакше – на наступну ітерацію, крок 1.

3. Знаходимо індекси рядків, де є пропуски в *j*-му стовпці, та зберігаємо до масиву *indexes_nan*.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

4. Знаходимо підмножину *full_data* датасету, де в стовпцях немає жодного пропуску. Це буде наш датасет, за рахунок якого ми будемо прогнозувати пропущені значення в стовпці *j*.

5. Якщо *full_data* пустий, заповнюємо пропускі на основі стандартного методу: кількісну ознаку заповнюємо медіаною, якісну – модою. Якщо *full_data* непустий, переходимо на крок 6.

6. Розділяємо *full_data* на тренувальну та тестову вибірки за такою ознакою: в тренувальній вибірці залишаємо всі рядки, де в *j*-му стовпці є дані, тобто всі рядки, окрім зазначених в *indexes_nan*.

7. Якщо *j*-й стовпець містить якісні дані, використовуємо класифікатор, інакше – регресор.

8. Навчаємо модель на тренувальній вибірці. Отримаємо прогнозні значення для *j*-го стовпця за даними тестової вибірки.

9. Переходимо на крок 1.

Наступний алгоритм є модифікацією попереднього, де надається можливість спочатку імпутувати пропущені дані у стовпцях з найменшою кількістю пропусків, або навпаки – з найбільшою.

Алгоритм 2. Метод *fillna_k_sorted_columns*

1. Сортуємо масив колонок *columns* за кількістю пропусків в них. В залежності від параметру *ascending* будемо виконувати обхід стовпців за зростанням кількості пропусків або за спаданням.

2. Цикл по стовпцям датасету *for j in columns*:

3. Якщо в *j*-му стовпці є пропущені дані, переходимо на крок 3, інакше – на наступну ітерацію, крок 1.

4. Знаходимо індекси рядків, де є пропуски в *j*-му стовпці, та зберігаємо до масиву *indexes_nan*.

5. Знаходимо підмножину *full_data* датасету, де в стовпцях немає жодного пропуску. Це буде наш датасет, за рахунок якого ми будемо прогнозувати пропущені значення в стовпці *j*.

6. Якщо *full_data* пустий, заповнюємо пропускі на основі стандартного методу: кількісну ознаку заповнюємо медіаною, якісну – модою. Якщо *full_data* непустий, переходимо на крок 7.

7. Розділяємо *full_data* на тренувальну та тестову вибірки за такою ознакою: в тренувальній вибірці залишаємо всі рядки, де в *j*-му стовпці є дані, тобто всі рядки, окрім зазначених в *indexes_nan*.

8. Якщо *j*-й стовпець містить якісні дані, використовуємо класифікатор, інакше – регресор.

9. Навчаємо модель на тренувальній вибірці. Отримаємо прогнозні значення для *j*-го стовпця за даними тестової вибірки.

10. Переходимо на крок 1.

Наступний алгоритм є двох-етапним. На першому етапі шукаємо підмножину датасету, що складається з рядків, де є повністю всі дані за всіма ознаками. Доєднуємо до нього ті рядки, де є лише один пропуск. Виконуємо імпутування даних в цих додаткових рядках за аналогією попередніх алгоритмів. На другому етапі повторюємо дії, описані в алгоритмі *fillna_k_sorted_columns*.

Алгоритм 3. Метод *fillna_2steps_rg_class*

Етап 1.

1. Формуємо масив ознак *cols_1nan*, для рядків, в яких є лише один пропуск саме в цих стовпцях.

2. Якщо такий масив пустий, переходимо на крок 11.

3. Цикл по стовпцям датасету *for j in cols_1nan*:

4. Знаходимо індекси рядків, де є тільки один пропуск, і він знаходиться в *j*-му стовпці, та зберігаємо до масиву *indexes_nan*.

5. Знаходимо підмножину *full_data* датасету, де в рядках немає жодного пропуску. До нього приєднуємо рядки *indexes_nan*. Це буде наш датасет, за рахунок якого ми будемо прогнозувати пропущені значення в стовпці *j* рядків *indexes_nan*.

6. Якщо *full_data* пустий, переходимо достроково до наступної ітерації, тобто на крок 3.

7. Розділяємо *full_data* на тренувальну та тестову вибірки за такою ознакою: в тренувальній вибірці залишаємо всі рядки, де в *j*-му стовпці є дані, тобто всі рядки, окрім зазначених в *indexes_nan*.

8. Якщо *j*-й стовпець містить якісні дані, використовуємо класифікатор, інакше – регресор.

9. Навчаємо модель на тренувальній вибірці. Отримаємо прогнозні значення для *j*-го стовпця за даними тестової вибірки.

10. Переходимо на крок 3.

Етап 2.

11. Викликаємо алгоритм *fillna_k_sorted_columns*.

Останній алгоритм використовує в будь-якому разі регресор, але у випадку якісної ознаки виконує коригування отриманого значення одним з двох спо-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

собів: обирається звичайне найближче значення зі словника за цією ознакою (*measure = 'value'*) або найближче за середньо зваженою оцінкою з урахуванням частоти значень (*measure = 'weight'*).

Алгоритм 4. Метод *fillna_2steps_rg*

Етап 1.

1. Формуємо масив ознак *cols_1nan*, для рядків, в яких є лише один пропуск саме в цих стовпцях.

2. Якщо такий масив пустий, переходимо на крок 11.

3. Цикл по стовпцям датасету *for j in cols_1nan*:

4. Знаходимо індекси рядків, де є тільки один пропуск, і він знаходиться в *j*-му стовпці, та зберігаємо до масиву *indexes_nan*.

5. Знаходимо підмножину *full_data* датасету, де в рядках немає жодного пропуску. До нього приєднуємо рядки *indexes_nan*. Це буде наш датасет, за рахунок якого ми будемо прогнозувати пропущені значення в *j*-му стовпці рядків *indexes_nan*.

6. Якщо *full_data* пустий, переходимо достроково до наступної ітерації, тобто на крок 3.

7. Розділяємо *full_data* на тренувальну та тестову вибірки за такою ознакою: в тренувальній вибірці залишаємо всі рядки, де в *j*-му стовпці є дані, тобто всі рядки, окрім зазначених в *indexes_nan*.

8. В будь-якому випадку використовуємо регресор.

9. Навчаємо модель на тренувальній вибірці. Отримаємо прогнозні значення для *j*-го стовпця за даними тестової вибірки.

10. Якщо *j*-й стовпець містить якісні дані, виконуємо коригування отриманого значення за одним з двох алгоритмів (*value* або *weight*).

11. Переходимо на крок 3.

Етап 2.

12. Сортуємо масив колонок *columns* за кількістю пропусків в них. В залежності від параметру *ascending* будемо виконувати обхід стовпців по зростанню кількості пропусків або по спаданню.

13. Цикл по стовпцям датасету *for j in columns*:

14. Якщо в *j*-му стовпці є пропущені дані, переходимо на крок 15, інакше – достроково на наступну ітерацію, тобто на крок 13.

15. Знаходимо індекси рядків, де є пропуски в *j*-му стовпці, та зберігаємо до масиву *indexes_nan*.

16. Знаходимо підмножину *full_data* датасету, де в стовпцях немає жодного пропуску. Це буде наш датасет, за рахунок якого ми будемо прогнозувати пропущені значення в стовпці *j*.

17. Якщо *full_data* пустий, заповнюємо пропускі на основі стандартного методу: кількісну ознаку заповнюємо медіаною, якісну – модою. Якщо *full_data* непустий, переходимо на крок 18.

18. Розділяємо *full_data* на тренувальну та тестову вибірки за такою ознакою: в тренувальній вибірці залишаємо всі рядки, де в *j*-му стовпці є дані, тобто всі рядки, окрім зазначених в *indexes_nan*.

19. В будь-якому випадку використовуємо регресор.

20. Навчаємо модель на тренувальній вибірці. Отримаємо прогнозні значення для *j*-го стовпця за даними тестової вибірки.

21. Якщо *j*-й стовпець містить якісні дані, виконуємо коригування отриманого значення за одним з двох алгоритмів (*value* або *weight*).

22. Переходимо на крок 13.

Аналіз результатів застосування методів. Для аналізу запропонованих підходів використовується два типи тестів. В першому тесті в датасет штучно вносяться пропуски, далі проводиться імпутування пропущених значень різними методами, оцінюється середньо-квадратична похибка та час виконання алгоритмів. Розглядаються 10% пропусків, 20%, 30% та 40%. В другому тесті для вихідних датасетів виконують імпутування пропущених даних різними методами, навчають моделі бінарної класифікації та порівнюють точність класифікації на тестовій вибірці. В якості методу класифікації використовується випадковий ліс (Random Forest). Датасет Framingham Heart Study попередньо балансується методами SMOTE [17]. В таблиці 1 наведено результати тесту 1 для датасетів UCI Heart Disease Data та Framingham Heart Study, в таблиці 2 – результати тесту 2 відповідно. Виконуємо тести для таких ситуацій:

1. Алгоритм *le_non_missing, SimpleImputer* – стандартне заповнення значенням з найбільшою частотою

2a. Алгоритм *le_non_missing, IterativeImputer(sample_posterior = False)*

2b. Алгоритм *le_non_missing, IterativeImputer(sample_posterior = True)*

3. Алгоритм *le_non_missing, kNNImputer*

4. Алгоритм *le_non_missing, nonaImputer*

5. Алгоритм *le_non_missing, fillna_k_columns*

6a. Алгоритм *le_non_missing, fillna_k_sorted_columns(ascending = True)*

6b. Алгоритм *le_non_missing, fillna_k_sorted_columns(ascending = False)*

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

7. Алгоритм *le_non_missing, fillna_2steps_rg_class*

8a. Алгоритм *le_non_missing_frequent(ascending=False), fillna_2steps_rg_class*

8b. Алгоритм *le_non_missing_frequent(ascending=True), fillna_2steps_rg_class*

9a. Алгоритм *le_non_missing_frequent(ascending=False), fillna_2steps_rg(measure = 'value')*

9b. Алгоритм *le_non_missing_frequent(ascending=False), fillna_2steps_rg(measure = 'weight')*

Таблиця 1

Результати тесту 1

	Датасет UCI HDD					**Датасет Framingham FHS**				
	10%	20%	30%	40%	Час, с	10%	20%	30%	40%	Час, с
1	3.23	3.9	5.48	5.73	0.037210	2.75	4.32	5.41	6.32	0.054037
2	**2.42**	**3.35**	4.72	4.82	0.642131	**2.16**	**3.24**	4.89	5.5	3.245387
3	2.63	3.59	4.44	4.82	0.614188	2.44	3.82	4.86	5.54	9.091372
4	2.49	3.51	4.31	4.66	4.917326	2.33	3.43	4.44	5.12	4.973262
5	2.49	3.51	**4.31**	**4.66**	4.990149	2.33	3.43	**4.44**	**5.12**	5.004628
6a	2.44	3.45	4.32	4.68	4.468355	2.34	3.5	4.49	5.26	5.805581
6b	2.55	3.51	4.35	4.75	4.465862	2.4	3.63	4.65	5.24	5.847364
7	2.55	3.51	4.35	4.75	7.567574	2.4	3.63	4.65	5.24	9.386514
8a	2.56	3.82	4.35	4.68	7.706980	2.37	3.62	4.64	5.24	9.479322
8b	2.56	3.65	4.35	4.65	7.583769	2.37	3.62	4.64	5.24	9.406302
9a	2.55	3.77	4.37	4.71	0.605775	2.36	3.61	4.63	5.24	**1.187455**
9b	2.55	3.76	4.35	4.73	**0.596536**	2.36	3.61	4.63	5.24	1.241692

Жирним шрифтом виділено переможців по кожному рядку за точністю та часом виконання. Запропоновані методі показують приблизно однакову точність, однак останній метод *fillna_2steps_rg* дає значний виграш у часі виконання. І якщо порівняти запропоновані методи з результатами методу *IterativeImputer* з бібліотеки *scikit-learn*, то слід зауважити, що цей метод використовує регресор без корегування значень якісних даних, а значить, він не на-

дає можливості декодувати отримані результати для подальшого використання. На відміну від нього будь-який з запропонованих чотирьох алгоритмів отримує значення з потрібної множини для кожної якісної ознаки, а тому допускає декодування та подальше використання.

Далі розглянемо тест 2. Додамо в таблицю ще нульовий рядок з базовим результатом – класифікація за датасетами, з яких видалено рядки з пропусками. В таблиці 2 жирним шрифтом виділено переможців за точністю.

Таблиця 2

Результати тесту 2

	Датасет UCI HDD		**Датасет Framingham FHS**	
	Найкращі гіперпараметри моделі	Точність	Найкращі гіперпараметри моделі	Точність
0	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 50}	0.83	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.89
1	{'max_depth': None, 'min_samples_leaf': 2, 'min_samples_split': 2, 'n_estimators': 100}	0.80	{'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}	0.89
2a	{'max_depth': None, 'min_samples_leaf': 2, 'min_samples_split': 5, 'n_estimators': 150}	0.88	{'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	**0.90**
2b	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}	0.82	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.88
3	{'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 10, 'n_estimators': 50}	0.82	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}	**0.90**
4	{'max_depth': None, 'min_samples_leaf': 4, 'min_samples_split':	0.89	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split':	0.89

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

	10, 'n_estimators': 50}		2, 'n_estimators': 150}	
5	{'max_depth': None, 'min_samples_leaf': 4, 'min_samples_split': 10, 'n_estimators': 50}	0.89	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.89
6a	{'max_depth': None, 'min_samples_leaf': 4, 'min_samples_split': 10, 'n_estimators': 100}	0.88	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.88
6b	{'max_depth': None, 'min_samples_leaf': 2, 'min_samples_split': 5, 'n_estimators': 50}	0.91	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.87
7	{'max_depth': 10, 'min_samples_leaf': 2, 'min_samples_split': 2, 'n_estimators': 100}	0.91	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.88
8a	**{'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 50}**	**0.92**	**{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}**	**0.90**
8b	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 50}	0.90	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.88
9a	{'max_depth': None, 'min_samples_leaf': 4, 'min_samples_split': 10, 'n_estimators': 150}	0.90	{'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}	0.88
9b	{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 10, 'n_estimators': 100}	0.91	**{'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150}**	**0.90**

З аналізу результатів випливає, що запропоновані методи дозволяють покращити точність класифікації, а на датасеті UCI Heart Disease Data навіть суттєво покращити порівняно з 83% до 92%. Серед переваг запропонованих методів слід віднести можливість декодування імпутованих даних та виграш в часі для останнього методу імпутування.

Висновки та перспективи подальших досліджень. У даній роботі запропоновано два алгоритми кодування та декодування якісних порядкових ознак та чотири алгоритми імпутування пропусків даних:

1. Алгоритм кодування / декодування *le_non_missing*, який дозволяє працювати з даними, що містять пропуски та не втратити інформацію про пропуски.

2. Алгоритм кодування / декодування *le_non_missing_frequent*, який дозволяє працювати з даними, що містять пропуски та не втратити інформацію про пропуски, і враховує частоту значень.

3. Метод *fillna_k_columns*, який виконує імпутування пропущених даних за *k* повними стовпцями. Використовується регресор або класифікатор в залежності від типу стовпця.

4. Метод *fillna_k_sorted_columns*, який виконує обхід стовпців у порядку, що відповідає кількості пропущених значень. Використовується регресор або класифікатор в залежності від типу стовпця.

5. Метод *fillna_2steps_rg_class*, який виконується у 2 етапи: спочатку за повними рядками, а потім за повними стовпцями. Використовується регресор або класифікатор в залежності від типу стовпця.

6. Метод *fillna_2steps_rg*, який виконується у 2 етапи: спочатку за повними рядками, а потім за повними стовпцями. Використовується тільки регресор з коригуванням значення для якісних стовпців за двома критеріями.

7. Проведено два типи тестів на двох наборах даних. Аналіз показав виграш у часі для методу *fillna_2steps_rg* та покращення точності моделі класифікації у випадках використання методу кодування з урахуванням частоти та методу імпутування *fillna_2steps_rg_class*.

Таким чином, запропоновані методи показали гарні результаті, які можуть стати альтернативами існуючих методів та надати досліднику додатковий інструментарій для підвищення точності прийняття рішень.

Надалі планується оформити запропоновані методи в архітектурі бібліотеки *scikit-learn* для уніфікованого використання дослідниками.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЛІТЕРАТУРА / REFERENCE

1. Janosi, Andras, Steinbrunn, William, Pfisterer, Matthias, and Detrano, Robert. (1988). Heart Disease. UCI Machine Learning Repository. https://doi.org/10.24432/C52P4X.

2. UCI Heart Disease Data. Heart Disease Data Set from UCI data repository. – [Електронний ресурс]. – Режим доступу: https://www.kaggle.com/datasets/redwankarimsony/heart-disease-data

3. Framingham Heart Study-Cohort (FHS-Cohort). – [Електронний ресурс]. – Режим доступу: https://biolincc.nhlbi.nih.gov/studies/framcohort/

4. Framingham heart study dataset. – [Електронний ресурс]. – Режим доступу: https://www.kaggle.com/datasets/aasheesh200/framingham-heart-study-dataset

5. Roderick J. A. Little, Donald B. Rubin. Statistical Analysis with Missing Data, 3rd Edition. -Wiley, 2019. - 464 p. ISBN: 978-0-470-52679-8.

6. Sefidian A. M., Daneshpour N. Estimating missing data using novel correlation maximization based methods // Applied Soft Computing. Volume 91. 2020. 106249. DOI: 10.1016/j.asoc.2020.106249

7. A.Barrios, G. Trincado, René Garreaud. Alternative approaches for estimating missing climate data: application to monthly precipitation records in South-Central Chile // Forest Ecosystems. 2018. 5(1). Pp. 1-10. DOI 10.1186/s40663-018-0147-x

8. Kamakura W.A., Wedel M. Factor Analysis and Missing Data // Journal of Marketing Research. 2000. Vol. 37. No. 4: Nov. P. 490–498.

9. van Buuren, S. and Groothuis-Oudshoorn, K. 2011. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 45, 3 (Dec. 2011), 1–67. DOI:https://doi.org/10.18637/jss.v045.i03.

10. A complete guide on how to handle missing data with IterativeImputer in Python. – Learning AI, 2023. – Режим доступу: https://justlearnai.com/a-complete-guide-on-how-to-handle-missing-data-with-iterativeimputer-in-python-6b224cf0896c

11. Imputation of missing values in scikit-learn 1.4.1. – [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/ modules/impute.html

12. NoNa: Missing Data Imputation Algorithm. – Medium, 2023. – [Електронний ресурс]. – Режим доступу: https://medium.com/@abdualimov/nona-missing-data-imputation-algorithm-d6ff92f70ab8

13. nona: Python gap filling toolkit. [Електронний ресурс]. – Режим доступу: https://pypi.org/project/nona/

14. Земляний О.Д., Ізмайлова М.К., Антоненко С.В. Методи поповнення пропусків даних гідрологічного моніторингу // Актуальні проблеми автоматизації та

інформаційних технологій: Зб. наук. пр. / наук. ред. О.Г. Байбуз. – Дніпро, 2020. – Т. 24. – С. 3 – 15

15. Feature Encoding. – Medium, 2023. – [Електронний ресурс]. – Режим доступу: https://medium.com/@denizgunay/feature-encoding-f099a6c1abe8

16. sklearn.preprocessing.LabelEncoder– [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/modules/generated/sklearn.
preprocessing.LabelEncoder.html

17. N.V. Chawla, K.W. Bowyer, L. O.Hall, W.P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of artificial intelligence research, 321-357, 2002.

Received 27.02.2024.
Accepted 29.02.2024.

Methods for imputing missing data on coronary heart disease

Preliminary analysis is an important stage of data analysis. A significant problem is the detection of missing values, and the most difficult part is that there is no universal algorithm to resolve this problem. For each specific task, known methods, their combinations, modifications, or completely new approaches have to be selected.

Most machine learning models cannot handle missing values, so we cannot simply ignore gaps in the data. The problem of missing data needs to be addressed during preprocessing. The simplest solution is to delete each observation containing missing values. This solution is implemented in well-known Python programming language libraries such as NumPy or Pandas. However, this approach is extreme because we lose all the useful information that may be important for data analysis. There are several main strategies for imputing missing data: replacing missing values with mean/median or mode; replacing with the most frequently occurring value or a constant; data imputation using the kNN algorithm; multiple imputation of missing data (MICE algorithm); data imputation using deep learning.

We suppose several modifications of algorithms for iterative multiple imputing of mixed data represented by quantitative and qualitative features. To convert qualitative features into numerical ones, we propose our own algorithms that work with missing data and allow for the conversion back to qualitative features. Two well-known datasets on observations of coronary heart disease are considered.

The following is a brief description of the data imputation algorithms. The fillna_k_columns method, which performs data imputation based on k complete columns. It uses a regressor or classifier depending on the column type. The fillna_k_sorted_columns method, which traverses columns in the order corresponding to the number of missing values. It uses a regressor or classifier depending on the column type. The fillna_2steps_rg_class method, which is executed in 2 steps: first by complete rows, then by complete columns. It uses a regressor or classifier depending on the column

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

type. The fillna_2steps_rg method, which is executed in 2 steps: first by complete rows, then by complete columns. It only uses a regressor with value adjustment for qualitative columns based on two criteria.

Two types of tests are used to analyse the approaches. In the first test, a dataset is artificially filled with gaps at random positions, imputed using different methods, and the mean square error and execution time of the algorithms are estimated. In the second test, binary classification models are trained on datasets imputed with different methods and the classification accuracy is compared. The analysis showed a time advantage for the fillna_2steps_rg method and improved classification model accuracy in cases of using encoding method considering frequency and the fillna_2steps_rg_class imputation method.

Thus, the proposed methods have shown promising results, which can serve as alternatives to existing methods and provide researchers with additional tools to enhance decision-making accuracy.

Further, the plan is to formalize the proposed methods in the scikit-learn library architecture for unified use by researchers.

Keywords: data imputation, iterative multiple data imputation, mixed data processing, regression, binary classification, transformation of qualitative features into quantitative ones, python.

Земляний Олексій Дмитрович – аспірант 1-го року навчання Дніпровського національного університету імені Олеся Гончара.

Zemlianyi Oleksii - postgraduate student of the 1st year of study at the Dnipro National University named after Oles Honchar.

Байбуз Олег Григорович - доктор технічних наук, професор, завідувач кафедри математичного забезпечення ЕОМ Дніпровського національного університету імені Олеся Гончара.

Baibuz Oleh - doctor of technical sciences, professor, head of the department of computer mathematical support of Dnipro National University named after Oles Honchar.

DOI 10.34185/1562-9945-2-151-2024-05
УДК 004.942:

О.М. Поляков, Б.Ю. Жураковський

ПРОТОТИПУВАННЯ ПРИСТРОЇВ КЕРУВАННЯ СИСТЕМ З ПРОМИСЛОВИМИ КОНТРОЛЕРАМИ

Анотація. Скорочення термінів проектування пристрою керування системою залишається актуальним завданням розробників цих систем. Проблемою проектування пристроїв керування на основі програмованих логічних контролерів (ПЛК) є їхня висока вартість і, як правило, недоступність на початковому етапі проектування. Метою дослідження є скорочення термінів та зниження витрат на проектування системи шляхом прототипування пристроїв керування з програмною реалізацією алгоритмів керування мовами стандарту МЕК 61131 3, але виконанням програм у платі Ардуїно. Метод дослідження полягає в декомпозиції проектних моделей пристрою управління та реалізації їх у середовищі програми OpenPLC у вигляді компонентів організації програми (POU) мовами SFC, FBD та LD. Результатом дослідження є методика створення типових POU операційних та керуючих автоматів системи керування, які виконуються у платі Ардуїно. Наведено приклад застосування запропонованої методики проектування прототипу, який може бути корисним для навчання програмуванню ПЛК. Проведено тестування розробленого прототипу за допомогою логічного ПЛК та фізичного макета, яке підтвердило їхню функціональну відповідність оригіналу та зниження, як мінімум на порядок витрат на обладнання.
Ключові слова: програмовані логічні контролери, мови програмування контролерів, прототипування систем управління.

Постановка проблеми. Для керування обладнанням та виробничими процесами широко та ефективно застосовуються промислові контролери [1-3]. Ці контролери ще називають програмованими логічними контролерами (ПЛК).

Слід зазначити, що ПЛК це дороге обладнання. Тому на ранніх стадіях проектування системи доцільно застосування прототипів ПЛК – функціональних аналогів реального ПЛК на базі більш дешевих мікроконтролерів, наприклад плат Ардуїно [4]. Але для програмування ПЛК використовуються спеціальні мови за стандартом МЕК 61131-3 [5], а найбільш поширені середовища програмування, наприклад Arduino IDE [4] підтримують мову програмування С і тому не застосовні при прототипуванні завдань керування систем з ПЛК.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Відомі програмні продукти для трансформації програм, написаних мовами стандарту МЕК 61131-3 в код виконаний у мікроконтролері плати Ардуїно. Наприклад, середовище програмування OpenPLC, що вільно розповсюджується [6]. Компонент цього середовища OpenPLC Editor призначений для введення, редагування, компіляції програмного коду мовами стандарту МЕК 61131-3 та симуляції виконання цього коду в логічному ПЛК. Компонент OpenPLC Runtime транслює код, створений в OpenPLC Editor у здійсненний код і завантажує його в реальний ПЛК, який підключений до комп'ютера через USB-порт.

Водночас використання OpenPLC для прототипування систем керування утруднене через відсутність методики трансформації проектних моделей системи керування в компоненти організації програм (POU – Program Organization Unit, англ. мова) у цьому середовищі. Таким чином, створення методики прототипування систем керування на основі ПЛК за допомогою мікроконтролерних плат типу Ардуїно є актуальним не вирішеним науково-технічним завданням.

Аналіз останніх досліджень і публікацій. Найбільш узагальнена модель системи керування на теоретико-множинному рівні визначає виділення об'єкта керування та керуючого пристрою [7]. Деталізація моделі керуючого пристрою залежить від виду системи керування (безперервна, дискретна, гібридна), розміщення елементів системи в просторі (зосереджена, розподілена), наявності ієрархії елементів, що керують, паралельності в часі різних процесів керування та інших факторів [8]-[10].

Об'єкт досліджень: процес прототипування систем керування з на базі ПЛК. У роботі розгляд питань прототипування обмежено дискретними зосередженими системами з програмною реалізацією алгоритму управління. Для таких систем у керуючому пристрої виділяють операційні (ОА) та керуючі (КА) автомати [7]. Моделлю керуючого автомата є кінцевий автомат (FSM – Finite State Machine) типу перетворювач. Класичний FSM описується кортежом виду [7]:

$$A = < S, X, Y, s0, \delta, \lambda >,$$
$$\delta: S \times X \to S,$$
$$\lambda: S \times X \to Y (\text{для автоматів Мілі}),$$
$$\lambda: S \to Y (\text{для автоматів Мура}),$$

де S — скінченна непорожня множина (станів); X — скінченна непорожня множина входів (вхідний алфавіт); Y — скінченна непорожня множина виходів (ви-

хідний алфавіт); *s0* — початковий стан; δ— функція переходів; λ— функції виходів автоматів Мілі та Мура.

Розрізняють вхідні (ВхОА), вихідні (ВихОА) операційні автомати. Структурну схему системи керування наведено на рис.1.

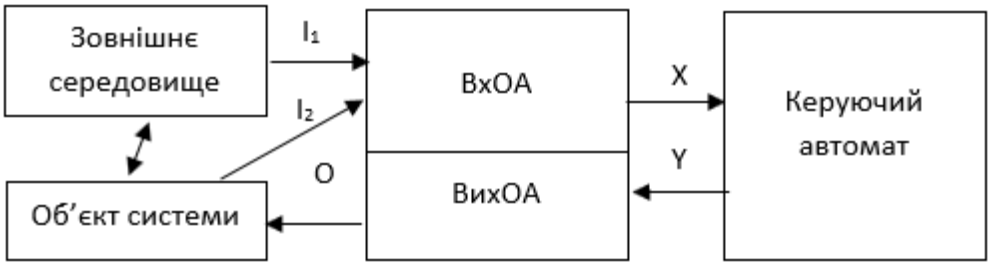


Рисунок 1 – Структура системи керування

Вхідний операційний автомат формує вектор входів *X* автомата, що управляє, як функцію інформації від об'єкта (I_1) і зовнішнього середовища (I_2) системи. Вихідний операційний автомат формує вектор *O* впливів на об'єкт системи як функцію виходів *Y* керуючого автомата. Описана декомпозиція елементів пристрою керування системи спрощує їхню програмну реалізацію в процесі прототипування.

Наведені теоретико-множинні моделі елементів керування є основою для побудови програмно-апаратних моделей прототипу базової системи керування. У складніших системах ці моделі відображають можливості параметричної та структурної адаптації елементів керуючого пристрою, небінарні властивості множин *X*, *Y*. І це також вимагає відображення в методиці прототипування.

Мета роботи: скорочення термінів та зниження витрат на проектування системи шляхом прототипування пристроїв керування з програмною реалізацією алгоритмів керування мовами стандарту МЕК 61131-3, але виконанням програм у платі Ардуїно.

Основний матеріал дослідження. Під прототипом зосередженої системи управління на основі ПЛК розумітимемо її функціональний аналог, в якому ПЛК заміщений спрощеною мікроконтролерною платою, при цьому:

• Вихідний код керуючої програми, яка виконується в мікроконтролері плати, написаний мовами програмування за стандартом МЕК 61131-3 та функціонально повторює програмний код системи оригіналу.

• До входів та виходів мікроконтролерної плати підключено всі датчики та виконавчі механізми об'єкта керування, які підключені до ПЛК у реальній сис-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

темі. У прототипі системи сигнали від датчиків та до виконавчих механізмів узгоджені за рівнями та полярністю з каналами мікроконтролерної плати.

• Є таблиця відповідності входів/виходів ПЛК, каналів мікроконтролерної плати та їх адреси у програмі.

• Тимчасові константи у програмі мікроконтролерної плати залежні від тактової частоти процесора узгоджені з константами керуючої програми ПЛК.

• Зв'язку мікроконтролерної плати з персональним комп'ютером через віртуальний COM-порт достатньо для перевірки функціональності системи за допомогою її прототипу.

У разі невиконання хоча б однієї з перелічених вимог говоритимемо про не повну функціональну відповідність прототипу оригіналу. Але, у багатьох випадках це також корисно, наприклад при навчанні програмування ПЛК.

Можливо два варіанти використання прототипування системи керування:

• Спроектовано систему керування на базі ПЛК, але фізично не доступний обраний ПЛК. І тут прототипування спрощує перевірку правильності алгоритму керування реальним об'єктом чи його макетом.

• Спроектований та протестований прототип системи керування на базі мікроконтролерної плати та макета об'єкта керування. Ця робота спрощує вибір ПЛК.

Вихідними даними для прототипування системи керування є електрична схема системи та моделі операційних та керуючого автоматів, які треба реалізувати у програмно-апаратному комплексі прототипу.

На першому етапі проводиться вибір мікроконтролерної плати з урахуванням необхідної кількості цифрових та аналогових каналів введення/виводу, каналів з ШІМ, оцінюється потреба у додаткових вузлах, які відсутні у платі прототипу. Це можуть бути, наприклад, вузли узгодження входів/виходів прототипу за рівнем та потужністю, додаткові джерела живлення. Багато ПЛК працюють у діапазоні напруги 24 В, а мікроконтролерні плати – у діапазоні 5 В.

На другому етапі проводиться вибір структури та елементів POU для проекту прототипу керуючої програми. Програма є типом POU верхньому рівні організації проекту прототипу. У проекті може бути кілька програм. Операційні автомати реалізуються як функції чи функціональні блоки, а управляючі – як функціональні блоки. Для зв'язку компонентів програми з апаратною частиною мікроконтролерної плати використовуються змінні класів вхід та вихід. Функціональні блоки у програмі описуються як локальні класи.

На третьому етапі обираються мови програмування для POU проекту прототипу програми, що управляє. Різні мови стандарту МЕК 61131-3 мають різні ефективності при реалізації компонентів проекту прототипу системи керування. При виборі мови слід перевірити її доступність у планованому ПЛК. Якщо є можливість вибору, то реалізації ОА вибираються ефективні мови ST, FBD, рідше – мова IL. У той же час, КА простіше реалізувати мовою SFC, але можливо й іншими мовами. Таким чином, типова керуюча програма для прототипу системи є мультимовною.

Приклад прототипування. Розглянемо найпростіший приклад дискретної системи стабілізації температури об'єкта. Структура системи відповідає моделі, наведеній на рис. 1. У цьому випадку об'єкт системи містить датчик температури та нагрівач. За відсутності нагрівання об'єкт остигає, передаючи теплову енергію у зовнішнє середовище. Зовнішнє середовище формує завдання I_1 на температуру об'єкта. Операційний автомат ВхОА порівнює фактичну I_2 і задану I_1 температури та формує події (входи X): «температура менша за завдання», «температура більша за завдання». Ці входи визначають поточний стан керуючого автомата («нагрів» або «охолодження»), в якому формується вихід Y. Операційний автомат ВихОА перетворює цей вихід сигнал керування нагрівачем.

У цій системі як ПЛК, так і мікроконтролерна плата повинні мати по одному аналоговому входу (температура) та по два дискретні виходи (ввімкнення/вимкнення нагрівача). Цим вимогам задовольняють контролер Micrologix 1000 [3] та плата Arduino Uno [4]. Схема підключення датчика та виконавчих механізмів до цієї плати наведена на рис.2.

Датчик приєднано до аналогового входу A0. Двійковий код температури надано змінній *temp* (адреса %IW0). Управління нагрівачем виконується через цифрові виходи 7 і 8. Управляюча дія «включити нагрівання» виконується змінною *outon* (вихід 7, адреса %QX0.0), а вплив «вимкнути нагрівання» – змінної *outoff* (вихід 8, адреса %QX0.1).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

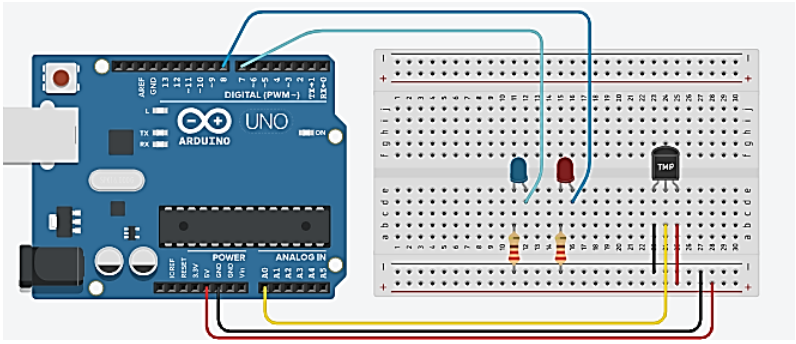


Рисунок 2 – Схема підключення об'єкта до мікроконтролерної плати прототипу

Елементи POU проекту прототипу системи керування у середовищі програми OpenPLC Editor наведено на рис. 3.

Проект *Prototype* включає функціональні блоки ОА і КА, які виконуються в програмі *LimTemp*. Ця програма реалізує алгоритм керування нагрівом об'єкта системи за якого нагрівання відбувається за умови, що температура об'єкта менша за граничну. У проекті *Prototype* значення коду температури прийнято рівним 40.

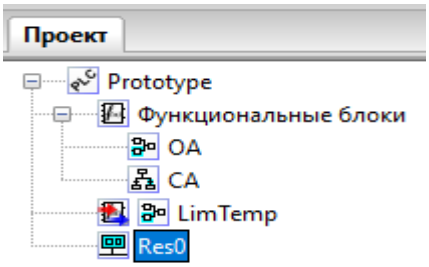


Рисунок 3 – Елементи POU проекту прототипу системи керування у середовищі програми OpenPLC Editor

Змінні та логіка функціонального блоку ОА наведені на рис. 4. Цей блок формує входи *outheat* і *outcold* керуючого автомата КА, тобто X = {*outheat*, *outcold*}.

Функціональний блок КА реалізує керуючий автомат обмеження нагріву об'єкта, граф якого представлений на рис. 5. Він містить два стани та два переходи. Початковий стан автомата – *Cooling*. У цьому стані виконується дія *outoff* – вимкнути нагрівання об'єкта.

#	Имя	Класс	Тип	Исходное значение
1	temperature	Вход	DINT	
2	outheat	Выход	BOOL	
3	outcold	Выход	BOOL	

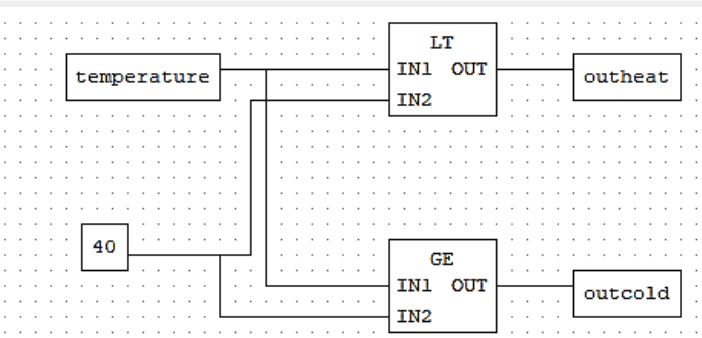


Рисунок 4 – Графічне уявлення функціонального блоку ОА

Графічний варіант реалізації функціонального блоку КА мовою SFC наведено на рис. 6. Описані вище функціональні блоки сприймаються програмою *LimTemp* як користувальницькі POU. Графічний варіант цієї програми на мові FBD наведено на рис. 7.

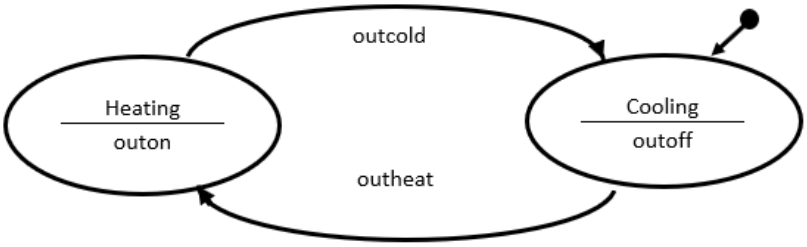


Рисунок 5– Граф автомата обмеження нагріву

#	Имя	Класс	Тип	Исходное значение	Квалификатор
1	heat	Вход	BOOL		
2	cold	Вход	BOOL		
3	outon	Выход	BOOL		
4	outoff	Выход	BOOL		

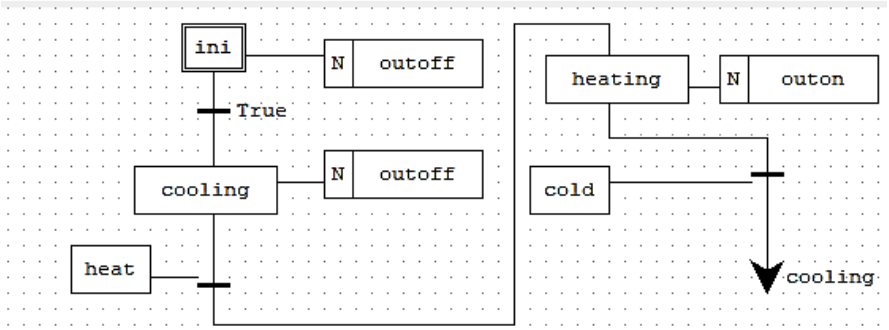


Рисунок 6 – Графічне представлення функціонального блоку КА мовою SFC

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

#	Имя	Класс	Тип	Адрес	Исходное значение	Квалификатор
1	temp	Вход	DINT			
2	outon	Выход	BOOL			
3	outoff	Выход	BOOL			
4	OA0	Локальный	OA			
5	CA0	Локальный	CA			
6	outheat	Вход/Выход	BOOL			
7	outcold	Вход/Выход	BOOL			
8	heatcontrol	Выход	BOOL			

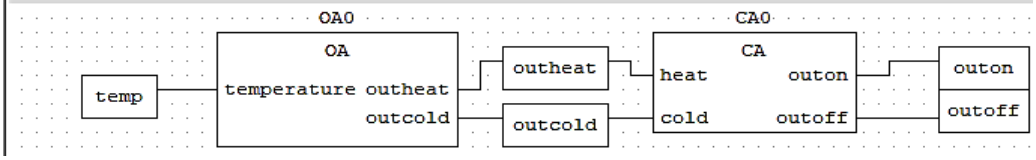


Рисунок 7 – Графічне представлення програми *LimTemp* мовою FBD

Функціональний блок КА був спроектований мовою LD. Програму цього блоку наведено на рис. 8. Вона містить 6 рангів.

#	Имя	Класс	Тип	Исходное значение	Квалификатор	Описание
1	firstscan	Локальный	BOOL	1		
2	heat	Вход	BOOL	0		
3	cold	Вход	BOOL	0		
4	outon	Выход	BOOL	0		
5	outoff	Выход	BOOL	0		
6	cooling	Локальный	BOOL	0		

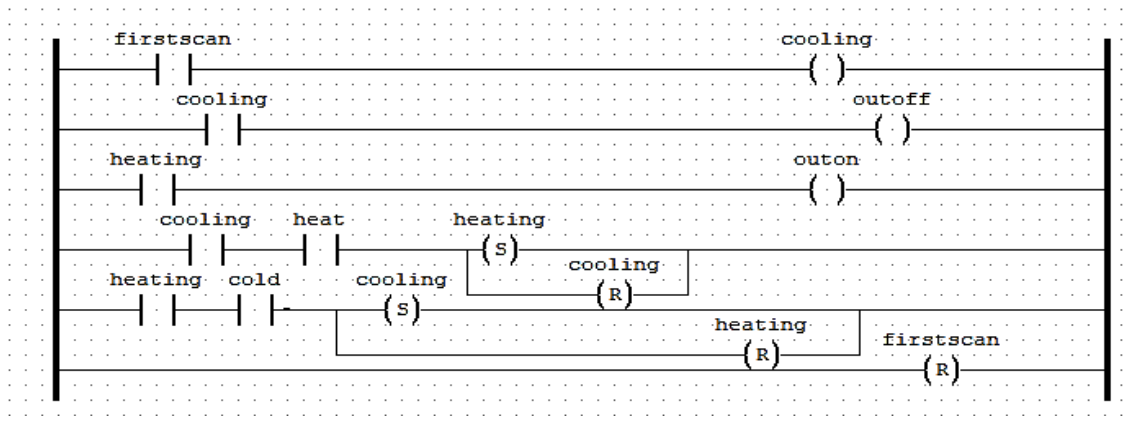


Рисунок 8 – Графічне представлення функціонального блоку КА мовою LD

Перший та останній ранги забезпечують перехід у стан *cooling* при ініціалізації. Другий та третій ранги визначають дії (виходи автомата) в активному стані, а третій та четвертий – переходи з одного стану до іншого.

Експерименти Перевірка проекту Prototype у середовищі програми OpenPLC Editor з використанням логічного ПЛК для обох версій автомата КА підтвердила відповідність проекту вимогам системи керування. Бажане зна-

чення змінної *temp* встановлювалося у налагоджувачі у вікні "форсувати значення змінної".

Перевірка проекту *Prototype* із виконанням програмного коду в платі Ардуїно, до якої підключено за схемою рис. 2 прототипи датчика температури та нагрівача проводилася з використання програм OpenPLC Editor та OpenPLC Runtime. Фактична температура задавалася шляхом нагрівання датчика. При цьому сигнали керування нагріванням індикували світлодіодами і відповідали вимогам системи керування.

Оцінимо зниження витрат під час використання прототипу керуючого пристрою системи керування. Прототип - плату Arduino Uno, на момент написання статті, можна було придбати приблизно за 10 у. о. У той час як промисловий контролер з аналогічним набором входів/виходів оцінюється від 300 до 500 у.о. Наприклад, ПЛК Micrologix 1000 1761–L20BWB–5A (компанія Rockwell Automation [3]) має 20 дискретних входів/виходів та 4 аналогові входи. Його можна придбати в Україні за 480 євро.

Висновки. Основні проблеми проектування та навчання проектуванню систем керування на базі ПЛК – висока вартість та недоступність ПЛК на початковому етапі проектування.

Запропонована методика прототипування дозволяє трансформувати операційні та керуючі автоматів проектних моделей системи керування на компоненти організації програм проекту прототипу системи.

Додаток OpenPLC дозволяє ефективно розробляти та тестувати проекти прототипів у мультимовному середовищі сімейства мов програмування ПЛК за стандартом МЕК 61131-3.

Приклад прототипування системи обмеження нагріву виконано з використанням доступних мікроконтролерних плат Ардуїно та реалізації операційних та керуючих автоматів графічними мовами LD, SFC, FBD. Проведене тестування розробленого прототипу за допомогою логічного ПЛК та фізичного макета підтвердило його функціональну відповідність оригіналу та зниження, як мінімум, на порядок витрат на обладнання.

Подальше дослідження передбачається проводити у бік прототипування адаптивних ієрархічних систем керування.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЛІТЕРАТУРА

1. Parr, E. A. Programmable Controllers. An engineer's guite / E. A. Parr. 3rd ed. – Oxford : Newnes, 2003. – 429 p.

2. Віддалений та віртуальний інструментарій в інжинірингу : монографія / за заг. ред. Хенке К. – Запоріжжя: Дике поле, 2015. – 250 с.

3. 1761-UM003B-EN-P MicroLogix 1000 Programmable Controllers User Manual [Electronic resource] – Access mode: https://literature.rockwellautomation.com/idc/groups/literature/documents/um/1761-um003_-en-p.pdf

4. Arduino - Home [Electronic resource] – Access mode: https://www.arduino.cc/.

5. IEC 61131-3, Revision 3.0, February 2013 - Programmable controllers – Part 3: Programming languages. Published By: International Electrotechnical Commission (IEC). – 468 p.

6. OpenPLC Overview – Autonomy. [Electronic resource] – Access mode: https://autonomylogic.com/docs/openplc-overview/.

7. Глушков В. М. Синтез цифрових автоматів/В. М. Глушков. – М. Фізматіздат, 1962. - 456 с.

8. Поляков М. А. Комплекс математичних моделей функціональних елементів та структур інтегрованих та когнітивних систем /М. А. Поляков // Матеріали міжнародної науково-технічної конференції «Інформаційні технології в металургії та машинобудуванні» імені професора Михальова А.В. І. (НМетАУ, 17 – 19 березня 2020 року) . – Дніпро, 2020. – с. 228–233.

9. Behavioral Typesin Programming Languages Foundations and Trends Rin Programming Languages / [D. Anconaetal.]. – 2016. – Vol. 3. – No. 2–3. – p. 95–230.

10. Poliakov, O. Performance indicators of models of non-binary control automates / M. Poliakov, S. Subbotin, O. Poliakov // Proceeding of 2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM) (22-26 February, 2021 Lviv, Ukraine). – P. 38–42.

REFERENSIS

1. Parr, E. A. Programmable Controllers. An engineer's guide / E. A. Parr. 3rd ed. – Oxford: Newnes, 2003. – 429 p.

2. Viddalenyy ta virtual′nyy instrumentariy v inzhynirynhu: monohrafiya / za zah. ed. Khenke K. - Zaporizhzhya: Dyke pole, 2015. - 250 c.

3. 1761-UM003B-EN-P MicroLogix 1000 Programmable Controllers User Manual [Electronic resource]– Access mode:

https://literature.rockwellautomation.com/idc/groups/literature/documents/um/1761-um003_-en-p.pdf
4. Arduino - Home [Electronic resource]– Access mode: https://www.arduino.cc/
5. IEC 61131-3, Revision 3.0, February 2013 - Programmable controllers – Part 3: Programming languages. Published By: International Electrotechnical Commission (IEC). – 468 p.
6. OpenPLC Overview – Autonomy (autonomylogic.com) [Electronic resource]– Access mode: https://autonomylogic.com/docs/openplc-overview/
7. Hlushkov V. M. Syntez tsyfrovykh avtomativ/V. M. Hlushkov. - M. Fizmatizdat, 1962. -456 p.
8. Polyakov, M. O. Kompleks matematychnykh modeley funktsional'nykh elementiv ta struktur intehrovanykh ta kohnityvnykh system / M. O. Polyakov // Materialy mizhnarodnoyi naukovo-tekhnichnoyi konferentsiyi "Informatsiyni tekhnolohiyi u metalurhiyi ta mashynobuduvanni" imeni profesora Mikhal'ova O. I. (NMetAU, 17-19 bereznya 2020). - Dnipro, 2020. - S. 228-233.
9. Behavioral Types in Programming Languages Foundations and Trends R in Programming Languages / [D. Ancona et al.]. – 2016. – Vol. 3. – No. 2–3. – P. 95–230.
10. Poliakov, O. Performance indicators of models of non-binary control automates / M. Poliakov, S. Subbotin, O. Poliakov // Proceeding of 2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM) (22-26 February, 2021 Lviv, Ukraine). – P. 38–42.

Received 04.03.2024.
Accepted 06.03.2024.

Prototyping of control units for systems with industrial controllers

Reducing the design time of the system control unit remains an urgent task for the developers of these systems. The problem of designing control units based on programmable logic controllers (PLCs) is their high cost and, as a rule, unavailability at the initial design stage. The aim of the research is to reduce the time and cost of designing the system by creating prototypes of control units with the software implementation of the control algorithms of the languages of the IEC 61131-3 standard and the execution of programs in the Arduino board. The research method consists in the decomposition of project models of operating and control automata of the control device and their implementation in the OpenPLC application environment in the form of program organization components (POU) in Ladder Diagram, Function Block Diagram and Sequential Function Chart languages. The result of the study is a method of creating typical POU operating and control automata of the control system, which are executed in the Arduino board. An

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

example of the application of the proposed methodology for the design of a prototype of the object's temperature control system, which can be useful for teaching PLC programming, is given. The developed prototype was tested using a logical PLC and a physical prototype, which confirmed their functional compliance with the original and a reduction in the cost of the equipment by at least an order of magnitude.

Key words: programmable logic controllers, controller programming languages, control system prototyping.

Поляков Олексій Михайлович – інж. програміст компанії LineUp, студент каф. «Інформаційні системи та технології», Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Жураковський Богдан Юрійович – проф. каф. «Інформаційні системи та технології», Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Poliakov Oleksii – Software Developer LineUp company, student dep. ”Information systems and technology”, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute".

Zhurakovskyi Bogdan – professor, dep. ”Information systems and technology”, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute".

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-06
UDC 621.771.068

O.M. Hrechanyi, T.O. Vasilchenko, A.O. Vlasov, P.O. Vypryzhkin, D.I Yakymchuk

SIMULATION MODELING IN THE RESEARCH OF METALLURGICAL EQUIPMENT OPERATION

Annotation. Rolling production refers to the final link of the metallurgical cycle, the quality of products shipped to the consumer depends on the consistency of the work of all units. The wear and tear of the main production facilities of enterprises of the metallurgical complex requires not only updating, but also constant modernization of dated equipment in the conditions of active production. The main equipment of the production lines of wide scale rolling mills in cludes coilers, the quality of which depends not only on the rhythm of the rolling equipment, but also on the quality of the material shipped to the consumer. Simulation modeling of the winding process of hot rolled billot made it possible to establish the possibility of torsional oscillations in the coiler drum drive. Further analysis of the form of the resulting oscillations made it possible to establish that the elastic deformations from the resistance forces of the electric motor and the rotating parts of the coiler drum are in antiphase. The performed calculations create prerequisites for the study of forced oscillations occurring in the coiler drum drive.
Key words: simulation modeling, coiler, strip, torsional oscillations, elastic deformations, free oscillations

Statement of the problem

Modern production lines of rolling shops have a lot of additional and basic equipment, such as scissors, plate straightening machines, coilers, uncoilers, turn-over devices, manipulators, conveyors [1-3]. Not only the rhythmicity of the work of the entire workshop depends on how well and efficiently each unit from the rolling mill performs its functions, but also the timeliness of the shipment of the final product to the consumer.

The main equipment of the rolling mill includes coilers of hot-rolled billot, the quality of the billot's winding depends on rolling mill's operation, which affects not only the next sheet cutting operation, but also the shipment of a high-quality, tightly wound roll according to the requests.

Today, coilers are precisely the vulnerable technological equipment that prevents the further improvement of the productivity of the rolling line. They are often the reason for the lack of rolling. The creation of new high-speed models or

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

the improvement of the design of existing ones require thorough studies of the loads on their units in the form of a full-scale experiment [4].

Analysis of recent research and publications

Modernization of existing equipment poses the task of mechanical designers to develop a sufficiently reliable design that can provide quick and easy adjustment during the flow of the technological process. Correctly determined technological loads and coefficients of the durability reserve allow not only to extend the life cycle of the equipment, but also to significantly reduce its metal content [5].

Work [6] considered in detail the issue of the impact of dynamic loads on the interaction of mechanical rolling equipment, while the impact of changing the technological modes of operation of coilers on the dynamics of the staff winding process was not covered.

Any calculation related to technological equipment is a prediction of its performance. There are calculations designed to ensure the functioning of mechanical systems and the possibility of performing the technological operations entrusted to them (1st group), as well as calculations for strength and reliability, designed to confirm the resource and quality of works due to the equipment (2nd group). The first group of calculations includes calculations of operating parameters of processes and structures for their implementation (productivity, speed, technological supports, power). The second group includes calculations that determine the load on individual structural elements and their stress-strained state, limit state equations, durability reserves, and the probability of failure-free operation [7].

Due to uncertainty during the martial law in the country, it is quite difficult to conduct full-scale experiments that would allow to confirm or disprove certain aspects of the technological process, therefore the problem of using simulation modeling and experiment in the design of metallurgical equipment becomes urgent [8].

Simulation modeling allows to replace the researched model with a mathematical one and conduct experiments on it by means of statistical modeling using numerical methods in specialized programs of computational experiments [9].

Purpose of the study

Taking into account the importance of the technological process of winding hot-rolled billot, the purpose of the study was to develop a simulation model of winding billot into a roll depending on the change in technological parameters, for its further use in the conditions of current production in order to optimize not only

the technological cycle, but also when modernizing the structural elements of technological equipment.

Statement of the main research material

The nature of the movement of the coiler drum when winding the hot-rolled billot is rotational, that is, all dynamic processes in this case will be characterized by the moment of inertia. If we accept the case when the staff is tightly wound and all structural gaps have already been selected, then the drum together with the wound coiler can be considered as a rotating cylinder for which the moment of inertia [10]:

$$J_б = \frac{m_т \cdot D_б^2}{8}, \text{ kg·m}^2 \quad (1)$$

where $m_т$ – technological load taking into account the weight of the drum;

$D_б$ – coiler's diameter;

Under the condition of a constant linear speed of winding the billot, the instantaneous weight of the roll can be calculated by the formula:

$$m_р = V \cdot t \cdot m_{м.п.}, \text{ kg} \quad (2)$$

where V – winding's speed (linear);

t – time of the technological winding operation;

$m_{м.п.}$ – the mass of one meter of winding billot:

$$m_{lm.} = \rho \cdot t \cdot b, \text{ kg/m} \quad (3)$$

where ρ – density of the material of the winding billot;

t – thickness of the winding material;

b – width of the winding material.

Taking it as the technological load in the expression (1) – roll's mass, drum's diameter – instantaneous roll's diameter, expression (1) will be:

$$J_б = \frac{(m_б + m_р) \cdot D_р^2}{8}, \text{ kg·m}^2 \quad (4)$$

where $m_б$ – directly the mass of the drum;

$m_г$ – instantaneous roll's mass calculated by (2);

$D_г$ – instantaneous roll's diameter, m;

During the flow of the technological process of winding hot-rolled stock, the diameter of the roll is a variable value over time. The technological process should be considered for a different assortment of billot, wound in the thickness range t=1.5÷4 mm with a roll width of 1250 mm, for the most complex technological mode, when the mass of the product roll reaches a maximum, i.e. 16 tons.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

At the start of winding, the diameter of the roll will be equal to the diameter of the drum. The results of the calculation according to formulas (1)-(4) are entered in table 1 and we plot the dependence of the mass and diameter of the roll as a function of time (Fig. 1-2).

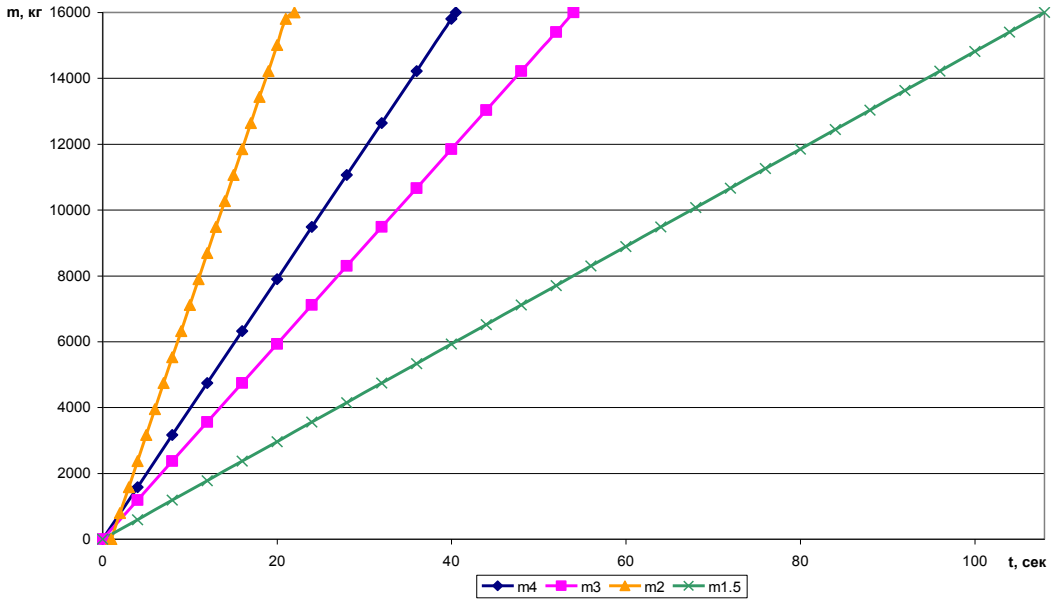


Figure 1 – Graph of the dependence of the mass of the roll in a function of time

Table 1

The results of determining the moment of inertia of the coiler drum or different conditions of flow of the technological process

The time of the technical operation of roll formation t, sec	Billot's thickness											
	4 mm			3 mm			2 mm			1,5		
	m_p kg	D_p m	$J_б$ kg·m^2	m_p, kg	D_p m	$J_б$ kg·m^2	m_p kg	D_p, m	$J_б$ kg·m^2	m_p kg	D_p m	$J_б$ kg·m^2
0	0	0,75	1177	0	0,75	1177	0	0,75	1177	0	0,75	1177
4	1580	0,90	1871	1185	0,87	1688	790	0,83	1511	593	0,81	1425
8	3160	1,04	2666	2370	0,97	2256	1580	0,90	1871	1185	0,87	1688
12	4740	1,15	3562	3555	1,07	2881	2370	0,97	2256	1778	0,92	1965
16	6320	1,26	4558	4740	1,15	3562	3160	1,04	2666	2370	0,97	2256
20	7900	1,36	5655	5925	1,23	4299	3950	1,10	3101	2963	1,02	2561
24	9480	1,45	6853	7110	1,31	5094	4740	1,15	3562	3555	1,07	2881
28	11060	1,53	8151	8295	1,38	5945	5530	1,21	4047	4148	1,11	3214

32	12640	1,61	9550	9480	1,45	6853	6320	1,26	4558	4740	1,15	3562
36	14220	1,69	11049	10665	1,51	7817	7110	1,31	5094	5333	1,19	3924
40	15800	1,76	12649	11850	1,57	8838	7900	1,36	5655	5925	1,23	4300
40,5	15998	1,77	12857	11998	1,58	9570	7999	1,36	5727	5999	1,24	4348
44				13035	1,63	9915	8690	1,40	6241	6518	1,27	4690
48				14220	1,69	11049	9480	1,45	6853	7110	1,31	5094
52				15405	1,75	12240	10270	1,49	7489	7703	1,34	5513
54				15998	1,77	12857	10665	1,51	7817	7999	1,36	5727
56							11060	1,53	8151	8295	1,38	5945
60							11850	1,57	8838	8888	1,41	6392
64							12640	1,61	9550	9480	1,45	6853
68							13430	1,65	10287	10073	1,48	7328
72							14220	1,69	11049	10665	1,51	7817
76							15010	1,73	11837	11258	1,54	8320
80							15800	1,76	12649	11850	1,57	8838
81							15998	1,77	12857	119988	1,58	8969
84										12443	1,60	9369
88										13035	1,63	9915
92										13628	1,66	10475
96										14220	1,69	11049
100										14813	1,72	11638
104										15405	1,75	12240
108										15998	1,77	12857

By analyzing the graphs obtained in fig. 1-2, it can be concluded that the mass of the winding roll is directly proportional to the winding time, and the diameter of the roll does not change linearly, which can cause oscillations in the coiler's drive; to study the nature of these oscillations, it is worth studying their shape.

Concentrated masses, which during rotational movement at different moments of time either lead or lag behind each other, create torsional oscillations. From the point of view of the strength of the machine nodes, this can be very dangerous, because with this phenomenon moments of elastic forces can significantly exceed the calculated loads, especially if cyclical technological loads are added. Therefore, there is an important issue of taking into account and controlling changes in the nature and magnitude of elastic forces during the flow of the technological process [11].

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

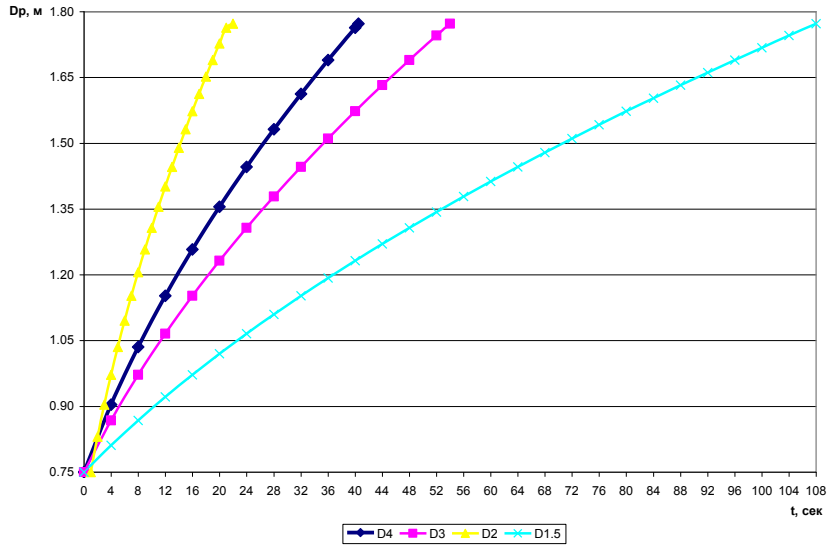


Figure 2 – Graph of dependence of roll diameter in a function of time

When calculating the oscillations of a multi-mass system, the most important stage is the compilation of the calculation scheme [10], while it is worth considering that the more elements in the system, the more complicated the calculation, therefore it is worth moving from multi-mass to two-mass calculation schemes (Figure 3).

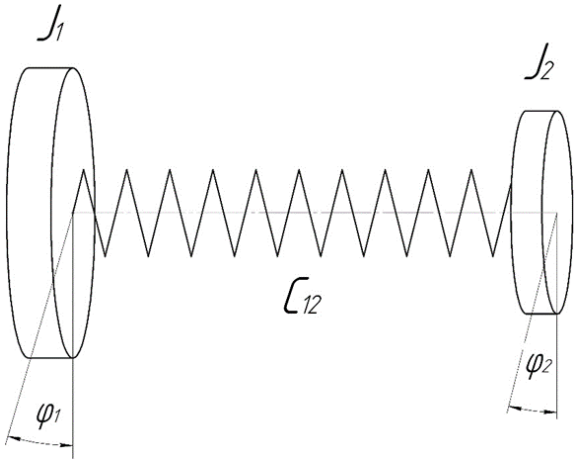


Figure 3 – Scheme for calculating free oscillations occurring in the coiler's drive of hot-rolled staff: J_1 – moment of inertia of the roll with the coiler drum, J_2 – moment of inertia of the armature electric motor of the drive, C_{12} – the stiffness of the intermediate shaft

The task of calculations on oscillations is to find, first of all, the spectrum of natural frequencies and the shape of oscillations.

When studying the oscillations of mechanical systems with a limited number of degrees of freedom, Lagrangian equation of motion of the second kind is used [10].

For the case of free oscillations without taking into account resistance forces, the Lagrange equation has the form [12]:

$$\frac{d}{dt}\cdot\left(\frac{\partial E}{\partial \dot{q}_i}\right)-\frac{\partial E}{\partial q_i}+\frac{\partial E_{П}}{\partial q_i}=0 \tag{5}$$

For a two-mass system, the Lagrange equation can be written in the form [13]:

$$\frac{d}{dt}\cdot\left(\frac{\partial E}{\partial \dot{q}_2}\right)-\frac{\partial E}{\partial q_2}+\frac{\partial E_{П}}{\partial q_2}=0\ ;\frac{d}{dt}\cdot\left(\frac{\partial E}{\partial \dot{q}_1}\right)-\frac{\partial E}{\partial q_1}+\frac{\partial E_{П}}{\partial q_1}=0 \tag{6}$$

After the necessary transformations of equations (5)-(6), we obtain the formulas for determining, respectively, the spectrum of natural frequencies and the form of oscillations for the calculation scheme of Fig. 3:

$$p^2=\frac{C_0(j_1+j_2)}{j_1 j_2}\qquad \text{or}\qquad p=\sqrt{\frac{C_0(j_1+j_2)}{j_1\cdot j_2}} \tag{7}$$

$$\mu=\frac{C_0-j_1p^2}{C_0}. \tag{8}$$

where $C_0=C_{12}$ – stiffness of the intermediate shaft:

$$C_{12}=\frac{\pi\cdot G\cdot d^4}{64\cdot g\cdot l},\ \text{kg·m} \tag{9}$$

where l – length of the intermediate shaft;

G – modulus of rigidity;

d – diameter oft he shaft;

g – gravitational acceleration.

Next, we find the amplitude of oscillations:

$$A_1=\frac{\omega}{p(1-\mu)} \tag{10}$$

where ω – angular velocity of the drum;

Now that the frequency, amplitude and form of oscillations are known, it is possible to determine the elastic deformations that occur during the winding of the billot in a function of time using the formulas:

$$\varphi_1=A_1\sin(pt+\beta);\qquad \varphi_2=\mu A_1\sin(pt+\beta), \tag{11}$$

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

where β – initial phase of oscillations (the angular displacement of the beginning of oscillations relative to the coordinate reference point is β=0);

p – own circular frequency of oscillations;

μ – shape coefficient of oscillations;

A_1 – amplitude of oscillations.

Using the known parameters of the technological process of winding billot and formulas (7)-(11), we obtain the necessary data (table 2-3) for constructing a graph of free oscillations occurring in the drive of the hot billot coiler (Fig. 4).

Table 2

Results of calculating the circular frequency of oscillations, shape factor and amplitude of oscillations

The time of the technical operation of roll formation t, sec	Billot's thickness											
	4 mm			3 mm			2 mm			1,5		
	p	μ	A	p	μ	A	p	μ	A	p	μ	A
0	162	-0,02150	0,16137	162	-0,02150	0,16137	162	-0,02150	0,16137	162	-0,02150	0,16137
4	161	-0,01352	0,13546	161	-0,01499	0,14076	161	-0,01674	0,14674	161	-0,01775	0,15002
8	161	-0,00949	0,11898	161	-0,01121	0,12643	161	-0,01352	0,13546	161	-0,01499	0,14076
12	161	-0,00710	0,10733	161	-0,00878	0,11572	161	-0,01121	0,12643	161	-0,01288	0,13302
16	161	-0,00555	0,09854	161	-0,00710	0,10733	161	-0,00949	0,11898	161	-0,01121	0,12643
20	160	-0,00447	0,09159	161	-0,00588	0,10053	161	-0,00816	0,11271	161	-0,00988	0,12072
24	160	-0,00369	0,08593	160	-0,00497	0,09487	161	-0,00710	0,10733	161	-0,00878	0,11572
28	160	-0,00310	0,08120	160	-0,00426	0,09007	161	-0,00625	0,10265	161	-0,00787	0,11129
32	160	-0,00265	0,07718	160	-0,00369	0,08593	161	-0,00555	0,09854	161	-0,00710	0,10733
36	160	-0,00229	0,07369	160	-0,00324	0,08231	160	-0,00497	0,09487	161	-0,00645	0,10377

40	160	-0,00200	0,07064	160	-0,00286	0,07911	160	-0,00447	0,09159	161	-0,00588	0,10053
40,5	160	-0,00197	0,07028	160	-0,00282	0,07874	160	-0,00442	0,09121	161	-0,00582	0,10015
44				160	-0,00255	0,07626	160	-0,00405	0,08863	160	-0,00539	0,09758
48				160	-0,00229	0,07369	160	-0,00369	0,08593	160	-0,00497	0,09487
52				160	-0,00207	0,07137	160	-0,00338	0,08347	160	-0,00459	0,09238
54				160	-0,00197	0,07028	160	-0,00324	0,08231	160	-0,00442	0,09121
56							160	-0,00310	0,08120	160	-0,00426	0,09007
60							160	-0,00286	0,07911	160	-0,00396	0,08793
64							160	-0,00265	0,07718	160	-0,00369	0,08593
68							160	-0,00246	0,07537	160	-0,00345	0,08406
72							160	-0,00229	0,07369	160	-0,00324	0,08231
76							160	-0,00214	0,07212	160	-0,00304	0,08066
80							160	-0,00200	0,07064	160	-0,00286	0,07911
81							160	-0,00197	0,07028	160	-0,00282	0,07874
84										160	-0,00270	0,07765
88										160	-0,00255	0,07626
92										160	-0,00242	0,07494
96										160	-0,00229	0,07369
100										160	-0,00217	0,07250

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

104										160	-0,00207	0,07137
108										160	-0,00197	0,07028

Table 3

Calculation results of elastic deformations resulting from free oscillations and moments of elastic forces

The time of the technical operation of roll formation t, sec	Billot's thickness							
	4 mm		3 mm		2 mm		1,5	
	φ_1	φ_2	φ_1	φ_2	φ_1	φ_2	φ_1	φ_2
0	0	0	0	0	0	0	0	0
4	-0,131	0,00177	-0,136	0,00204	-0,141	0,00237	-0,144	0,00256
8	-0,053	0,00050	-0,059	0,00066	-0,066	0,00089	-0,070	0,00106
12	0,085	-0,00060	0,090	-0,00079	0,095	-0,00106	0,097	-0,00125
16	0,073	-0,00041	0,082	-0,00058	0,095	-0,00090	0,104	-0,00117
20	-0,048	0,00022	-0,049	0,00029	-0,049	0,00040	-0,047	0,00047
24	-0,080	0,00030	-0,090	0,00045	-0,104	0,00074	-0,113	0,00100
28	0,016	-0,00005	0,014	-0,00006	0,008	-0,00005	0,001	-0,00001
32	0,077	-0,00020	0,086	-0,00032	0,098	-0,00054	0,106	-0,00075
36	0,011	-0,00003	0,016	-0,00005	0,027	-0,00013	0,037	-0,00024
40	-0,067	0,00013	-0,074	0,00021	-0,082	0,00037	-0,086	0,00051
40,5	0,011	-0,00002	0,016	-0,00004	0,026	-0,00012	0,036	-0,00021
44			-0,040	0,00010	-0,053	0,00022	-0,065	0,00035
48			0,055	-0,00013	0,059	-0,00022	0,059	-0,00029
52			0,056	-0,00012	0,070	-0,00024	0,082	-0,00038
54			0,014	-0,00003	0,024	-0,00008	0,035	-0,00015
56					-0,031	0,00010	-0,027	0,00012
60					-0,077	0,00022	-0,087	0,00035
64					0,004	-0,00001	-0,004	0,00002
68					0,075	-0,00018	0,082	-0,00028
72					0,022	-0,00005	0,032	-0,00010
76					-0,064	0,00014	-0,068	0,00021
80					-0,043	0,00009	-0,054	0,00015
81					0,021	-0,00004	0,031	-0,00009

84							0,047	-0,00013
88							0,068	-0,00017
92							-0,022	0,00005
96							-0,073	0,00017
100							-0,004	0,00001
104							0,069	-0,00014
108							0,028	-0,00006

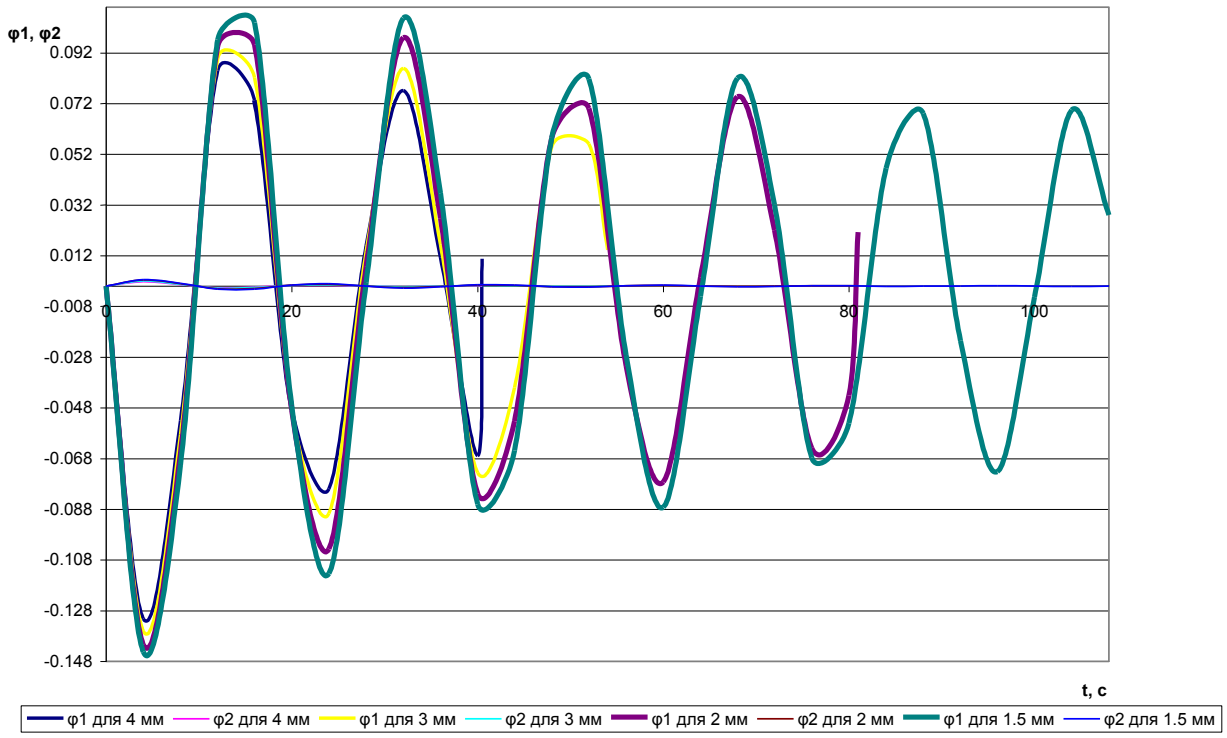


Figure 4 – The graph of free oscillations occurring in the drive of the winding of the hot billot

Conclusions Simulation modeling of the technological process of winding hot-rolled billot indicates that the mass of the winding roll is directly proportional to the winding time, and the diameter of the roll does not change lineary, which causes the need to study the form of possible oscillations in the coiler drum drive. According to the obtained schedule of free oscillations that occur in the coiler drive during the technological process of winding the billot, it can be concluded that the elastic deformations from the resistance forces of the electric motor and the rotating parts of the coiler drum are in antiphase, therefore the breakdowns associated with resonant oscillations can be disregarded. The obtained results create prerequisites for the study of forced oscillations occurring in the coiler drum drive and will allow

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

to determine the most unfavorable technological mode from the point of view of dynamic component loads.

REFERENCE

1. Analysis of possible ways to increase the productivity of the equipment production line of rolling shops / O. M. Hrechanyi та ін. *Visnyk of Kherson National Technical University*. 2021. T. 78, № 3. C. 36–42.
URL: https://doi.org/10.35546/kntu2078-4481.2021.3.4

2. Hrechanyi O. M. Rationale for the choice of technical parameters of guillotine shears rolling mill. *Metallurgy: scientific works of the Zaporizhia State Engineering Academy*. 2017. Vol. 38, no. 2. P. 126–130.

3. Belodedenko S., Grechany A., Yatsuba A. Prediction of operability of the plate rolling rolls based on the mixed fracture mechanism. *Eastern-European Journal of Enterprise Technologies*. 2018. Vol. 1, no. 7 (91). P. 4–11.
URL: https://doi.org/10.15587/1729-4061.2018.122818

4. Ivanchenko F. K., Grebeniy V. M., Shiryaev V. I. Rozrahunki mashin i mehanizmiv prokatnih tsehiv. Kiyiv : Vischa shk., 1994. 455 c.

5. Vykorystannya chastotnykh modeley v tekhnichniy diahnostytsi nespravnostey metalurhiynoho obladnannya / O. Hrechanyi ta in. *Metalurhiya*. 2019. № 1. C. 95–100.

6. Doslidzhennya dynamiky, mitsnosti i tekhnolohichnosti mekhanichnykh system : monohrafiya / L. M. Mamayev ta in. Kam″yans′ke : DDTU, 2017. 183 c.

7. Belodedenko S. V., Ibragimov M. S. Models for optimization the preventive maintenance schedules of mechanical systems in metallurgy. *Metallurgical and Mining Industry*. 2017. No. 2. P. 26–37.

8. Murashko V., Kulik D., Hrechanyy O. M. Perspektyva vykorystannya imitatsiynoho modelyuvannya pry konstruyuvanni metalurhiynoho obladnannya. *Zbirnyk naukovykh prats' studentiv, aspirantiv, doktorantiv i molodykh vchenykh «MOLODA NAUKA-2023»* Zaporizhzhia: ZNU, 2023. Vol. 5 C. 363-365.

9. Nerush V. B., Kurdecha V. V. Imitatsiyne modelyuvannya system ta protsesiv: elektronne navchal′ne vydannya : konspekt lektsiy. K.: NN ITS NTUU "KPI", 2012. 115 c.

10. Zhuk A. YA., Zhelyabina N. K. Osnovy rozrakhunkiv pryvodiv mashyn : navch. posib. Zaporizhzhia: ZDIA, 1996. 145 c.

11. Zhuk A. YA., Zhelyabina N. K., Malyshev H. P. Osnovy naukovykh doslidzhen′. Knyha 1. Teoretychni doslidzhennya : navch. posib. Zaporizhzhia: ZDIA, 2008. 195 c.

12. Pysarenko H. S., Kvitka O. L., Umans′kyy E. S. Opir materialiv : pidruchnyk. K.: Vyshcha shkola, 2004. 655 c.

13. Shvab″yuk V. I. Opir materialiv : navch. posib. Kyyiv : Znannya, 2009. 380 c.

Received 05.03.2024.
Accepted 08.03.2024.

Імітаційне моделювання при дослідженні роботи металургійного обладнання

Сучасні потокові лінії прокатних цехів в своєму складі мають багато допоміжного та основного обладнання. Від того на скільки безвідмовно та якісно виконує свої функції кожен агрегат зі складу прокатного стану залежить не тільки ритмічність роботи всього цеху, а й своєчасність відвантаження кінцевого продукту споживачеві. До основного обладнання прокатного стана можна віднести моталки, що намотують основну продукцію широкоштабових станів – штабу, у рулон. Від роботи моталок залежить якість намотки штаби, що впливає не тільки на наступну операцію різання листа, а й відвантаження якісного, щільнонамотаного рулону згідно заявок.

На сьогодні моталки є саме тим вразливим технологічним устаткуванням, що перешкоджає подальшому підвищенню продуктивності потокової лінії прокатки. Вони нерідко є причиною браку прокату. Створення нових моделей, або удосконалення конструкції діючих, вимагають прискіпливих досліджень навантажень на їхні вузли у формі натурного експерименту.

В умовах воєнного стану в країні досить важко провести натурні експерименти, які б дозволили підтвердити або спростувати ті чи інші аспекти технологічного процесу, тому актуальним постає питання використання імітаційного моделювання при конструюванні металургійного обладнання. Імітаційне моделювання дозволяє досліджувану модель замінювати математичною і над нею вже проводити досліди шляхом статистичного моделювання чисельними методами в спеціалізованих програмах обчислювальних експериментів.

Характер руху барабана моталки при намотуванні штаби – обертальний, тому всі динамічні процеси при цьому процесі характеризуються моментом інерції.

При імітації процесу намотування штаби різного типорозміру встановлено, що маса намотуємого рулону прямо пропорційна часу намотки, а діаметр рулону має не прямолінійну зміну, що може викликати у приводі моталки коливання, для дослідження природи цих коливань в першу чергу необхідно встановити спектр власних частот та форму коливань. Визначені частота, амплітуда і форма коли-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

вань, дозволили дослідити пружні деформації, що виникають в процесі намотування штаби, у функції часу.

Графічний аналіз форми виникаючих коливань дозволив встановити, що пружні деформації від сил опору електродвигуна та обертових частин барабана моталки знаходяться в протифазі. Виконані розрахунки створюють передумови для дослідження вимушених коливань, що виникають в приводі барабана моталки.

Гречаний Олексій Миколайович - доктор філософії (PhD**),** ст. викладач, кафедра металургійного обладнання, Запорізький національний університет.

Васильченко Тетяна Олександрівна - к.т.н.**,** доцент, доцент кафедри металургійного обладнання, Запорізький національний університет.

Власов Андрій Олександрович - к.т.н.**,** доцент, доцент кафедри металургійного обладнання, Запорізький національний університет.

Виприжкін Павло Олександрович - магістр кафедри металургійного обладнання, Запорізький національний університет.

Якимчук Денис Іванович - магістр кафедри металургійного обладнання, Запорізький національний університет.

Hrechanyi Oleksii - Ph.D., lecturer, Departament Metallurgical Equipment, Zaporizhzhia National University.

Vasilchenko Tetyana - Candidate of Technical Sciences, Associate Professor of Departament Metallurgical Equipment, Zaporizhzhia National University.

Vlasov Andrii - Candidate of Technical Sciences, Associate Professor of Departament Metallurgical Equipment, Zaporizhzhia National University.

Vypryzhkin Pavlo - magister of Departament Metallurgical Equipment Zaporizhzhia National University.

Yakymchuk Denys Ivanovich - Master of the Department of Metallurgical Equipment, Zaporizhzhia National University.

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-07
УДК 519.8

К.Д. Караваєв

ПРО НЕОБХІДНІ УМОВИ ІСНУВАННЯ ЩІЛЬНИХ УПОРЯДКУВАНЬ В КЛАСИЧНІЙ ЗАДАЧІ ПАРАЛЕЛЬНОГО УПОРЯДКУВАННЯ

Анотація. Побудова схем паралелізації обчислень на ЕОМ при обробці великих масивів даних нерозривно пов'язана з задачами паралельного упорядкування. Розглядається класична задача мінімізації довжини упорядкування при заданій ширині, в якій шукане упорядкування є щільним. Досліджуються необхідні умови існування щільного упорядкування при застосуванні методу гілок та меж, пов'язані з обмеженістю потужності місць та можливістю їх заповнення. Отримані умови були зведені до однієї, запропоновано ефективні алгоритми її перевірки в загальному випадку та для графів, всі вершини яких знаходяться на критичних шляхах. В результаті дослідження також були отримані нові уточнені оцінки знизу довжини упорядкування та запропоновані узагальнення спеціальних упорядкувань, які враховують ширину упорядкування.
Ключові слова: графи, комбінаторна оптимізація, теорія розкладів, метод гілок та меж, розмітка графів, щільні упорядкування, оптимальний розподіл вершин.

Постановка проблеми. Бурхливий розвиток теорії розкладів у середині минулого століття був пов'язаний з різноманітністю важливих практичних застосувань задач, які в ній розглядаються. До них належать планування у багатьох сферах виробництва, таких як будівництво, хімічна та автомобільна промисловість, сільське господарство тощо; питання побудови розкладів транспорту і оптимізації логістики, побудова схем розпаралелювання обчислень на ЕОМ та багато іншого.

Окрема увага дослідників була приділена задачам, де на порядок виконання робіт накладені деякі технологічні обмеження. Однією з розповсюджених математичних моделей цих задач є задачі паралельного упорядкування. Вони формулюються як комбінаторні оптимізаційні задачі на орієнтованих ациклічних графах, вершини яких відповідають роботам, а дуги задають частковий порядок їх виконання.

Класичною вважається наступна постановка [1]. Нехай задано ациклічний орграф $G(V,U)$ та параметр-ширина h, який відповідає наявній кількості ви-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

конавців. Необхідно побудувати таке упорядкування S його вершин, тобто розподілити їх між місцями, розташованими в лінію, в якому:

- дотримано виконання технологічних обмежень, тобто, якщо між деякими вершинами v_i та v_j існує шлях у G, то v_i знаходиться в S лівіше за v_j;
- на кожному місці знаходиться не більше h вершин;
- кількість непорожніх місць у ньому l (його довжина) є мінімальною.

Довжина упорядкування є формалізацією загального часу виконання усіх завдань. Вважається, що виконавці є універсальними, а час виконання всіх завдань є однаковим та дорівнює 1. Таку задачу прийнято позначати $S(G, h, l)$. Отримане упорядкування називають щільним, якщо на кожному місці розташовано рівно h вершин (таке упорядкування завжди є оптимальним).

Окрім задачі $S(G, h, l)$ також розглядається задача $S(G, l, h)$, в якій за заданою довжиною упорядкування потрібно побудувати упорядкування з мінімальним значенням ширини.

В загальному випадку ці задачі виявилися NP-важкими, тому для знаходження точних розв'язків застосовуються методи направленого перебору, зокрема метод гілок та меж. Точні поліноміальні алгоритми відомі лише для деяких спеціальних випадків.

Аналіз останніх досліджень і публікацій. Оскільки поліноміальна розв'язність задачі $S(G, h, l)$ при фіксованих $h > 2$ є невідомою, то основними напрямками досліджень цієї задачі є пошук класів графів, для яких існують точні поліноміальні алгоритми, та побудова наближених алгоритмів.

Значне просування протягом останніх років було здійснено в напрямку побудови $(1+\varepsilon)$-наближених алгоритмів, що мають квазіполіноміальну складність [2-5]. Найкращою відомою до того відносною похибкою наближених алгоритмів для $h = 3$ було значення $4/3$ [6].

Також ведеться активна розробка алгоритмів, які засновані на метаевристиках, таких як Баєсова оптимізація та генетичні алгоритми [7,8]. Окрім цього проводяться дослідження в напрямку узагальнення точних відомих поліноміальних алгоритмів для розширення класу задач, для яких вони можуть знаходити точні розв'язки [9,10].

Окрім класичної задачі науковцями також розглядаються її узагальнення, які мають ускладнені структури завдань та виконавців, додаткові умови на виконання робіт, інші цільові функції, тощо [11]. В зв'язку з розвитком туманних

обчислень протягом останніх років, багато робіт присвячені дослідженню таких задач в рамках саме цієї сфери застосування [12].

В рамках дослідження методу гілок та меж ведеться робота щодо побудови схем перебору для узагальнених задач і скорочення перебору шляхом уточнення оцінок знизу довжини та виключення гілок, що відповідають ізоморфним підграфам [13,14].

Мета дослідження. Дослідити обмеження, що накладає на проміжні графи у методі гілок та меж умова щільності шуканого упорядкування. Отримати необхідні умови існування щільного упорядкування та запропонувати ефективні методи їх перевірки.

Виклад основного матеріалу дослідження. У роботі [9] було показано, що будь-яка класична задача паралельного упорядкування може буде зведена до задачі, в якій шукане упорядкування є щільним. Очевидною перевагою такої задачі є те, що довжина оптимального упорядкування відома і потрібно лише його побудувати.

Щільність упорядкування накладає ряд додаткових обмежень на проміжні упорядкування, що будуються при використанні методу гілок та меж [13], тому їх врахування може дозволити значно скоротити кількість гілок, що підлягають розгляду.

Відзначимо спершу, що через те, що довжина упорядкування відома, можемо визначити місця, які в упорядкуванні може займати кожна з вершин. Розглянемо узагальнену схему визначення крайнього лівого та правого місця, які в упорядкуванні може займати вершина.

Для визначення цих місць в задачі $S(G,l,h)$ використовуються спеціальні упорядкування $\underline{S}$ та $\overline{S}$ [1]. Для кожної вершини v_i визначаються місця $\underline{\mu_i}$ та $\overline{\mu_i}$, які вона займає в цих упорядкуваннях відповідно. Тоді проміжок допустимих місць, що може займати ця вершина, визначається наступним чином: $\left[\underline{\mu_i}; \overline{\mu_i} + l - \underline{l}\right]$, де $\underline{l}$ – довжина критичного шляху у графі G.

Відмітимо, що цей алгоритм неможливо застосувати до задачі $S(G,h,l)$ в загальному випадку, оскільки довжина упорядкування є невідомою величиною. Помітимо також, що в наведених міркуваннях не враховується інформація про допустиму ширину упорядкування h, бо для $S(G,l,h)$ вона є шуканою величиною. Проте для щільного упорядкування в задачі $S(G,h,l)$ обидві величини є

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

відомими, тому зможемо більш точно визначити межі можливого розташування вершин.

Значення $\underline{\mu_i}$ та $\underline{l} - \overline{\mu_i} + 1$ можна розглядати як елементарні оцінки знизу довжини упорядкування для підграфів $\underline{G_i}$ та $\overline{G_i}$, можливо порожніх, вершини яких мають шляхи до вершини v_i та з цієї вершини відповідно. Дійсно, $\underline{\mu_i}$ та $\underline{l} - \overline{\mu_i} + 1$ на одиницю більші за довжини критичних шляхів у цих графах (необхідність додавати одиницю пов'язана з відсутністю вершини v_i у графах). Враховуючи, що всі вершини графів $\underline{G_i}$ та $\overline{G_i}$ мають бути розташованими в упорядкуванні ліворуч та праворуч від v_i відповідно, можемо замість довжини критичного шляху використати будь-яку іншу відому оцінку знизу для довжини упорядкування, оскільки це є теоретично мінімально можливою кількістю місць, на яких можуть бути розташовані всі вершини графу. Таким чином отримаємо деякі уточнені значення $\underline{\mu_i^h}$ та $\overline{\mu_i^h}$, які можемо використовувати аналогічним чином.

Відмітимо також, що такі ж міркування можна застосовувати й при обчислені оцінок, які використовують спеціальні упорядкування $\underline{S}$ та $\overline{S}$. Таким чином зможемо побудувати упорядкування $\underline{S^h}$ та $\overline{S^h}$, в яких кожна вершина v_i графу займає крайнє ліве та крайнє праве місце у відповідності з деякою оцінкою довжини для графів $\underline{G_i}$ та $\overline{G_i}$. Ті оцінки, що використовують упорядкування $\underline{S^h}$ та $\overline{S^h}$ замість $\underline{S}$ та $\overline{S}$ будуть уточненими оцінками довжини знизу.

Зауважимо, що упорядкування $\underline{S^h}$ та $\overline{S^h}$ мають дві ключові відмінності від упорядкувань $\underline{S}$ та $\overline{S}$: по-перше, вони можуть мати порожні місця, по-друге, їх довжини можуть бути різними. Так, наприклад, для графу $K_{1,4}$, $h = 3$ та базової оцінки довжини упорядкування $\overline{S^h}$ має порожнє друге місце, його довжина дорівнює 3, в той же час довжина упорядкування $\underline{S^h}$ дорівнює 2. З цієї причини при визначені крайнього правого допустимого місця в упорядкуванні замість довжини критичного шляху у графі потрібно використовувати довжину упорядкування $\overline{S^h}$.

Усі наведені міркування підсумовано у наступному алгоритмі.

Алгоритм визначення поміток $\underline{\mu_i^h}$

1. Вершин без вхідних дуг отримують помітки $\underline{\mu_i^h} = 1$.

2. Розглянемо множину вершин $\underline{V}$, які ще не мають поміток, а усі їх попередники вже помічені.

3. Для кожної вершини $v_i \in \underline{V}$ побудуємо спеціальне упорядкування $\underline{S_i^h}$, в якому попередники v_j вершини v_i будуть розташовані відповідно на місцях $\underline{\mu_j^h}$.

4. Вершина v_i отримує помітку $\underline{\mu_i^h} = 1 + f\left(\underline{S_i^h}\right)$, де $f(\cdot)$ – деяка відома оцінка знизу довжини упорядкування.

5. Якщо у графі залишилися вершини без поміток повертаємося на 2, інакше кінець.

Аналогічні помітки $\overline{\mu_i^h}$ можна отримати, якщо рухатись від вершин, що не мають вихідних дуг.

Перейдемо тепер до розгляду обмежень, що накладає вимога щільності на проміжні упорядкування. Тож нехай на деякому етапі побудови дерева розгалужень вже заповненими в упорядкуванні є перші k місць та нерозташованими залишаються вершини підграфу $\tilde{G}$. Будемо розглядати обмеження в порядку складності їх перевірки.

Очевидно, якщо у графі $\tilde{G}$ менше ніж h відкритих вершин, то побудувати щільне упорядкування вже не вдасться, тому поточну гілку можна відсікти.

Помітимо, що в процесі розташування вершин підграфи $\overline{G_i}$ залишаються незмінними, а отже значення поміток $\overline{\mu_i^h}$ можна порахувати один раз для початкового графу G, враховуючи у подальшому, що їх значення потрібно розглядати відносно вихідного графу, а не графу $\tilde{G}\left(\tilde{V}, \tilde{U}\right)$ (вони будуть відрізнятися на значення різниці між довжинами упорядкувань $\overline{S^h}$ для цих двох графів). Для зручності подальшого викладення будемо використовувати зведені помітки крайнього правого місця, яке вершина може займати в упорядкуванні: $\overline{\lambda_i^h} = \overline{\mu_i^h} + l - \overline{l^h}$, де $\overline{l^h}$ – довжина упорядкування $\overline{S^h}$ для початкового графу G.

Розглянемо тепер множини вершин $\overline{V_j^h} = \left\{v_i \in \tilde{V} : \overline{\lambda_i^h} \le j\right\}$, $j = k+1, \ldots, l$, тобто тих вершин, які мають бути розташовані у шуканому упорядкуванні до місця j включно. З іншого боку, кількість вершин, що можна розташувати на місця з $k+1$ по j дорівнює $h * (j - k)$. Зрозуміло, що у випадку, коли потуж-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ність відповідної множини перевищує цю кількість, тобто $\left|\overline{V_j^h}\right| > h * (j - k)$, побудувати щільне упорядкування також вже не вдасться і подальші розгалуження можна не проводити. Відзначимо, що додатково можна було б перевіряти, що всі множини $\overline{V_j^h}$, $j = 1,\dots,k$ є порожніми, проте достатньо перевірити лише множину $\overline{V_k^h}$, за побудовою. Ця перевірка вписується у загальну схему, оскільки маємо $h * (k - k) = 0$. Маємо для цього випадку $\overline{\lambda_i^h} \le k \Rightarrow l - k \le \overline{l^h} - \overline{\mu_i^h} < \overline{l^h} - \overline{\mu_i^h} + 1$, тобто довжина критичного шляху у $\tilde{G}$ буде перевищувати кількість місць доступних для розташування.

Зауважимо, що тут і далі, довжина критичного шляху розуміється у дещо узагальненому сенсі як мінімальна кількість місць, необхідна для розташування всіх вершин графу відповідно до деякого критерію.

Для перевірки наступних обмежень потрібно обчислити помітки $\underline{\mu_i^h}$ для вершин графу $\tilde{G}$, які на відміну від $\overline{\mu_i^h}$ постійно змінюються внаслідок розташування вершин. Також будемо використовувати зведені помітки крайнього лівого допустимого місця: $\underline{\lambda_i^h} = \underline{\mu_i^h} + k$.

Порівняємо тепер для кожної вершини помітки $\underline{\lambda_i^h}$ та $\overline{\lambda_i^h}$. Якщо знайдеться така вершина v_i, для якої $\underline{\lambda_i^h} > \overline{\lambda_i^h}$, то щільне упорядкування побудувати також не вдасться, оскільки для розміщення всіх нащадків та попередників цієї вершини знадобиться більше ніж $l - k$ місць: $\underline{\lambda_i^h} > \overline{\lambda_i^h} \Rightarrow \underline{\mu_i^h} + \left(\overline{l^h} - \overline{\mu_i^h}\right) > l - k$. Перейдемо тепер до вершин, для яких ці помітки співпадають: $V_j^h = \left\{v_i \in \tilde{V} : \underline{\lambda_i^h} = \overline{\lambda_i^h} = j\right\}$, $j = k + 1,\dots,l$. Всі вершини множин V_j^h можуть бути розташовані лише на місці j, тоді, якщо хоча б одна з них містить більше ніж h елементів, то також не отримаємо щільне упорядкування. Отже, отримали необхідну умову: $\left|V_j^h\right| \le h$.

Залишилося розглянути вершини, для яких $\underline{\lambda_i^h} < \overline{\lambda_i^h}$. Побудуємо множини $V_{j_1,j_2}^h = \left\{v_i \in \tilde{V} : j_1 \le \underline{\lambda_i^h}, \overline{\lambda_i^h} \le j_2\right\}$, $k + 1 \le j_1 \le j_2 \le l$, всі вершини цих множин можуть бути розміщенні лише на місцях з відрізку $[j_1; j_2]$. З іншого боку цей діапазон вміщує $h * (j_2 - j_1 + 1)$ вершин, а отже у випадку, коли потужність якоїсь з

цих множин перевищує наведене значення, також не вдасться побудувати щільне упорядкування, тобто має виконуватись: $\left|V^{h}_{j_1, j_2}\right| \le h * (j_2 - j_1 + 1)$.

Підсумовуючи розглянуті випадки, бачимо, що неможливо отримати щільне упорядкування через те, що для розміщення вершин графу $\tilde{G}$ необхідно більше ніж $l - k$ місць. А отже, якби вдалося побудувати оцінку знизу довжини, яка б враховувала всі зазначені характеристики у структурі графу, то не потрібно було б перевіряти кожен пункт окремо.

Відмітимо, що з виконання останньої необхідної умови існування щільного упорядкування випливає виконання майже всіх попередніх більш простіших умов. По-перше, легко побачити, що $V^{h}_{j} = V^{h}_{j, j}$, $j = k + 1, \dots, l$ та $h * (j - j + 1) = h$, а отже умова для вершин з рівними зведеними помітками виконується. По-друге, якщо $\left|\overline{V^{h}_{k}}\right| = 0$, тоді $\overline{V^{h}_{j}} = V^{h}_{k+1, j}$, $j = k + 1, \dots, l$ і $h * (j - (k + 1) + 1) = h * (j - k)$, тобто виконується умова для вершин, права зведена помітка яких не перевищує j. По-третє, розглянемо множину $V^{h}_{k+2, l}$: вона не містить відкритих вершин, оскільки всі відкриті вершини мають помітки $\underline{\lambda^{h}_{i}} = k + 1$. Тоді, якщо їх кількість менша за h, то матимемо $\left|V^{h}_{k+2, l}\right| > h * (l - k) - h = h * (l - (k + 2) + 1)$, а отже відповідна умова буде порушена. Останні дві умови, пов'язані з $\overline{V^{h}_{k}}$ та вершинами, для яких $\underline{\lambda^{h}_{i}} > \overline{\lambda^{h}_{i}}$, можуть бути замінені перевіркою довжини критичного шляху у $\tilde{G}$, як показано вище.

Відзначимо також, що майже в усіх попередніх необхідних умовах перевірялося обмеження потужності місць, а не можливість їх заповнення. Так, наприклад, можна розглядати множини вершин, які можуть бути розташовані на заданому проміжку: $\dot{V}^{h}_{j_1, j_2} = \left\{v_i \in \tilde{V} : [j_1; j_2] \cap \left[\underline{\lambda^{h}_{i}}; \overline{\lambda^{h}_{i}}\right] \neq \varnothing\right\}$, $k + 1 \le j_1 \le j_2 \le l$. Очевидно, що для щільного заповнення проміжку місць необхідно, щоб потужності всіх цих множин були не меншими за $h * (j_2 - j_1 + 1)$. Розглянемо тепер множини $V^{h}_{k+1, j_1 - 1}$ та $V^{h}_{j_2 + 1, l}$. За побудовою, ці множини, якщо вони непорожні, містять вершини, які можна розташувати виключно на проміжках $[k + 1; j_1 - 1]$ та $[j_2 + 1; l]$ відповідно, а отже вони не містять спільних елементів. Більш того, перетин $\dot{V}^{h}_{j_1, j_2}$ з цими множинами також є порожнім, оскільки проміжки не перетинаються, і також можна побачити, що кожна вершина належить принаймні

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

одній з цих множин. Тоді у випадку недостатньої кількості вершин для розміщення $\left|V^{h}_{j_1, j_2}\right| < h * (j_2 - j_1 + 1)$ отримаємо, що

$$\left|V^{h}_{k+1, j_1\text{-}1} \cup V^{h}_{j_2+1, l}\right| = \left|V^{h}_{k+1, j_1\text{-}1}\right| + \left|V^{h}_{j_2+1, l}\right| > h * (l - k) - h * (j_2 - j_1 + 1) = h * (l - k - j_2 + j_1 - 1) =$$
$$= h * (l - (j_2 + 1) + 1) + h * ((j_1 - 1) - (k + 1) + 1),$$

а тоді, за принципом Діріхле, принаймні одна з необхідних умов для множин $V^{h}_{k+1, j_1 - 1}$ та $V^{h}_{j_2+1, l}$ буде порушена. А отже, цю необхідну умову також можна не перевіряти окремо.

Насправді, обидві зазначені необхідні умови можна об'єднати у одну. Для цього проведемо спочатку перетворення над умовою для діапазонів місць, аби звести її до більш зручного вигляду. Розділимо обидві частини формули на h та скористаємося тим, що для $\forall a, c \in \mathrm{Z}, b \in \mathrm{N}$ виконується $a \le b * c \Leftrightarrow \frac{a}{b} \le c \Leftrightarrow \left\lceil \frac{a}{b} \right\rceil \le \lceil c \rceil = c$, оскільки $\lceil \cdot \rceil$ є неспадною функцією, тоді отримаємо $\left\lceil \frac{\left|V^{h}_{j_1, j_2}\right|}{h} \right\rceil \le j_2 - j_1 + 1$ для $k + 1 \le j_1 \le j_2 \le l$. Перенесемо тепер ліворуч праву частину нерівності та додамо l до обох частин, тоді для тих саме значень j_1, j_2 матимемо $(j_1 - 1) + \left\lceil \frac{\left|V^{h}_{j_1, j_2}\right|}{h} \right\rceil + (l - j_2) \le l$. Оскільки остання нерівність має виконуватись для всіх значень $k + 1 \le j_1 \le j_2 \le l$, то вона є еквівалентною наступній:

$$\max_{k+1 \le j_1 \le j_2 \le l} \left((j_1 - 1) + \left\lceil \frac{\left|V^{h}_{j_1, j_2}\right|}{h} \right\rceil + (l - j_2) \right) \le l.$$

Розглянемо детальніше доданки, що утворюють цей вираз. Другий доданок відповідає мінімальній кількості місць, необхідних для розташування вершин, що можуть розташовуватись виключно на місцях з діапазону $[j_1; j_2]$. Перший та третій доданок – мінімальній довжині критичних шляхів зліва та справа для цих вершин відповідно.

Розглянемо тепер випадок, коли довжина критичного шляху перевищує $l - k$. У цьому випадку знайдуться такі значення $j_1 > j_2$, для яких множина $V^{h}_{j_1, j_2}$ не буде порожньою. Тоді для цих вершин маємо $\left\lceil \frac{\left|V^{h}_{j_1, j_2}\right|}{h} \right\rceil \ge 1$, і тоді

$(j_1 - 1) + \left\lceil \frac{\left| V^{h}_{j_1, j_2} \right|}{h} \right\rceil + (l - j_2) \geq l + (j_1 - j_2) > l$, що вписується у загальну перевірку. А отже зможемо об'єднати цю умову з попередньою, замінивши допустиму область на множину пар (j_1, j_2), для яких множина $V^{h}_{j_1, j_2}$ не є порожньою. Це не зменшує загальності, оскільки у випадку $j_1 \leq j_2$ та $V^{h}_{j_1, j_2} = \varnothing$ матимемо $(j_1 - 1) + \left\lceil \frac{\left| V^{h}_{j_1, j_2} \right|}{h} \right\rceil + (l - j_2) = l - 1 - (j_2 - j_1) < l$, тобто умова буде виконана.

Відмітимо окремо, що для двох пар значень (j_1, j_2) та (j_1', j_2') таких, що $j_1' \leq j_1, j_2 \leq j_2', V^{h}_{j_1, j_2} = V^{h}_{j_1', j_2'}$, матимемо наступне співвідношення:

$$(j_1' - 1) + \left\lceil \frac{\left| V^{h}_{j_1', j_2'} \right|}{h} \right\rceil + (l - j_2') = (j_1 - 1) + \left\lceil \frac{\left| V^{h}_{j_1, j_2} \right|}{h} \right\rceil + (l - j_2) - (j_1 - j_1') - (j_2' - j_2) \leq$$

$$\leq (j_1 - 1) + \left\lceil \frac{\left| V^{h}_{j_1, j_2} \right|}{h} \right\rceil + (l - j_2),$$ а отже будь-яке розширення діапазонів не призводить до порушення умови і розглядати має сенс лише найкоротші з них.

З іншого боку, у випадку $j_1 = k + 1, j_2 = l$ множина $V^{h}_{j_1, j_2}$ міститиме всі вершини, за побудовою позначок $\underline{\lambda}^{h}_{i}$ та $\overline{\lambda^{h}_{i}}$, тобто $\left| V^{h}_{j_1, j_2} \right| = h * (l - k)$, оскільки шукане упорядкування є щільним. Тоді отримаємо $(j_1 - 1) + \left\lceil \frac{\left| V^{h}_{j_1, j_2} \right|}{h} \right\rceil + (l - j_2) = (k + 1 - 1) + \left\lceil \frac{h * (l - k)}{h} \right\rceil + (l - l) = l$, а отже шуканий максимум обмежений зверху та знизу значенням довжини упорядкування, і необхідна умова існування щільного упорядкування приймає остаточний вигляд:

$$l = \max_{(a,b): V^{h}_{a,b} \neq \varnothing} \left((a - 1) + \left\lceil \frac{\left| V^{h}_{a,b} \right|}{h} \right\rceil + (l - b) \right).$$

Розглянемо тепер як зсуви значень $\underline{\lambda}^{h}_{i}, \overline{\lambda^{h}_{i}}$ на константи $\underline{\Delta}$ та $\overline{\Delta}$ відповідно впливають на отриманий вираз. Зафіксуємо деяку довільну пару значень $(a, b): V^{h}_{a,b} \neq \varnothing$. Розглянемо тепер пару $\left(a + \underline{\Delta}, b + \overline{\Delta}\right)$, зрозуміло, що в рамках зсунутих позначень множина вершин $V^{h}_{a + \underline{\Delta}, b + \overline{\Delta}}$ співпаде з множиною $V^{h}_{a,b}$ у вихідних позначеннях, а отже другий доданок виразу не зміниться. Для третього

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

доданку маємо $\left(l+\overline{\Delta}\right)-\left(b+\overline{\Delta}\right)=l-b$, а отже зсув також не впливає і на його значення. Для першого доданку – $(a+\underline{\Delta}-1)=(a-1)+\underline{\Delta}$, тобто підсумкове значення правої частини зміниться на $\underline{\Delta}$.

Відзначені спостереження дозволяють перейти від розгляду поміток $\underline{\lambda_i^h}, \overline{\lambda_i^h}$ до $\underline{\mu_i^h}, \overline{\mu_i^h}$. При цьому допустима множина прийме вигляд $X=\left\{(a,b): 1\le a\le \underline{l^h}, 1\le b\le \overline{l^h}, V_{a,b}^h\ne\varnothing\right\}$, де $\underline{l^h}$, $\overline{l^h}$ – довжини упорядкувань $\underline{S^h}$ та $\overline{S^h}$ для графу $\tilde{G}$, тут і далі $V_{a,b}^h$ будуються, використовуючи значення $\underline{\mu_i^h}, \overline{\mu_i^h}$ замість $\underline{\lambda_i^h}, \overline{\lambda_i^h}$, без зміни позначень. Тоді умова прийме вигляд:

$$l=k+\max_{(a,b)\in X}\left((a-1)+\left\lceil\frac{\left|V_{a,b}^h\right|}{h}\right\rceil+\left(\overline{l^h}-b\right)\right).$$

Така форма є більш звичною для методу гілок та меж, оскільки тепер всі елементи, які приймають участь в пошуку максимуму пов'язані лише з проміжним графом $\tilde{G}$. Введена допустима множина також є більш зручною для опрацювання випадку $\underline{l^h}\ne\overline{l^h}$.

У загальному випадку, отриманий максимум буде оцінкою знизу довжини упорядкування для проміжного графу $\tilde{G}$, що випливає з наведеного вище аналізу доданків, з яких він складається. Можна також переконатися, що він є узагальненням вже відомих оцінок [14]. Це дозволяє очікувати, що використання такої оцінки знизу також дозволить скоротити кількість розгалужень у методі гілок та меж у загальному випадку.

Перейдемо тепер до питання, як можна ефективно обчислювати значення цього максимуму за відомих значень $\left(\underline{\mu_i^h}, \overline{\mu_i^h}\right), i=\overline{1,n}$.

Скористаємося тим, що маємо перевірити всі проміжки $[a;b]$ по черзі. Зафіксуємо деякі значення a та b. Легко побачити, що $V_{a,b+1}^h=V_{a,b}^h\cup\left\{v_i\in\tilde{V}:\overline{\mu_i^h}=b+1, \underline{\mu_i^h}\ge a\right\}$ та $V_{a+1,b}^h=V_{a,b}^h/\left\{v_i\in\tilde{V}:\overline{\mu_i^h}\le b, \underline{\mu_i^h}=a\right\}$.

Обчислимо спочатку числовий масив V_h, де на кожному місці $k=1,\dots,\overline{l^h}$ буде знаходитись кількість вершин, для яких $\overline{\mu_i^h}=k$, тоді $\left|V_{1,b}^h\right|=\sum_{k=1}^{b}V_h[k]$. Це дає змогу обчислити значення виразу для всіх значень b за лінійний час.

Для переходу до $a = 2$ маємо виключити з розгляду всі вершини, для яких $\underline{\mu_i^h} = 1$. Для цього достатньо у масиві V_h зменшити відповідні значення на кількість вершин $v_i : \overline{\mu_i^h} = k, \underline{\mu_i^h} = 1$. Після цього матимемо $\left|V_{2,b}^h\right| = \sum_{k=1}^{b} V_h[k]$. Продовжуючи таким чином знайдемо шуканий максимум.

Помітимо також, що якщо попередньо відсортуємо масив позначок за зростанням $\underline{\mu_i^h}$, то кожне значення масиву потрібно буде використати лише 1 раз і складність алгоритму складатиме $O\left(\underline{l^h} * \overline{l^h} + n\right)$.

Псевдокод запропонованого алгоритму наведено на рис. 1.

Вхідні дані: ширина упорядкування h,

лексикографічно відсортований масив пар $L = \left\{\left(\underline{\mu_i^h}, \overline{\mu_i^h}\right), i = \overline{1,n}\right\}$.

Вихідні дані: значення оцінки знизу $\tilde{l}$.

```
procedure enh_lower_estimate(L, h)
    μ̄_min := min_i(μ̄_i^h); μ̄_max := max_i(μ̄_i^h);
    for k := μ̄_min to μ̄_max
        V_h[k] := 0;
    end for
    for i := 1 to n
        V_h[μ̄_i^h] := V_h[μ̄_i^h] + 1;
    end for

    l̃ := 0; j := 1;
    while j ≤ n
        cnt := 0;
        for k := μ̄_min to μ̄_max
            cnt := cnt + V_h[k];
            if cnt > 0 then
                l̃ := max(l̃, μ_j^h − 1 + ⌈cnt/h⌉ + μ̄_max − k);
            end if
        end for
        while j ≤ n and (j = 1 or μ_{j−1}^h = μ_j^h)
            V_h[μ̄_j^h] := V_h[μ̄_j^h] − 1;
            j := j + 1;
        end while
    end while

    return l̃;
end procedure
```

Рисунок 1 – Псевдокод обчислення оцінки для довільного графу

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

У випадку, коли упорядкування $\underline{S}^h$ та $\overline{S^h}$ співпадають (всі вершини знаходяться на критичних шляхах), обчислення може бути пришвидшене до лінійної складності $O\left(\underline{l}^h\right)$, оскільки $\underline{\mu}_i^h = \overline{\mu_i^h}, i = \overline{1,n}$ і кількість вершин у $V_{a,b}^h$ можна обчислити, підсумувавши кількості вершин на місцях з a по b в упорядкуванні $\underline{S}^h$ ($\overline{S^h}$).

Цільова функція приймає вигляд $\max\limits_{1\le a\le b\le \underline{l}^h}\left((a-1)+\left\lceil\frac{1}{h}\sum\limits_{i=a}^{b}\left|\underline{S}^h[i]\right|\right\rceil+\left(\overline{l^h}-b\right)\right)$. Помітимо, що елемент $\overline{l^h}$ є сталим, а $l=(b-a+1)$ дорівнює кількості місць, що приймає участь у підсумуванні, тоді пошук максимуму зводиться до наступної задачі:

$$\max_{1\le a\le b\le \underline{l}^h}\left(\left\lceil\frac{1}{h}\sum_{i=a}^{b}\left|\underline{S}^h[i]\right|\right\rceil-l\right)=\max_{1\le a\le b\le \underline{l}^h}\left\lceil\frac{1}{h}\sum_{i=a}^{b}\left|\underline{S}^h[i]\right|-l\right\rceil=\max_{1\le a\le b\le \underline{l}^h}\left\lceil\frac{1}{h}\left(\sum_{i=a}^{b}\left|\underline{S}^h[i]\right|-l*h\right)\right\rceil=$$

$$=\max_{1\le a\le b\le \underline{l}^h}\left\lceil\frac{1}{h}\sum_{i=a}^{b}\left(\left|\underline{S}^h[i]\right|-h\right)\right\rceil=\left\lceil\frac{1}{h}\max_{1\le a\le b\le \underline{l}^h}\sum_{i=a}^{b}\left(\left|\underline{S}^h[i]\right|-h\right)\right\rceil.$$

Тобто отримали задачу пошуку безперервної послідовності з елементами $\left|\underline{S}^h[i]\right|-h$, яка має максимальну суму. Така задача ефективно розв'язується за допомогою алгоритму Кадана [15], складність якого і складає $O\left(\underline{l}^h\right)$.

Псевдокод цього алгоритму наведено на рис. 2.

Вхідні дані: ширина упорядкування h,
масив потужностей місць упорядкування $\underline{S_i^h}$: $C=\left\{\left|\underline{S_i^h}\right|, i=1,\dots,\underline{l^h}\right\}$.
Вихідні дані: значення оцінки знизу $\tilde{l}$.

```
procedure enh_lower_estimate_cp(C, h)
    cnt := 0; l̃ := l̲ʰ; l := 0;
    for i := 1 to l̲ʰ
        cnt := cnt + C[i];
        l := l + 1;
        if cnt ≥ l * h then
            l̃ := max(l̃, ⌈cnt/h⌉ + l̲ʰ − l);
        else
            cnt := 0; l := 0;
        end if
    end for

    return l̃;
end procedure
```

Рисунок 2 – Псевдокод обчислення оцінки для графу, всі вершини якого знаходяться на критичних шляхах

Висновки. Було розглянуто і досліджено низку необхідних умов існування щільного упорядкування, пов'язаних з обмеженістю потужності місць та можливістю їх заповнення. Ці умови вдалося об'єднати в одну та запропонувати ефективний алгоритм її перевірки.

В результаті також були отримані нові уточнені оцінки знизу довжини упорядкування та запропоновані узагальнення спеціальних упорядкувань $\underline{S^h}$ та $\overline{S^h}$, які враховують значення ширини h.

Подальшого дослідження потребує експериментальна оцінка прискорення методу гілок та меж при використанні запропонованих оцінок довжини упорядкування, як для графів з щільним упорядкуванням, так і в загальному випадку. Цікавим також є пошук класів графів, для яких отримані оцінки дають точне значення довжини, та вивчення алгоритмів, заснованих на методі гілок та меж з обмеженою глибиною пошуку.

ЛІТЕРАТУРА

1. Бурдюк В. Я., Турчина В. А. Алгоритмы параллельного упорядочения: учебное пособие. Днепропетровск: ДГУ, 1985. 83 с.

2. Levey E., Rothvoss T. A (1+epsilon)-Approximation for Makespan Scheduling with Precedence Constraints Using LP Hierarchies. *SIAM Journal on Computing*. 2019. P. 201–217.

3. Garg Sh. Quasi-PTAS for scheduling with precedences using LP hierarchies. *In 45th International Colloquium on Automata, Languages, and Programming, ICALP*. 2018.

4. Li S. Towards PTAS for Precedence Constrained Scheduling via Combinatorial Algorithms. *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2021. P. 2991–3010.

5. Nederlof J., Swennenhuis C. M. F., Węgrzycki K. A Subexponential Time Algorithm for Makespan Scheduling of Unit Jobs with Precedence Constraints. 2023. 25 p. (Preprint. arXiv; arXiv:2312.03495).

6. Gangal D., Ranade A. Precedence constrained scheduling in (2–7/3p+1)optimal. *Journal of Computer and System Sciences*. 2008. Vol. 74, no. 7. P. 1139–1146.

7. Muhuri P. K., Biswas S. K. Bayesian optimization algorithm for multi-objective scheduling of time and precedence constrained tasks in heterogeneous multiprocessor systems. *Applied Soft Computing*. 2020. Vol. 92. P. 106–274.

8. Yoo M. Real-time task scheduling by multiobjective genetic algorithm. *Journal of Systems and Software*. 2009. Vol. 82, no. 4. P. 619–628.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

9. Turchyna V., Karavaiev K. Analysis of algorithms for constructing dense sequencing of digraphs vertices. *Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020)*, Zaporizhzhia, 2020. P. 690–703.
10. Chardon M., Moukrim A. The Coffman-Graham Algorithm Optimally Solves UET Task Systems with Overinterval Orders. *SIAM Journal on Discrete Mathematics*. 2005. Vol. 19, no. 1. P. 109–121.
11. Vásquez Ó. C. The Scheduling Zoo. URL: http://schedulingzoo.lip6.fr/ (date of access: 01.03.2024).
12. Li K. Scheduling Precedence Constrained Tasks for Mobile Applications in Fog Computing. *IEEE Transactions on Services Computing*. 2022. P. 1–14.
13. Караваєв К.Д., Турчина В.А. Аналіз впливу автоморфізму графу на схеми направленого перебору. *Збірник наукових праць «Питання прикладної математики і математичного моделювання»*, м. Дніпро, 2021. Вип. 21. С. 94–104.
14. Турчина В.А., Караваєв К.Д. Дослідження оцінок довжини паралельного упорядкування вершин графу. *Збірник наукових праць «Питання прикладної математики і математичного моделювання»*, м. Дніпро, 2018. Вип. 18. С. 186–195.
15. Bentley J. Programming pearls. *Communications of the ACM*. 1984. Vol. 27, no. 9. P. 865–873.

REFERENCES

1. Burdiuk V. Ya., Turchyna V. A. Parallel sequencing algorithms: training manual. Dnipropetrovsk: DSU, 1985. 83 p.
2. Levey E., Rothvoss T. A (1+epsilon)-Approximation for Makespan Scheduling with Precedence Constraints Using LP Hierarchies. *SIAM Journal on Computing*. 2019. P. 201–217.
3. Garg Sh. Quasi-PTAS for scheduling with precedences using LP hierarchies. *In 45th International Colloquium on Automata, Languages, and Programming, ICALP*. 2018.
4. Li S. Towards PTAS for Precedence Constrained Scheduling via Combinatorial Algorithms. *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2021. P. 2991–3010.
5. Nederlof J., Swennenhuis C. M. F., Węgrzycki K. A Subexponential Time Algorithm for Makespan Scheduling of Unit Jobs with Precedence Constraints. 2023. 25 p. (Preprint. arXiv; arXiv:2312.03495).

6. Gangal D., Ranade A. Precedence constrained scheduling in (2–7/3p+1)optimal. *Journal of Computer and System Sciences*. 2008. Vol. 74, no. 7. P. 1139–1146.
7. Muhuri P. K., Biswas S. K. Bayesian optimization algorithm for multi-objective scheduling of time and precedence constrained tasks in heterogeneous multiprocessor systems. *Applied Soft Computing*. 2020. Vol. 92. P. 106–274.
8. Yoo M. Real-time task scheduling by multiobjective genetic algorithm. *Journal of Systems and Software*. 2009. Vol. 82, no. 4. P. 619–628.
9. Turchyna V., Karavaiev K. Analysis of algorithms for constructing dense sequencing of digraphs vertices. *Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020)*, Zaporizhzhia, 2020. P. 690–703.
10. Chardon M., Moukrim A. The Coffman-Graham Algorithm Optimally Solves UET Task Systems with Overinterval Orders. *SIAM Journal on Discrete Mathematics*. 2005. Vol. 19, no. 1. P. 109–121.
11. Vásquez Ó. C. The Scheduling Zoo. URL: http://schedulingzoo.lip6.fr/ (date of access: 30.03.2024).
12. Li K. Scheduling Precedence Constrained Tasks for Mobile Applications in Fog Computing. *IEEE Transactions on Services Computing*. 2022. P. 1–14.
13. Karavaiev K. D., Turchuna V.A. Analysis of the effect of graph automorphism on state search schemes. *Problems of applied mathematics and mathematical modelling*, Dnipro, 2021. Vol. 21. P. 94–104.
14. Turchuna V. A., Karavaiev K. D. Investigation of the estimates of the length of parallel alignment of the vertices of the graph. *Problems of applied mathematics and mathematical modelling*, Dnipro, 2018. Vol. 18. P. 186–195.
15. Bentley J. Programming pearls. *Communications of the ACM*. 1984. Vol. 27, no. 9. P. 865–873.

Received 12.03.2024.
Accepted 14.03.2024.

On the necessary conditions for the existence of dense sequencing in the classical parallel sequencing problem

The rapid development of the scheduling theory in the middle of the last century was linked to the variety of important practical applications of the problems it considers. Special attention was paid to problems in which the order of job execution is subject to certain technological constraints. One of the common mathematical models of these problems is the parallel sequencing problem.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

We consider the classical problem of minimizing the length of a sequencing for a given width, in which the target sequencing is dense. Since the polynomial tractability of these problems for fixed width > 2 is unknown, the main areas of research on this problem include searching for classes of graphs for which exact polynomial algorithms exist, developing approximate algorithms and ways to prune state space search schemes.

Substantial progress has been made in recent years in the development of approximate algorithms with quasi-polynomial complexity and algorithms based on metaheuristics.

In addition to the classical problem, scientists also consider its generalizations, which have more complex structures of jobs and workers, additional constraints on the job execution, other objective functions, etc. Due to the development of fog computing in recent years, many articles have been devoted to the study of such problems within this particular application area.

The aim of this study was to investigate the constraints imposed on intermediate graphs by the condition of density of the target sequencing in the branch-and-bound method, to derive the necessary conditions for the existence of a dense sequencing and to propose methods to test them.

The necessary conditions for the existence of a dense sequencing when using the branch-and-bound method, related to the limited capacity of places and the possibility of filling them, are investigated. The obtained conditions were reduced to a single one, and efficient algorithms to test it in general and for graphs with all vertices on critical paths were proposed. In addition, the study also resulted in new improved lower bound estimates of the sequencing length and generalization of special sequencings in which the vertices occupy the leftmost and rightmost possible places, that take into account the sequencing width.

Keywords: graphs, combinatorial optimization, scheduling theory, branch-and-bound method, graph labeling, dense sequencings, optimal vertex distribution.

Караваєв Костянтин Дмитрович – аспірант кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара (м. Дніпро).

Karavaiev Kostiantyn – post-graduate student of the Department of Calculating Mathematics and Mathematical Cybernetics, Oles Honchar Dnipro National University, Dnipro, Ukraine.

DOI 10.34185/1562-9945-2-151-2024-08
UDC 519.85:004.42

E. Koshel

NEURAL NETWORK-ASSISTED CONTINUOUS EMBEDDING OF UNIVARIATE DATA STREAMS FOR TIME SERIES ANALYSIS

Abstract. Univariate time series analysis is a universal problem that arises in various science and engineering fields and the approaches and methods developed around this problem are diverse and numerous. These methods, however, often require the univariate data stream to be transformed into a sequence of higher dimensional vectors (embeddings). In this article, we explore the existing embedding methods, examine their capabilities to perform in real time, and propose a new approach that couples the classical methods with the neural network based ones to yield results that are better in both accuracy and computational performance. Specifically, the Broomhead King inspired embedding algorithm implemented in a form of an autoencoder neural network is employed to produce unique and smooth representation of the input data fragments in the latent space.
Keywords: time series analysis, time series embedding, dimensionality reduction

Introduction. Time series analysis is the process of extracting information from the series of data points in an attempt to obtain useful information about the system that produced it or to predict the next values in the series. The simplest form of the time series is a univariate series or just a sequence of one-dimensional measurements. If the system that the data was obtained from is sufficiently complex, the time series may exhibit periodic, quasi-periodic, chaotic behavior, drift, etc. It is easy for a human to see whether these behaviors are present or not simply by looking at the plotted sequence of data points, but this can hardly qualify as proper analysis. Luckily, there are methods developed in different fields that can be used for identification and quantification of different properties of the time series data. Most of these methods, however, have a hard requirements for the dimensionality of the problem and can yield nonsensical results if the dimension of the problem they are applied to is too low or too high. This issue can be solved by embedding the available time series data into a higher-dimensional space where the analysis methods can be applied.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Time series embedding is a mapping between the segments of the time series data and some vector space $\mathbb{R}^n$. This mapping is expected to be an injective function that produces a smooth trajectory of embedding vectors for adjacent time series data segments. In the following sections, we examine some of the methods for producing time series embeddings, see how well they are able to uphold the requirements listed earlier, compare their ability to preserve the qualitative characteristics of their source system, their resistance to noise, and the overall computational cost of each method.

Literature Review. The broad definition of time series embedding provides a lot of flexibility in selecting the mapping functions. Below is the review of some of the classical transformation methods.

One of the most well-known ways of embedding the time series data is the short-time Fourier transform (STFT) and related wavelet transform-based methods (WTs) which are widely used for feature extraction from a variety of signals [7, 10, 17] by transitioning from temporal to frequency domain representation of signal frames. The STFT algorithm is defined as follows:

1. Break down the signal into short overlapping frames.

2. Apply a window function to each frame to suppress the high frequency artifacts at the edges.

3. Pad the windowed frame with zeros to achieve the required frequency resolution and/or achieve the length required for fast Fourier transform (FFT) to have maximum efficiency.

4. Compute the discrete Fourier transform of the padded windowed frame.

5. Compute the squares of the first half of the resulting vector to obtain the relative magnitude of the frequencies that make up the original signal frame.

The STFT's result is a sequence of *n*-dimensional vectors where *n* is the number of "frequency bins" that depends on the chosen length of the frame. Selecting longer frames will result in more available frequency bins and thus higher resolution in the frequency domain representation of the signal. But this will not only introduce latency but also lower the temporal resolution of the STFT because the frequency information will be spread out over the longer stretches of the signal.

Another well-known class of embedding methods is time-delay based and is intended for phase-space reconstruction of dynamical systems but also used for feature extraction [3, 5] and data mining applications [14]. These methods operate under the assumption that there exists a twice-differentiable *observation function* $\alpha: \mathbb{R}^m \to \mathbb{R}$ that maps all points on the *m*-dimensional attractor of the smooth map

$f: \mathbb{R}^m \to \mathbb{R}^m$ to real numbers so that an embedding function can be constructed in the following form:

$$\varphi(X) = \left(\alpha(X), \alpha\left(f(X)\right), \ldots, \alpha\left(f^{k-1}(X)\right)\right),$$
$$\alpha\left(f^i(X)\right) = x_{t-\tau_i}, \quad (1)$$

where X is a point on the attractor in $\mathbb{R}^m$ space, x_t is the time-series observation at time t, and τ_i are lag values that specify the distance in time between the current point and its neighbors.

When performing a time-delay embedding, the researchers must solve two problems. First one is selecting a proper embedding dimension k which, according to the Takens's theorem [16], would ensure that the dynamics of the embedded trajectory will be completely equivalent to the dynamics of the original system. The second problem is selecting the proper lag values τ_i which is especially difficult considering that the Takens's theorem doesn't provide any guidance regarding this and assumes that the time series is not noisy and only considers the minimal lag time of 1.

To solve these problems, either separated or unified approaches can be taken. Separated approaches start by selecting the time delay value τ and use it to estimate the optimal embedding dimension employing one of the False Nearest Neighbors algorithm variations [2, 8, 9, 13]. The optimal embedding dimension is usually selected based on values of some statistic. Unified methods recognize the interdependence between the delay values and the embedding dimensions and employ methods that allow them to select the τ_i values dynamically as they increase the dimension of the embedding during the "embedding cycles", essentially optimizing the delay times and the dimensions in parallel. One prominent example of the unified approaches is the PECUZAL algorithm [15] which uses two statistics to objectively evaluate the quality of the embedding and decide whether or not to continue increasing the number of dimensions and select appropriate delay values for each dimension.

There is also an alternative to the delay embedding which applies more complex transformations to the time series to obtain trajectories in a higher-dimensional space. One such alternative is the Broomhead-King method for extracting qualitative dynamics from experimental data [1]. It applies the Takens's theorem directly to the time series with the smallest possible delay and an

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

embedding dimension d. Then singular value decomposition (SVD) is performed on the resulting embedding:

$$X = \frac{1}{\sqrt{N}} \begin{pmatrix} x_1 & x_2 & \cdots & x_d \\ x_2 & x_3 & \cdots & x_{d+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-d+1} & x_{N-d+2} & \cdots & x_N \end{pmatrix} = U \cdot S \cdot V^{tr}, \quad (2)$$

where N is the length of the time series, x_i is the value from time series adjusted to have the mean of 0, U is the new embedding space, S is a diagonal matrix that holds the "importance values" of each column of U. Since by the definition of SVD the matrix U is a unitary one, its columns form an orthonormal basis which means that the components of the high-dimensional trajectory it describes are guaranteed to not contain correlations and redundant information. The properties of the matrix S also allow the researchers to judge the level of importance of each of the columns of U and reduce its dimensionality by dropping columns with importance lower than some specified threshold.

Research methodology and results. The methods discussed above provide a reliable set of tools for time series embedding, dimensionality estimation and reduction. To make the further discussion easier, let us introduce some definitions and generalizations.

First, observe how the definitions of all the methods from the previous section can be rewritten in terms of matrix multiplication. The STFT is based on the discrete Fourier transform which can be expressed as a square complex-valued transformation matrix that the signal is multiplied by:

$$W = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{N-1} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix}, \quad (3)$$

where $\omega = e^{i\pi/N}$.

Similarly, the delay embedding methods can be reduced to a problem of finding a rectangular matrix that yields an embedding vector given a signal segment:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ a_{21} & a_{22} & \cdots & a_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ a_{\tau_{max}1} & a_{\tau_{max}2} & \cdots & a_{\tau_{max}d} \end{pmatrix}, \quad (4)$$

where columns of A are unitary vectors and $\forall j \exists i: a_{ij} = 1$. The number of columns d is the embedding dimension and the number of rows is given by the largest selected

delay value. As for the Broomhead-King approach, the transformation matrix can be trivially derived from its definition in (2):

$$A = (V^{tr})^{-1} \cdot S^{-1} = V \cdot S^{-1}. \quad (5)$$

Notably, the matrices (4) and (5) both describe which values of the time series segment contribute to the embedding vector by assigning coefficients to them. The difference is that the matrix (4) is very rigid and can only "select" one value from the time series per its column. The matrix (5), however, is a lot more flexible and can incorporate multiple time series values into a single embedding vector component, which is very similar to how the discrete Fourier transformation matrix operates.

Second, let's examine the equation (1) more closely. It introduces the assumption that there exists some unknown observation function α that associates real numbers to high-dimensional points on the attractor. Its definition does not constrain it to any particular structure and only requires it to be twice-differentiable. In practice, however, it is almost always assumed to be a linear transformation that selects a single coordinate from each point on the attractor. This assumption also informs the $\alpha\left(f^i(X)\right) = x_{t-\tau_i}$ equation which in the real world applications almost never the case due to the presence of noise and the fact that neither the map f nor the function α are known. Considering (3) and (5), the second part of the equation (1) can be rewritten as follows:

$$\alpha\left(f^i(X)\right) = \beta(x_t, x_{t-1}, \ldots, x_{t-l}) \quad (6)$$

where $\beta: \mathbb{R}^l \to \mathbb{R}$ is the *feature extraction function* that maps time series segments to the codomain of the unknown observation function α.

All of the above implies that if an optimal embedding can be achieved using matrices of type (4) and the properties of the SVD transformation show that an equally valid alternative embedding can be achieved with matrices of type (5), then an optimization problem can be formulated that optimizes the equation (6) with respect to a loss function designed to enforce the objective statistics used to evaluate the quality of the embedding in unified embedding methods.

Since both α and β functions are unknown, a data-driven approach for their modeling can be employed. More specifically, an autoencoder neural network architecture can be used to find a model that accurately maps the time series data to the codomains of α and β and back. The main idea of an autoencoder is to couple two functions – encoder E and decoder D – such that $x \approx D\left(E(x)\right)$ with the goal of

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

either increasing or decreasing the dimensionality of x in some latent space. Applying this idea with the context of equations (1) and (4), we get the following:

$$\begin{gathered} E([x_t, x_{t-1}, \ldots, x_{t-l}]) = \beta(x_t, x_{t-1}, \ldots, x_{t-l}) = Y, \\ D(Y) = \alpha\left(f^{k-1}(X)\right) = x_t. \end{gathered} \tag{7}$$

The advantage of using this neural network-based approach is that the structure of functions E and D is not limited to linear transformations like all the methods discussed in the previous section and can contain any number of nonlinearities that can be tailored to a specific time series being analyzed.

Let's now examine the performance of each of the discussed embedding methods using some synthetic data. It's best to chose a well-known system to make the benchmarking of it easier to reproduce in different computing environments, which is why the chaotic regime of the Lorenz system will be examined in the examples below. However, to acknowledge the previous remarks about the observation function α, it will be altered and instead of picking a specific coordinate of the points on the attractor, it will compute the norm of each point. To evaluate the preservation of the qualitative characteristics of the original attractor in the embedding space, their Lyapunov spectra will be compared.

All computations and simulations will be done using the Julia programming language and its various packages.

The system and the observation function that is used to generate the synthetic time series data are as follows:

$$\begin{gathered} \dot{x} = \sigma(y - x) \\ \dot{y} = x(\rho - z) - y, \\ \dot{z} = xy - \beta z \\ \sigma = 10, \rho = 28, \beta = 8/3, \\ \alpha(X) = \|X\|_2 \end{gathered} \tag{8}$$

The shape of the time series generated by (7) and modified to have a mean of 0.

To obtain the STFT embedding of the test time series, we set the matrix (3) size to 10 and apply it to windows of the time series with same length which have 9 overlapping samples. This transformation is different from all the others because it yields complex numbers in its output which makes it difficult to visualize. For the purposes of demonstration, we will avoid plotting the norms of the embeddings and instead simply omit their imaginary parts. This will help us to keep the output smooth and comparable to other embedding methods.

Next, the PECUZAL algorithm is applied to the time series and is used to select the embedding dimension and the delay values. In this example, the Julia implementation of the algorithm from the DelayEmbeddings.jl [4] package was used. No extra parameters were supplied to the algorithm aside from the time series itself.

The Broomhead-King algorithm was supplied two parameters for constructing the embedding – the time series data itself and the target embedding dimension of 20. The matrix (5) was constructed using the algorithm's output and all columns of the resulting matrix were discarded except the first 3. In this example, the Julia implementation of the algorithm from the ChaosTools.jl [4] package was used.

Finally, for constructing the autoencoder model, two simple linear models were selected and trained:

$$\begin{aligned} E(x) &= A_e x + b_e, \\ D(y) &= A_d y + b_d, \end{aligned} \tag{9}$$

where A_e is a N by m matrix that maps N points of the trajectory into the m dimensional latent space, A_d is a m by 1 matrix that maps the points in the latent space to the last point of the converted segment of the time series as shown in (7). The b_e and b_d are the bias vectors.

To compare the quality of each embedding, let's calculate their respective Lyapunov spectra to see how well they preserve the dynamics of the original system. Since the embedded trajectories do not have analytical representations, numerical approach to calculating the spectrum will be used. The algorithm that is used is a QR decomposition-based one described in [6] but instead of using the analytical representation of the original map, the local Jacobian matrix is computed numerically for each point based on its nearest neighbors using the algorithm from [11]. The Lyapunov spectra values are given in Table 1.

Another embedding quality measure that we can use is the L statistic that the PECUZAL algorithm uses to evaluate the quality of its own embeddings, the higher the value of the statistic, the less correlation there is between the components of the embedding vector The values of the L statistic for the original system trajectory and each of the embeddings are presented in the Table 2.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Table 1

Lyapunov spectrum values for the actual attractor compared to the values calculated for the trajectories in the embedding spaces

	True values	STFT	PECUZAL	Broomhead-King	Autoencoder
λ_1	0.91	3.9	2.18	0.04	1.9
λ_2	0	1.08	0.5	0	0.22
λ_3	-14.57	-28.02	-13	-0.19	-12.07

As evident from both the Table 1 and Table 2, most embedding methods were able to accurately reconstruct the trajectory in a way that preserved the original dynamics of the system. The only outlier is the Broomhead-King approach that turned out to not have the amount of information in its first few columns required to create reliable embeddings.

Table 2

L statistic values for the actual attractor compared to the values calculated for the trajectories in the embedding spaces

	True values	STFT	PECUZAL	Broomhead-King	Autoencoder
	-0.09	1.41	0.43	-0.48	0.4

Another thing to point out is that the autoencoder can sometimes produce intersecting trajectories if the loss function is not constructed properly or its non-linearities do not allow for the valid behavior to be achieved [12].

Conclusions. In this article, we demonstrated and evaluated the performance of different embedding methods and provided a basic framework for constructing new embedding algorithms based on the autoencoder neural network approach.

REFERENCES

1. Broomhead D.S. and King G. P. (1986). Extracting qualitative dynamics from experimental data. Physica D: Nonlinear Phenomena, Volume 20 (2-3), Pages 217-236. https://doi.org/10.1016/0167-2789(86)90031-X.

2. Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. Physica D: Nonlinear Phenomena, Volume 110 (1-2), 43-50, https://doi.org/10.1016/S0167-2789(97)00118-8.

3. Chakraborty, B. (2014).A Proposal for Classification of Multisensor Time Series Data based on Time Delay Embedding. International Journal on Smart Sensing and Intelligent Systems,7(5) 1-5. https://doi.org/10.21307/ijssis-2019-120
4. Datseris, G. (2018). DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics. Journal of Open Source Software, 3(23), Pages 598, https://doi.org/10.21105/joss.00598
5. Frank, J., Mannor, S., & Precup, D. (2010). Activity and Gait Recognition with Time-Delay Embeddings. Proceedings of the AAAI Conference on Artificial Intelligence, 24(1), Pages 1581-1586. https://doi.org/10.1609/aaai.v24i1.7724
6. Geist K., Parlitz U., and Lauterborn W., (1990). Comparison of Different Methods for Computing Lyapunov Exponents, Progress of Theoretical Physics, Volume 83 (5), Pages 875–893, https://doi.org/10.1143/PTP.83.875
7. Gupta, V. and Mittal, M. (2019). QRS Complex Detection Using STFT, Chaos Analysis, and PCA in Standard and Real-Time ECG Databases. J. Inst. Eng. India Ser. B 100, 489–497. https://doi.org/10.1007/s40031-019-00398-9
8. Hegger R. and Kantz H. (1999). Improved false nearest neighbor method to detect determinism in time series data. Phys. Rev. E, Volume 60, Pages 4970-4973, https://doi.org/10.1103/PhysRevE.60.4970
9. Kennel M. B., Brown R., and Abarbanel H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A, Volume 45, Pages 3403–3411, https://doi.org/10.1103/PhysRevA.45.3403.
10. Kıymık M. K., Güler İ., Dizibüyük A., and Akın M. (2005). Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application, Computers in Biology and Medicine, Volume 35, Issue 7, 603-616, ISSN 0010-4825,
https://doi.org/10.1016/j.compbiomed.2004.05.001.
11. Koshel, E. (2020). Local jacobian estimation for delay embedded time series data. In XVIII International scientific and practical conference, Dnipro (pp. 152-153).
12. Koshel Y. V. and Belozyorov V. Y. (2023). Univariate Time Series Analysis with Hyper Neural ODE. Journal of Optimization, Differential Equations and Their Applications (JODEA), 31(2), 50-66.
13. Krakovská A., Mezeiová K., and Budáčová H. (2015). Use of False Nearest Neighbours for Selecting Variables and Embedding Parameters for State Space Reconstruction. Journal of Complex Systems, Volume 2015,
https://doi.org/10.1155/2015/932750.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

14. Nalmpantis, C., Vrakas, D. (2019). Signal2Vec: Time Series Embedding Representation, Engineering Applications of Neural Networks. EANN 2019. Communications in Computer and Information Science, vol 1000. https://doi.org/10.1007/978-3-030-20257-6_7
15. Pecora L. M., Moniz L., Nichols J., and Carroll T. L. (2007). A unified approach to attractor reconstruction. Chaos, Volume 17 (1). https://doi.org/10.1063/1.2430294
16. Takens, F. (1981). Detecting strange attractors in turbulence. In: Rand, D., Young, LS. (eds) Dynamical Systems and Turbulence, Warwick 1980. Lecture Notes in Mathematics, vol 898. Springer, Berlin, Heidelberg.
https://doi.org/10.1007/BFb0091924
17. Tüske, Z., Golik, P., Schlüter, R., & Drepper , F. R. (2011). Non-stationary feature extraction for automatic speech recognition, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 5204-5207, doi: 10.1109/ICASSP.2011.5947530.

Received 12.03.2024.
Accepted 15.03.2024.

Нейронно-мережевий підхід до неперервного вкладення одновимірних потоків даних для аналізу часових рядів в реальному часі

Задача вкладення одновимірних часових рядів у багатовимірні простори дуже розповсюджена і зустрічається у багатьох галузях досліджень. Методи, які розв'язують цю задачу, зазвичай покладаються на ту чи іншу форму вкладання з затримкою, яке реконструює фазовий простір невідомої системи шляхом асоціювання значень часового ряду з історичними даними. Ми пропонуємо більш гнучкий метод, який використовує правила для комбінування значень часового ряду, щоб створити вкладення, які є більш репрезентативними щодо взаємозв'язку даних часового ряду одне з одним.

Кошель Євген - аспірант кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара.

Koshel Eugene - postgraduate student, department of computer technologies, Oles Honchar Dnipro National University.

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-09
УДК 669.15-194.53

О.І. Бабаченко, Р.В. Подольський, Г.А. Кононенко,
О.Є. Меркулов, О.А. Сафронова, С.О. Дудченко

АНАЛІЗ ВПЛИВУ ШВИДКОСТІ ОХОЛОДЖЕННЯ НА ТВЕРДІСТЬ СТАЛЕЙ ДЛЯ ЗАЛІЗНИЧНИХ РЕЙОК ПЕРЛІТНОГО ТА БЕЙНІТНОГО КЛАСУ

Анотація. Процес експлуатації транспортних засобів визначає взаємодію колеса і рейки. Від параметрів цього процесу багато в чому залежать безпека руху та основні техніко економічні показники господарств колії та рухомого складу. Результатом є вплив, що виникає від тертя кочення і особливо від тертя ковзання колеса по рейці при гальмуван ні, відносно цих змін відбувається істотне зростання інтенсивності зношування коліс рухомого складу, яке, в свою чергу може призвести до катастрофічних результатів для локомотивного господарства. Також в процесі експлуатації рейки в більшості випадків утворюються дефекти, що мають характер складнонавантаженого стану: її головка піддається зношуванню, зминанню, розтріскуванню і викришуванню, в металі можуть розвиватися контактно втомні пошкодження. В перлітних сталях зносостійкість за безпечується за рахунок високого вмісту вуглецю і малої відстані між пластинами перлі ту (що досягається за рахунок процесу загартування головки рейки), які обидва підви щують твердість. Виходячи з досліджень останніх років відомо, що міцність перлітних рейкових сталей досягла межі. Крім того, подальше збільшення вмісту вуглецю вплине на ударну в'язкість та зварюваність матеріалів рейок. Отже, існує гостра потреба в інших альтернативних матеріалах. Бейнітна сталь, що забезпечує як високу міцність, так і відмінну пластичність, вважається одним з найбільш перспективних напрямків. Встано влено, що дослідна сталь при швидкості охолодження від 0,2 ˚С/с до 0,52 ˚С/с має бейніт ну структуру з невеликою кількістю мартенситу та аустеніту залишкового; при збіль шені швидкості охолодження від 1,3 ˚С/с структуру мартенситу з аустенітом залишко вим.

Ключові слова: залізнична рейка, перліт, бейніт, зносостійкість, колія.

Вступ

Процес експлуатації транспортних засобів визначає взаємодія колеса і рейки. Від параметрів цього процесу багато в чому залежать безпека руху та основні техніко-економічні показники господарств колії та рухомого складу. Результатом є вплив, що виникає від тертя кочення і особливо від тертя ковзання колеса по рейці при гальмуванні, відносно цих змін відбувається істотне

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

зростання інтенсивності зношування коліс рухомого складу [1], яке, в свою чергу, може призвести до катастрофічних результатів для локомотивного господарства. Також в процесі експлуатації рейки [2] в більшості випадків утворюються дефекти, що мають характер складнонавантаженого стану: її головка піддається зношуванню, зминанню, розтріскуванню і викришуванню, в металі можуть розвиватися контактно-втомні пошкодження [3].

У перлітних сталях зносостійкість зростає в міру збільшення вмісту вуглецю і зменшення міжпластинчастої перлітної відстані. Перліт складається з пластин перлітного фериту і карбіду заліза що чергуються, відстань між пластинами змінюється в залежності від температури утворення відповідно до умов охолодження рейки після прокатки. При збільшенні швидкості охолодження відстань між рейками перліту зменшується, отже, збільшується твердість, яка забезпечується в рейках з термообробленою головкою. Збільшення вмісту вуглецю збільшує об'ємну частку карбідів заліза, які є твердими і мають тенденцію приймати орієнтацію, паралельну зі зношеною поверхнею.

Товщина карбідних пластин впливає на спосіб їх деформації в контакті кочення [4]. Товсті пластини карбідів мають тенденцію до розстріскування при високих деформаціях; в той час як тонкі пластини карбіду деформуються пластично, без руйнування [5]. Однак відстань між пластинами перліту не є постійною і, ймовірно, буде варіюватися приблизно нормальним чином щодо цих середніх відстаней. Отже, тонкі пластини карбіду в термічно обробленій голівці рейки з більшою ймовірністю деформуються без утворення тріщин, ніж більш товсті карбіди в гарячекатаній сирій рейковій сталі. Точний спосіб утворення частинок зносу невідомий, але можна припустити, що мікроструктури, в яких пластинки карбіду розтріскуються при деформації (з утворенням порожнин), ймовірно, будуть мати більш низьку зносостійкість, ніж мікроструктура, карбіди яких деформуються пластично.

Таким чином, в перлітних сталях зносостійкість забезпечується за рахунок високого вмісту вуглецю і малої відстані між пластинами перліту (що досягається за рахунок процесу загартування головки рейки), які обидва підвищують твердість. На підтвердження вторинного впливу твердості на знос, Хіракава і інші підвищили твердість зразків з перлітною структурою за рахунок зниження температури відпуску, але в лабораторних випробуваннях дане твердження не підтвердилось (не збільшили зносостійкість) [6-7].

Виходячи з досліджень останніх років [8-9] відомо, що міцність перлітних рейкових сталей досягла межі [10]. Крім того, подальше збільшення вмісту вуг-

лецю вплине на ударну в'язкість та зварюваність матеріалів рейок [11]. Наприклад, у порівнянні з доевтектоїдною рейкою R200 відносне подовження рейки R400НТ з заевтектоідної сталі знижується на 6%. Отже, існує гостра потреба в інших альтернативних матеріалах. Бейнітна сталь, що забезпечує як високу міцність, так і відмінну пластичність, вважається одним з найбільш перспективних напрямків.

Низьковуглецеві бейнітні сталі відрізняються від звичайних перлітних сталей тим, що в них мало карбідів, якщо вони взагалі є. Бейнітні низьковуглецеві сталі, міцність яких більше 1200 МПа при цьому володіють високим рівнем ударної в'язкості, трибологічними властивостями, сприятливою реакцією на великі швидкості деформації, стійкістю до втоми і дешеві у виробництві [12]. Такий комплекс властивостей досягається за рахунок дуже дрібної і сильно зміцненої мікроструктури.

Звичайний верхній бейніт складається з неламеллярної суміші пластин бейнітного фериту з проміжними частинками цементиту. Мікроструктура створюється послідовно: пластини ростуть без дифузії, але потім надлишковий вуглець розподіляється в аустеніті залишковому [13]. Згодом цей збагачений вуглецем аустеніт розпадається на суміш цементиту і фериту. Цементит погіршує властивості, але його виділенню можна запобігти, додавши в сталь достатню кількість кремнію [14], щоб залишити мікроструктуру з пластин бейнітного фериту і збагаченого вуглецем аустеніту.

Як відомо, в ході бейнітного перетворення формування пакету бейнітних рейок відбувається на границі аустенітного зерна і подальше зростання пакета відбувається вглиб зерна. Пакет бейніту складається з рейок фериту, розділених в основному малокутовими границями [15], в той час як голчастий ферит гетерогенно зароджується на неметалевих включеннях всередині аустенітного зерна і росте в різних напрямках, не утворюючи виражених пакетів, в іншому ж залишаючись подібним з бейнітом [16]. Енергія зародження на границі аустенітного зерна практично завжди залишається нижче, ніж енергія зародження на частці включення і таким чином зародження на границі аустенітного зерна більш вигідне з енергетичної точки зору [14,16]. З цього випливає, що структуру сталі, що містить виключно голчастий ферит отримати практично неможливо. У той же час при відносно великому аустенітному зерні ймовірність інтрагранулярного зародження бейніту як в зварних з'єднаннях, так і в прокаті підвищується [15].

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Постановки мети і завдань дослідження

На основі аналізу перспектив застосування нових матеріалів для високоміцних залізничних рейок, запропонувати хімічний склад дослідної сталі та виконати дослідження впливу швидкості охолодження на твердість дослідної сталі бейнітного класу у порівнянні з відомою сталлю перлітного класу для залізничних рейок вищої категорії.

Матеріал і методика досліджень

В умовах Інституту чорної металургії була виконана виплавка дослідної плавки. Виплавку виконували за допомогою комплексної установки, що складається з плавильного агрегату ІТПЕ-0,01 закритого типу і високочастотного джерела струму ВТГ-20-22, що має вбудовану станцію автономного охолодження. Загальний вигляд комплексної експериментальної установки показано на рис. 1. Ця установка дозволяє виплавляти в лабораторних умовах дослідні сталі, в тому числі спеціально леговані марганцем, кремнієм, хромом, молібденом та ін.

Рисунок 1 - Процес розплавлення сталі

Для забезпечення необхідних показників в досліджуваних варіантах і відсутності небажаних домішок присутніх в рядовому брухті в якості вихідної сировини вибрали метал марки Ст3 - вуглецева якісна сталь з вмістом вуглецю 0,1%, мас.

Плавки проводили методом переплавлювання без примусового окислення домішок. З огляду на малий обсяг печі (до 12 кг), і той факт, що діаметр тигля становить всього близько 80 мм шихту подрібнювали до розмірів Ø3х12 мм. Фактичний хімічний склад представлений в таблиці 1.

Таблиця 1

Фактичний хімічний склад лабораторної плавки

C	Si	Mn	Cr	Ni	Mo	Al	Cu	V
0,37	1,11	1,35	0,74	0,03	0,17	0,03	0,03	0,15

В рамках роботи для порівняння було застосовано сталь для виробництва високоміцних рейок (марка К76Ф ДСТУ 4344), хімічний склад, яких представлено в таблиці 2.

Таблиця 2

Фактичний хімічний склад сталі К76Ф

Марка	C	Si	Mn	Cr	Ni	Mo	V	Джерело
К76Ф	0,80	0,25	0,97	0,04	0,03		0,05	[17]

Дослідження мікроструктури дослідних зразків проводилось за допомогою витравлення в 4% розчині азотної кислоти з застосуванням оптичного мікроскопа Axiovert 200 MMat та електронних мікроскопів Gemini та РЕМ-106.

З результатів аналізу мікроструктури злитку (рис. 2) бейнітного класу встановлено, що дослідна сталь поблизу поверхні та 1/2 радіусу має структуру гранулярного бейніту з різноорієнотованими бенітними ламеліями (458…495 кГ/мм^2). Дослідний злиток має структуру бейніту, голчастого фериту (331…344 кГ/мм^2) та залишкового аустеніту по границях зерен, середня мікротвердість міждендрітного простору – 495 кГ/мм^2.

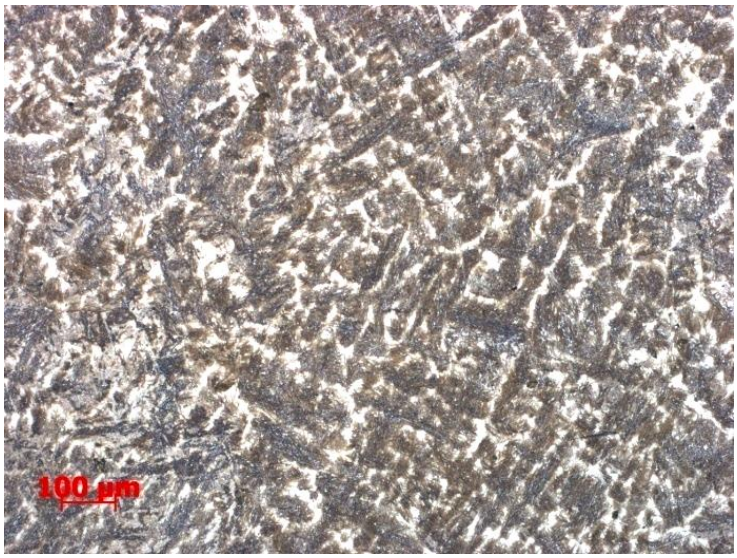


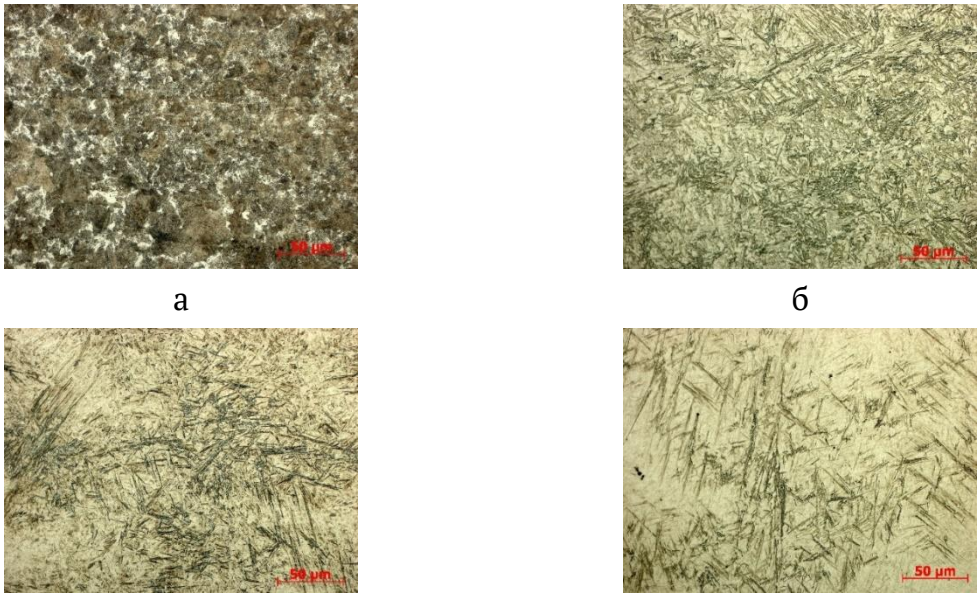
Рисунок 2 - Мікроструктура лабораторного злитку після лиття, ×100

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

В лабораторних умовах були виконані експерименти з відпрацюванням параметрів термічної обробки зразків дослідної сталі та порівняльної відомої сталі з різним вмістом легуючих елементів.

Зразки нагрівали до температури 900 °С, витримували протягом ≈30 хвилин при даній температурі та охолоджували у різних середовищах з фіксацією швидкості охолодження. На підставі технічних джерел були обрані для досліджень швидкості охолодження від 0,2…5,1°С/с. Для їх досягнення при охолодженні даної температури виконано: охолодження з піччю – 0,2˚С/с; нагрів в печі з термостійким покривалом та послідуючим охолодження в ньому на повітрі – 0,52˚С/с; нагрів зразків в печі з охолодженням на повітрі з термостійким покривалом – 1,3˚С/с; на спокійному повітрі – 5,1°С/с.

За результатами мікроструктурних досліджень (рис. 3) сталі бейнітного класу встановлено тип структури та співвідношення фазового складу. Встановлено, що дослідна сталь при швидкості охолодження 0,2˚С/с має феритоперлітну структуру; при збільшенні швидкості охолодження до 0,52˚С/с структура має бейнітну структуру з невеликою кількістю мартенситу та аустеніту залишкового; при збільшені швидкості охолодження від 1,3˚С/с до 5,1˚С/с - структура мартенситу з аустенітом залишковим.



а б

в г

Рисунок 3 – Мікроструктура дослідної сталі після різних швидкостей охолодження: а – 0,2 ˚С/с, б – 0,52 ˚С/с, в – 1,3 ˚С/с, г-5,1 ˚С/с

При порівнянні механічних властивостей, а саме значень твердості, з відомою сталлю перлітного класу К76Ф було побудовано залежність зміни твердості від швидкості охолодження (рис. 4).

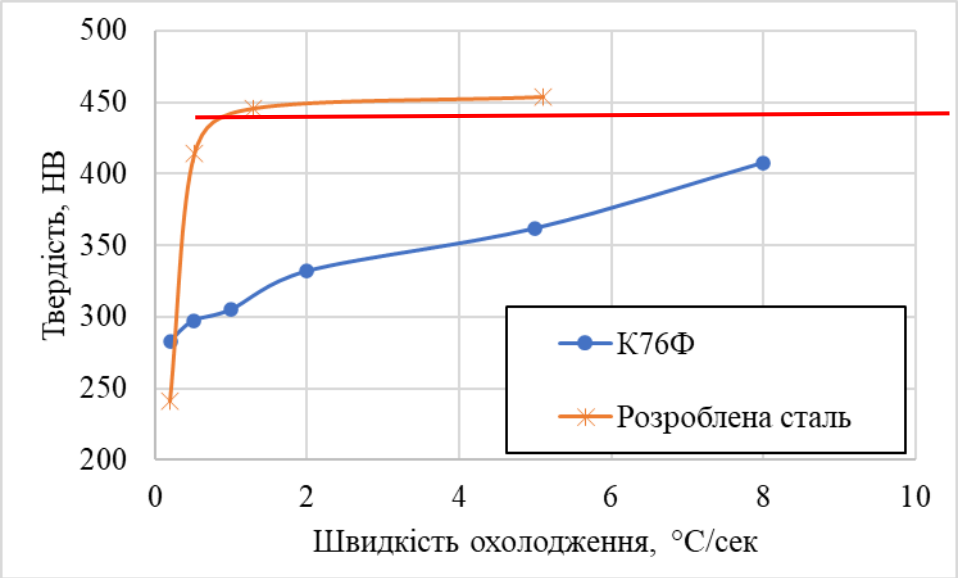


Рисунок 4 – Залежність зміни твердості від швидкості охолодження та рівень вимог згідно ДСТУ EN 13674:2019 розробленої сталі та сталі К76Ф

В результаті аналізу отриманої залежності було встановлено, що при швидкості охолодження ⩾0,52°С/с мікроструктура дослідної сталі має структуру безкарбідного бейніту з твердістю, що перевищує вимоги регламентованих стандартів та літературних даних за рівнем значень твердості.

Висновок

1. За результатами аналітичного огляду систематизовані відомі шляхи підвищення зносостійкості залізничних рейок. Показано, що найбільш ефективним способом підвищення твердості сталей є управління їх хімічним складом шляхом легування, мікролегування і зменшення кількості шкідливих домішок. Перспективним є підхід до підвищення опору через виготовлення залізничних рейок з бейнітною структурою.

2. Встановлено, що дослідна сталь при швидкості охолодження від 0,2˚С/с до 0,52˚С/с має бейнітну структуру з невеликою кількістю мартенситу та аустеніту залишкового; при збільшені швидкості охолодження від 1,3˚С/с - структура мартенситу з аустенітом залишковим.

3. Встановлено вплив швидкості охолодження на твердість сталей для залізничних рейок перлітного та бейнітного класу. Показано, що твердість сталі

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

бейнітного класу значно перевищує твердість перлітної сталі для високоміцних рейок.

ЛІТЕРАТУРА

1. R. Harder. Creep Force – Creepage and Frictional Work Behaviour in Non-Hertzian Counter formal Rail/Wheel Contacts. Proceedings of IHHA'99 STS-Conference on Wheel/Rail Interface. 1999. V. 1. p. 207 – 214.

2. B. Paul. J. Hashemi. User's Manual for Program CONTACT. Technical Report No. 4. FRA/ORD-78/27/PB286097. NTIS. Springfield. VA. Sept. 1977.

3. Большаков В.И.. Долженков И.Е.. Зайцев А.В. «Оборудование термических цехов. технологии термической и комбинированной обработки металлопродукции». Изд.2-е. Днепропетровск. РИА Днепр-VAL. 2010 г. -619с

4. H. de Boer et al.. "Naturally Hard Bainitic Rails with High Tensile Strength." Stahl und Eisen. Vol. 115. No. 2.1995. pp. 93-98.

5. N. Jin. "Mechanical Properties and Wear Performance of Bainitic Steels." Ph.D. Thesis. Oregon Graduate Institute. Portland. OR. 1995.

6. N. Jin and P. Clayton. "Effect of Microstructure on Rolling/Sliding Wear of Low Carbon Bainitic Steels." Wear. Vol. 202.1997. pp. 202-207

7. W. Heller and R. Schweitzer. "Hardness. Microstructure and Wear Behavior of Steel Rails." 2nd. International Heavy Haul Railway Conference. Colorado Springs. Colorado. 1982. pp. 282-286.

8. Розробка сталей для металопродукції залізничного призначення. Бабаченко О.І.. Кононенко Г.А.. Рослик О.В.. Майстренко К.М. Подольський Р.В. Дніпро. «Домінанта-принт». .2021. 298 стор.

9. Бабаченко О.І.. Кононенко Г.А.. Подольський Р.В.. Сафронова О.А. Сталь для залізничних рейок з поліпшеними експлуатаційними властивостями. Фізико-хімічна механіка матеріалів. 56(6). 2020. с 82-87.

10. S. Sharma. S. Sangal. K. Mondal. Wear behaviour of bainitic rail and wheel steels. Mater. Sci. Technol. 32 (4) (2016) 266–274.

11. P. Pointner. High strength rail steels-The importance of material properties in contact mechanics problems. Wear 265 (2008) 1373–1379.

12. Bhadeshia. H K. D.H.. ``Hyperbolic Tangents and Alloys of Iron". Materials World. 7 . pp. 643-645. 1996.

13. Merkulov, O., Podolskyi, R., Kononenko, A. et al. Development of Promising Steels for Railway Rails of a New Generation Using Modeling of Phase-Structural Transformations. Trans Indian Inst Met (2024). https://doi.org/10.1007/s12666-024-03265-4.

14. Takahashi. M. and Bhadeshia. H. K. D. H.. ``A Model for the Transition from Upper to Lower Bainite. Materials Science and Technology. 6. pp. 592-603. 1990.
15. J.R. Yang. H.K.D.H. Bhadeshia Advances in Welding Science and Technology ed. S. David. ASM. Metals Park. ohio. USA (1986). p. 187-191.
16. R.A. Ricks. P.R. Howell. G.S. Barrite International Conference on Solid-Solid Phase Transformation. eds. H.I. Aaronson et.al.. TMS-AIME. Warrendale. PA. USA. 1981. p. 463-468.
17. Подольський Р. В. Розробка хімічного складу та режимів термічної обробки високоміцних рейок сталей перлітного класу : дис. Докт. філософ. : 132-Матеріалознавство – Дніпро, 2023. – 181 с.

REFERENCE

1. R. Harder. Creep Force – Creepage and Frictional Work Behaviour in Non-Hertzian Counter formal Rail/Wheel Contacts. Proceedings of IHHA'99 STS-Conference on Wheel/Rail Interface. 1999. V. 1. p. 207 – 214.
2. B. Paul. J. Hashemi. User's Manual for Program CONTACT. Technical Report No. 4. FRA/ORD-78/27/PB286097. NTIS. Springfield. VA. Sept. 1977.
3. Bolshakov V.I.. Dolzhenkov I.E.. Zajcev A.V. Oborudovanie termicheskih cehov. tehnologii termicheskoj i kombinirovannoj obrabotki metalloprodukcii. Izd.2-e. Dnepropetrovsk. RIA Dnepr-VAL. 2010 g. -619s
4. H. de Boer et al.. "Naturally Hard Bainitic Rails with High Tensile Strength." Stahl und Eisen. Vol. 115. No. 2.1995. pp. 93-98.
5. N. Jin. "Mechanical Properties and Wear Performance of Bainitic Steels." Ph.D. Thesis. Oregon Graduate Institute. Portland. OR. 1995.
6. N. Jin and P. Clayton. "Effect of Microstructure on Rolling/Sliding Wear of Low Carbon Bainitic Steels." Wear. Vol. 202.1997. pp. 202-207
7. W. Heller and R. Schweitzer. "Hardness. Microstructure and Wear Behavior of Steel Rails." 2nd. International Heavy Haul Railway Conference. Colorado Springs. Colorado. 1982. pp. 282-286.
8. Rozrobka stalej dlya metaloprodukciyi zaliznichnogo priznachennya. Babachenko O. I.. Kononenko G. A.. Roslik O.V.. Majstrenko K.M.. Podolskij R.V. Dnipro. «Dominanta-print». .2021. 298 stor.
9. Babachenko O. I.. Kononenko G. A.. Podolskij R. V.. Safronova O. A. Stal dlya zaliznichnih rejok z polipshenimi ekspluatacijnimi vlastivostyami. Fiziko-himichna mehanika materialiv. 56(6). 2020. s 82-87.
10. S. Sharma. S. Sangal. K. Mondal. Wear behaviour of bainitic rail and wheel steels. Mater. Sci. Technol. 32 (4) (2016) 266–274.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

11. P. Pointner. High strength rail steels-The importance of material properties in contact mechanics problems. Wear 265 (2008) 1373–1379.
12. Bhadeshia. H. K. D. H.. ``Hyperbolic Tangents and Alloys of Iron". Materials World. 7 . pp. 643-645. 1996.
13. Merkulov, O., Podolskyi, R., Kononenko, A. et al. Development of Promising Steels for Railway Rails of a New Generation Using Modeling of Phase-Structural Transformations. Trans Indian Inst Met (2024). https://doi.org/10.1007/s12666-024-03265-4.
14. Takahashi. M. and Bhadeshia. H. K. D. H.. ``A Model for the Transition from Upper to Lower Bainite. Materials Science and Technology. 6. pp. 592-603. 1990.
15. J.R. Yang. H.K.D.H. Bhadeshia Advances in Welding Science and Technology ed. S. David. ASM. Metals Park. ohio. USA (1986). p. 187-191.
16. R.A. Ricks. P.R. Howell. G.S. Barrite International Conference on Solid-Solid Phase Transformation. eds. H.I. Aaronson et.al.. TMS-AIME. Warrendale. PA. USA. 1981. p. 463-468.
17. Podolskyi R. V. Rozrobka himichnogo skladu ta rezhimiv termichnoyi obrobki visokomicnih rejok stalej perlitnogo klasu : dis. Dokt. filosof. : 132-Materialoznavstvo – Dnipro, 2023. – 181 s.

Received 14.03.2024.
Accepted 15.03.2024.

Analysis of the influence of the cooling rate on the hardness of steel for railway rails of the pearlite and bainetic classes

The process of operating vehicles determines the interaction between the wheel and the rail. Traffic safety and the main technical and economic indicators of track management and rolling stock largely depend on the parameters of this process. The result is the effect arising from the rolling friction and especially from the friction of the wheel sliding on the rail during braking, relative to these changes there is a significant increase in the intensity of wear of the wheels of the rolling stock, which, in turn, can lead to catastrophic results for the locomotive industry. Also, in the process of operation of the rail in most cases, defects are formed that have the character of a complicated state: its head is subject to wear, crumpling, cracking and buckling, contact fatigue damage can develop in the metal. In pearlite steels, the wear resistance is provided by the high carbon content and the small distance between the pearlite plates (achieved by the hardening process of the rail head), both of which increase hardness. Based on research in recent years, it is known that the strength of pearlite rail steels has reached its limit. In addition, a further increase in the carbon content will affect the impact strength and weldability of rail materials. Therefore, there is an urgent need for other alternative materials. Bainite steel,

which provides both high strength and excellent plasticity, is considered one of the most promising directions. It was established that the structure of the test steel at a cooling rate of 0.2˚C/s to 0.52˚C/s has a bainite structure with a small amount of martensite and residual austenite; with an increased cooling rate from 1.3˚C/s - martensite structure with residual austenite.

Keywords: railway rail, perlite, bainite, wear resistance, track.

Бабаченко Олександр Іванович - докт. техн.наук, директор Інституту чорної металургії ім. З.І. Некрасова НАН України.

Подольський Ростислав Вячеславович – докт. філософ., науковий співробітник Інституту чорної металургії ім. З.І. Некрасова НАН України.

Кононенко Ганна Андріївна - докт. техн. наук, ст. досл., вчений секретар Інституту чорної металургії ім. З.І. Некрасова НАН України.

Меркулов Олексій Євгеньович - докт. техн. наук, заст. директора Інституту чорної металургії ім. З.І. Некрасова НАН України.

Сафронова Олена Анатоліївна – науковий співробітник Інституту чорної металургії ім. З.І. Некрасова НАН України.

Дудченко Сергій Олександрович – канд. техн. наук, науковий співробітник Інституту чорної металургії ім. З.І. Некрасова НАН України.

Babachenko Alexander - Dr. Technical Sciences, Director of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine.

Podolskyi Rostyslav – PhD, researcher of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine.

Kononenko Ganna - Doct. technical science, sen.resear., scientific secretary of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine.

Merkulov Oleksii - Dr. technical sciences, assistant director of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine.

Safronova Olena - researcher of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine.

Dudchenko Serhii - PhD, researcher of the Institute of Iron and Steel of Z.I. Nekrasov NAS of Ukraine.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-10
UDC 004.4:004.8

V.V. Hnatushenko, I.V. Pavlenko

THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE IN SOFTWARE TESTING

Abstract. This article explores the potential of using generative artificial intelligence (AI) for software testing, reflecting on both the advantages and potential drawbacks of this emerging technology. Considering the vital role of rigorous testing in software production, the authors ponder whether generative AI could make the testing process more efficient and comprehensive, without the need to increase resources. The article delves into the current limitations of this technology, emphasizing the need for continuous exploration and adaptation. It concludes with a summation of potential innovative solutions and avenues for future investigation. The paper encourages discussions surrounding the question of fully automated testing and the role of hu man specialists in the future of QA. It ultimately provides a thought provoking reflection on the intersection of emerging technologies, and their societal impacts.
Keywords: generative AI, GenAI, software testing, large language model, test automation, AI augmented testing, AI assistant, autonomous test creation

Statement of the problem. In the previous few years there was a significant development in the pre-trained generative AI based on large language models (LLMs) with transformers and the services built upon them, OpenAI's ChatGPT, Google's Gemini, Anthropic's Claude, GitHub Copilot, Amazon CodeWhisperer and others. However, there remain several challenges and uncertainties regarding the effective integration of generative AI into software testing frameworks. This includes issues such as ensuring the reliability and accuracy of generated test cases, adapting AI algorithms to diverse software environments, and addressing ethical concerns surrounding autonomous testing systems. The following questions remain unanswered. At this stage of their development, what can GenAI on LLMs with transformers bring to software testing, and what can be expected from their further advancement? For which software testing tasks can GenAI on LLMs with transformers be used?

Analysis of recent studies and publications. Considering that GenAI on LLMs with transformers is an emerging topic [1–3], the number of scientific articles on the use of this type of GenAI in software testing is relatively small compared to

more developed topics in the field of information systems. In the work [3] the authors note that application of AI/ML has a long history in software engineering (SE) research. However, the use of GenAI specifically, is a more emerging topic. While the promise of GenAI has been acknowledged for some time, progress in the research area has been rapid. GenAI was not a prominent research area in SE until 2020. Following the recent improvements in the performance of these systems, especially the release of services such as GitHub Copilot and ChatGPT-3, research interest has now surged across disciplines, including SE. At the present stage of development of the considered models, there is a relatively rapid (a couple of times a year) growth in the number of model parameters, the volume of data used for their training, and the size of the model context. This leads to rapid growth in the capabilities of the models. Thus, this topic can be characterized as fast-changing. In connection with this, in a work [3] devoted to the analysis of existing publications for mid-2023, it is noted that: "A comprehensive and systematic review might not be suitable for research on GenAI in software development at the time this research being conducted". The authors further justify the use of preprints and other non-peer-reviewed works: "Firstly, research work is rapidly conducted and published on the topic, hence rendering the findings of a comprehensive review probably outdated shortly after publication. Secondly, we found a lot of relevant work as gray literature from non-traditional sources such as preprints, technical reports, and online forums. These sources may not be as rigorously reviewed or validated as peer-reviewed academic papers, making it difficult to assess their quality and reliability. Thirdly, we would like to publish the agenda as soon as possible to provide a reference for future research. A systematic literature review would consume extensive effort and time, which might be obsolete by the time the review is complete". They suggest conducting a literature review as follows: "Our strategy is to conduct focused, periodic reviews to capture the most current and relevant information without the extensive resource commitment of a comprehensive review. This approach allows for agility in keeping up with the latest developments without claiming comprehensiveness and repeatability" [3]. In this paper, in addition to the literature review, focus group surveys are conducted: "We conducted four structured working sections as focus groups to identify, refine, and prioritize Research Questions on GenAI for SE… …All participants are SE researchers who have experience or interest in the topic…". It is noted: "Almost all of the existing studies on the topic (e.g. "[13][14][15]") are mostly experimental studies and thus do not take into consideration the industrial context. Therefore, how GenAI models deal with real-world software quality issues remains a mystery. We need more real-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

world case studies and industrial examples to understand the effectiveness of various tasks of SQA with GenAI. Moreover, it would be interesting to study how well GenAI enhances the productivity of SQA professionals" [3]. Because of what was stated above and the practical nature of the selected theme (the topic is mostly practical, so focus groups from industry specialists would be much better), I used meetings and webcasts of testing industry professionals as a kind of focus groups to explore ideas for potential opportunities and challenges when adopting GenAI in software testing activities [10 - 12].

Purpose of the study. The purpose of this study is to explore the capability of GenAI integrated with transformers within Large Language Models to tackle distinct challenges encountered in software testing, such as test case generation, bug prediction, and test data synthesis. Through an examination of existing literature and insights from domain experts regarding the implementation of GenAI with transformers in software testing contexts, this research aims to formulate pertinent inquiries. These critical questions will serve as a framework for discussions concerning the contemporary role of advanced Generative AI in augmenting and refining software testing methodologies.

Statement of the main research material. As an introduction in [4] authors state that software testing, an integral part of the development cycle. However, automated software testing can be challenging, and necessitates a high level of technical acumen. There is significant expertise required to appropriately test software, as evidenced by the existence of test engineers/architects. Furthermore, software testing and the writing of software tests can be repetitive, as Hass et al. note [16]. It's important to acknowledge the crucial role of testing in software production. Testing isn't just vital for ensuring software quality; it also plays a significant role in refactoring or transitioning between different project types (proof of concept, MVP, etc.). Having test coverage is especially crucial when working on projects using an Agile approach. During refactoring and migrations, it's essential to verify that functionality and data remain intact throughout the process. In the case of software, testing serves as a similar verification tool. To ensure reliable verification, having a sufficient level of test coverage is necessary. Insufficient test coverage can lead to limitations in the ability to make changes, consequently affecting project development. It might also demand extensive and comprehensive manual testing, thereby slowing down project progress. Creating adequate test coverage for a project, especially maintaining the functionality of tests while the project undergoes continuous changes during its active development phase, is a labor-intensive task.

Software testing plays a crucial role. Despite this, many projects either aren't tested or are tested inadequately. This is attributable to the inherent complexity of testing itself. Approaching testing as "exhaustive," meant to cover all paths in the code or all possible input data, has highlighted that achieving complete software testing is impossible under these conditions. The sheer volume of potential input data, code execution paths, and system states is immense, rendering “exhaustive” testing theoretically impossible. As a result, in testing, it's necessary to establish a "stopping criterion," reaching which is regarded as a tradeoff between achieving as comprehensive testing as possible and the resources allocated for testing within the project's scope. However, the task of testing a project remains complex and unattainable in its entirety. There's hope that generative AI might assist in conducting more comprehensive testing without increasing the resources dedicated to testing.

Generative AI based on large language models with transformers is a new technology. We're only beginning to understand its capabilities and potential applications.

The primary tasks of testing can be generally described as follows:

1. Requirements analysis
2. Creating meaningful tests
3. Achieving code coverage targets
4. Maintaining the test suites
5. Documenting

When performing the aforementioned tasks, a human specialist encounters the following difficulties:

1. Requires a lot of effort
2. Requires a lot of time
3. Overwhelming for a person

Often, tests and test data created for testing are very similar from one test to another. During active project development and intensive changes to its codebase, a significant amount of effort and time are required to maintain the functionality of tests. Such work appears preferable for automation. However, automation based on predefined procedures does not yield the desired effect. It seems possible that generative AI based on LLM with transformers at this stage of development could help solve these tasks. It can be assumed that generative AI might assist in the following tasks:

- Requirements analysis.
- Generating test plans.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

• Generating test scenarios.

• Generating mocks.

• Generating test data.

• Extending and supplementing existing test sets.

• Maintaining the relevance of tests.

• Defect detection and prediction.

• Reducing maintenance cost by eliminating redundant tests.

• Bug reports writing.

• Generating documentation.

What advantages (capabilities) of applying generative AI can be expected:

• "Self-healing" of test suites when changes are made to the source code.

• Generating tests in the background or without using resources designated for development.

• Analysis of the user interface.

• Improving the productivity of human specialists.

• Generating tests from models or data description languages.

• Generating tests from informal descriptions, documentation written in natural languages, or audio descriptions.

• Optimizing test suites.

However, at this stage of generative AI development, attempting to use it brings some challenges:

• The field is rapidly evolving. Approaches and tools developed might become outdated before their completion with the publishing of more advanced models.

• High variability in generated tests (Deterministic vs. Probabilistic).

• Incorrect code generation.

• Inconsistent/Unreliable output.

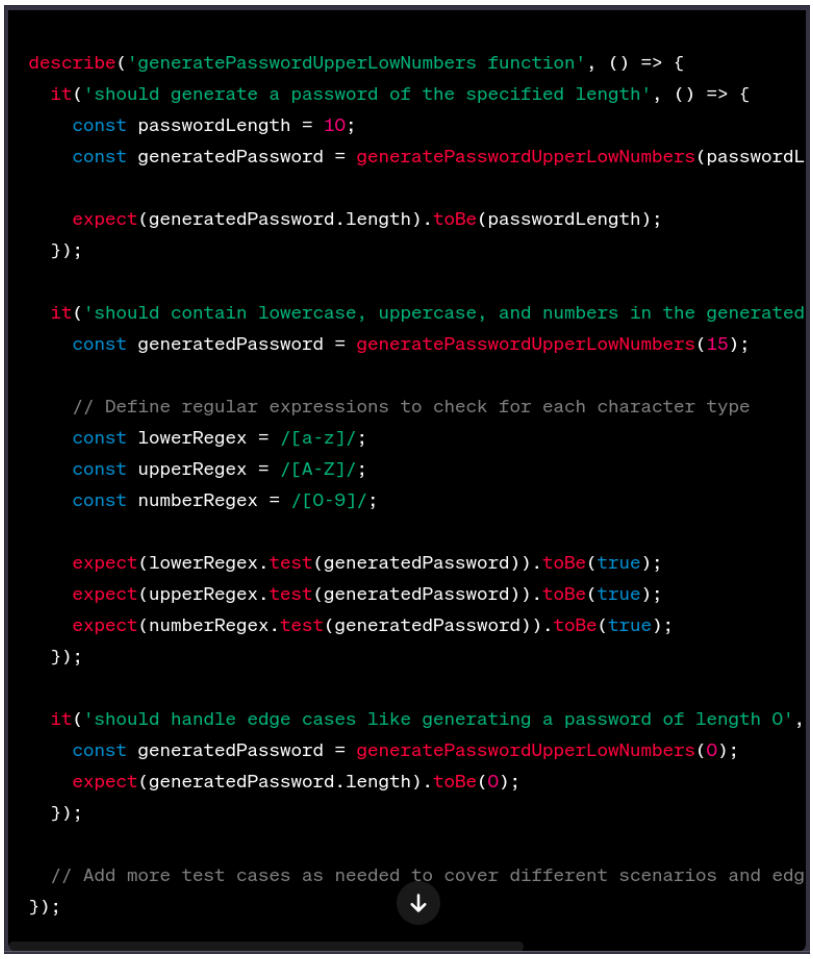
• Performance issues due to reliance on external systems.

• Costs associated with using generative AI services.

• Project code exposure when using external services, which might be unsuitable for certain cases.

One of the problems in generating tests using LLM is the probabilistic nature of content generation. While test cases and the source code of applications and tests themselves have a deterministic nature, LLMs built on transformers rely on a probabilistic model. This inherent probabilistic nature poses complexity in using this type of generative AI in tasks related to code generation. Additional efforts are required in post-processing the results either by a human or possibly even using the same

generative AI, not for generating but for analyzing the outcome, compiling, or checking for functionality. There might be a need for an additional step in code generation, asking the generative AI model to correct identified errors or mismatches. This process demands the development of additional criteria for the generated code to meet the given task. It involves work on selecting prompts for the model (prompt engineering), constructing a multi-step content generation system. However, due to the probabilistic model [5], possible model hallucinations [6], obtaining the expected result remains unguaranteed and requires verification by a human. You can see an example below where Figure 1 and Figure 2 contain different tests generated by ChatGPT 3.5 from the same prompt.



```
describe('generatePasswordUpperLowNumbers function', () => {
  it('should generate a password of the specified length', () => {
    const passwordLength = 10;
    const generatedPassword = generatePasswordUpperLowNumbers(passwordL

    expect(generatedPassword.length).toBe(passwordLength);
  });

  it('should contain lowercase, uppercase, and numbers in the generated
    const generatedPassword = generatePasswordUpperLowNumbers(15);

    // Define regular expressions to check for each character type
    const lowerRegex = /[a-z]/;
    const upperRegex = /[A-Z]/;
    const numberRegex = /[0-9]/;

    expect(lowerRegex.test(generatedPassword)).toBe(true);
    expect(upperRegex.test(generatedPassword)).toBe(true);
    expect(numberRegex.test(generatedPassword)).toBe(true);
  });

  it('should handle edge cases like generating a password of length 0',
    const generatedPassword = generatePasswordUpperLowNumbers(0);
    expect(generatedPassword.length).toBe(0);
  });

  // Add more test cases as needed to cover different scenarios and edg
});
```

Figure 1 - Case 01

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

```
describe('generatePasswordUpperLowNumbers function', () => {
  it('should generate a password of the specified length', () => {
    const passwordLength = 10;
    const generatedPassword = generatePasswordUpperLowNumbers(passwordL

    expect(generatedPassword.length).toBe(passwordLength);
  });

  it('should generate a non-empty password by default', () => {
    const generatedPassword = generatePasswordUpperLowNumbers();

    expect(generatedPassword.length).toBeGreaterThan(0);
  });

  it('should only contain characters from the defined character pools',
    const generatedPassword = generatePasswordUpperLowNumbers(20);

    const lowercase = 'abcdefghijklmnopqrstuvwxyz';
    const uppercase = 'ABCDEFGHIJKLMNOPQRSTUVWXYZ';
    const numbers = '0123456789';
    const pool = lowercase + uppercase + numbers;

    for (let i = 0; i < generatedPassword.length; i++) {
      expect(pool.includes(generatedPassword[i])).toBe(true);
    }
  });

  it('should return an empty string for zero password length', () => {
    const generatedPassword = generatePasswordUpperLowNumbers(0);
```

Figure 2 - Case 02

The paper [2], however, notes that software testing appears to be in a better position regarding the impact of inaccuracies in the generated test code, compared to business logic code generation: “However, there have been concerns about the correctness and reliability of the code generated by LLMs, as some studies have shown that the code generated by LLMs may not always be correct, or may not meet the expected software requirements. By comparison, when LLMs are used for software testing tasks, such as generating test cases or validating the correctness of software behavior, the impact of this problem is relatively weaker. This is because the primary goal of software testing is to identify issues or problems in the software system, rather than to generate correct code or meet specific software requirements. At worst, the only consequence is that the corresponding defects are not discovered. Furthermore, in some cases, the seemingly incorrect outputs from LLMs may be beneficial for testing corner cases in software and can help uncover defects. Taken in this sense, we think the LLMs are a natural match with software testing”.

Thus, the question arises whether, at this stage of generative AI development, it is possible to achieve automatic generation of a sufficiently complete set of tests and whether it should be pursued. Or should the current goal be to build a generative AI – human system using AI as an assistant? Some projects declare and attempt to move towards automated testing. For example, projects like Pythagora [7], TestCraftApp [8]. It can be expected that advancing the use of generative AI for testing will enhance test coverage, improve the cost-effectiveness ratio of testing, and expand the implementation of testing practices among companies and projects.

At the current level of development, automated creation of acceptable, not to mention fully covered, tests are rather impossible. It seems more appropriate to aim for test creation in cooperation with a human (developer) in assistant mode. Attempting complete automation might result in significant resource expenditure without achieving an acceptable outcome. It seems reasonable to start with a human-driven AI assistant, while further continuing to create a framework for as complete and comprehensive test automation as possible. Such paid platforms exist, created even before the use of AI on LLM.

At present, it can be stated that generative AI does not replace a tester but rather serves as a force multiplier, a "new electricity," as noted by the renowned machine learning expert Andrew Ng [9].

Efforts should also be directed towards creating a bridge (synergy) between AI and the practices and knowledge already established by humans (human specialists) in QA. Generative AI may help achieve greater efficiency in testing, but human involvement and knowledge remain critical for defining testing objectives and evaluating results.

It's interesting to note that many anticipated that the emergence of AI would allow humans to engage in creativity by eliminating routine tasks. However, it appears that, on the contrary, LLMs handle creative tasks better than tasks requiring attention to details and precision. It's possible that even at the highest available level of generative AI development, we might end up with an analog of a human specialist with inherent issues in human-to-human interaction, such as precise task setting (already emerging in prompt engineering and might remain unsolved in the future), understanding the task and its execution process by the performer in their own way. There might be incomplete or low-quality task execution (reports are already emerging that existing models gradually provide less comprehensive answers over time, indicating a sort of laziness in task execution).

Below are emerging questions that seem worth exploring.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Utilizing for context and processing not only the source code of a program but also existing documentation, both human-readable and formal documentation, DSLs (e.g., OpenAI) or formal service descriptions (e.g., protobuf). It might even involve the preliminary generation of documentation using AI as an additional supporting step. Perhaps an additional source in some generalized, processed and less formal form than the application source code, which the documentation is, will be useful in the current implementation of generative AI (LLM with transformers) to achieve better results.

Prompt engineering.

Fine-tuning the model specifically as a testing specialist.

How to best feed source code into the model?

Is it preferable to use and develop models with larger prompt token limits? One significant issue with using a large prompt size is that the model may forget parts of the prompt.

Alternatively, using retrieval augmented generation (RAG), or perhaps models fine-tuning on project artifacts, might yield a more suitable result.

Another practical question regarding generative AI is whether to lean towards more powerful proprietary models or use even smaller open-source models, possibly with fine-tuning or RAG, to achieve the desired outcome.

Findings. This article explores the potential of using generative AI for software testing, reflecting on both the advantages and potential drawbacks of this emerging technology. To provide the above analysis a review of existing publications was conducted. For initial assessments of interactions with LLM, the ChatGPT version 3.5 service and API access to the GPT 3.5 and GPT 4 models were used.

REFERENCES

1. Fan Angela, Gokkaya Beliz, Harman Mark, Lyubarskiy Mitya, Sengupta Shubho, Yoo Shin, Zhang Jie M. Large Language Models for Software Engineering: Survey and Open Problems // *arXiv. 2023. DOI:* https://doi.org/10.48550/arXiv.2310.03533.

2. Wang Junjie, Huang Yuchao, Chen Chunyang, Liu Zhe, Wang Song, Wang Qing. Software Testing with Large Language Model: Survey, Landscape, and Vision // arXiv. 2023. DOI: https://doi.org/10.48550/arXiv.2307.07221.

3. Nguyen-Duc Anh, Cabrero-Daniel Beatriz et al. Generative Artificial Intelligence for Software Engineering – A Research Agenda // arXiv. 2023. DOI: https://doi.org/10.48550/arXiv.2310.18648.

4. Feldt Robert, Kang Sungmin, Yoon Juyeon, Yoo Shin. Towards Autonomous Testing Agents via Conversational Large Language Models // arXiv. 2023. DOI: https://doi.org/10.48550/arXiv.2306.05152.
5. Keen M. How Large Language Models Work // IBM Technology. 2023. URL: https://www.youtube.com/watch?v=5sLYAQS9sWQ.
6. Keen M. Why Large Language Models Hallucinate // IBM Technology. 2023. URL: https://www.youtube.com/watch?v=cfqtFvWOfg0.
7. Pythagora project Github repository // Pythagora. 2023. URL: https://github.com/Pythagora-io/pythagora.
8. TestCraftApp project Github repository // TestCraft. 2023. URL: https://github.com/TestCraft-App/test-craft-api-v1.
9. Ng Andrew. The Near Future of AI // Stanford eCorner. 2023. URL: https://www.youtube.com/watch?v=KDBq0GqKpqA.
10. Kirilenko Igor, Jakubiak Nathan. Pros & Cons of Generative AI in Software Testing // Parasoft. 2023. URL: https://www.youtube.com/watch?v=57HTehOnll0.
11. Kirilenko Igor, Jakubiak Nathan. Use AI to Achieve Continuous Software Quality // Parasoft. 2023. URL: https://www.youtube.com/watch?v=mUPL7qJFtKA.
12. Kirilenko Igor, Jakubiak Nathan. Generative AI: the future of testing? - Test Smarter, Not Harder Event // Magnifai. 2023.
URL: https://www.youtube.com/watch?v=9Hfb4nW4uSg.
13. Liu H., Liu L., Yue C., Wang Y., Deng B. Autotestgpt: A system for the automated generation of software test cases based on chatgpt. 2023.
URL: https://ssrn.com/abstract=4584792.
14. Yuan Z., Lou Y., Liu M., Ding S., Wang K., Chen Y., Peng X. No More Manual Tests? Evaluating and Improving ChatGPT for Unit Test Generation. 2023 // arXiv. DOI: https://doi.org/10.48550/arXiv.2305.04207.
15. Ma W., Liu S., Wang W., Hu Q., Liu Y., Zhang C., Nie L., Liu Y. The scope of chatgpt in software engineering: A thorough investigation. 2023 // arXiv *preprint arXiv:2305.12138*.
16. Haas R., Elsner D., Juergens E., Pretschner A., Apel S. How can manual testing processes be optimized? developer survey, optimization guidelines, and case studies // 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ser. ESEC/FSE 2021, Athens, Greece: Association for Computing Machinery, 2021, P. 1281–1291.

Received 14.03.2024.
Accepted 18.03.2024.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Використання генеративного штучного інтелекту в тестуванні програмного забезпечення

У роботі досліджується потенціал використання генеративного штучного інтелекту (GenAI) на основі великих мовних моделей (LLM) з трансформерами для покращення різних аспектів тестування програмного забезпечення. Акцент робиться на можливих практичних застосуваннях і проблемах, що виникають у цих нових підходах. Визначено проблеми тестування і потенціал генеративного ШІ для можливого їх вирішення або зниження їх впливу на ведення проектів програмних систем. Хоча генеративний ШІ на даному етапу розвитку не є повною заміною тестувальникам-людям, він пропонує значні перспективи як потужний допоміжний інструмент, який може трансформувати практики тестування. Очікуваними перевагами є "самовиліковні" тести, які адаптуються до змін коду; генерація тестів у фоновому режимі; можливість генерувати тести з різних джерел (моделей, описів природною мовою, неофіційної документації), і в решті-решт – підвищення продуктивності праці тестувальників. Зазначено виклики використання генеративного ШІ на великих мовних моделях з трансформерами.

Гнатушенко Володимир Володимирович - д.т.н., професор, завідувач кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка» (м. Дніпро).

Павленко Єгор Вікторович - аспірант кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка» (м. Дніпро).

Hnatushenko Volodymyr - Doctor of Technical Sciences (Dr.-Ing. habil.), Professor, Head of Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

Pavlenko Iegor - doctoral student of Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-11
УДК 51-74

К.Г. Гук, А.Є. Шевельова

НЕЙРОМЕРЕЖЕВИЙ ПІДХІД ДО ІДЕНТИФІКАЦІЇ ЗАПОВНЮВАНОСТІ ПРИМІЩЕНЬ ЗА ПАРАМЕТРАМИ ПОВІТРЯ

Анотація. У роботі запропоновано підхід до визначення кількості людей у приміщенні на основі даних спостережень за параметрами повітря із застосуванням багатошарової нейронної мережі. Досліджено питання вибору архітектури та параметрів нейронної мережі для розв'язання задачі прогнозування. Надано рекомендації для підвищення продуктивності моделі.
Ключові слова: концентрація CO_2, нейронна мережа, точність прогнозування, часовий ряд.

Постановка проблеми. Останнім часом питання дослідження параметрів повітря в шкільних класах, лікарнях, офісних та промислових приміщеннях стають поширеними через суттєвий вплив якості повітря на здоров'я та комфорт людей. Наявність у повітрі значної концентрації вуглекислого газу (CO_2), оксиду азоту (NO_2), дрібних часток пилу ($PM_{2.5}$, PM_{10}), летючих органічних сполук (VOC), бактерій та грибків призводить до проблем зі здоров'ям людини, тому моніторинг цих параметрів, можливість прогнозувати їх значення та керувати системами вентиляції є важливими для забезпечення здорового та комфортного середовища у приміщеннях. Таким чином, постає наукова проблема розроблення математичних моделей та методів прогнозування параметрів повітря за результатами моніторингу в реальному часі та суміжні задачі, зокрема керування системами вентиляції на основі спостережень, ідентифікація наповненості приміщень тощо.

Аналіз останніх досліджень. Підвищена увага до якості повітря у приміщеннях під час пандемії COVID-19, промислового забруднення, лісових пожеж та інших екологічних катастроф сприяли розробці та застосуванню моделей на основі даних для прогнозування рівнів забруднюючих речовин в приміщенні і їх впливу на організм людини. У фундаментальному дослідженні [1] наводяться відомості про вплив параметрів повітря на комфорт людини, мето-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ди вимірювання параметрів повітря, відомості про розподіл джерел забруднення і їх вплив на якість повітря.

Останнім часом багато досліджень присвячується застосуванню штучних нейронних мереж (ШНМ) для прогнозування рівня забруднюючих речовин в приміщенні. В роботі [2] штучна нейронна мережа використовується для прогнозування щогодинних концентрацій $PM_{2.5}$, PM_{10}. Для навчання нейронної мережі в якості вхідних даних обираються значення концентрацій $PM_{2.5}$, PM_{10}, час та метеорологічні параметри, а значення відповідних концентрацій, що вимірюються наступної години, використовуються як вихідні дані. Розглядаються різні архітектури нейронних мереж та досліджується вплив вибору тренувальної функції пакету прикладних програм MatLab на результат прогнозування. Порівнюються результати прогнозування з використанням 30 тренувальних функції, серед яких байєсова регуляризація зворотнього поширення (Trainbr) та метод Левенберга-Марквардта (Trainlm) виявились найкращими за значеннями середньоквадратичної та середньої абсолютної похибок. Доводиться, що обидві функції мають здатність до узагальнення та задовольняють потреби погодинного прогнозування наявності часточок пилу.

В [3] досліджуються три моделі ШНМ, а саме ШНМ зворотного поширення, багатошарова ШНМ, а також ШНМ довготривалої та короткочасної пам'яті (LSTM). За результатами масштабного обчислювального експерименту встановлено, що найкращою для прогнозування концентрації $PM_{2.5}$, PM_{10} виявилась багатошарова ШНМ та ансамблевий метод машинного навчання – випадковий ліс, який працює за допомогою побудови численних дерев прийняття рішень під час тренування моделі й здійснює усереднений прогноз.

Автори [4] наводять огляд типів та джерел забруднювачів повітря в приміщенні, досліджують відомі в літературі прогностичні моделі ШНМ, зокрема процеси моделювання та алгоритми їх навчання. За результатами аналізу встановлено, що найчастіше використовуються традиційні ШНМ, тоді як моделі рекурентних нейронних мереж (RNN) краще розв'язують проблему прогнозування часових рядів.

Поточні дослідження найчастіше використовують значення концентрацій $PM_{2.5}$, PM_{10}, CO_2 та не враховують вплив характеристик самого приміщення, параметрів його заповнюваності на результати прогнозування якості повітря через складність включення цих показників в прогнозні моделі. Також актуальною залишається задача визначення кількості людей у приміщенні виходячи

зі значень параметрів повітря для подальшого розрахунку повітрообміну у приміщеннях та керування системами вентиляції.

Мета дослідження. Метою дослідження є створення моделі для прогнозування кількості людей у приміщенні на основі поточних даних про якість повітря, зокрема концентрації CO_2.

Викладення основного матеріалу досліджень. Одним з підходів, який можна використовувати для визначення кількості людей у приміщенні, є оцінка концентрації CO_2 в повітрі. Вважаємо, що люди видихають CO_2, і його рівень у повітрі збільшується зі збільшенням кількості людей у приміщенні. Таким чином, вимірюючи рівень CO_2, можна зробити приблизну оцінку кількості присутніх людей. Однак, визначення кількості людей у приміщенні за допомогою параметрів повітря є складним завданням, оскільки вимагає врахування різноманітних факторів, зокрема, призначення приміщення, умов вентиляції, активності людей, наявності додаткових чинників, таких як відкривання вікон, дверей, наявності інших джерел надходження CO_2, наприклад, джерел відкритого вогню тощо. Характер та кількісні показники деяких з перелічених факторів не піддаються обліку та не можуть бути додані до параметрів моделі. Математичні моделі, що визначають параметри повітрообміну приміщень відповідно до кількості людей, також не дозволяють врахувати перелічені фактори. Використовується лише незначна кількість параметрів в якості коефіцієнтів моделі, а їх значення обираються з деякого інтервалу.

Тому у дослідженні пропонується використання ШНМ для прогнозування кількості людей за результатами спостереження за параметрами повітря у приміщенні. Для побудови моделі пропонується використовувати дані вмісту CO_2 у повітрі, що одержуються з датчиків аналізаторів якості повітря та збираються в офісних приміщеннях різного призначення – конференц-залах, стандартних офісних кімнатах, які обладнано сучасними меблями та оргтехнікою. Данні про вміст CO_2 разом з даними про кількість людей у приміщенні, що також одержуються за допомогою лічильників або камер відеоспостереження, пропонується використати для навчання ШНМ так, щоб навчену на таких даних нейронну мережу в подальшому можна було використовувати для прогнозування кількості людей у приміщенні лише за даними аналізаторів якості повітря.

Для спостереження за параметрами повітря та формування навчальної вибірки у приміщеннях було розташовано сенсори, що вимірюють концентрації CO_2 щогодини протягом року. Вибір зазначеного часового інтервалу пояс-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

нюється тим, що модель має враховувати зміни обумовлені порою року. Також в приміщеннях було розташовано лічильники, які вимірюють кількість людей в приміщенні в ті ж самі проміжки часу. Експериментально створений набір даних, який складався з близько 7000 записів у вигляді часових рядів, використовувався для навчання та тестування ШНМ. Дані, зібрані в приміщеннях А та В було використано для навчання та тестування моделі. Тестування відбувалось на випадково відібраної вибірці, яка складалася з 20 % записів з числа загальних даних для навчання по приміщенням А, В. Дані, що було зібрано в приміщенні С, не брали участь у налаштуванні та тестуванні моделі, їх було використано виключно для перевірки здатності моделі до узагальнення.

Попередньо, для мінімізації обчислювальних витрат набір даних було нормалізовано в такий спосіб:

$$P = (a - a_{\min}) / (a_{\max} - a_{\min}),$$

де P – нормалізоване значення параметра; a – значення параметра з вибірки; $a_{\min}$, $a_{\max}$ – максимальне та мінімальне значення параметра a з вибірки.

Модель використовує виміряні концентрації CO_2 та відомі значення про кількість людей в моменти часу від $(t-1)$ до $(t-i)$, де $i = \overline{2,m}$, щоб спрогнозувати кількість людей в приміщенні на момент *t*. Надалі модель бере прогнозовані значення концентрації CO_2 у приміщенні та кількість людей у момент часу *t* і використовує їх як вхідні дані для прогнозування кількість людей у момент часу $t+1$. В такий спосіб можна прогнозувати кількість людей у приміщенні в часі.

Для оцінювання якості прогнозування моделі застосовувались метрики середньоквадратичної помилки прогнозу (RMSE) та середньої абсолютної помилки прогнозу (MAE):

$$RMSE = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}\left(y(t) - \hat{y}(t)\right)^2}\ ;\ MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y(t) - \hat{y}(t)\right|,$$

де N – загальна кількість використаних даних; $y(t)$ – експериментально виміряне значення, а $\hat{y}(t)$ – прогнозоване значення параметра.

Для програмної реалізації ШНМ було використано бібліотеку програмного забезпечення з відкритим кодом для числових розрахунків з використанням графів потоку даних TensorFlow. Було побудовано багатошарову нейронну мережу, яка складається з вхідного, вихідного та кількох прихованих шарів. Багатошарові ШНМ з прихованими шарами здатні до апроксимації складних нелінійних залежностей, якими є часові ряди. Модель ШНМ оновлює вагову матри-

цю прихованого шару за допомогою алгоритму зворотнього поширення помилки і використовує вихід попереднього прихованого шару в якості входу поточного прихованого шару, щоб поступово змінювати внутрішні параметри мережі. Слід зазначити, що використання значної кількості прихованих шарів призводить до ускладнення навчання та зниження якості прогнозування. Вибір архітектури мережі, зокрема визначення кількості шарів, кількості нейронів у шарі помітно впливає на точність прогнозування кількості людей у приміщенні. Варіанти конфігурацій, що розглядались, та якість результатів прогнозування на тестових вибірках наведено в Табл. 1, оптимальні результати та відповідний варіант конфігурації мережі виділено жирним шрифтом. Для обох приміщень оптимальним виявився один і той самий варіант конфігурації ШНМ. Оцінювання якості прогнозування, що здійснюється мережею, відбувалось за критерієм Краскала – Уолліса, що дозволило перевірити рівність медіан кількох навчальних вибірок.

Таблиця 1

Вибір конфігурації ШНМ

Приміщення	Варіант конфігурації ШНМ	Кількість нейронів у шарах мережі	MAE	RMSE
приміщення А	1	[20, 10, 1]	0.2597475	0.1976358
	2	[20, 10, 5, 1]	0.0923175	0.1259638
	3	**[20, 10, 5, 5, 1]**	**0.0349712**	**0.0463273**
	4	[20, 10, 5, 5, 5, 1]	0.1158294	0.1078490
	5	[20, 10, 5, 5, 5, 5, 1]	0.1435618	0.1096385
приміщення В	1	[20, 10, 1]	0.2749025	0.2112699
	2	[20, 10, 5, 1]	0.1479261	0.1611635
	3	**[20, 10, 5, 5, 1]**	**0.0477911**	**0.0586487**
	4	[20, 10, 5, 5, 5, 1]	0.1247318	0.1414523
	5	[20, 10, 5, 5, 5, 5, 1]	0.1636910	0.1866812

На рис. 1 наводяться результати прогнозування кількості людей у приміщенні А по днях, прогнозні значення позначено пунктирною лінією, сполошною лінією позначено дійсну кількість людей (дані одержано з всієї тестової вибірки).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

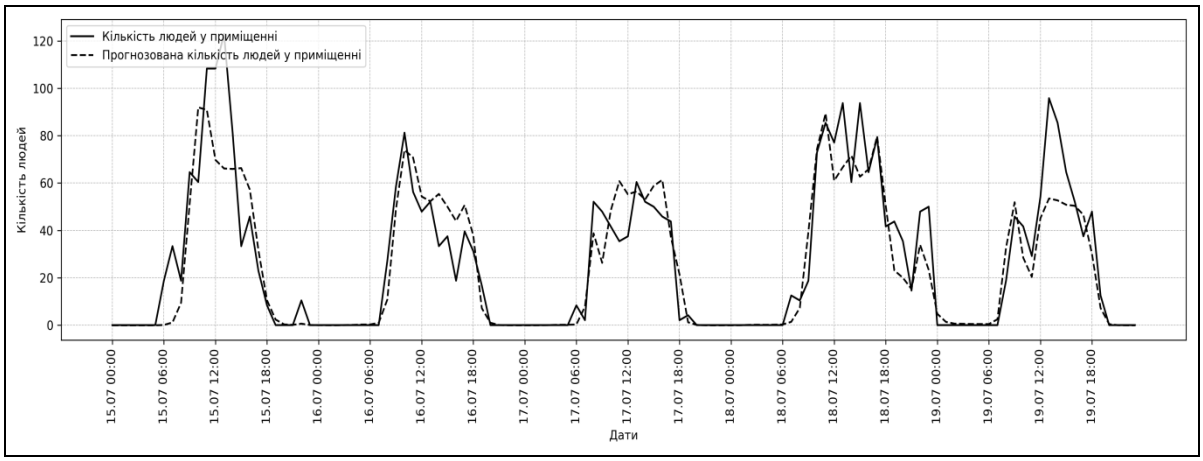


Рисунок 1 – Порівняння прогнозованих значень кількості людей у приміщенні А із спостережуваними значеннями для всієї тестової вибірки

На рис. 2 наведено усереднені значення прогнозованої кількості людей у приміщенні С (відображено пунктирною лінією) та порівняно із усередненими спостережуваними значеннями для всієї навчальної вибірки. Слід зазначити, що модель успішно відтворює характер розподілу прогнозованого параметру, хоча спостережувані данні з приміщення С не брали участь у навчанні моделі. За результатами навчання модель виявилась здатною охопити складні нелінійні залежності між різними факторами, що впливають на результат прогнозування.

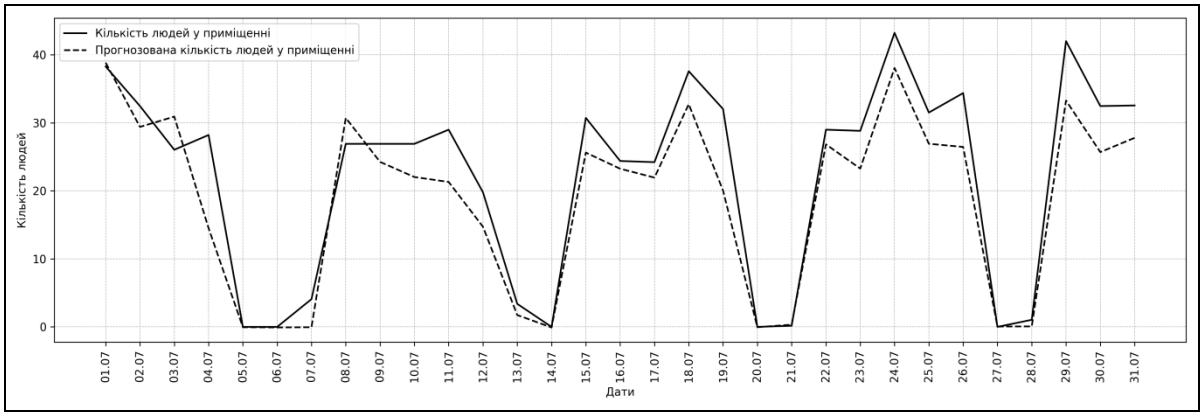


Рисунок 2 – Порівняння усереднених прогнозованих значень кількості людей у приміщенні С із усередненими спостережуваними значеннями для всієї навчальної вибірки

Незначні похибки прогнозованих значень, що залишаються в результатах прогнозу, можна пояснити обмеженістю початкових даних, похибкою самої моделі та наявністю факторів невизначеності. Слід зазначити, що продуктив-

ність моделі можна значно покращити, якщо до набору даних додати параметри впливу системи механічної вентиляції, інтенсивність людської діяльності, частоту відкривання та закривання дверей, вікон тощо.

Висновки. У роботі досліджено можливості багатошарових нейронних мереж для розв'язання задачі прогнозування. За результатами аналізу літературних джерел досліджено стан проблеми, визначено основні підходи до розв'язання задач, пов'язаних з аналізом якості повітря у приміщеннях. Побудовано математичну модель багатошарової нейронної мережі для визначення кількості людей у приміщенні за результатами спостережень за параметрами повітря та її програмну реалізацію. За результатами проведених обчислювальних експериментів із застосуванням статистичного аналізу сформульовано рекомендації щодо вибору архітектури мережі та її параметрів. Надано рекомендації для підвищення продуктивності моделі.

ЛІТЕРАТУРА

1. Pipal A.S. Measurements of Indoor Air Quality // A.S. Pipal, A. Taneja. – Handbook of Metrology and Applications. Springer, Singapore. – 2023. – Pp 1-35. DOI:10.1007/978-981-19-1550-5_90-1

2. Guo Q. Prediction of Hourly $PM_{2.5}$ and PM_{10} Concentrations in Chongqing City in China Based on Artificial Neural Network // Q. Guo, Z. He, Z. Wang // Aerosol and Air Quality Research. – 2023.– Volume 23, issue 6. DOI: 10.4209/aaqr.220448

3. Gao-wa S. Using Artificial Neural Networks to Predict Indoor Particulate Matter and Tvoc Concentration in an Office Building: Model Selection and Method Development // S. Gao-wa, Z. Zhen, N. Jianchun, L. Linxiao, A. Han, Y. Zhili // Energy and Built Environment. 2024. DOI: 10.1016/j.enbenv.2024.03.001.

4. Dong J. A Review of Artificial Neural Network Models Applied to Predict Indoor Air Quality in Schools // J. Dong, N. Goodman, P. Rajagopalan // Int. J. Environ. Res. Public Health. – 2023. – Vol. 20. – Pp. 1-18. DOI: 10.3390/ijerph20156441

REFERNCES

1. Pipal A.S. Measurements of Indoor Air Quality // A.S. Pipal, A. Taneja. – Handbook of Metrology and Applications. Springer, Singapore. – 2023. – Pp 1-35. DOI:10.1007/978-981-19-1550-5_90-1

2. Guo Q. Prediction of Hourly $PM_{2.5}$ and PM_{10} Concentrations in Chongqing City in China Based on Artificial Neural Network // Q. Guo, Z. He, Z. Wang // Aerosol and Air Quality Research. – 2023.– Volume 23, issue 6. DOI: 10.4209/aaqr.220448

3. Gao-wa S. Using Artificial Neural Networks to Predict Indoor Particulate Matter and Tvoc Concentration in an Office Building: Model Selection and Method

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Development // S. Gao-wa, Z. Zhen, N. Jianchun, L. Linxiao, A. Han, Y. Zhili // Energy and Built Environment. 2024. DOI: 10.1016/j.enbenv.2024.03.001.
Dong J. A Review of Artificial Neural Network Models Applied to Predict Indoor Air Quality in Schools // J. Dong, N. Goodman, P. Rajagopalan // Int. J. Environ. Res. Public Health. – 2023. – Vol. 20. – Pp. 1-18. DOI: 10.3390/ijerph20156441.

Received 13.03.2024.
Accepted 18.03.2024.

A neural network approach to the identification of room occupationalness according to air parameters

The paper introduces an approach to determining the number of people in a room based on data from observations of air parameters using a multilayer neural network. Monitoring of air parameters, the ability to predict their values and manage ventilation systems are important to ensure a healthy and comfortable indoor environment. The purpose of the research is to develop mathematical models and methods of forecasting air parameters based on the results of real-time monitoring. Different approaches to predicting air parameters and the number of people in rooms using mathematical models in the form of equations and artificial neural networks with different architectures and types of training functions are considered. The paper proposes an approach to forecasting with the help of a multilayer neural network, which allows taking into account various factors, the nature and quantitative values of which cannot be taken into account and cannot be added to the model parameters. The CO_2 data together with the indoor occupancy data from the meters are used to train the neural network. In the future, a neural network trained on such data can be used to predict the number of people in a room based only on data from air quality analyzers.

The issue of choosing the architecture of a multilayer neural network and its parameters for solving the forecasting problem has been investigated. Neural network training is carried out by the method of error back propagation. To evaluate the forecasting quality of the model, the metrics of mean square error of forecast and mean absolute error of forecast are used. The Kruskal-Wallis criterion is used to take into account the results of forecasting on several samples. Based on the results of the computational experiment, the optimal network architecture is determined. The model successfully reproduces the nature of the distribution of the predicted parameter, as it captures the complex nonlinear dependencies between the various factors of the model. Recommendations are given to improve the performance of the model.

Гук Костянтин Григорович – магістрант кафедри ракетно-космічних та інноваційних технологій, Дніпровський національний університет імені Олеся Гончара, Дніпро.

Шевельова Алла Євгенівна – д-р фіз.-мат. наук, професор, професор кафедри обчислювальної математики та математичної кібернетики, Дніпровський національний університет імені Олеся Гончара, Дніпро.

Huk Kostyantin – master's student of the Department of Rocket-space and Innovative Technologies, Oles Honchar Dnipro National University, Dnipro.

Sheveleva Alla – Doctor of Physical and Mathematical Sciences, Professor, Professor of the Department of Computational Mathematics and Mathematical Cybernetics, Oles Honchar Dnipro National University, Dnipro.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

«Системні технології» 2 (151) 2024 «System technologies»
DOI 10.34185/1562-9945-2-151-2024-12
УДК 004.021, 004.042, 519.688

В. Жушман, Т. Зайцева

КОМПЛЕКСНИЙ ПІДХІД ДО РОЗВ'ЯЗАННЯ ЗАДАЧІ ВЗАЄМОДІЇ АБСОЛЮТНО ЖОРСТКОГО ДВОЗВ'ЯЗНОГО ШТАМПУ ТА ПРУЖНОГО ПІВПРОСТОРУ

Анотація. У роботі представлено комплексний підхід що базується на принципах системного аналізу для розв'язання контактних задач. Розглядаються задачі про вдавлення в однорідний та ізотропний пружний півпростір абсолютно жорстких плоских однозв'язних і двозв'язних штампів у формі некругового кільця. Для отримання аналітичного рішення застосовується метод, що базується на використанні розвинення потенціалу простого шару для областей близьких до кільцевих. Для візуалізації і аналізу результатів на мові С++ створено програмне забезпечення. Скінченно елементні моделі для відтворення взаємодії жорсткого штампу з пружним півпростором будуються у програмному середовищі ANSYS. Важливим етапом є перевірка адекватності моделей, яка відбувається у томі числі і за рахунок порівняння отриманих чисельних результатів з аналітичними. Досягнуто задовільне узгодження результатів чисельного моделювання з отриманими раніше аналітичними. У разі перебування системи штамп пружній півпростір в складних природних умовах або в агресивному середовищі за певний час модельного застосування, враховуються можливі випадкові пошкодження або такі, що виникають за певним законом, наприклад, корозія. Тобто, за таких умов, з часом, розміри зон контакту можуть змінюватися і стають невідомими. Створюється числова база розрахунків системи штамп пружний півпростір для різних форм поперечних перерізів штампів, поєднуючи їх у спеціальні групи. Для розробки і підтримки експертної системи використано програмний інструмент CLIPS. Розрахункова база передається в нього за допомогою спеціально створеного програмного додатку на мові С++. На основі набору правил та знань, які були створені і використані для розв'язання конкретних завдань відбувається автоматизація процесу прийняття рішень. Для кожної окремої комп'ютерної моделі, розраховуються масиви даних – нормальні і дотичні напруження в певних точках. Проводиться ідентифікація форми поперечного перерізу штампу відповідно до визначених у базі знань критеріїв. Процес генерації форми поперечного перерізу штампу відбувається за допомогою спеціально розробленого програмного забезпечення в OpenGL. У якості математичного інструменту, використано кубічну сплайн інтерполяцію.

Ключові слова: аналітичний розв'язок, контактна задача, штамп, математична модель, метод скінченних елементів, напруження, теорія пружності, область контакту, експертна система, системний підхід.

Постановка проблеми. Застосування системного підходу до вирішення задач контактної механіки дозволяє враховувати різні аспекти системи, такі як геометрія, матеріальні властивості, граничні умови та навантаження; допомагає уникнути спрощень; дозволяє створювати математичні моделі, які відображають взаємодію всіх частин конструкцій та механізмів.

Аналіз останніх досліджень і публікацій. Важливою задачею сьогодення є підвищення якості, надійності, економічності та продуктивності машин, обладнання та іншої продукції машинобудування, зниження їх матеріаломіскості. Тому розвиток математичної теорії та підвищення ефективності її використання в прикладних цілях завжди залишається актуальним питанням. Саме шляхом застосування аналітичних методів у роботі [1,2] отримано розв'язки контактних задач для класичних областей. У роботі [3] авторами запропоновано новий підхід до пошуку рішення задачі для кільцевого штампа у вигляді подвійного ряду, в якому коефіцієнти обчислюються точно з простих рекурентних співвідношень.

Мають розвиток дослідження у напрямку взаємодію пружних тіл з початковими залишковими напруженнями з урахування сил тертя та без. В роботах [4, 5] розглядається взаємодія жорсткого циліндричного кільцевого штампа та пружного півпростору вивчався.

У роботах [6,7] авторами вирішуються квазістатичні контактні задачі, які мають непересічне значення у задачах промисловості.

Широко застосовується дослідниками метод скінченних елементів, як інструмент для знаходження розв'язків інтегральних та диференціальних рівнянь у частинних похідних [8, 9]. Треба сказати, що ефективність застосування методу скінченних елементів продемонстрована авторами для вирішення обернених контактних задач в роботах [8, 9]. Також достатньо популярним є і метод граничних елементів.

Авторами у роботі [10], застосовується розкладання потенціалу простого шару для отримання відповіді, щодо оптимізації розподілу тиску в контактній зоні. Задача ділиться на дві послідовно розв'язувані задачі.

У роботах [11-13] досліджено проблеми вдавлювання штампів з плоскою основою, обмеженою двозв'язними площадками контакту. Для розв'язання цих

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

задач застосовується і має своє розвинення аналітичний підхід запропонований у роботі [3] і також широко використовуються інформаційні технології.

Необхідно відзначити, що комбінування аналітичних і чисельних методів є важливим і сучасним підходом при розв'язанні багатьох задач контактної механіки. Тобто, поєднання різних підходів до рішення задач та надання комплексних оцінок допомагають знайти шляхи до вирішення складних проблемних питань, що давно стоять перед інженерами і науковцями. Такі підходи стають потужним інструментом для аналізу та моделювання різних явищ і процесів.

Метою дослідження. Метою дослідження є розробка комплексного підходу що базується на принципах системного аналізу для розв'язання задач створення та дослідження математичних і комп'ютерних моделей контактної взаємодії абсолютно жорсткого циліндричного штампа із плоскою основою з пружнім півпростором під дією стискаючої сили.

Викладення основного матеріалу дослідження.

Розглядається абсолютно жорсткий плоский штамп, що займає в плані область, контур якої обмежений двома подібними кривими. Початок координат розташовано в точці, що співпадає з центром тяжіння перерізу основи штампа. Штамп перебуває в рівновазі під дією прикладених до нього зовнішніх сил і реакції пружного середовища. Площина, що обмежує півпростір, містить область контакту (основа штампу) і область, що вільна від впливу. Поверхня штампу, що контактує з пружним середовищем описується рівнянням $z = f(x, y)$, де $(x, y) \in \Omega_f$. Передбачається відсутність сил тертя в області контактної взаємодії штампу та пружного середовища $z \geq 0$. Будемо вважати, що форма штампу $f(x, y)$ є неперервною і гладкою функцією координат, а тиск буде $p = p(x, y) \geq 0$ [11]

Було розроблено програмне забезпечення для аналізу та візуалізації аналітичного рішення. Вибір мови програмування C++ для реалізації програмного забезпечення визначився низкою обґрунтованих причин, які сприяли отриманню ефективного вирішенню поставленої задачі. Треба зазначити, що мова програмування C++ відома своєю високою продуктивністю та можливістю ефективно працювати на низькорівневому рівні, що особливо важливо для вирішенні складних математичних задач. Також, наявність багатофункціональних бібліотек C++ сворює можливості для вдалого проведення математичних обчислень. Об'єктно-орієнтований підхід сприяє створенню структурованих та легко збережених програм, що є важливим для обчислювальних наукових досліджень. Переносимість коду дозволяє створює

умови для кросплатформеості. Тому програмне забезпечення для аналізу та візуалізації аналітичного рішення, отриманого раніше [11], було розроблено на мові С++.

Було побудувано скінченно-елементні моделі взаємодії жорстких штампів з пружним півпростору у формі однозв'язних і двозв'язних штампів за допомогою програмного комплексу ANSYS [14]. Було створено тривимірні моделі геометрії пружного півпростору та штампу в програмному модулі SpaceClaim. Для доступа до функціоналу, що відсутній в поточній версії інтерфейсі ANSYS було розроблено програмний додаток на мові APDL. Для створення організації за особими правилами масивів даних, відбувалося зчитування напружень і переміщень зі спеціально розташованих умовних датчиків на границях штампів. Проаналізувано напружено-деформований стан системи та оцінено результати шляхом порівняння з отриманими раніш аналітичними. Для випадків коли похибка перевищує допустиму, то проводився уточнювальний процес генерації скінченно-елементної моделі.

Розгладалися випадки коли вважалося, що системв штамп- пружний півпростір перебуває в складних природних умовах або в агресивному середовищі, або відбуваются процеси зношення за певний час модельного застосування. Враховано можливі випадкові пошкодження або такі, що виникають за певним законом, наприклад, корозія. Побудована значна кількість скінченно-елементних моделей для різних випадків пошкоджень штампів, для яких нові розрахункові форми поперечних перерізів було змінено від початкових за визначеним правилом.

Нижче в процесі пошуку контактного тиску $p(x,y)$, $(x,y)\in\Omega_f$, розглядаються рішення двох крайових задач теорії пружності для півпростору $z\geq 0$ [10].

Пошук розподілу тиску $p(x,y)$ в області Ω_f при завданні нормальних переміщень. $w(x,y)=f(x,y)$ на Ω_f. Розв'язок даної крайової задачі теорії пружності для півпростору $z\geq 0$ з граничними умовами $\sigma_{xz}=\sigma_{yz}=0$, $w=f(x,y)$, $(x,y)\in\Omega_f$, та $\sigma_{zz}=\sigma_{xz}=\sigma_{yz}=0$, $(x,y)\in\Omega_0$, існує і єдино [10], а розподіл тиску, що визначається, $p(x,y)=\sigma_{zz}(x,y)$, $(x,y)\in\Omega_f$ може бути представлено у вигляді

$$p=p(f)=L_a f, \qquad (1)$$

де $\sigma_{ij}, i,j=x,y,z,$ - компоненти тензора напружень, w - компонента вектора переміщень, що нормальна до площини xy, а L_a - лінійний оператор.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Пошук розподілу нормальних зміщень $w(x,y) = f(x,y)$ в області Ω_f при прикладенні до цієї області заданих невід'ємних центральних тисків $p(x,y) \geq 0$. Розв'язок відповідної крайової задачі для півпростору $z \geq 0$ з граничними умовами $\sigma_{xy} = \sigma_{yz} = 0$, $\sigma_{zz} = p(x,y)$ при $(x, y) \in \Omega_0$ існує, єдиний і знаходиться в вигляді потенціалу простого шару. Розподіл нормальних прогинів $w(x,y) = f(x,y)$ в області Ω_f надається виразом

$$f = f(p) = L_b p, \tag{2}$$

з лінійним оператором L_b.

При відомому розподілі тиску $p(x,y),\ (x,y) \in \Omega_f$, результуюча сила P і моменти M_x, M_y відносно осей x і y, що прикладені до штампу визначаються виразами [10,11]

$$P = \int_{\Omega_f} p d\Omega_f,\ M_x = \int_{\Omega_f} y p d\Omega_f,\ My = \int_{\Omega_f} x p d\Omega_f. \tag{3}$$

Введемо в розгляд функцію $p_g = p_g(x,y) \geq 0$, що описує деякий заданий розподіл тиску в області контакту Ω_f. Як оптимізуючий функціонал розглянемо інтеграл вигляду

$$J_f = J(p(f)) = \int_{\Omega_f} (p - p(f))^2 d\Omega_f. \tag{4}$$

Функціонал (4) характеризує неузгодженість між розподілом тиску $p(x,y)$, що відповідає певній формі штаму $f(x,y)$, і заданим (цільовим) розподілом тиску $p_g(x,y)$ [10].

Експертну систему було реалізовано на базі існуючого відкритого для загального використання програмного продукту CLIPS [15]. Експорт даних щодо напружень і переміщень з умовних датчиків в базу знань експертної системи було реалізовано в автоматизованому форматі через розроблене спеціальне програмне забезпечення на мові програмування C++, та за допомогою ручної передачі даних через імпорт інформації до Excel файлу, та подальшу його обробку авторським додатком на мові програмування JavaScript. База знань системи зберігає інформацію щодо геометричних характеристик штампів та властивостей матеріалів півпростору і штампу, а також набір правил і фактів, що використовуються для визначення відповідності напружено-деформівного стану системи абсолютно жорсткий штамп – пружний півпростір координатам точок за заданими критеріями.

Після обробки вхідних данних експертна система видає інформацію щодо координат точок контуру штампу, яка передається на вхід до розробленого спеціального програмного забезпечення для візуалізації контуру штампів. Для відтворення контурів штампів було використано кубічну сплайн-інтерполяцію. На завершальному етапі відбувається порівнянні побудованих контурів штампів з модельними. На рис. 1. наведено загальну схема будови комплексного підходу до розв'язання задач створення та дослідження математичних і комп'ютерних моделей контактної взаємодії абсолютно жорсткого циліндричного штампа із плоскою основою з пружнім півпростором під дією стискаючої сили.

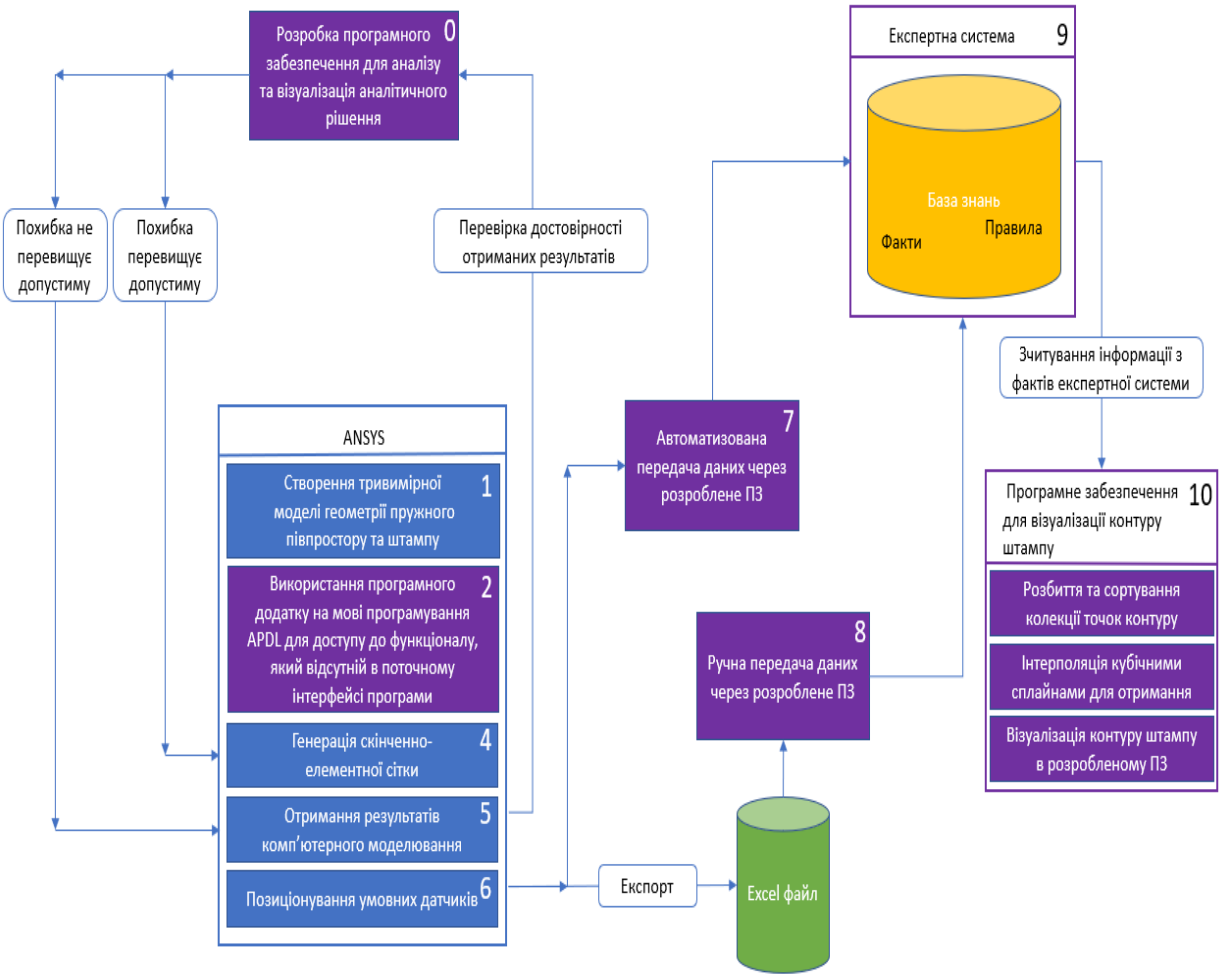


Рисунок 1 - Загальна схема будови комплексного підходу

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Висновки

В роботі застосовано комплексний підхід що базується на принципах системного аналізу для розв'язання задач створення та аналізу математичних і комп'ютерних моделей контактної взаємодії абсолютно жорсткого циліндричного штампа із плоскою основою з пружнім півпростором під дією стискаючої сили. В процесі виконання роботи було розроблено програмне забезпечення для розрахункових формул аналітичного рішення задач щодо плоских штампів різної форми в плані близьких до кільцевих для проведення ефективних розрахунків і аналізу отриманих результатів; створення скінченно-елементні моделі взаємодії плоского абсолютно жорсткого двозв'язного в плані штампу з пружним півпростором за допомогою програмного комплексу ANSYS; створено програмне забезпечення для передачі масивів даних з програмного комплексу ANSYS до програмної системи CLIPS; побудовано експертну систему програмному середовищі CLIPS для автоматизації прийняття рішення щодо вибору найбільш вдалого варіанту побудованої границі плоского штампу, у випадку зміни його форми під впливом дії агресивного середовища або випадкових пошкоджень.

ЛІТЕРАТУРА

1. Galin L. A., Gladwell G. M. L. Contact Problems – the legacy of L.A. Galin. Solid Mechanics and Its Applications. 2008. Vol. 155. Springer, 318 p. https://doi.org/10.1007/978-1-4020-9043-1.

2. Моссаковский В. И. Контактные задачи математической теории упру гости / В. И. Моссаковский, Н. Е. Качаловская, С. С. Голикова.– К.: Наукова думка, 1985. – 176 с.

3. Roitman A. B., Shishkanova S. F. The solution of the annular punch problem with the aid of recursion relations. Soviet Applied Mechanics. 1973. 9 (7), 725–729. https://doi.org/10.1007/BF00882996.

4. Babych, S.Y., Yarets'ka, N.O Contact Problem for an Elastic Ring Punch and a Half-Space with Initial (Residual) Stresses*. International Applied Mechanics. 2021. 57, 297–305. https://doi.org/10.1007/s10778-021-01081-7.

5. Бабич, С. Ю., Ярецька, Н. О., Лазар, В. Ф., Щекань, Н. П. Аналітичні розв'язки статичної задачі про тиск попередньо напружених півпросторів та пружного циліндра з початковими напруженнями. Науковий вісник Ужгородського університету. Серія «Математика і інформатика». 2022. 41 (2), 91–102. https://doi.org/10.24144/2616-7700.2022.41(2).91-102.

6. Li, B., Li P., Zhou R., Feng X., Zhou, K. Contact mechanics in tribological and contact damage-related problems: A review, Tribology International. 2022. Volume 171, 107534, https://doi.org/10.1016/j.triboint.2022.107534.
7. Guz A. N., Zozulya, V. V. Elastodynamic unilateral contact problems with friction for bodies with crack. International Applied Mechanics. 2002. 38 (8), 895–932 p. https://doi.org/10.1023/A:1021266113662.
8. Obodan, N. I., Zaitseva, T. A., Fridman, O. D. Contact Problem for a Rigid Punch and an Elastic Half Space as an Inverse Problem. Journal of Mathematical Sciences. 2019. 240, 184–193 p. https://doi.org/10.1007/s10958-019-04346-2.
9. Guk, N. A., Kozakova, N. L. Delamination of a Three-Layer Base Under the Action of Normal Loading. Journal of Mathematical Sciences. 2021. 254(1), 89–102 p. https://doi.org/10.1007/s10958-021-05290-w.
10. Banichuk N V, Ivanova S Yu Optimal Structural Design: Contact Problems and High-Speed Penetration. Berlin, Boston: De Gruyter. 2017. 206 p. https://doi.org/10.1515/9783110531183.
11. Зайцева Т. А., Шишканова Г. А. Розв’язання просторових контактних задач для некласичних багатозв’язних областей. Дніпро: Вид-во Дніпропетр. нац. ун-ту., 2011. 192 с.
12. Shyshkanova G. A., Zaytseva T. A., Frydman A. D. The analysis of manufacturing errors effect on contact stresses distribution under the ring parts deformed asymmetrically. Metallurgical and Mining Industry. 2015. 7. 352–357. www.metaljournal.com.ua/assets/Journal/englishedition/MMI_2015_7/055Shyshkanova-352-357.pdf
13. Shyshkanova G. A., Zaytseva T. A., Zhushman V. V., Levchenko N. M., Korotunova O. V. Solving three-dimensional contact problems for foundation design in green building. Journal of Physics: Conference Series. 2023. 2609 (1), 012001. http://doi.org/10.1088/1742-6596/2609/1/012001
14. Ansys Free Student Software Downloads.
URL: https://www.ansys.com/academic/free-student-products (дата звернення: 02.04.2024).
15. CLIPS A Tool For Building Expert Systems. URL: https://www.clipsrules.net (дата звернення: 02.04.2024).

REFERENCES

1. Galin L. A., Gladwell G. M. L. Contact Problems – the legacy of L.A. Galin. Solid Mechanics and Its Applications. 2008. Vol. 155. Springer, 318 p. https://doi.org/10.1007/978-1-4020-9043-1.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

2. Mossakovskiy V I. Kontaktnye zadachi matematicheskoy teorii uprugosti / V.I. Mossakovskiy, N.E. Kachalovskaya, S.S. Golikova. – K.: Naukova dumka,1985.– 176 s.
3. Roitman A. B., Shishkanova S. F. The solution of the annular punch problem with the aid of recursion relations. Soviet Applied Mechanics. 1973. 9 (7), 725–729. https://doi.org/10.1007/BF00882996.
4. Babych, S.Y., Yarets'ka, N.O Contact Problem for an Elastic Ring Punch and a Half-Space with Initial (Residual) Stresses*. International Applied Mechanics. 2021. 57, 297–305. https://doi.org/10.1007/s10778-021-01081-7.
5. Babych, S. Yu., Yaretska, N. O., Lazar, V. F., Shchekan, N. P. Analitychni rozvyazky statychnoyi zadachi pro tysk poperedno napruzhenykh pivprostoriv ta pruzhnogo tsylindra z pochatkovymy napruzheniyamy. Naukovyy visnyk Uzhhorodskoho universytetu. Seriya "Matematyka i informatyka". 2022. 41 (2), 91–102 s. https://doi.org/10.24144/2616-7700.2022.41(2).91-102.
6. Li, B., Li P., Zhou R., Feng X., Zhou, K. Contact mechanics in tribological and contact damage-related problems: A review, Tribology International. 2022. Volume 171, 107534, https://doi.org/10.1016/j.triboint.2022.107534.
7. Guz A. N., Zozulya, V. V. Elastodynamic unilateral contact problems with friction for bodies with crack. International Applied Mechanics. 2002. 38 (8), 895–932 p. https://doi.org/10.1023/A:1021266113662.
8. Obodan, N. I., Zaitseva, T. A., Fridman, O. D. Contact Problem for a Rigid Punch and an Elastic Half Space as an Inverse Problem. Journal of Mathematical Sciences. 2019. 240, 184–193 p. https://doi.org/10.1007/s10958-019-04346-2.
9. Guk, N. A., Kozakova, N. L. Delamination of a Three-Layer Base Under the Action of Normal Loading. Journal of Mathematical Sciences. 2021. 254(1), 89–102 p. https://doi.org/10.1007/s10958-021-05290-w.
10. Banichuk N V, Ivanova S Yu Optimal Structural Design: Contact Problems and High-Speed Penetration. Berlin, Boston: De Gruyter. 2017. 206 p. https://doi.org/10.1515/9783110531183.
11. Zaytseva T. A., Shyshkanova H. A. Rozvyazannia prostorovykh kontaktnykh zadach dlia neklasichnykh bahatozv'iaznykh oblastey. Dnipro: Vyd-vo Dnipropetr. nats. un-tu., 2011. 192 s.
12. Shyshkanova G. A., Zaytseva T. A., Frydman A. D. The analysis of manufacturing errors effect on contact stresses distribution under the ring parts deformed asymmetrically. Metallurgical and Mining Industry. 2015. 7. 352–357. www.metaljournal.com.ua/assets/Journal/englishedition/MMI_2015_7/055Shyshkanova-352-357.pdf

13. Shyshkanova G. A., Zaytseva T. A., Zhushman V. V., Levchenko N. M., Korotunova O. V. Solving three-dimensional contact problems for foundation design in green building. Journal of Physics: Conference Series. 2023. 2609 (1), 012001. http://doi.org/10.1088/1742-6596/2609/1/012001
14. Ansys Free Student Software Downloads.
URL: https://www.ansys.com/academic/free-student-products (date of application: 02.04.2024).
15. CLIPS A Tool For Building Expert Systems. URL: https://www.clipsrules.net (date of application: 02.04.2024).

Received 12.03.2024.
Accepted 19.03.2024.

A complex approach to solving the problem of interaction between a rigid double-connected punch and an elastic half-space

The paper presents an integrated approach based on the principles of system analysis for solving contact problems. We consider the problems of pressing rigid plane single- and double-connected punches in the form of a non-circular ring into a homogeneous and isotropic elastic half-space. To obtain an analytical solution, we apply a method based on the use of the development of the simple layer potential for regions close to the ring. Software was developed using C++ to visualize and analyze the results. Finite-element models to reproduce the interaction of a rigid punch with an elastic half-space are built in the ANSYS software environment. An important step is to verify the adequacy of the models, which is carried out, among other things, by comparing the numerical results with the analytical ones. A satisfactory agreement of the numerical modeling results with the analytical ones obtained earlier was achieved. If the punch-elastic half-space system is exposed to difficult natural conditions or an aggressive environment during a certain time of modeling, possible accidental damage or damage that occurs according to a certain law, such as corrosion, is taken into account. That is, under such conditions, the dimensions of the contact zones may change over time and become unknown. A numerical base for calculating the punch-elastic half-space system is created for various shapes of punch cross-sections, combining them into special groups. The CLIPS software tool was used to develop and maintain the expert system. The calculation base is transferred to it using a specially created C++ software application. Based on a set of rules and knowledge that have been created and used to solve specific problems, the decision-making process is automated. For each individual computer model, data sets are calculated - normal and tangential stresses at certain points. The cross-sectional shape of the punch is identified in accordance with the criteria defined in the knowledge base. The

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

process of generating the cross-sectional shape of the punch is performed using specially developed software in OpenGL. The cubic spline interpolation is used as a mathematical tool.

Keywords: analytical solution, contact problem, punch, mathematical model, finite-element method, stresses, elasticity theory, contact zone, expert system, systematic approach.

Зайцева Тетяна Анатоліївна - к.т.н., доцент, зав. каф. комп'ютерних технологій, Дніпровський національний університет імені Олеся Гончара, ORCID: https://orcid.org/0000-0002-6346-3390

Жушман Владислав Вікторович – аспірант, кафедра комп'ютерних технологій, Дніпровський національний університет імені Олеся Гончара, ORCID: https://orcid.org/0009-0000-7292-5879

Tetyana Zaytseva – PhD, Associate Professor, Head of Department, Department of Computer Technologies, Oles Honchar Dnipro National University, ORCID: https://orcid.org/0000-0002-6346-3390

Vladyslav Zhushman – PhD student, Department of Computer Technologies, Oles Honchar Dnipro National University, ORCID: https://orcid.org/0009-0000-7292-5879

DOI 10.34185/1562-9945-2-151-2024-13
УДК 378.14:629.331

О.Ю. Рудик, О.В. Диха, К.Е. Голенко

ІННОВАЦІЙНІ ПІДХОДИ ПРИ ВИКЛАДАННІ ДИСЦИПЛІН АВТОМОБІЛЬНОГО СПРЯМУВАННЯ

Анотація. Для навчальної дисципліни «Моделювання технологічних процесів підприємств автомобільного транспорту» у середовищах SolidWorks Simulation і Ansys Workbench сформовані основні принципи та положення автоматизованого проектування в області комп'ютерного моделювання агрегатів, вузлів і деталей транспортних засобів, а також пристосувань для їх ремонту (піднімачів, домкратів, стендів, знімачів тощо). Основна увага приділена теорії й практичному використанню методів скінченних елементів та набуття навичок у проєктуванні та розрахунках деталей автомобільного транспорту. Визначені обов'язкові елементи досліджень у SolidWorks та практичні навички моделювання різних режимів навантажень дорожніх та спеціальних транспортних засобів у Ansys Workbench. Для подовження ресурсу роботи конструктивних елементів і деталей автомобільного транспорту визначені методи їх відновлення та підвищення зносостійкості.
Ключові слова: автомобільний транспорт, комп'ютерне моделювання, SolidWorks Simulation, Ansys Workbench, методика.

Навчальна дисципліна «Моделювання технологічних процесів підприємств автомобільного транспорту» є обов'язковою складовою фахової підготовки здобувачів вищої освіти для спец. 274 «Автомобільний транспорт» (бакалавр). Студент, який успішно завершив вивчення дисципліни, повинен: вміло використовувати понятійний апарат з фаху; розуміти структуру й динаміку технологічних процесів (ТП) автотранспортних підприємств; уміти використовувати чисельні методи розв'язку статичних і динамічних задач механіки твердого тіла; працювати з пакетом прикладних програм для плоского й твердотільного моделювання SolidWorks; володіти інженерними розрахунками твердотільних моделей на міцність у машинобудуванні (SolidWorks Simulation і Ansys Workbench).

Основна увага приділяється теорії й практичному використанню методів скінченних елементів (МСЕ) та набуття навичок у проєктуванні та розрахунках

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

деталей і вузлів автомобільного транспорту (АТ), а також пристосувань для їх ремонту.

Мета дисципліни: формування сучасних уявлень і практичних навичок з використання методів комп'ютерного проєктування деталей, вузлів та агрегатів транспортних засобів з наступною імітацією їх натурних випробуван у розрахункових середовищах МСЕ; дати уявлення про основи чисельних методів розв'язку статичних і динамічних задач механіки твердого тіла, а також про алгоритми та особливості їх чисельної реалізації; навчити застосовувати наближені методи для вирішення конкретних задач, які виникають у науково-технічній практиці; оволодіти навичками використання систем автоматизованого проєктування (САПР) в області ТП автотранспортних підприємств.

Для навчальної дисципліни «Моделювання технологічних процесів підприємств АТ» необхідно сформувати основні принципи та положення автоматизованого проєктування в області комп'ютерного моделювання деталей, вузлів та агрегатів транспортних засобів, а також імітації різних режимів їх натурних випробувань та умов експлуатації у середовищах SolidWorks Simulation і Ansys Workbench.

Обов'язковими елементами досліджень у SolidWorks Simulation є [1, 2]: вибір додатку SolidWorks Simulation та кроків роботи у ньому [3]; призначення виду дослідження (статичний аналіз) моделі; обрання матеріалу деталі з бібліотеки матеріалів SolidWorks; вибір кріплення деталі; прикладення навантаження до певних площин, граней чи елементів деталі;створення сітки МСЕ; вплив якості сітки на точність розрахунків (рис. 1 – [4]);

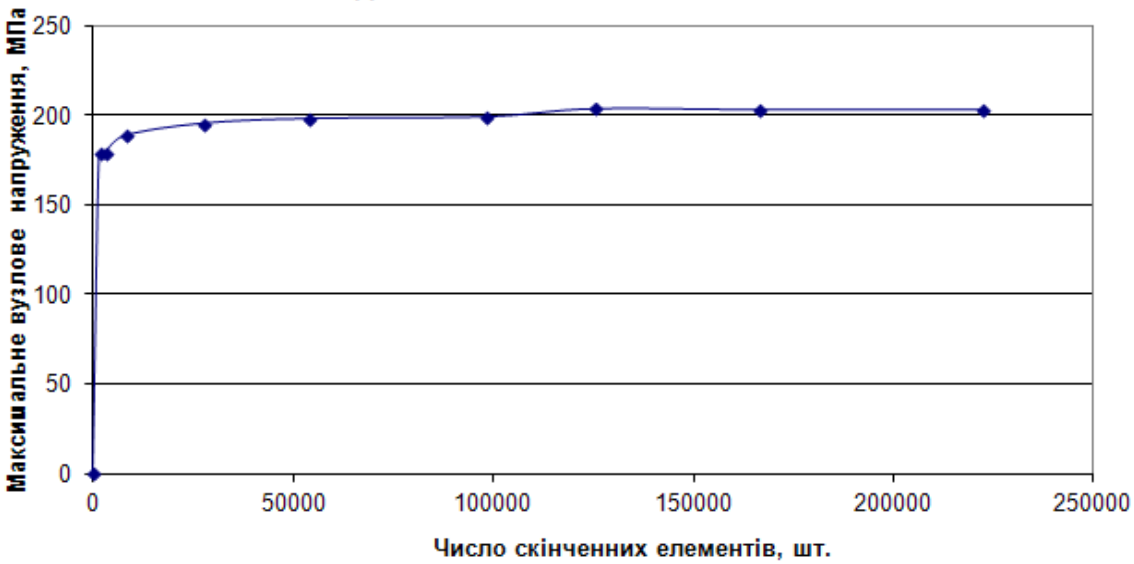


Рисунок 1 – Залежність вузлового напруження
від числа скінченних елементів моделі

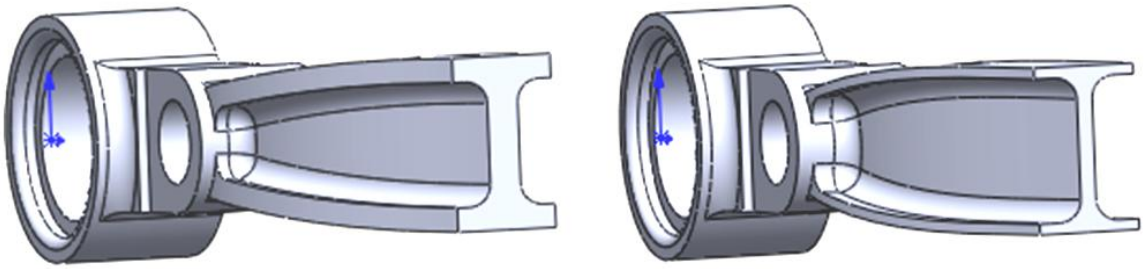
– можливість заміни матеріалу деталі (рис. 2 - [5]);

Результати дослідження вал-шестерні

Сталь	Напруження (макс.), МПа	Переміщення (макс.), мм	Деформація (макс.), мм	Запас міцн. (мінім.)
12Х14Г14Н	89.9025	0.0713178	0.000333314	3.05887
20	89.670	0.071990	0.00030610	2.7880

Рисунок 2 – Результати заміни матеріалу деталі

– економія матеріалу деталі на основі аналізу напруженого стану її моделі (рис. 3 – [6]);



а б

Рисунок 3 – Деталь до (а) і після (б) оптимізації

– зондування напружень у критичних точках деталі (рис. 4 – [7]);

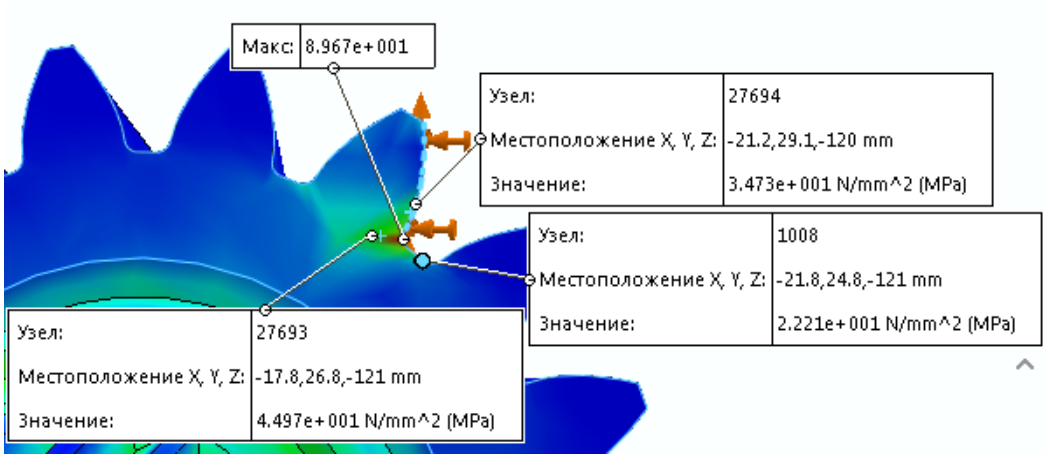
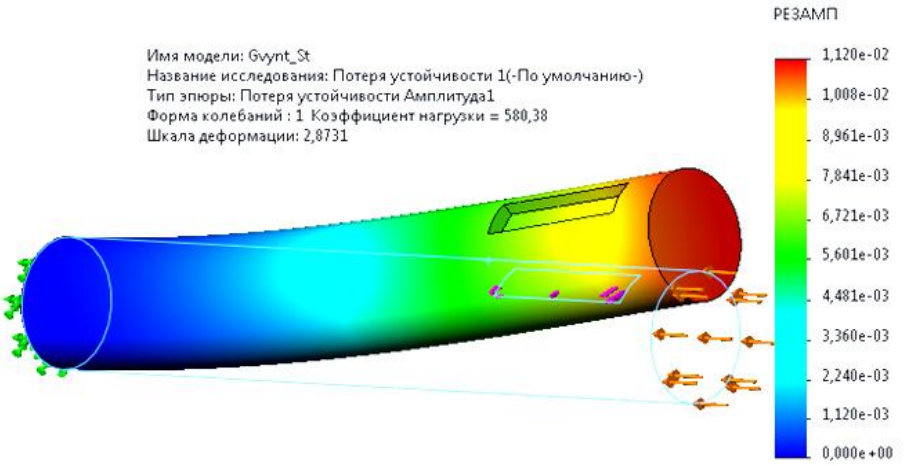


Рисунок 4 – Зондування напружень у небезпечному перерізі поблизу їх максимальних значень

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

– можлива втрата стійкості деталі (рис. 5 – [8]);



Имя	Тип	Мин	Макс
Амплитуда1	РЕЗАМП: результирующая амплитуда Эпюра - для формы колебаний: 1(Коэффициент нагрузки = 580,383)	0,000e+00 Узел: 1	1,120e-02 Узел: 513

Рисунок 5 – Результуюча амплітуда та запас міцності при втраті стійкості

– максимальне навантаження, яке може витримати змодельована деталь без руйнування, при заданому коефіцієнті запасу міцності [9];

– вплив кріплень деталей на їх працездатність (рис. 6 – [10]);

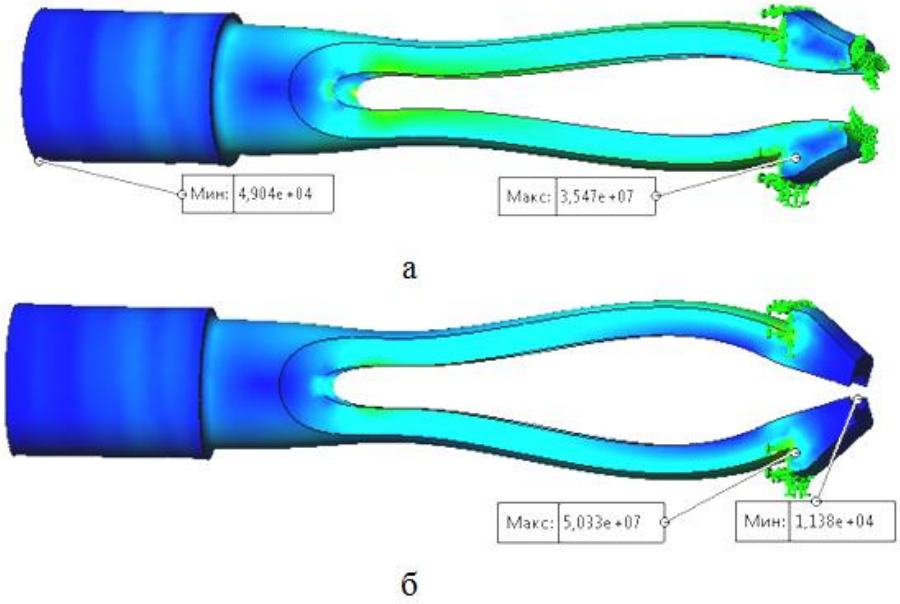
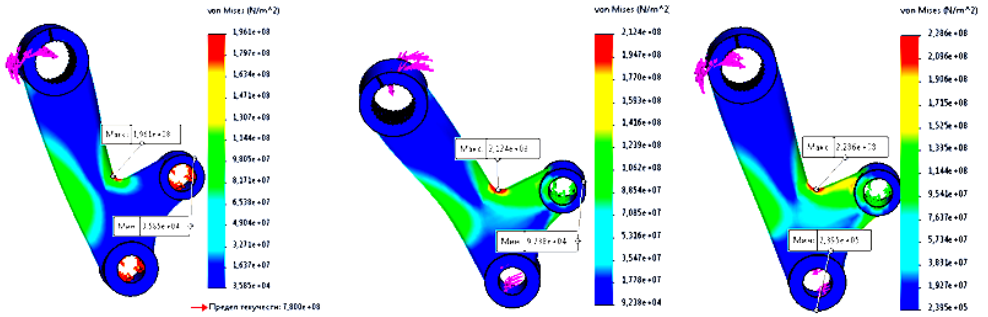


Рисунок 6 – Епюри розподілу сумарних напружень деталі:
а – із закріпленою кромкою захвату; б – без її закріплення

– вплив удару на стійкість деталей колісних машин (рис. 7 – [11]);



а б в

Рисунок 7 – Розподіл еквівалентних напружень у моделі рульової сошки
а – поворот на місці; б – удар лівим колесом; в – удар правим колесом)

Метою індивідуального завдання дисципліни є здобуття практичних навичок з особливостей моделювання різних режимів навантажень дорожніх та спеціальних транспортних засобів, а також їх складових, з наступним аналізом напружено-деформованого стану, термічних навантажень, аеродинаміки, теплопередачі, конвекції, тертя, тощо. Завдання полягає у моделюванні напружено-деформованого стану каркасних елементів транспортних засобів у середовищі Ansys Workbench: засвоєння методики оцінки міцності каркасу кузова міського автобуса при статичних випробуваннях у режим «згинання» та «кручення».

Студенти в результаті виконання індивідуального завдання [12]:

– досліджують особливості формування крайових умов імітації натурних випробувань; набувають досвіду роботи з гібридними моделями, представленими стрижневими (beam) та оболонковими (shell) елементами, експериментують з функціоналом Ansys Workbench при їх імпорті та обробці; отримують практичні навички з формування сітки скінченних елементів, досліджуючи особливості розбиття моделі; проводять розрахунки на режим кручення (рис. 8); проводять розрахунки на згинання (рис. 9-10); аналізують отримані результати, встановлюють закономірності та залежності показників.

До особливості освітньо-професійної програми «Автомобільний транспорт» також відносять можливість отримати поглиблену професійну підготовку з аналізу причин виходу з ладу конструктивних елементів і деталей автомобільного транспорту та розробку прогресивних технологій їх технічного

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

обслуговування, ремонту, відновлення та підвищення зносостійкості з метою подовження ресурсу роботи.

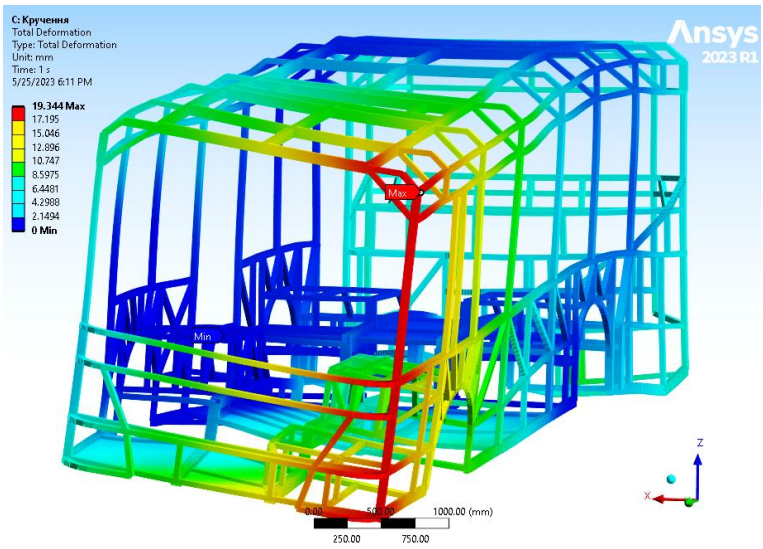


Рисунок 8 – Карта деформацій моделі каркасу кузова автобуса при крученні

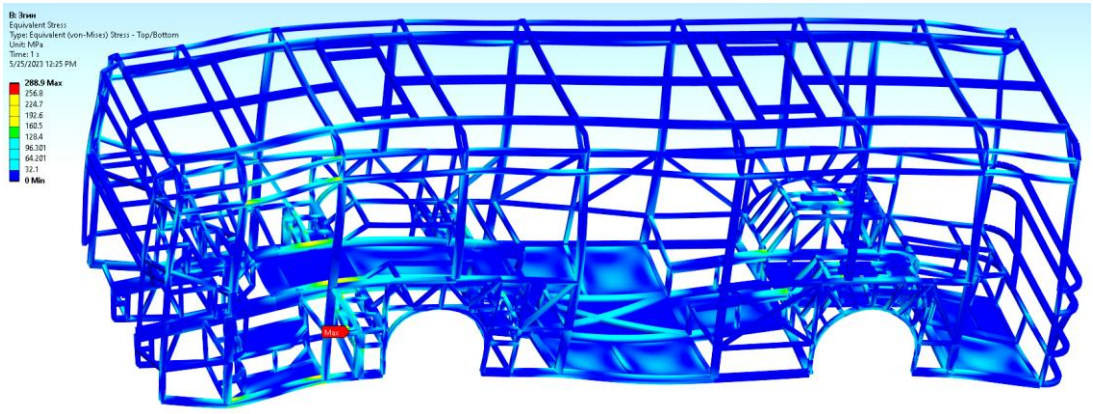


Рисунок 9 – Карта напружень каркасу кузова автобуса при згинанні

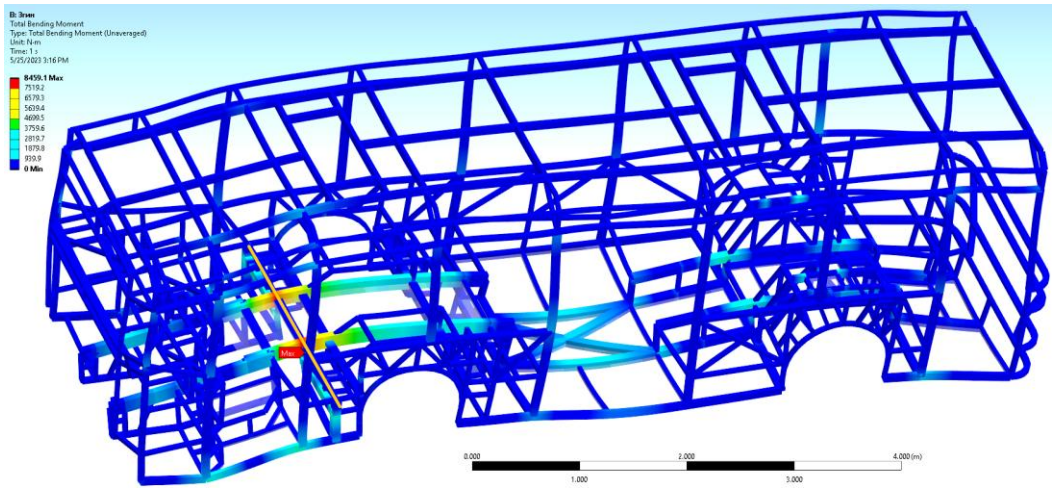


Рисунок 10 – Карта моментів згинання каркасу кузова автобуса

Отже, впровадження інформаційних технологій у вивчення розглянутої дисципліни: професіонально орієнтує слухачів; дає ґрунтовну підготовку за фахом, яка пов'язана із спадкоємністю у викладанні; активізує науково-дослідницьку діяльність, підвищуючи гарантії працевлаштування випускників.

Вивчення дисципліни і виконання індивідуального завдання сприяє поглибленню та розширенню фахових компетентностей, зокрема, таких здатностей як: складати, оформлювати й оперувати технічною документацією ТП на підприємствах АТ; застосовувати спеціалізоване програмне забезпечення для розв’язання складних задач АТ; визначати основні механізми зношування деталей та агрегатів систем автомобіля; розробляти інженерні заходи з підвищення ресурсу деталей автомобіля за критерієм зношування; проводити розрахункову та експериментальну оцінку технологічних, конструкторських та експлуатаційних заходів підвищення зносостійкост деталей автомобіля.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Диха О. В. Моделювання технологічних процесів підприємств автомобільного транспорту: методичні вказівки до лабораторних робіт для студентів напряму підготовки “Автомобільний транспорт” / О. В. Диха, О. Ю. Рудик. – Хмельницький : ХНУ, 2018. – 102 с.

2. Псьол Сергій. Застосування комп’ютерного моделювання для розрахунків автомобільного транспорту / Сергій Псьол, Олександр Диха, Олександр Рудик, Костянтин Голенко // Збірник наукових праць Національної академії Державної прикордонної служби України. Серія: педагогічні науки / гол. ред. О. В. Діденко. Хмельницький : Видавництво НАДПСУ, 2023. № 1(32). – С. 148-170.

3. Рудик О. Ю. SolidWorks як інноваційний засіб вивчення дисциплін автомобільного профілю / О. Ю. Рудик, О. В. Диха // «Системні технології» 3 (128) 2020. – С. 21-35.

4. Трасковецька Л. М. Математичне моделювання працездатності деталей методом скінченних елементів / Л. М. Трасковецька, О. Ю. Рудик, О. В. Бірюков // Інформатика, інформаційні системи та технології: тези доповідей шістнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 19 квітня 2019 р. – Одеса: ОНУ. – С. 96-97.

5. Рудик О. Ю. Дослідження можливості заміни матеріалу деталі автомобіля за допомогою SolidWorks Simulation / О. Ю. Рудик, А. А. Тарашевський // Матеріали IV Всеукраїнської науково-практичної Інтернет-конференції «Ре-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

сурсно-орієнтоване навчання в «3D»: доступність, діалог, динаміка» / укл. Н. В. Кононец, В. О. Балюк. – Полтава : КУЕП ПДАА, 2020. – С. 100-102.
6. Rudyk O. Yu. Optimization of the steering bipod of the vehicle / O. Yu. Rudyk, V. V. Gerega, N. R. Tymchenko // Scientific achievements of modern society. Abstracts of the 7th International scientific and practical conference. Cognum Publishing House. Liverpool, United Kingdom. 2020. – Pp. 187-192.
7. Рудик О. Ю. Комп'ютерне моделювання напружено-деформованого стану вал-шестерні головної передачі заднього моста автомобіля ГАЗ-53 / О. Ю. Рудик, В. В. Гончар // Матеріали II Всеукраїнської науково-практичної Інтернет-конференції «Ресурсно-орієнтоване навчання в «3D»: доступність, діалог, динаміка» / укл. Н. В. Кононец, В. О. Балюк. – Полтава : КУЕП ПДАА, 2018. – С. 158-162.
8. Psol S. V. Using SolidWorks to ensure passability of automotive equipment / S. V. Psol, Y. Leshchak, O. Yu. Rudyk // Achievements and prospects of modern scientific research. Abstracts of the 2nd International scientific and practical conference. Editorial EDULCP. – Buenos Aires, Argentina. 2021. – Pp. 140-146.
9. Rudyk O. Investigation of a universal puller of bearings with SolidWorks /O. Rudyk, P. Kaplun, V. Honchar // Інформаційні технології в металургії та машинобудуванні. ІТММ'2022: тези доповідей міжнародної науково-практичної конференції (Дніпро, 18 травня 2022 р.) / Міністерство освіти і науки України, Український державний університет науки і технологій, Дніпропетровський національний університет імені О. Гончара, Національний технічний університет «Дніпровська політехніка» та ін. – Дніпро : НМетАУ, 2022. – С. 248-251.
10. Псьол С. В. Вплив кріплень у SolidWorks Simulation на працездатність деталей / С. В. Псьол, О. Ю. Рудик, Б. В. Андрійчук // Ресурсно-орієнтоване навчання в «3D»: доступність, діалог, динаміка : збірник тез доповідей І Міжнародної науково-практичної інтернет-конференції (22–23 лютого 2021 року, м. Полтава). – Полтава : ПУЕТ, 2021. – С. 88-91.
11. Psol S. V. Impact of blow on the stability of details of wheeled machines / S. V. Psol, S. Gramenko, O. Yu. Rudyk // The world of science and innovation. Abstracts of the 6th International scientific and practical conference. Cognum Publishing House. – London, United Kingdom. 2021. – Pp. 217-224.
12. Голенко К. Е. Моделювання технологічних процесів підприємств автомобільного транспорту: Методичні рекомендації до індивідуального завдання «Моделювання напружено-деформованого стану каркасних елементів»

для здобувачів першого (бакалаврського) рівня вищої освіти спеціальності 274 «Автомобільний транспорт» / К. Е. Голенко, О. Ю. Рудик. – Хмельницький : ХНУ, 2023. – 82 с.

REFERENCES

1. Dykha O. V. Modeling of technological processes of road transport enterprises: methodical instructions for laboratory work for students of the field of preparation for road transport / O. V Dykha, O. Yu. Rudyk // Khmelnitsky : KhNU, 2018. – 102 p.

2. Serhiy Psol. Application of computer modeling for road transport calculations / Serhiy Psol, Oleksandr Dykha, Oleksandr Rudyk, Kostyantyn Golenko // Collection of scientific works of the National Academy of the State Border Service of Ukraine. Series: pedagogical sciences / head. ed. O. V. Didenko. Khmelnytskyi : NADPSU Publishing House, 2023. No. 1(32). – P. 148-170.

3. Rudyk O. Yu. SolidWorks as an innovative means for studying the disciplines of automobile profile / O. Yu. Rudyk, O. V. Dykha // «System technologies» 3 (128) 2020.

4. Traskovetska L. M. Mathematical modeling of the performance of parts by the method of finite elements / L. M. Traskovetska, O. Yu. Rudyk, O. V. Biryukov // Informatics, information systems and technologies: abstracts of reports of the sixteenth all-Ukrainian conference of students and young scientists. Odesa, April 19, 2019. – Odesa: ONU. – P. 96-97.

5. Rudyk O. Yu. Study of the possibility of replacing the material of a car part using SolidWorks Simulation / O. Yu. Rudyk, A. A. Tarashevskyi // Materials of the 4th All-Ukrainian scientific and practical Internet conference "Resource-oriented learning in "3D": availability, dialogue, dynamics" / incl. N. V. Kononets, V. O. Balyuk. – Poltava : KUEP PDAA, 2020. – P. 100-102.

6. Rudyk O. Yu. Optimization of the steering bipod of the vehicle / O. Yu. Rudyk, V. V. Gerega, N. R. Tymchenko // Scientific achievements of modern society. Abstracts of the 7th International scientific and practical conference. Cognum Publishing House. Liverpool, United Kingdom. 2020. – Pp. 187-192.

7. Rudyk O. Yu. Computer modeling of the stress-strain state of the gear shaft of the main drive of the rear axle of the GAZ-53 car / O. Yu. Rudyk, V. V. Gonchar // Proceedings of the II All-Ukrainian Scientific and Practical Internet Conference "Resource- oriented learning in "3D": accessibility, dialogue, dynamics" / incl. N. V. Kononets, V. O. Balyuk. – Poltava : KUEP PDAA, 2018. – P. 158-162.

8. Psol S. V. Using SolidWorks to ensure passability of automotive equipment / S. V. Psol, Y. Leshchak, O. Yu. Rudyk // Achievements and prospects of modern sci-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

entific research. Abstracts of the 2nd International scientific and practical conference. Editorial EDULCP. – Buenos Aires, Argentina. 2021. – Pp. 140-146.
9. Rudyk O. Investigation of a universal puller of bearings with SolidWorks /O. Rudyk, P. Kaplun, V. Honchar // Information technologies in metallurgy and mechanical engineering. ITMM'2022: abstracts of reports of the international scientific and practical conference (Dnipro, May 18, 2022) / Ministry of Education and Science of Ukraine, Ukrainian State University of Science and Technology, Dnipropetrovsk National University named after O. Honchar, National Technical University "Dnipro Polytechnic" and others – Dnipro : NMetAU, 2022. - P. 248-251.
10. Psyol S. V. The influence of fasteners in SolidWorks Simulation on the performance of parts / S. V. Psyol, O. Yu. Rudyk, B. V. Andriychuk // Resource-oriented learning in "3D": accessibility, dialogue, dynamics: collection of theses reports of the 1st International Scientific and Practical Internet Conference (February 22-23, 2021, Poltava). – Poltava : PUET, 2021. – P. 88-91.
11. Psol S. V. Impact of blow on the stability of details of wheeled machines / S. V. Psol, S. Gramenko, O. Yu. Rudyk // The world of science and innovation. Abstracts of the 6th International scientific and practical conference. Cognum Publishing House. – London, United Kingdom. 2021. – Pp. 217-224.
12. Golenko K. E. Modeling of technological processes of road transport enterprises: methodological recommendations for the individual task "Simulation of the stress-strain state of frame elements" for students of the first (bachelor's) level of higher education, specialty 274 "Automotive transport" / K. E. Golenko, O. Yu. Rudyk. – Khmelnytskyi : KhNU, 2023. – 82 p.

Received 12.03.2024.
Accepted 19.03.2024.

Innovative approaches in teaching automotive disciplines

For the educational discipline "Modeling of technological processes of road transport enterprises" in the SolidWorks Simulation and Ansys Workbench environments, the basic principles and provisions of automated design in the field of computer modeling of units, assemblies and parts of vehicles, as well as devices for their repair (lifts, jacks, stands, puller etc.). Complemented following program results: install specialized software, information and information-communication technologies to track object models and vehicle processes on automobile transport (AT), operational authorities of AT functions, construction engineering and technical and economic developments, development of design documentation and development of other AT tasks; find necessary information in scientific and technical literature, databases and other sources; analyze and evaluate this information; make effective decisions, analyze and compare alternative op-

tions taking into account goals and constraints, quality assurance issues, as well as technical, economic, legislative and other aspects; analyze the information obtained as a result of research, generalize, systematize and use it in professional activities; develop and implement technological processes, technological equipment and technological equipment, means of automation and mechanization in the process of operation, repair and maintenance of JSC facilities, their systems and elements; to analyze the technical-operational and technical-economic indicators of AS means, their systems and elements; apply mathematical and statistical methods for building and researching models of objects and processes of AT, calculating their characteristics, forecasting and solving other complex tasks of AT; to present the results of research and professional activities, to argue one's position. The main attention is paid to the theory and practical use of finite element methods and the acquisition of skills in the design and calculations of AT details. Mandatory elements of research in SolidWorks and practical skills of modeling various load modes of road and special vehicles in Ansys Workbench are defined. In order to extend the service life of structural elements and parts of AT, methods of their restoration and increase in wear resistance are defined.

Рудик Олександр Юхимович – доцент, кандидат технічних наук, кафедра трибології, автомобілів та матеріалознавства, Хмельницький національний університет.

Диха Олександр Володимирович – професор, доктор технічних наук, завідувач кафедри трибології, автомобілів та матеріалознавства, Хмельницький національний університет.

Голенко Костянтин Едуардович – викладач, кандидат технічних наук, кафедра трибології, автомобілів та матеріалознавства, Хмельницький національний університет.

Rudyk Oleksandr Yuhymovych – associate professor, candidate of technical sciences, department of tribology, automobiles and materials science, Khmelnitsky national university.

Dykha Oleksandr Volodymyrovych – professor, doctor of technical sciences, head of the department of tribology, automobiles and materials science, Khmelnitsky national university.

Golenko Konstiantyn Eduardovych – teacher, candidate of technical sciences, department of tribology, automobiles and materials science, Khmelnitsky national university.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-2-151-2024-14
УДК 004.6: 007.5: 004.94

В.В. Скалозуб, В.М. Горячкін, І.В. Клименко, І.А. Терлецький

ДОСЛІДЖЕННЯ ІНТЕЛЕКТУАЛЬНИХ МОДЕЛЕЙ УПРАВЛІННЯ НА ОСНОВІ ПРОЦЕДУР КЛАСИФІКАЦІЇ НЕВИЗНАЧЕНИХ ДАНИХ ЗІ ВСТАНОВЛЕНИМИ ВИМОГАМИ ДОСТОВІРНОСТІ РЕЗУЛЬТАТІВ

Анотація. Стаття присвячена дослідженням властивостей і розвитку інтелектуальних моделей управління складними системами за умов невизначеності даних на основі процедур класифікації на основі методів редукції та статистики каппа Коена. Застосування цих методів забезпечує достовірне вирішення завдань з урахуванням оцінки граничної розмірності моделей класифікації. В роботі були досліджені можливості удосконалення нейронних мереж Хеммінга для класифікації даних у форматах нечітких величин і certainty factor CF(A). Також були визначені особливості удосконаленої математичної моделі завдань нечіткої класифікації на основі набору шаблонів ознак. Представлено структуру програмного комплексу інформаційної технології управління призначенням/відбором виконавців на основі класифікації наборів шаблонів із певних нечітких ознак, який використовує процедури редукції і каппа статистики.
Ключові слова: класифікація наборів шаблонів, невизначені дані, редукція розмірності, статистика каппа Коена, нечіткі величини, certainty factor CF(A), модифікована мережа Хеммінга, інформаційна технологія, програмне забезпечення, завдання визначення авторства та призначення/відбір виконавця.

Вступ та постановка проблеми. Натепер засобами широкого кола технологій класифікації та діагностування неповно визначених даних щодо станів та умов функціонування складних систем вирішуються завдання вибору варіантів керування різноманітними технологічними процесами, відбору виконавців встановлених завдань тощо [1, 5, 6, 9]. При цьому актуальними залишаються завдання формування адекватних математичних моделей процедур класифікації та встановлення їх коректності, повноти та достовірності результатів. У статті [1] отримано розвиток моделей класифікації в умовах невизначеності даних з використанням процедур редукції та статистики «каппа Коена» [4, 8]. В ній розглянуті завдання достовірності результатів класифікації. А саме, забезпечення узгодженості ймовірнісних вимог щодо достовірності наступних результатів, розмірності параметрів моделей

(шаблонів), а також кількості даних (класів) за якими виконується діагностування. Для реалізації процедур класифікації запропоновані модифіковані мережі Хеммінга (МХН) [1, 2, 10], а для забезпечення встановлених вимог достовірності результатів – методи редукції, пристосовані для завдань класифікації за рахунок використання статистики каппа Коена. Коректність, достовірність та ефективність моделей і процедур класифікації при невизначених даних [1, 2] визначалась на основі результатів числового моделювання. Разом з тим потребують подальшого дослідження завдання щодо уточнення і розширення метрик нечіткої подібності шаблонів класифікації, а також застосування метрик формату CF(A)ϵ[-1; 1] (біполярні коефіцієнти, certainty factor), формування інтелектуальних засобів для управління на основі класифікації невизначених даних складних систем і процесів тощо. Таким завданням присвячена представлена стаття.

Аналіз останніх досліджень і публікацій. У статтях [1, 2, 5, 9] і багатьох інших відзначаються можливості значної ступені існування джерел і умов невизначеності щодо характеристик та станів сучасних складних систем, наявності різноманітність метрик невизначеності, повноти отримання даних тощо. Серед інформаційної технології (ІТ), які використовують моделі класифікації та діагностування, відзначимо процеси оптимізації потоків замовлень у сервісних системах (С&С) [2, 6], розформування-формування вантажних залізничних составів та керування певними процесами [3, 9].

Розвитку моделей та багатопараметричних інтелектуальних процедур класифікації та діагностування за неповними і збуреними даними присвячена стаття [1]. В ній подано нові постановки, математичні моделі та реалізації завдань класифікації за недетермінованими даними. Одним із прикладів моделі класифікації за нечіткими даними являється окреме завдання, в якому на основі нечіткої класифікації встановлюються автори україномовних текстів. Також на основі моделей класифікації даних у форматі коефіцієнтів упевненості CF(A) [5, 6] виконується оптимальний відбір кандидата на певну виробничу діяльність (ВД). При тому модель ознак кандидатів на ВД має удосконалену багатовекторну структуру. У всіх зазначених процедурах передбачене застосування методу редукції простору моделі [3] за рахунок дослідження статистики каппа Коена [4, 8], що гарантує отримання результатів із встановленою достовірністю результатів вибору.

Завдання нечіткої класифікації (також і класифікації за ознаками метрики CF(A)) реалізується шляхом застосування модифікованої моделі асоціативної

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

пам'яті Хеммінга (МХН). В МХН величини відстані між нечіткими елементами (μX) зразків моделі класифікації (w_{ik}) і вхідним вектором X = {xi: i=0...n-1} визначаються на основі нечіткого відношення [9] виду

$$R(W, X) = \{\mu R(w_i, x_i)/\{w_i, x_i\}=(1 - abs(w_i-x_i))/\{w_i, x_i\}\}, i=0, 1, ..., n-1. \quad (1)$$

Відношення (1) дозволяє вимірювати відстані між нечіткими векторами X = {x_i: i=0...n-1} та відповідно моделлю шаблонів класифікації (w_{ik}). Зауважимо, що при однакових значеннях ступенів приналежності величин {w_i, x_i} значення $\mu R(w_i, x_i)$ =1. Якщо одна з величин {w_i, x_i} дорівнює 0, а інша 1, тоді значення $\mu R(w_i, x_i)$ = 0, інакше $\mu R(w_i, x_i) \rightarrow$ [0; 1]. Відношення (1) дозволяє модифікувати МХН для нечітких моделей класифікації, а також моделей метрики CF(A) [5].

Проблематика щодо визначення розмірності параметрів математичних моделей із застосуванням процедур індукції і редукції представлена у [3]. В якості процедури, за якою виконується формування простору математичних моделей, призначених для завдань представлення «закономірностей рівності» ($y=f(X)$) , використовують α-процедуру. Відзначимо, що завдання класифікації відноситься до визначення «закономірностей схожості», для реалізації яких і була запропонована спеціалізована процедура на основі використання статистики каппа Коена [1]. При тому оцінки граничної розмірності параметрів моделі класифікації, що синтезується «n_0», визначаються відповідно до рівнянь

$$n_0 = (\varepsilon * L+\ln(h))/\ln(m). \quad (2)$$

Відповідність даних рівнянням (2) визначає умову щодо виконання заданих вимог достовірності «ε, h»; (1 – h) оцінка ймовірності безпомилкового розділення випадкової і незалежної вибірки довжини «L» при заданій граничній величині помилкової класифікації «ε» [3].

Призначення статистики «каппа Коена» [4, 8] (у пропонованій процедурі редукції розмірності математичної моделі класифікації) полягає у встановленні «подібності», або належності конкуруючих багатопараметричних математичних моделей до одного класу, щодо відповідності вимогам шаблонів, сформованих із ознак властивостей контрольованих процесів. Для такої перевірки «подібності» необхідно утворити певну множину математичних моделей, шаблонів із різними наборами змінних, для яких в подальшому виконувати порівняльний аналіз їх відповідності між собою. Наприклад, в [1] за оцінками каппа статистики перевірялась «подібність» результатів класифікації для всіх пар змінних шаблонів (X/X_j) та (X/X_i). При цьому за каппа-оцінками наборам змінних (X/X_j) для кожного шаблону приписувалися значення «+» або

«-» в залежності від результатів класифікації. У табл. 1 приведено структуру моделі класифікації та її умовні результати для шаблонів з шести змінних. Далі за порівняльним аналізом результатів класифікації наборів змінних формуються таблиці розбіжностей та розраховують статистичні оцінки «каппа Коена» [4]:

$$K = (P_0 - P_e) / (1 - P_e), \quad (3)$$

де P_0 – імовірнісна оцінка що показує наскільки спостережувана узгодженість краща за випадкову, а P_e – результат підрахунку максимально можливої узгодженості за винятком випадкової узгодженості конкуруючих шаблонів.

Таблиця 1

Класифікації шаблонів з параметрів (X1, X2, …, X6).

K_i	L_i	X	$X/\{X_1\}$	$X/\{X_2\}$	$X/\{X_3\}$	$X/\{X_4\}$	$X/\{X_5\}$	$X/\{X_6\}$
K_1	L_{11}		+					+
	L_{12}	+		+		+		
	L_{13}					+		+
K_2	L_{21}					+		+
	L_{22}		+	+			+	
	L_{23}				+			
	L_{24}	+				+		
K_3	L_{31}		+	+				
	L_{32}	+					+	
K_4	L_{41}							
	L_{42}		+		+	+		
	L_{43}							
	L_{44}	+		+		+		+
	L_{45}					+		
K_P	L_{P1}	+				+		
	L_{P2}	+	+	+				

У таблиці позначено: K_i – класи/зразки моделі (можливі кілька екземплярів L_i); X – параметри моделі на поточному етапі редукції; $X/\{X_j\}$ – скорочені набори моделей класифікації (без множини параметрів $\{X_j\}$); знаки «+» – визначення шаблонів переможців (класів) за моделлю МХН. У стовпці X позначаються шаблони-переможці для X наборів параметрів. Множина варіантів моделей класифікації (КМ) містить всі комбінації пар скорочених мо-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

делей X/{X_j}. Завдання полягало у визначенні серед скорочених моделей шаблонів X/{X_j} такого набору параметрів {X_j}, які необхідно видалити на наступному етапі процедури редукції. Разом з тим можливості формування множин {X_j} у процедурах редукції залишилися не розкритими.

Наведені вище результати свідчать про те, що формування коректних і ефективних інтелектуальних засобів технологій і систем управління на основі класифікації невизначених даних являється актуальним. Також актуальними для цього є дослідження можливостей метрик нечіткої класифікації та метрик формату CF(A), як і розробка автоматизованих засобів реалізації процедур редукції.

Мета дослідження – розвиток інтелектуальних моделей управління складними системами за умов невизначеності даних на основі процедур класифікації, а також шляхом формування програмних засобів, які забезпечують достовірне вирішення завдань класифікації з урахуванням рівнянь (2) за даними метрик нечітких величин і коефіцієнтів впевненості CF(A).

Результати та основний матеріал дослідження. Серед завдань цієї статті визначено дослідження можливостей метрик нечіткої класифікації та формату certainty factor CF(A) щодо їх застосування для модифікування мережі Хеммінга (МХ) із областю значень векторів ознак {– 1; + 1} для їх переходу до мереж МХН із областями нечітких значень μ_X (X → [0; 1]), а також показників CF(A) із значеннями множини [–1; + 1]. В якості обґрунтування запропонованого нами рішення щодо МХН (1) розглянемо основні типи нечітких метрик Хеммінга відповідно [2], а також їх співвідношення з мережею МХ з ознаками {– 1; + 1}. В статті [2] метрика нечіткої відстані Хеммінга (з додатним відхиленням при введенні α_r) між нечіткими множинами $A\sim$ і $B\sim$, визначається за формулою:

$$d^+(A\sim, B\sim) = \sum_{r=1}^{q} \alpha_r . |\mu A\sim(u_r) - \mu B\sim(u_r)| \qquad (4)$$

де $u_r \in U$, $\mu A\sim(u_r)$, $\mu B\sim(u_r) \in [0,1]$, $r = \underline{1,q}$, коли α_r – індикатор:

$$\alpha_r = \{1, \text{якщо } \mu A\sim(u_r) \geq \mu B\sim(u_r)\ 0, \text{якщо } \mu A\sim(u_r) < \mu B\sim(u_r) \qquad (5)$$

Зауважимо, що метрика (4) з індикатором (5) дозволяє також розрізняти вимоги щодо переваги показників $\mu A\sim(u_r)$, $\mu B\sim(u_r) \in [0,1]$, тобто, вводити певні вимоги/умови стосовно оцінок порівнюваних нечітких векторів $A\sim$ і $B\sim$.

Метрика (4) не відповідає векторам ознак мережі Хеммінга МХ з областю значень {– 1; + 1}, тому в [6] була запропонована процедура перекодування показників μX: X → [0; 1] до множини значень моделі МХ. При цьому відзначається принципова потенціальна можливість представлення однаковими кодами певних різних нечітких величин. Також відзначається необхідність поєднувати кодування з метою застосування моделі МХН, з процедурою класифікації зразків. Числові розрахунки на основі МХ показали існування таких же питань щодо реалізації завдань класифікації даних формату certainty factor CF(A) на основі використання подібних (4) метрик.

Дослідження процедур МХН на основі (1) підтвердили коректність і ефективність його застосування для модифікування МХН, як для нечітких метрик моделей класифікації, так і для моделей метрики CF(A). При тому також встановлені можливості визначення додаткових вимоги/умови щодо показників близькості порівнюваних нечітких векторів $A \sim$ і $B \sim$. Простий змістовний приклад таких додаткових умов полягає в наступному. Нехай існують дві пари векторів нечітких оцінок певних ознак системи, які мають наступні значення $(\mu A \sim (u_r) = 0.2,\ \mu\text{A} \sim (u_\text{к}) = 0.8)$ та $(\mu\text{B} \sim (u_r)0.3,\ \mu B \sim (u_\text{к}) = 0.9)$). Тут абсолютні відстані між u_r та $u_\text{к}$ дорівнюють 0.1, але відносні – $(\mu A \sim (u_r) - \ \mu\text{B} \sim (u_r))/(\mu A \sim (u_r); \mu\text{B} \sim (u_r))$ суттєво різні для u_r та $u_\text{к}$. Для забезпечення можливості встановлення нових вимог до метрики відстані відношення виду (1) треба замінити наступним

$$R(W, X)=\{\mu R(w_i, x_i)/\{w_i, x_i\}=(1-abs(w_i-x_i))/max\{abs(w_i);abs(x_i)\}\},\ i=0, 1, \ldots, n-1 \quad (6)$$

Як і (1), відношення (6) придатне для нечітких моделей класифікації, а також моделей типу certainty factor CF(A).

Приведемо деякі результати застосування модифікованих моделей класифікації МХН для зазначених типів даних. На рис. 1 подані приклади завдань

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

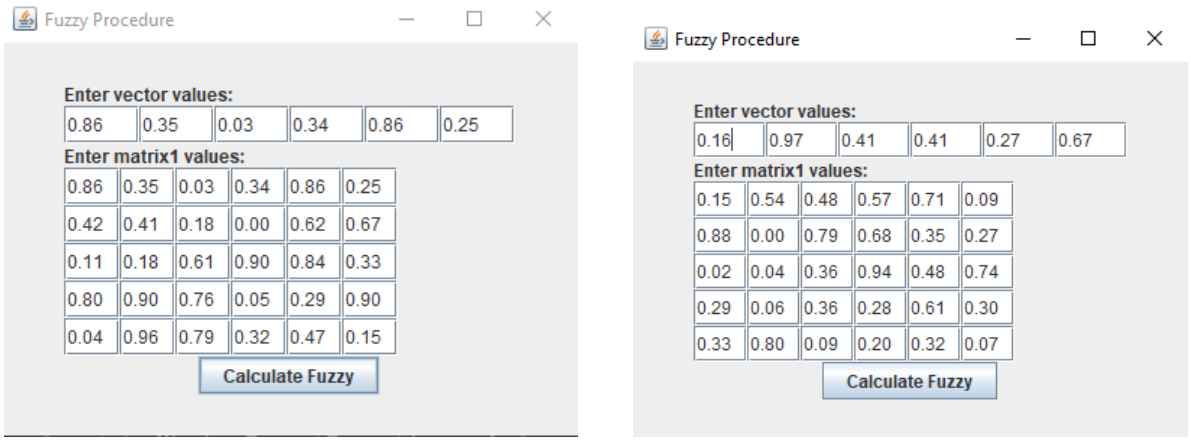


Рисунок 1 - Постановки завдань нечіткої класифікації

нечіткої класифікації, а на рис. 2 їх реалізація. В них, рис. 1 ліворуч, вхідний вектор відповідає шаблону (перший рядок матриці, індекс 1), а праворуч приведена реалізація моделі досить невизначеної нечіткої структури. Перше завдання було реалізоване за одну ітерацію, як у стандартній МХ, а друге – через кілька кроків процедури (індекс 5). При тому показник «value», що означає оцінку ступеня порівняльної достовірності результату класифікації, у другому завданні був досить близький до нуля (границі припинення відбору шаблонів МХ). На рис. 3 приведено приклад реалізації завдання нечіткої класифікації за МХН, з високим ступенем достовірності (показник «value»).

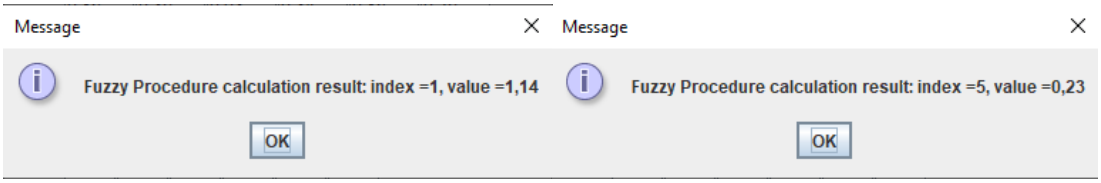


Рисунок 2 - Результати реалізації завдань нечіткої класифікації рис. 1

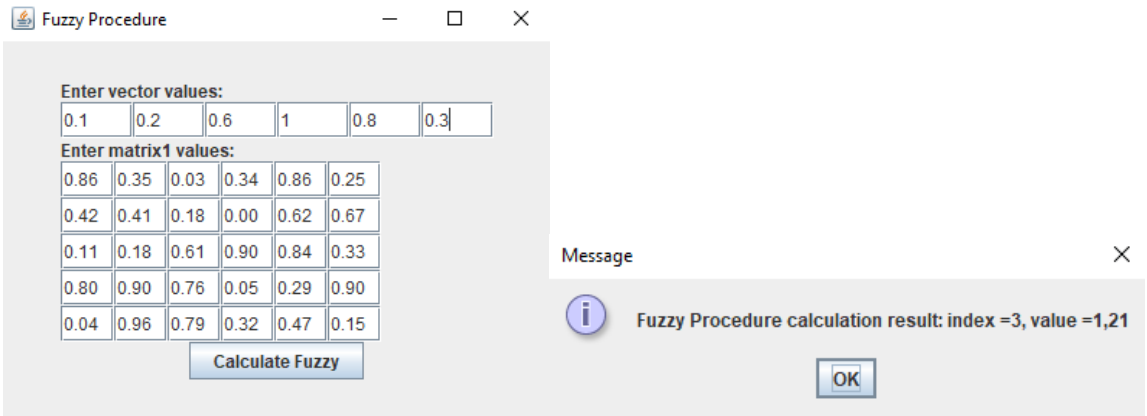


Рисунок 3 - Результати реалізації нечіткої класифікації за МНХ для шаблону «3»

На рис. 4 та рис. 5 подано подібні результати класифікації за даними моделей типу certainty factor CF(A). Рис. 4 демонструє результат, коли вхідний вектор відповідає шаблону «5», а на рис. 5 досліджується довільний вектор ознак.

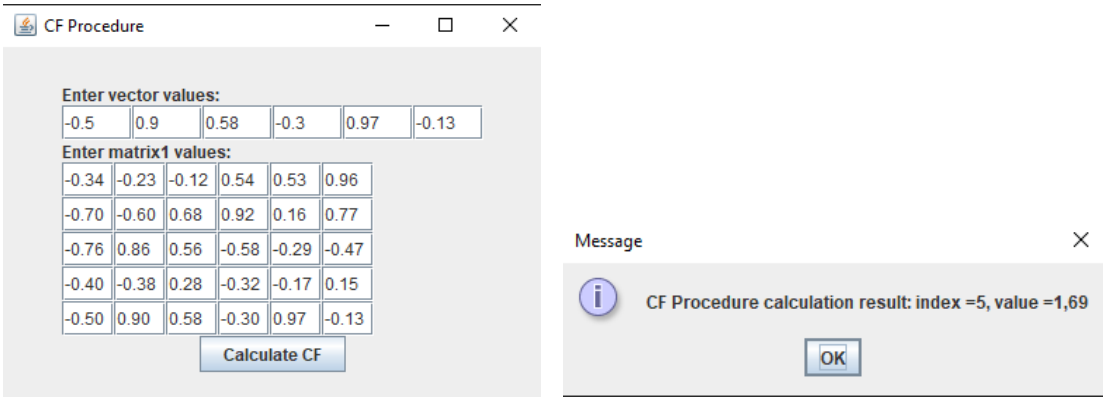


Рисунок 4 - Результати реалізації CF(A) класифікації за МНХ шаблону «5»

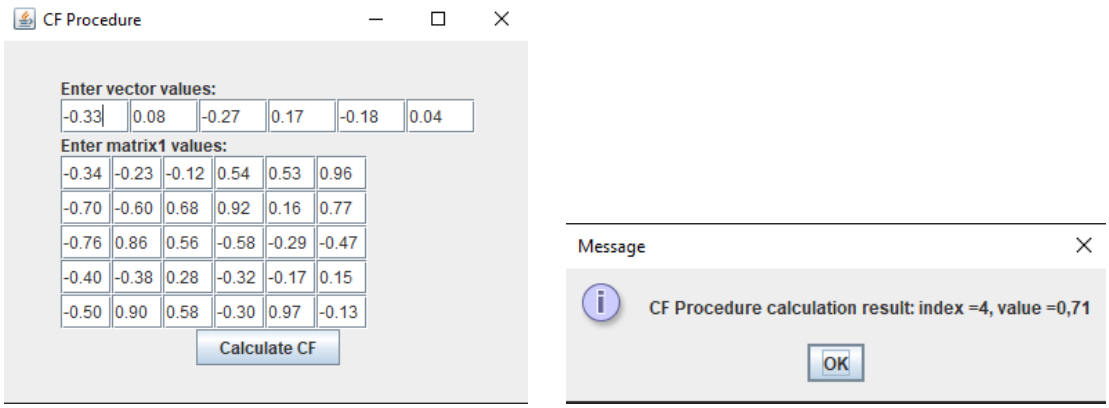


Рисунок 5 -Результати реалізації моделей завдань CF(A) класифікації

Приведемо сутність та результати виконання процедури редукції на основі каппа статистики ПКР (3), призначеної для формування достовірних математичних моделей класифікації за МХН, яка має такі етапи:

- 1) на основі моделей шаблонів з різним числом або складом параметрів визначити оцінку показника «каппа» (3),

- 2) якщо результати класифікації для різних моделей-шаблонів відповідають вимогам «подібності», скоротити модель, залишити один із шаблонів,

- 3) враховуючи значення «n_0» (2), визначити і видалити найменше значимі чи найбільше «подібні» між собою параметри моделі класифікації.

Для пояснення сутності процедури редукції, запропонованій у [1, 6], використаємо умовний приклад табл. 2 зі структурою моделі класифікації

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

шаблонів з параметрами (X_1, X_2, …, X_6) табл. 1. У табл.2 приведені значення (3), отримані при співставленні узгодженості приведених у таблиці пар змінних.

Таблиця 2

Розрахунки при m=5 за зразками з урахуванням відомих даних X ($КМ_2$, C_1)

К (X_2/X_3) 0,150		Так	ні
	так	12	5
	ні	5	4

К (X_2/X_4) 0,103		Так	ні
	так	13	4
	ні	6	3

К (X_2/X_5) 0,150		Так	ні
	так	12	5
	ні	5	4

К (X_2/X_6) 0,023		Так	ні
	так	8	9
	ні	4	5

К (X_3/X_4) 0,103		Так	ні
	так	13	4
	ні	6	3

К (X_3/X_5) 0,320		Так	ні
	так	13	4
	ні	4	5

К (X_3/X_6) 0,324		Так	ні
	так	10	7
	ні	2	7

К (X_4/X_5) 0,283		Так	ні
	так	14	5
	ні	3	4

К (X_4/X_6) 0,331		Так	ні
	так	11	8
	ні	1	6

К (X_5/X_6) 0,173		Так	ні
	так	9	8
	ні	3	6

Найвища в табл. 2 оцінка коефіцієнту «каппа» (3) у пари параметрів К(X_4/X_6) = 0,331 (помірний рівень узгодженості), а у пар К(X_3/X_6) = 0,324 і К(X_5/X_6) = 0,173. При цьому видаляється параметр X_6 через те, що «вплив» фактору X_6 на модель класифікації враховують та опосередковано представляють фактори X_5, X_3 і X_4. Розрахунки кроку процедури аналізу математичної моделі класифікації за даними табл. 2 демонструють зміст ПКР формування моделей класифікації на основі редукції та каппа статистики.

Розглянемо особливість математичної моделі класифікації табл. 1, з урахуванням вимог завдання із визначення авторства україномовних творів (ЗАТ), приведеного у [6, 7]. У нашому варіанті завдання ЗАТ на підставі сукупності ознак, які характеризують набори текстів і утворюють класи або зразки моделі (представляють особистий авторський профіль), необхідно визначити клас, до якого належить новий текст (представлений закодованим вектором ознак). У

завданні система ознак у всіх авторів однакова, а число наборів (творів авторів) може бути різним, не досить значним. При тому певні дані профілю можуть бути «збуреними» або навіть відсутніми. Через те, що окремі твори можуть відноситися до різних періодів, жанрів тощо, формування одного окремого шаблону для кожного із авторів представляє певну проблему. Тому в наведеній моделі класифікації текстів кожному автору відповідає кілька зразків/шаблонів у формі векторів нечітких величин (значень функцій приналежності), що демонструє табл.1. В ній класи K_i – це набори закодованих ознак шаблонів для окремих творів одного автора, серед яких також може бути і шаблон узагальнених показників, «профілю». За результатами досліджень було встановлено [6, 7], що для вибору найбільш ефективних і точних шаблонів моделей класифікації авторів україномовних текстів необхідні додаткові дослідження. В них в першу чергу необхідно визначити раціональні розмірності простору класифікації (число ознак текстів). У числових експериментах з аналізу шаблонів на існування надлишкових характеристик були встановлені можливості скоротити набори ознак X_1 – X_{14}, виконати процедуру редукції [1, 3]. Наприклад, для певних вхідних творів були визначені спрощені шаблони із 4-х параметрів, які за результатами класифікації відповідали сукупності параметрів-ознак X_1 – X_{14}. Дослідження показали необхідність формування шаблонів моделей завдань ЗАТ з урахуванням вимог методу спрощень [3].

Важливість врахування розмірності моделей класифікації показують оцінки граничної розмірності (3), визначені для умов роботи [1]. Так для умов (L=70, eps=0.2, h=0,1) отримано (m= 14, n_0 = 4,43)/(m= 7, n_0 = 6)/(m= 6, n_0 = 6,53). Далі при (L=200, eps=0,2, h=0,2) отримано (m= 14, n_0 = 14,6)/(m= 60, n_0 = 9,4); при eps=0,1 граничне значення лише n_0 = 4,5, а для забезпечення розмірності n_0 = 58 потрібно мати L =1200. Розрахунки показують, що підвищення достовірності результату забезпечується в основному обсягом вибірки даних (числом зразків моделі). Такий висновок визначає форми моделей ЗАТ, в яких безпосередньо використовуються закодовані тексти, без узагальнення даних в шаблонах.

Відзначимо особливості завдання ЗАТ і його реалізації на основі нашої моделі нечіткої класифікації. А саме, в ЗАТ немає вимог щодо числа етапів процедури із визначення автора твору. Разом з тим у [7] та подібних такі завдання передбачалося вирішувати за один розрахунок. Також відсутня необхідність формування єдиної моделі класифікації завдань ЗАТ при любих можливих вхідних творах, а також потреба перетворення моделей шаблонів при введення до моделі нових даних, творів. Зазначені особливості процедур

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

класифікації враховані у процедурах редукції та каппа Коена, приведених у статті.

На основі зазначених вище моделей класифікації і редукції була розроблена програма для інтелектуальної інформаційної технології (ІІТ) управління призначенням/відбором кандидатів на основі наборів їх певних ознак, рис. 6.

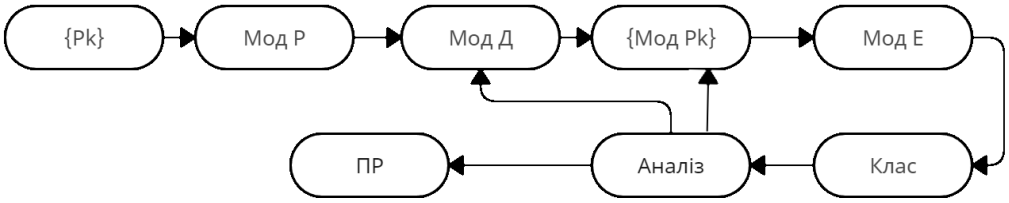


Рисунок 6 - Схема ІІТ управління процесами призначення кандидатів

Блоки структури ІТ визначають наступні процедури: - $\{P_k\}$ – отримання даних портфоліо/резюме кандидатів; - МодР – формування загальної моделі даних портфоліо/резюме конкурсу кандидатів; - МодД – вибір типу моделі представлення даних МодР (нечіткі величини/коефіцієнти впевненості); - $\{Мод\ P_k\}$ – формування моделей кожного кандидату у форматі МодД; - МодЕ – представлення еталону вимог у форматі МодД; - Клас – реалізація завдання класифікації на моделях $\{Мод\ P_k\}$, результат P^*_k; - Аналіз – визначення ступеня однозначності та достовірності результатів класифікації, прийняття результату / визначення необхідності коригування моделей кандидатів $\{Мод\ P_k\}$/ або коригування загальної моделі даних портфоліо/резюме МодР.

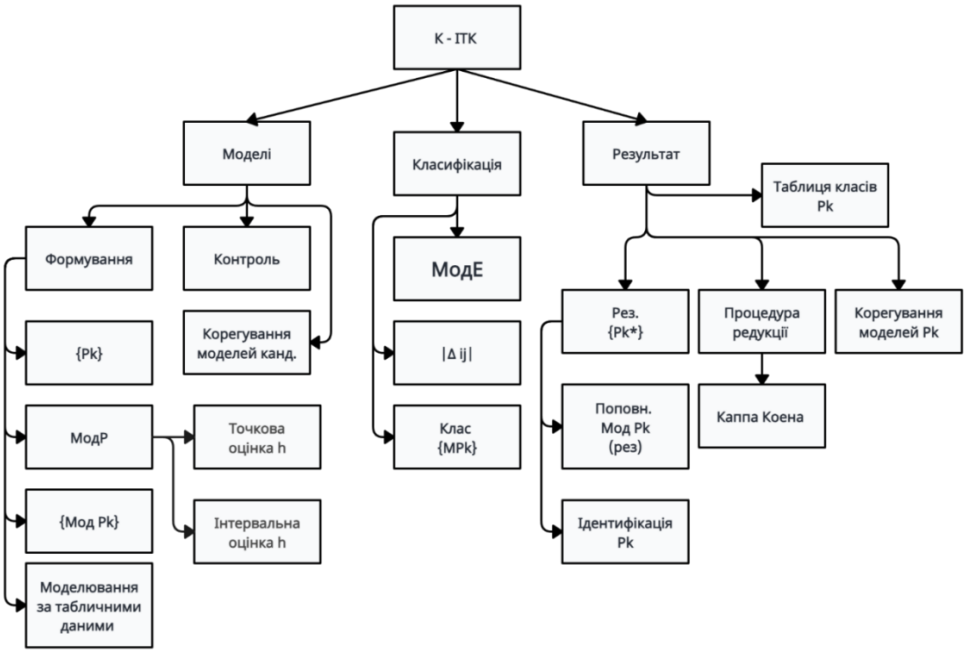


Рисунок 7 - Структура ІІТ програмного комплексу вибору кандидатів

Структурні елементи рис. 7 такі. Б1 Моделі – Формування, контроль, коригування моделей кандидатів {Мод P_k}; Б2 Класифікація – процедури визначення кандидата-переможця на основі модифікованої мережі Хеммінга; Б3 Результат – представлення результатів вибору кандидата-переможця; Б1.1 Формування - {P_k}, МодР, {Мод P_k}; Б1.2 Контроль; Б1.3 Коригування моделей кандидатів; Б2.1 МодЕ – представлення еталону вимог у форматі МодД; Б2.2 Моделі класифікації; Б2.3 Клас – реалізація завдання класифікації на моделях {Мод P_k}, результат P^*_k; Б3.1 Таблиця класів P_k; Б3.2 Результат P_k^*; Б3.3 Коригування Мод P_k; Б1.1.1 {P_k} – отримання даних портфоліо/резюме кандидатів; Б1.1.2 МодР – формування загальної моделі даних портфоліо/резюме конкурсу кандидатів; Б1.1.3 {Мод P_k} – формування моделей кожного кандидату у форматі МодД; Б3.2.1 Поповнення Мод P_k та вивід результатів; Модель і процедура отримання експертних оцінок показників моделей кандидатів {Мод P_k}. В якості вихідних даних до процедур класифікації на основі МХН надходять закодовані у формі нечітких величин (НВ) або коефіцієнтів впевненості CF(A) вектори ознак моделі класифікації. Такі вектори представляють властивості «шаблонів» {Мод Pk}, варіантів класифікації, а також векторів вимог – «Е-еталон». Завдання полягає у визначенні «найближчого» до вектору «Е-еталон» шаблону серед всіх {Мод Pk}. У роботі використано дві моделі та процедури отримання експертних оцінок показників моделей кандидатів {Мод Pk}. За ними кожному показнику загальної моделі даних портфоліо/резюме МодР призначається (експертом/менеджером) оцінка показника кандидата {P_k} у формі нечітких величин або коефіцієнтів впевненості CF(A). У першій моделі оцінювання МО1 значення показника вводиться безпосередньо, враховуючи область можливих значень НВ і CF(A).

У другій моделі оцінювання МО2 значення показника визначається на основі окремої нечіткої моделі «Показник» (МПок), призначеної для формування значень на основі опитування (експерта/менеджера).

Розроблений програмний комплекс забезпечує автоматизацію завдань інформаційної технології (ІІТ) управління призначенням/відбором кандидатів на основі наборів їх певних нечітких ознак, а також коефіцієнтів CF(A) при заданих ймовірнісних показниках достовірності результатів вибору. Зрозуміло, що комплекс рис. 7 реалізує також завдання ЗАТ, якщо дані портфоліо/резюме представляють закодовані україномовні тексти. На рис. 8 зображено діаграму варіантів використання програмного комплексу рис. 7. На діаграмі представлені наступні функції програми: Ф1 – процедури отриманні вихідних

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

даних; Ф2 – функції збереження даних та результатів досліджень; Ф3 – відображення графіків моделей процесів; Ф4 – формування шаблонів процесів; Ф5 - побудова класів шаблонів для різних типів невизначеності; Ф6 – формування нечітких моделей процедур Хеммінга; Ф7 – налаштування моделей класифікації невизначених даних; Ф8 - аналіз моделей класифікації даних; Ф9 – CF(X) процедури Хеммінга; Ф10- Оцінювання класів шаблонів; Ф11 – Відображення результатів; Ф12 – процедури порівняльного аналізу.

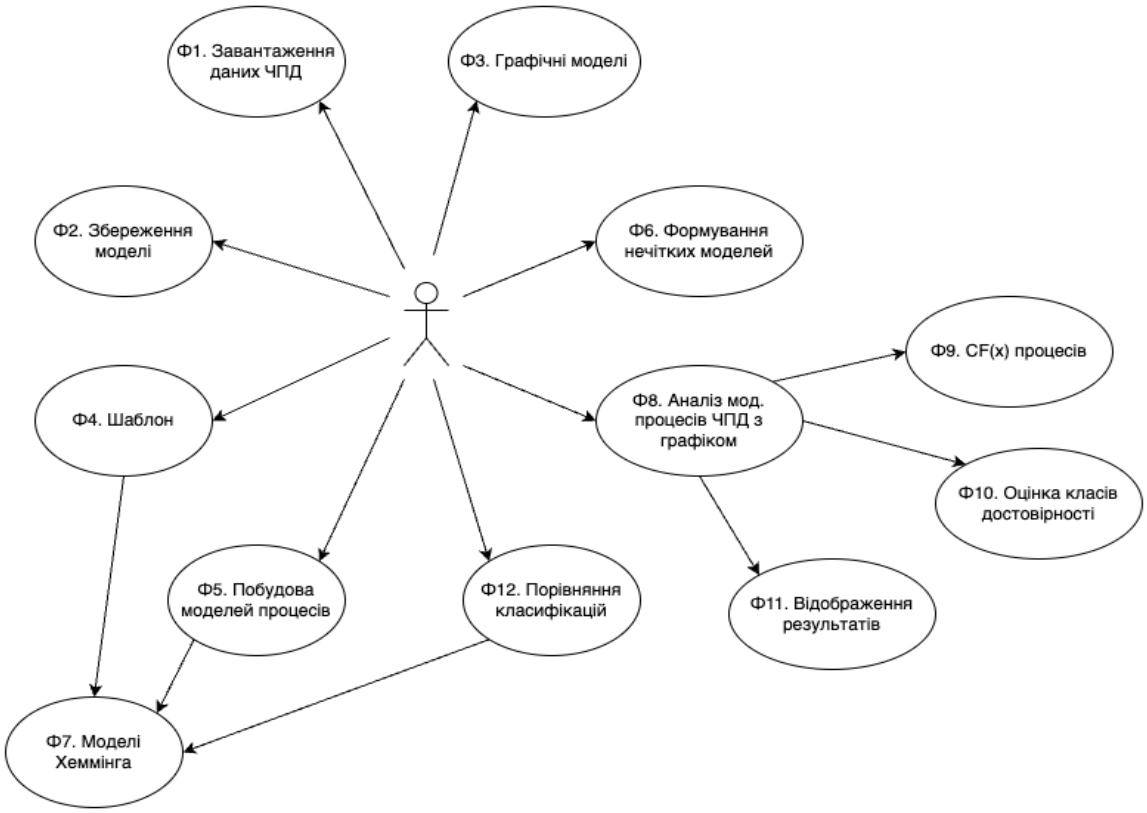


Рисунок 8 - Діаграма варіантів використання програмного комплексу із визначення авторства та призначення/відбір виконавця

Висновки. У статті приведені результати досліджень та розвитку інтелектуальних моделей управління складними системами за умов невизначеності даних на основі процедур, які забезпечують достовірне вирішення завдань з урахуванням оцінки граничної розмірності моделей. Досліджені можливості удосконалення нейронних мереж Хеммінга для класифікації даних у форматах нечітких величин і certainty factor CF(A). Визначені особливості математичної моделі завдань класифікації на основі набору шаблонів ознак. Приведено програмний комплекс інформаційної технології управління призначенням/відбором виконавців, а також визначення авторства україномовних творів на основі класифікації наборів шаблонів із

певних нечітких ознак. Програмний комплекс використовує процедури редукції і каппа статистики.

ЛІТЕРАТУРА

1. Скалозуб В.В., Горячкін В.М., Терлецький І.А., Дудник І.П. Формування моделей класифікації невизначених даних процедурами редукції і каппа статистики. Системні технології. – Дніпро, УДУНТ, 2023. – № 5 (148). – С. 141-155. DOI: 10.34185/1562-9945-5-148-2023-13
https://journals.nmetau.edu.ua/index.php/st/article/view/1401

2. Великоіваненко Г. І. Оцінювання рівня економічної безпеки на підґрунті відстані Хеммінга. 2018. URL: https://core.ac.uk/download/pdf/197269753.pdf

3. Васільєв В. І. Індукція и редукція в проблемах екстраполяції. Кібернетика і обчислювальна техніка. 1998. Вип. 116. С. 65–81.

4. Колесник А. С., Хайрова Н. Ф. Обґрунтування використання статистики каппа Коена в експериментальних дослідженнях NLP Text Mining. Кібернетика та системний аналіз. Т. 58, № 2. 2022. С. 143–153.

5. Li Min Fu, Shortliffe E. H. The application of certainty factors to neural computing for rule discovery. IEEE Transactions on Neural Networks. 2000. Vol. 11. Iss. 3. P. 647–657. DOI: https://doi.org/10.1109/72.846736

6. Скалозуб В. В., Горячкін В. М., Терлецький І. А. Багатопараметричні інтелектуальні процедури діагностування за неповними і збуреними даними // Логістика і транспортна безпека: Проблеми та перспективи розвитку в контексті аналізу сучасних викликів і загроз: матеріали доповідей ІІ Міжнародної науково-практичної конференції. — Дніпро: Середняк Т.К., 2023. С. 42 – 47.

7. Шинкаренко В. І., Демидович І. М. Визначення ознак авторства природномовних текстів. Штучний інтелект. 2018. № 3. С. 27–35.

8. Freitag R. M. Ko. Kappa statistic for judgment agreement in Sociolinguistics / Estatística Kappa para concordância de julgamento em Sociolinguística. Revista de Estudos da Linguagem. 2019. Vol. 27, No. 4. P. 1591–1612. DOI: https://doi.org/10.17851/2237-2083.0.0.1591-1612

9. Leszek Rutkowski Metody i techniki sztucznej inteligentcji. Naukove PWN, Warszava, 2005. – 520 p.

10. Haykin S. Neural networks: A Comprehensive Foundation. Prentice hall: New Jersey, 1999.1103 p.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

REFERENCES

1. Skalozub V.V., Horyachkin V.M., Terlets′kyy I.A., Dudnyk I.P. Formuvannya modeley klasyfikatsiyi nevyznachenykh danykh protseduramy reduktsiyi i kappa statys-tyky. Systemni tekhnolohiyi. – Dnipro, UDUNT, 2023. – № 5 (148). – С. 141-155. DOI: 10.34185/1562-9945-5-148-2023-13

2. Velykoivanenko, H. I. (2018). Otsiniuvannia rivnia ekonomichnoi bezpeky na pidgrunti vidstani Khemminha. Retrieved from https://core.ac.uk/download/pdf/197269753.pdf (in Ukrainian)

3. Vasilev V. I. Induktsiya i reduktsiya v problemakh ekstrapolyatsii. Cybernetics and Computer Engineering, 1998. Vyp.116, 65-81. (in Russian).

4. Kolesnyk A. S., Khairova N. F. Obgruntuvannia vykorystannia statystyky kappa Koena v eksperymentalnykh doslidzhenniakh NLP Text Mining Cybernetics and system analysis. Vol. 58, No. 2. 2022. P. 143–153.

5. Li Min Fu, Shortliffe E. H. The application of certainty factors to neural computing for rule discovery. IEEE Transactions on Neural Networks. 2000. Vol. 11. Iss. 3. P. 647–657. DOI: https://doi.org/10.1109/72.846736

6. Skalozub V.V., Goryachkin V.M., Terletskyi I.A. Bahatoparametrychni intelektualni protsedury diahnostuvannia za nepovnymy i zburenymy danymy// Logistics and transport safety: Problems and prospects of development in the context of analysis of modern challenges and threats: materials of reports II International scientific and practical conference. — Dnipro: Serednyak T.K., 2023. P. 42-47.

7. Shinkarenko V. I., Demidovych I. M. Determination of signs of authorship of natural language texts. Artificial Intelligence. 2018. No. 3. P. 27–35.

8. Freitag R. M. Ko. Kappa statistic for judgment agreement in Sociolinguistics / Estatística Kappa para concordância de julgamento em Sociolinguística. Revista de Estudos da Linguagem. 2019. Vol. 27, No. 4. P. 1591–1612. DOI: https://doi.org/10.17851/2237-2083.0.0.1591-1612

9. Leszek Rutkowski Metody i techniki sztucznej inteligentcji. Naukove PWN, Warsaw, 2005. – 520 p.

10. Haykin S. Neural networks: A Comprehensive Foundation. Prentice hall: New Jersey, 1999.1103 p.

Received 12.03.2024.
Accepted 19.03.2024.

Research of intellectual management models based on classification procedures of uncertain data with established requirements of result reliability

For a wide range of complex systems, tasks such as selection of control options for various technological processes, selection of performers for assigned tasks, and determination of authorship are resolved through classification and diagnosis of incomplete data regarding states and conditions of operation. The relevant problems include forming adequate mathematical models of classification procedures and establishing their correctness, completeness, and reliability of results. This article focuses on investigating the properties and development of intellectual management models for complex systems under conditions of data uncertainty based on classification procedures using reduction methods and Cohen's kappa statistics. It is noted that the application of these methods ensures reliable resolution of classification tasks considering the assessment of the maximum model dimensionality. Additionally, the possibilities of improving Hamming neural networks intended for data classification tasks in formats of fuzzy values and certainty factors CF(A) were explored. The features of the proposed enhanced mathematical model for fuzzy classification tasks based on a set of feature templates defining the classes of objects under analysis were identified.

The article also discusses the peculiarities of the mathematical model of classification designed for the task of determining the authorship of Ukrainian-language works (UAW). The characteristics of the UAW task and its implementation based on a fuzzy classification model include the absence of requirements regarding the number of stages in the authorship determination procedure, the unnecessary formation of a unified classification model for UAW tasks for any possible input works, and the absence of the need to transform template models when introducing new data or works into the model. The listed features of classification procedures are accounted for in the reduction and Cohen's kappa procedures outlined in the article.

To implement and study classification tasks of complex system parameters under conditions of uncertain data, appropriate software was developed. The article presents the structure of the software complex for information technology management of performer assignment/selection, as well as the task of determining authorship of Ukrainian-language works based on classification of sets of templates with certain fuzzy features. The software complex utilizes reduction and kappa statistics procedures.

Keywords: template set classification, uncertain data, dimensionality reduction, Cohen's Kappa statistics, fuzzy values, certainty factor CF(A), modified Hamming network, information technology, software, authorship determination, performer assignment/selection.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Скалозуб Владислав Васильович - професор, каф. «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, УДУНТ.

Горячкін Вадим Миколайович – зав. каф. «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, УДУНТ.

Клименко Іван Вікторович – кандидат економічних наук, доцент, доцент кафедри комп'ютерних інформаційних технологій, факультет комп'ютерних систем і технологій, Український державний університет науки і технологій.

Терлецький Ігор Андрійович – аспірант каф. «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, УДУНТ.

Skalozub Vladyslav Vasyliovych– professor, dep. "Computer and Information Technology", Ukrainian State University of Science and Technology, USUST.

Horiachkin Vadym Mykolayovych – Head of dep. "Computer and Information Technology", Ukrainian State University of Science and Technology, USUST.

Klymenko Ivan - ass. professor, dep. «Computer and Information Technology», Ukrainian State University of Science and Technology, USUST.

Terlitskyi Ihor Andriyovych– post-graduate student, Dep. "Computer and Information Technology", Ukrainian State University of Science and Technology, USUST.

DOI 10.34185/1562-9945-2-151-2024-15
УДК 330.3

А.В. Басько, Н.М. Єршова

МЕТОДОЛОГІЯ ПОЕТАПНОГО ПРОЄКТУВАННЯ ПОРТФЕЛЯ ІНВЕСТИЦІЙНИХ ПРОЄКТІВ

Анотація. Формування портфеля проєктів є ключовим завданням управління організацією. Аналіз життєвого циклу портфеля проєктів показує, що найважливіша є фаза вибору портфеля проєктів. Дотепер проблеми цієї фази не знайшли оптимального вирішення. Тому автори пропонують методологію поетапного проєктування портфеля інвестиційних проєктів. Перший етап формування портфеля проєктів на основі методів математичного програмування та моделювання. Другій етап оцінка ефективності відібраних проєктів методом аналізу ієрархій. Третій етап розподіл коштів інвесторів між проєктами портфеля проєктів на основі гри з природою. Виконана оцінка ефективності трьох проєктів методом аналізу ієрархій. Критеріями є показники ефективності: показник науково технічної ефективності, економічний показник, соціальний показник та показник забезпечення інформаційної безпеки. Кожен критерій має 4 підкритерія. В результати розрахунку визначені наступні ефективності проєктів: першого (44,36%), другого (22,95%) та третього (32,70%). Отже, в таких пропорціях необхідно розподіляти і ресурси між проєктами. Доведено, що проєктування портфеля інвестиційних проєктів є складним процесом, і виконувати його необхідно поетапно, використовуючи на кожному з них сучасні математичні методи прийняття рішень та технології.
Ключові слова: методологія проєктування, портфель проєктів, життєвий цикл, фаза циклу, метод аналізу ієрархій

Вступ та постановка проблеми. Білозеров А. у роботі «Управління портфелем проєктів. Нові методологічні підходи та інструменти» розглядає методологічні підходи до управління портфелем проєктів. На думку автора, різниця між поняттями «управління проєктами» і «управління портфелем проєктів» ось у чому: управління портфелем проєктів визначає проєкти, що мають максимальну цінність в організації, а управління проєктами дозволяє правильно управляти цими проєктами, тобто досягати проєктних цілей, не виходячи за рамки проєктних обмежень, тим самим забезпечуючи цю цінність.

Фази життєвого циклу портфеля проєктів представлені на рис.1.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

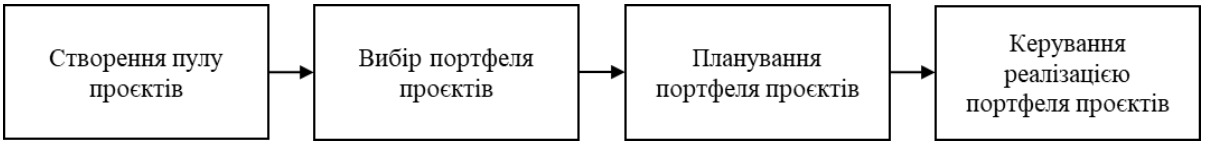


Рисунок 1 - Життєвий цикл портфеля проєктів

Основною метою фази створення є формування пулу проєктів, які потім можуть бути ініційовані та прийняті до реалізації. На цій фазі здійснюється збір проєктних (інвестиційних) ініціатив та заявок без урахування фінансових та інших обмежень організації.

Мета фази вибору портфеля проєктів - відбір проєктів у портфель з урахуванням фінансових та інших обмежень портфеля. На цій фазі з отриманого пулу потенційних проєктів обирається той портфель, який буде прийнято до реалізації.

У фазі планування портфеля проєктів здійснюється:

- запуск проєктів (призначення менеджерів проєктів, формування організаційних структур, випуск статутів проєктів);

- допланування (деталізація всіх видів планів бізнес-плану до ступеня, необхідного для успішної реалізації проєкту);

- виділення ресурсів (виділення конкретних людей, виробничих потужностей тощо).

Фаза керування реалізацією портфеля проєктів призначена для вирішення наступних задач:

- моніторинг виконання проєктів у портфелі, аналіз відхилень при реалізації проєктів та їх впливу на залежні проєкти та портфель загалом;

- координація ресурсів - під час реалізації деякі проєкти можуть призупинятися, а їх ресурси перекидатися на інші, більш пріоритетні проєкти.

Аналіз життєвого циклу портфеля проєктів показує, що найважливіша є фаза вибору портфеля проєктів. До сьогодні проблеми цієї фази не знайшли оптимального вирішення. Тому автори пропонують методологію поетапного проєктування портфеля інвестиційних проєктів, і в даній статті представляють вирішення задачі оцінки ефективності проєктів методом аналізу ієрархій.

Аналіз останніх досліджень і публікацій. Метод аналізу ієрархій (MAI) розроблений американським ученим Т. Саати на початку 1970 року [1]. Він забезпечує за допомогою простих і обґрунтованих правил вирішення багатокритерійних задач, які містять якісні і кількісні показники, при цьому кількісні показники можуть мати різну розмірність. Метод заснований на

декомпозиції задачі і представленні її у вигляді ієрархічної структури. Це дозволяє включити в ієрархію всі знання по вирішуваній проблемі. В результаті вирішення визначається відносний ступінь взаємодії елементів в ієрархії, що виражено у числовому форматі. МАІ використовується для вирішення слабоструктурованих і неструктурованих проблем.

Вирішення задачі з допомогою МАІ виконується поетапно.

Перший етап передбачає представлення проблеми у вигляді ієрархії. В найпростішому випадку ієрархія будується з мети, яка поміщається у вершину ієрархії, потім йдуть проміжні рівні, на яких розміщуються критерії і підкритерії, і найнижчий рівень містить перелік альтернатив.

Другий етап. На цьому етапі необхідно встановити пріоритети критеріїв і оцінити кожну альтернативу за критеріями для вибору з них найважливішій. В МАІ елементи порівнюються попарно по відношенню до їх впливу на загальну для них характеристику.

Для отримання позитивних результатів в порівняннях необхідно уміти:

- знаходити відповідну числову шкалу порівнянь;
- визначати ступінь неузгодженості наших суджень.

Головна вимога до шкали порівнянь – шкала повинна бути простою і природною. Шкала Т. Сааті містить числа від 1 до 9 і побудована на основі психометричних властивостей людини, які добре дозволяють проводити якісні порівняння властивостей двох об'єктів за одним критерієм. Крім того, в психології існує поняття психологічної межі здатності людини одночасно розрізняти якусь кількість меж за якоюсь властивістю. Ця межа дорівнює 7 ± 2, тобто для створення шкали необхідно не більше 9 точок.

Коли задача представлена у вигляді ієрархічної структури, то матриця складається для парного порівняння критеріїв на другому рівні по відношенню до загальної мети, розташованої на першому рівні. Такі ж матриці повинні бути побудовані для парних порівнянь кожної альтернативи на третьому рівні по відношенню до критеріїв другого рівня і так дали, якщо кількість рівнів більше трьох.

Третій етап. Після формування матриць парних порівнянь за всіма критеріями та альтернативами необхідно визначити власні вектори матриць, перевірити узгодженість матриць за допомогою їх власних чисел і провести синтез глобальних пріоритетів альтернатив щодо основної мети.

Т. Саати запропонував чотири алгоритми наближених методів визначення

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

нормованих власних векторів обернено симетричної матриці. Всі алгоритми дають один і той же власний вектор. Найбільш простій з них є алгоритм 1.

Розрахунок виконується по кроках. Спочатку визначається власний нормований вектор матриці парних порівнянь, сума елементів якого дорівнює 1. Це є контролем правильності розрахунків. Потім оцінюється узгодженість матриці парних порівнянь.

Існують два критерії оцінки узгодженості матриці парних порівнянь: індекс узгодженості і відношення узгодженості. Індекс узгодженості визначається у формулі 1.

$$I_y = \frac{\lambda_{max} - n}{n-1}, \quad (1)$$

де λ_{max} – максимальне власне значення матриці A; n – порядок матриці. Якщо $I_y \leq 0{,}1$, то практично вважається, що міра узгодженості знаходиться на прийнятному рівні. Відношення узгодженості визначається по формулі 2.

$$B_y = \frac{I_y}{M(I_c)}, \quad (2)$$

де $M(I_c)$ – середнє значення індексу узгодженості випадковим чином складеної матриці парних порівнянь, яке залежить від порядку матриці і береться з таблиці.

Відношення узгодженості указує, наскільки оцінюваний ступінь узгодженості сходиться із ступенем узгодженості самого неідеально проведеного експерименту. Вважається, що якщо $B_y \leq 0{,}1$, то можна бути задоволеним ступенем узгодженості думок.

Нині метод аналізу ієрархій завдяки своїй універсальності застосовується у багатьох сферах діяльності серед яких економічна, соціальна, політична технічна та технологічна сфера. В роботах [2-5] наведено постановки задач різного роду і вирішення простіших з них практично вручну методом аналізу ієрархій. Отже, з'являються висновки, що метод дуже складний і його не можна використовувати для вирішення складних проблем. Автор роботі [6] створив матричний метод аналізу ієрархій в середовищі Excel з допомогою матричних функцій. В цій роботі наведено вирішення багатьох прикладів з раніше приведених робіт. Метод використано в роботі [7] при виборі місця розташування підприємства для виробництва тротуарної плитки. На методику проєктування підприємства для випуску нової продукції отримано авторське свідоцтво [8].

Мета даного дослідження:

1. Розроблення методології поетапного проєктування портфеля інвестиційних проєктів.

2. Вирішення задачі оцінки ефективності проєктів методом аналізу ієрархій.

Викладення основного матеріалу. Білозеров А. у роботі «Управління портфелем проєктів. Нові методологічні підходи та інструменти» розглядає методологічні підходи до управління портфелем проєктів на усіх фазах його життєвого циклу. Найбільш складна фаза вибору портфеля проєктів, тобто відбір проєктів у портфель з урахуванням фінансових та інших обмежень портфеля. На цій фазі з отриманого пулу потенційних проєктів обирається той портфель, який буде прийнято до реалізації. Тут вирішуються задачі: формування оптимального портфеля проєктів, оцінка ефективності відібраних проєктів та розподіл коштів інвесторів між проєктами портфеля проєктів.

Нами пропонується методологія поетапного проєктування портфеля інвестиційних проєктів.

Перший етап. Формування портфеля проєктів на основі методів математичного програмування та моделювання.

Другій етап. Оцінка ефективності відібраних проєктів методом аналізу ієрархій.

Третій етап. Розподіл коштів інвесторів між проєктами портфеля проєктів на основі гри з природою.

Вирішення задачі оцінки ефективності проєктів методом аналізу ієрархій розглянемо на прикладі [5].

Постановка задачі. Є три проєкти А1, А2, А3 з портфеля проєктів, ефективність яких потрібно оцінити. Критеріями є показники ефективності: загальний показник (ЗП), показник науково-технічної ефективності (НТ), економічний показник (Е), соціальний показник (С) та показник забезпечення інформаційної безпеки (ІБ).

Критерій НТ залежить від:

- науково-технічного рівня (НТР);
- наявності нових проривних технологій (ПТ);
- технологічних засобів (СРТ);
- підвищенням рівня кваліфікації персоналу (РКП);

Економічний показник (Е) залежить від:

- нормою прибутковості (НД);
- перспективності (П);
- чистим дисконтованим доходом (ЧДД);
- коефіцієнтом фінансової автономності проєкту (КФ);

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Соціальний показник (С) залежить від:

- рівнем впливу на діяльність громадських та молодіжних організацій (ДО);
- сприянням розвитку малого та середнього бізнесу (МСБ);
- ступеня міжнародної кооперації (МК);
- збереження та збільшення робочих місць у галузі та державі (РМ);

Показник забезпечення інформаційної безпеки (ІБ) залежить від:

- сприяння розвитку технологій та та сприяння вирішенню проблем глобальної та регіональної безпеки (РПБ);
- ступінь важливості для вирішення завдань ІБ (СВ);
- потенційного масштабу практичного використання (МПІ);
- ймовірності досягнення позитивних результатів (ДР);

Створення ієрархічної моделі. Ієрархічна модель задачі представлена на рис. 2.

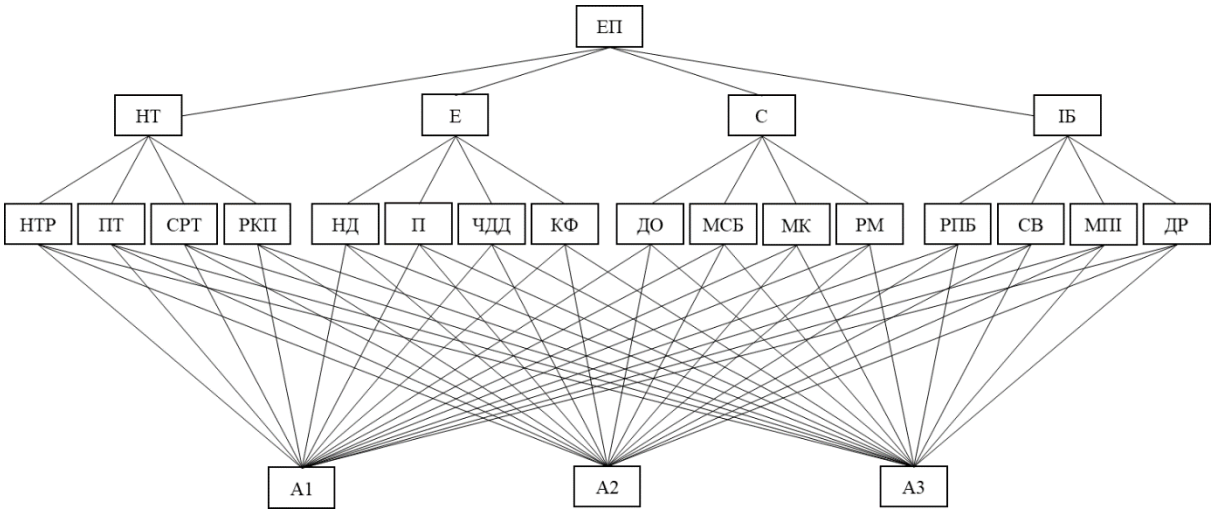


Рисунок 2 – Ієрархічна модель проблеми оцінки ефективності проєктів

Вирішення задачі. У цій задачі проєкти є альтернативами і утворюють четвертий рівень ієрархії. Критеріями є: показник науково-технічної ефективності, економічний показник, соціальний показник та показник забезпечення інформаційної безпеки.

Критерії утворюють другий рівень ієрархії. У свою чергу критерії залежать від підкритеріїв, які утворюють третій рівень ієрархії. Вершиною ієрархії є мета –ефективність проєктів (ЕП)

Для ієрархії складається 21 матриця: 1 – для другого рівня; 4 – для третього рівня та 16 – для четвертого рівня. Дослідження виконуються в Microsoft Excel, дослідження впливу критеріїв на загальну мету наведено на рис. 3.

Оцінка ефективності проектів				
ЕП	НТ	Е	С	ІБ
НТ	1	0,5	4	2
Е	2	1	6	2
С	0,25	0,166667	1	0,3333
ІБ	0,5	0,5	3,0003	1
Власний вектор		W		
алгоритм 1				
крок 1	7,5			
	11			
	1,749967			
	5,0003			
крок 2	25,25027			
крок 3			%	
W	0,297027	контроль	29,7	
	0,435639	1	43,6	
	0,069305		6,9	
	0,19803		19,8	
перевірка узгодженості				
	1,188125			
A*W	1,841581			
	0,282171			
	0,772298			
	4,000062			
лямбда	4,227309			
	4,071449			
	3,899911			
лямбда макс.	4,227309			
n	4			
Iy	0,07577			
M(Ic)	0,9			
By	0,084189			

Рисунок 3 – Вплив критеріїв на загальну мету

Отриманий результат свідчить про гарну узгодженість заданої матриці парних порівнянь критеріїв НТ, Е, С, ІБ, що забезпечує досягнення загальної мети ЕП. Нормований вектор пріоритетів W свідчить про те, що ефективність проєктів в цілому може залежати на 29,7% від показника науково-технічної ефективності, на 43,5% від економічного показника, на 6,9% соціального показника та на 19,8% від показника забезпечення інформаційної безпеки.

Далі виконаємо розрахунок другого рівня моделі. На рис. 4 приведено результати дослідження впливу підкритеріїв на показник науково-технічної ефективності та на економічний показник проєктів.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

НТ	НТР	ПТ	СРТ	РКП
НТР	1	2	3	0,7
ПТ	0,5	1	0,6	0,7
СРТ	0,333333	1,666667	1	1
РКП	1,428571	1,428571	1	1

Власний вектор	W1		
алгоритм 1			
крок 1	6,7		
	2,8		
	4		
	4,857143		
крок 2	18,35714		
крок 3			%
W1	0,364981	контроль	36,5
	0,152529	1	15,3
	0,217899		21,8
	0,264591		26,5
перевірка узгодженості			
	1,508949		
A1*W1	0,650973		
	0,858366		
	1,22179		
	4,134328		
лямбда	4,267857		
	3,939286		
	4,617647		
лямбда макс.	4,267857		
n	4		
Iy	0,089286		
M(Ic)	0,9		
By	0,099206		

Е	НД	П	ЧДД	КФ
НД	1	1,6	0,5	0,9
П	0,625	1	0,2	0,7
ЧДД	2	5	1	1,3
КФ	1,111111	1,428571	0,769231	1

Власний вектор	W2		
алгоритм 1			
крок 1	4		
	2,525		
	9,3		
	4,308913		
крок 2	20,13391		
крок 3			%
W2	0,19867	контроль	19,9
	0,12541	1	12,5
	0,461907		46,2
	0,214013		21,4
перевірка узгодженості			
	0,822891		
A2*W2	0,491769		
	1,764515		
	0,969228		
	4,142005		
лямбда	3,921283		
	3,820063		
	4,528833		
лямбда макс.	4,142005		
n	4		
Iy	0,047335		
M(Ic)	0,9		
By	0,052595		

Рисунок 4 – Вплив підкритеріїв на НТ та Е показники

Отриманий результат свідчить про гарну узгодженість заданої матриці парних порівнянь критерія НТ та Е.

Нормований вектор пріоритетів W1 свідчить про те, що рівень показника НТ в цілому може залежати на 36,5% від науково-технічної ефективності, на 15,3% від наявності нових проривних технологій, на 21,8% від сприяння розвитку технологій, на 26,5% від підвищення рівня кваліфікації персоналу.

Нормований вектор пріоритетів W2 свідчить про те, що рівень показника Е в цілому може залежати на 19,9% від внутрішньої норми прибутковості, на 12,5% від перспективності, на 46,2% від чистого дисконтованого доходу, на 21,4% від коефіцієнту фінансової автономності проєкту.

Отриманий результат свідчить про гарну узгодженість заданої матриці парних порівнянь критерія С та ІБ, що зображено на рис.5.

С	ДО	МСБ	МК	РМ
ДО	1	3	2	4
МСБ	0,333333	1	0,9	2
МК	0,5	1,111111	1	3
РМ	0,25	0,5	0,333333	1

Власний вектор	W3		
алгоритм 1			
крок 1	10		
	4,233333		
	5,611111		
	2,083333		
крок 2	21,92778		
крок 3			%
W3	0,456043	контроль	45,6
	0,193058	1	19,3
	0,255891		25,6
	0,095009		9,5
перевірка узгодженості			
	1,927033		
A3*W3	0,765391		
	0,983447		
	0,390845		
	4,225556		
лямбда	3,964567		
	3,843234		
	4,113778		
лямбда макс.	4,225556		
n	4		
Iy	0,075185		
M(Ic)	0,9		
By	0,083539		

ІБ	РПБ	СВ	МПІ	ДР
РПБ	1	5	7	3
СВ	0,2	1	1,6	1
МПІ	0,142857	0,625	1	0,5
ДР	0,333333	1	2	1

Власний вектор	W4		
алгоритм 1			
крок 1	16		
	3,8		
	2,267857		
	4,333333		
крок 2	26,40119		
крок 3			%
W4	0,606033	контроль	60,6
	0,143933	1	14,4
	0,0859		8,6
	0,164134		16,4
перевірка узгодженості			
	2,419398		
A4*W4	0,566713		
	0,344501		
	0,681878		
	3,992188		
лямбда	3,937343		
	4,010499		
	4,154396		
лямбда макс.	4,010499		
n	4		
Iy	0,0035		
M(Ic)	0,9		
By	0,003888		

Рисунок 5 – Вплив підкритеріїв на С та ІБ показники

Нормований вектор пріоритетів W3 свідчить про те, що рівень показника С в цілому може залежати на 45,6% від рівня впливу на діяльність громад та молоді, на 19,3% від сприяння розвитку малого та середнього бізнесу, на 25,6% від ступеня міжнародної кооперації, на 9,5% від збереження та збільшення робочих місць.

Нормований вектор пріоритетів W4 свідчить про те, що рівень показника ІБ в цілому може залежати на 60,6% від сприянню вирішення проблем безпеки, на 13,4% від ступеня важливості для вирішення задач, на 8,6% від масштабу практичного використання, на 16,4% від ймовірності досягнення позитивних результатів.

Наступним кроком є розрахунок третього рівня моделі, а саме розрахунок впливу альтернатив на підкритерії. Розрахунок виконується для всіх підкритеріїв НТР, ПТ, СРТ, РКП, НД, П, ЧДД, КФ, ДО, МСБ, МК, РМ, РПБ, СВ, МПІ, ДР. На рис. 6 відображено приклад розрахунку для підкритеріїв НТР та ПТ.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

НТР	А1	А2	А3
А1	1	4	3
А2	0,25	1	0,6
А3	0,333333	1,666667	1
Власний вектор		V1	
алгоритм 1			
крок 1	8		
	1,85		
	3		
крок 2	12,85		
крок 3			
V1	0,622568	контроль	
	0,143969	1	
	0,233463		
перевірка узгодженості			
	1,898833		
А1*V1	0,439689		
	0,680934		
	3,05		
лямбда	3,054054		
	2,916667		
лямбда ма	3,054054		
n	3		
Iy	0,027027		
M(Ic)	0,58		
By	0,046598		

ПТ	А1	А2	А3
А1	1	6	3
А2	0,166667	1	0,6
А3	0,333333	1,666667	1
Власний вектор		V2	
алгоритм 1			
крок 1	10		
	1,766667		
	3		
крок 2	14,76667		
крок 3			
V2	0,677201	контроль	
	0,119639	1	
	0,20316		
перевірка узгодженості			
	2,004515		
А2*V2	0,354402		
	0,628292		
	2,96		
лямбда	2,962264		
	3,092593		
лямбда ма	3,092593		
n	3		
Iy	0,046296		
M(Ic)	0,58		
By	0,079821		

Рисунок 6 – Вплив альтернатив на підкритерії НТР та ПТ

Оцінка впливу альтернатив на загальну мету виконується в три етапи:

- оцінка впливу альтернатив на підкритерії;
- оцінка впливу альтернатив на критерії;
- оцінка впливу альтернатив на загальну мету.

Оцінимо вплив альтернатив на підкритерії. Для цього складемо матрицю В (рис. 7)., стовпцями якої є власні вектори підкритеріїв. Стовпці цієї матриці оцінюють внесок альтернатив в кожний підкритерій, тобто перший стовпець в науково-технічний рівень, другий в перспективність і так далі. Оцінка впливу альтернатив на критерії виконується за матрицею В1 (рис. 8), стовпцями якої є стовпці - результат множення матриць власних векторів підкритеріїв із матриці В на відповідні власні вектори критеріїв.

матриця В															
НТ				Е				С				ІБ			
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16
0,622568	0,677201	0,540541	0,279057	0,606061	0,320665	0,320197	0,24811	0,06527	0,556166	0,271943	0,108633	0,559172	0,668844	0,45302	0,674473
0,143969	0,119639	0,174174	0,070281	0,128788	0,570071	0,203612	0,506144	0,676873	0,196026	0,034175	0,614596	0,14497	0,128968	0,211409	0,24356
0,233463	0,20316	0,285285	0,650662	0,265152	0,109264	0,47619	0,245747	0,257857	0,247808	0,693882	0,276772	0,295858	0,202188	0,33557	0,081967

Рисунок 7 – Оцінка впливу альтернатив на підкритерії

матриця B1			
НТ*W1	Е*W2	С*W3	ІБ*W4
0,522137	0,361621	0,217047	0,584764
0,127342	0,29945	0,413664	0,164556
0,35052	0,338929	0,369288	0,25068

Рисунок 8 – Оцінка впливу альтернатив на критерії

Аналіз результатів показує, що пріоритет за критерієм науко-технічної ефективності має перша альтернатива (52,21%), за критерієм економічності - перша альтернатива (36,16%), за соціальним показником – друга альтернатива (41,36%) та за інформаційною безпекою – перша альтернатива (58,47%).

Аналогічно для оцінки впливу альтернатив на загальну мету (рис. 9) необхідно помножити матрицю B1 на власний вектор матриці А.

матриця B2
B1*W
0,443467565
0,229532237
0,327000199

Рисунок 9 – Оцінка впливу альтернатив на загальну мету

У результаті отримали узагальнений (глобальний) вектор пріоритетів проєктів стосовно кінцевої мети – ефективності проєкту.

Таким чином, облік всіх факторів показав, що ефективність першого (44,36%), другого (22,95%) та третього (32,70%) проєкту. Отже, в таких пропорціях необхідно розподіляти і ресурси між проєктами.

Висновок. Отже аналіз життєвого циклу портфеля проєктів показує, що найважливішою фазою є вибір портфелю проєктів. Тому було запропоновано методологію поетапного проєктування портфеля інвестиційних проєктів, що включає в себе: формування портфеля проєктів на основі методів математичного програмування та моделювання, оцінка ефективності відібраних проєктів методом аналізу ієрархій та розподіл коштів інвесторів між проєктами портфеля на основі гри з природою.

Відповідно до запропонованої методології було приведено приклад практичної оцінки ефективності трьох проектів. Починаючи від постановки задачі, побудовою ієрархічної моделі і закінчуючи оцінкою впливу підкритеріїв на ефективність проектів. У результаті для кожного проекту була отримана ефективність, що дозволяє визначити в яких пропорціях необхідно розподілити ресурси між проєктами. Особливістю методики є використання

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

кількісних і якісних критеріїв з багатокритеріальний розглядом альтернатив, що дозволяє отримати кількісну характеристику розподілу ресурсів.

ЛІТЕРАТУРА

1. Саати Т. Принятие решений. Метод анализа иерархий. Москва: Радио и связь, 1993. 314 с.

2. Бадюл М. Г., Крамаренко В. А. Застосування методу аналізу ієрархій у проектуванні та будівництві. Строительство, материаловедение, машиностроение. 2013. № 70. С. 27-35.

3. Сєріков А. В., Білоцерківський О. В. Метод аналізу ієрархій у прийнятті рішень: навч. посіб. Харків: БУРУН КНИГА, 2006. 144 с.

4. Синюк В. Г., Шевырев А. В. Использование информационно-аналитических технологий при принятии управленческих решений: навч. посіб. Москва: Экзамен, 2003. 160 с.

5. Михалев А. И., Алпатов А. П., Баклан И. В. Структурный синтез систем управления проектами: навч. посіб., за ред. А. И. Михалева. Днепропетровск: ИК «Системные технологии», 2013. 144 с.

6. Ершова Н. М. Модели и методы теории принятия решений: навч. посіб. Дніпро. 2016. 246 с.

7. Development of the design method of the enterprise for the release of new products / N. Ostanina et al. Technology audit and production reserves. 2017. Vol. 1, no. 2(39). P. 61-68. URL: https://doi.org/10.15587/2312-8372.2018.123580 (date of access: 26.03.2024).

8. Єршова Н. М., Останіна А. О. Свідоцтво про реєстрацію авторського права на твір наукового характеру «Методика проектування підприємства для випуску нової продукції» , № 76491 від 01.02.2018 р.

REFERENCES

1. Saati T. Decision making. Hierarchical analysis method. Moscow: Radio and Communication, 1993. 314 p.

2. Badyul M. G., Kramarenko V. A. Application of the method of analysis of hierarchies in design and construction. Construction, materials science, mechanical engineering. 2013. No. 70. P. 27-35.

3. Syerikov A. V., Bilotserkivskyi O. V. The method of analyzing hierarchies in decision-making: teaching. manual Kharkiv: BURUN BOOK, 2006. 144 p.

4. Sinyuk V. G., Shevyrev A. V. The use of information and analytical technologies in making management decisions: training. manual Moscow: Exam, 2003. 160 p.

5. Mykhalev A. I., Alpatov A. P., Baklan I. V. Structural synthesis of project manage-

ment systems: training. manual, edited by A. I. Mykhalev. Dnipropetrovsk: IC "System Technologies", 2013. 144 p.
6. Ershova N.M. Models and methods of decision-making theory: teaching. manual Dnipro 2016. 246 p.
7. Development of the design method of the enterprise for the release of new products / N. Ostanina et al. Technology audit and production reserves. 2017. Vol. 1, no. 2(39). P. 61-68. URL: https://doi.org/10.15587/2312-8372.2018.123580 (date of access: 26.03.2024).
8. Ershova N. M., Ostanina A. O. Certificate of copyright registration for the scientific work "Methodology of enterprise design for the production of new products", No. 76491 dated 02.01.2018.

Received 12.03.2024.
Accepted 19.03.2024.

Methodology of step-by-step design of investment project portfolio

Forming a portfolio of projects is a key task of managing an organization. Analysis of the life cycle of the project portfolio shows that the phase of project portfolio selection is the most important. Until now, the problems of this phase have not found an optimal solution. Therefore, the authors propose a methodology for the step-by-step design of a portfolio of investment projects. The first stage is the formation of a portfolio of projects based on mathematical programming and modeling methods. The second stage is the evaluation of the effectiveness of the selected projects by the method of analysis of hierarchies. The third stage is the distribution of investors' funds between the projects of the project portfolio on the basis of playing with nature. The evaluation of the effectiveness of three projects was carried out using the method of hierarchy analysis. The criteria are indicators of efficiency: indicator of scientific and technical efficiency, economic indicator, social indicator and indicator of ensuring information security. Each criterion has 4 subcriteria. The results of the calculation determined the following efficiency of the projects: the first (44.36%), the second (22.95%) and the third (32.70%). Therefore, it is necessary to distribute resources between projects in such proportions. It has been proven that the design of a portfolio of investment projects is a complex process, and it must be carried out in stages, using modern mathematical decision-making methods and technologies for each of them.

Keywords: design methodology, project portfolio, life cycle, cycle phase, hierarchy analysis method.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Басько Артем Володимирович – магістрант кафедри комп'ютерних наук, інформаційних технологій та прикладної математики, Придніпровська державна академія будівництва та архітектури, Дніпро, Україна.

Єршова Ніна Михайлівна – д.т.н., професор кафедри комп'ютерних наук, інформаційних технологій та прикладної математики, Придніпровська державна академія будівництва та архітектури, Дніпро, Україна.

Basko Artem – Master's Degree of the Department of Computer Sciences, Information Technologies and Applied Mathematics, Prydniprovska State Academy of Civil Engineering and Architecture, Dnipro, Ukraine.

Ershova Nina - Doctor of Engineering Sciences, Professor of the Department of Computer Sciences, Information Technologies and Applied Mathematics, Prydniprovska State Academy of Civil Engineering and Architecture, Dnipro, Ukraine.

ЗМІСТ

CONTENTS

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

РЕФЕРАТИ

УДК 004.42:519.85

Мороз Б.І., Круглик А.С., Мороз Д.М., Мартиненко А.А. **Математична модель раціональної організації обробки інформаційних потоків в системі доставки літальними аппаратами** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.3 – 12.

Використання керованої дисципліни для раціональної організації обробки інформації з урахуванням її характеристик цінності і старіння має великі перспективи при доставці товару за допомогою літальних апаратів, особливо якщо таку модель розглядати як таку, де вартість доставки змінюється в залежності від термінів такої доставки. А тримати парк літальних апаратів для покриття усіх пікових навантажень є економічно недоцільно. Запропонована математична модель повинна дозволити керувати обробкою вхідних повідомлень і розподіляти час таким чином, щоб забезпечити виконання усіх заявок в межах часу, обумовленого для такої заявки.

Бібл. 6, іл. 3.

УДК 004.67 , 004.042, 004.622

Руденко К., Божуха Л. **Про інтеграцію рекомендаційних алгоритмів в систему мобільної торгівлі** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.13 – 20.

Розглянуто процес розроблення, впровадження та оптимізації рекомендаційних систем у сфері мобільної торгівлі з застосуванням хмарних сервісів. У роботі використано методи машинного навчання та рекомендаційні алгоритми з метою підвищення персоналізації та точності пропозицій. Результати роботи рекомендовані для підвищення персоналізації послуг у бізнесі (особливо в роздрібній торгівлі та онлайн платформах) для зростання задоволеності та лояльності клієнтів, а також для поліпшення маркетингових стратегій та аналізу трендів ринку.

Бібл. 11, іл. 0, табл. 0.

УДК 621.385.6.

Зимогляд А.Ю., Гуда А.І., Кліщ С.М. **Апаратний комплекс для вимірювання потужності UHF сигналів** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.21 – 32.

Розроблено апаратний комплекс, який складається з 2 блоків, а саме з підсилювача UHF сигналів на мікросхемі QPL9547 та демодулюючого логарифмічного підсилювача на AD8319. Результати перевірки, що були представленні у табл. 1 показують, як реагував UHF вимірювач потужності сигналів на низьку вхідних сигналів різної потужності.

Цей комплекс може бути використано у лабораторних дослідженнях вихідної потужності передаючих пристроїв, для якісної оцінки антен або порівнянь антен. За рахунок підсилювача на вході можливо дослідження потужності сигналу близько до –80 dBm. Описаний комплекс також має доволі помірну ціну, в порівнянні з промисловими аналогами.

Бібл. 2, іл. 8

УДК 004.42:519.25

Земляний О.Д., Байбуз О.Г. **Методи імпутування пропусків у даних про ішемічну хво робу серця** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.33 – 49.

Пропонуються декілька модифікацій алгоритмів ітеративного множинного імпутування пропусків у змішаних даних, що представлені кількісними і якісними ознаками. Серед кількісних є і неперервні, і дискретні. Серед якісних є порядкові та бінарні. Для аналізу підходів використовується два типи тестів. В першому тесті датасет з відомими даними штуч но заповнюється пропусками у випадкових позиціях, проводиться імпутування різними мето дами, оцінюється середньо квадратична похибка та час виконання алгоритмів. В другому тесті навчають моделі бінарної класифікації на датасетах, з імпутацією пропусків різними метода ми, та порівнюють точність класифікації на тестовій вибірці. Для перетворення якісних ознак на кількісні запропоновано власні алгоритми, які працюють з пропущеними даними та дозво ляють виконувати зворотне перетворення знов до якісних ознак. Розглядаються два відомих датасети про спостереження стосовно ішемічної хвороби серця.

Бібл.15, іл.0, табл.2.

УДК 004.942:

Поляков О.М., Жураковський Б.Ю. **Прототипування пристроїв керування систем з промисловими контролерами** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.50 – 61.

Запропоновано методику прототипування пристроїв керування систем із програмною реалізацією алгоритмів керування мовами SFC, FBD, LD стандарту МЕК 61131 3. У процесі про тотипування виконується декомпозиція проектних моделей операційних та керуючих автоматів пристрою управління, реалізація їх у середовищі програми OpenPLC у вигляді типо вих компонентів організації програм системи керування, які виконуються у платі Ардуїно. На ведено приклад застосування запропонованої методики , який може бути корисним для нав чання програмуванню промислових контролерів.

Бібл. 10, іл. 8.

УДК 621.771.068

Гречаний О.М., Васильченко Т.О., Власов А.О., Виприжкін П.О., Якимчук Д.І. **Імітаційне моделювання при дослідженні роботи металургійного обладнання** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.62 – 75.

Проаналізовані технологічні режими намотування гарячекатаної штаби різної товщини. Встановлено залежність діаметра рулону від часу намотки. Визначена форма вільних коли вань виникаючих в приводі моталки гарячекатаної штаби. Встановлено взаємозв'язок між пружними деформаціями від сил опору електродвигуна та обертових частин барабана моталки.

Бібл. 13, іл. 4, табл. 3

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

УДК 519.8

Караваєв К.Д. **Про необхідні умови існування щільних упорядкувань в класичній задачі паралельного упорядкування** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.76 – 91.

Розглядається класична задача мінімізації довжини упорядкування при заданій ширині, в якій шукане упорядкування є щільним. Досліджуються необхідні умови існування щільного упорядкування при застосуванні методу гілок та меж, пов'язані з обмеженістю потужності місць та можливістю їх заповнення. Отримані умови були зведені до однієї, запропоновано ефективні алгоритми її перевірки в загальному випадку та для графів, всі вершини яких зна ходяться на критичних шляхах. В результаті дослідження також були отримані нові уточнені оцінки знизу довжини упорядкування та запропоновані узагальнення спеціальних упорядку вань, в яких вершини займають крайнє ліве та праве можливе місця, які враховують ширину упорядкування.

Бібл. 15, іл. 2, табл. 0.

УДК 519.85:004.42

Кошель Є.В. **Нейронно мережевий підхід до неперервного вкладення одновимірних потоків даних для аналізу часових рядів в реальному часі** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.92 – 101.

Задача вкладення одновимірних часових рядів у багатовимірні простори дуже розпов сюджена і зустрічається у багатьох галузях досліджень. Методи, які розв'язують цю задачу, зазвичай покладаються на ту чи іншу форму вкладання з затримкою, яке реконструює фазо вий простір невідомої системи шляхом асоціювання значень часового ряду з історичними да ними. Ми пропонуємо більш гнучкий метод, який використовує правила для комбінування значень часового ряду, щоб створити вкладення, які є більш репрезентативними щодо взаємозв'язку даних часового ряду одне з одним.

Бібл. 17, іл. 2, табл. 2.

УДК 669.15 194.53

Бабаченко О.І., Подольський Р.В., Кононенко Г.А., Меркулов О.Є., Сафронова О.А., Дуд ченко С.О. **Аналіз впливу швидкості охолодження на твердість сталей для залізничних рейок перлітного та бейнітного класу** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.102 – 112.

Процес експлуатації транспортних засобів визначає взаємодію колеса і рейки. Від параметрів цього процесу багато в чому залежать безпека руху та основні техніко економічні показники господарств колії та рухомого складу. Результатом є вплив, що виникає від тертя кочення і особливо від тертя ковзання колеса по рейці при гальмуванні, відносно цих змін відбувається істотне зростання інтенсивності зношування коліс рухомого складу, яке, в свою чергу може призвести до катастрофічних результатів для локомотивного господарства. Також в процесі експлуатації рейки в більшості випадків утворюються дефекти, що мають характер складнонавантаженого стану: її головка піддається зношуванню, зминанню, розтріскуванню і викришуванню, в металі можуть розвиватися контактно втомні пошкодження.

Бібл. 17, іл. 4, табл. 2

УДК 004.4:004.8

Гнатушенко В.В., Павленко Є.В. **Використання генеративного штучного інтелекту в тестуванні програмного забезпечення** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.113 – 123.

У статті розглядаються можливості генеративного штучного інтелекту при застосуванні у тестуванні програмного забезпечення. Визнаючи стрімкий розвиток попередньо навчених генеративних моделей ШІ з трансформаторами, таких як ChatGPT від OpenAI, Gemini від Google та інші, досліджуються поточні та майбутні ролі цих технологій у вдосконаленні методологій тестування програмного забезпечення. Акцент робиться на перевагах, викликах і практичних аспектах використання генеративного ШІ при генерації, обслуговуванні та документуванні тестів.

Бібл. 16, рис. 2.

УДК 51 74

Гук К.Г., Шевельова А.Є. **Нейромережевий підхід до ідентифікації заповнюваності приміщень за параметрами повітря** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.124 – 132.

Розглядаються різні підходи до прогнозування параметрів повітря та кількості людей у приміщеннях за результатами моніторингу стану повітря в режимі реального часу. Запропо новано підхід до прогнозування за допомогою штучної нейронної мережі, який дозволяє вра хувати різноманітні фактори, характер та кількісні значення яких не піддаються обліку та не можуть бути додані до параметрів моделі. Досліджено питання вибору архітектури та параметрів нейронної мережі для розв'язання задачі прогнозування. Показано можливість використання запропонованого підходу до визначення кількості людей у офісних приміщеннях. Надано рекомендації для підвищення продуктивності моделі.

Бібл. 4, іл. 2, табл. 1.

УДК 004.021, 004.042, 519.688

Жушман В.В., Зайцева Т.А. **Комплексний підхід до розв'язання задачі взаємодії аб солютно жорсткого двозв'язного штампу та пружного півпростору** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.133 – 143.

Запропоновано комплексний підхід для розв'язання задач створення і аналізу матема тичних і комп'ютерних моделей контактної взаємодії абсолютно жорсткого циліндричного штампа з пружним півпростором. Побудовано скінченно елементні моделі взаємодії штампу з пружним півпростором у програмному комплексі ANSYS. Побудовано експертну систему за допомогою програмної системи CLIPS для ідентифікації форми штампу для випадків коли розміри зони контакту було змінено випадковим чином або за певним законом. Додатково розроблено програмне забезпечення на мові С++.

Бібл. 15, іл. 1.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

УДК 378.14:629.331

Рудик О. Ю., Диха О. В., Голенко К. Е. **Інноваційні підходи при викладанні дисциплін автомобільного спрямування** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.144 – 154.

Для навчальної дисципліни «Моделювання технологічних процесів підприємств автомобільного транспорту» у середовищах SolidWorks Simulation і Ansys Workbench сформовані основні принципи та положення автоматизованого проектування в області комп'ютерного моделювання агрегатів, вузлів і деталей транспортних засобів, а також пристосувань для їх ремонту (піднімачів, домкратів, стендів, знімачів тощо). Основна увага приділена теорії й практичному використанню методів скінченних елементів та набуття навичок у проєктуванні та розрахунках деталей автомобільного транспорту. Визначені обов'язкові елементи досліджень у SolidWorks та практичні навички моделювання різних режимів навантажень дорожніх та спеціальних транспортних засобів у Ansys Workbench. Для подовження ресурсу роботи конструктивних елементів і деталей автомобільного транспорту визначені методи їх відновлення та підвищення зносостійкості.

Бібл. 12, іл. 10.

УДК 004.6: 007.5: 004.94

Скалозуб В.В., Горячкін В.М., Клименко І.В., Терлецький І.А. **Дослідження інтелектуальних моделей управління на основі процедур класифікації невизначених даних зі встановленими вимогами достовірності результатів** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.155 – 171.

В роботі були досліджені можливості удосконалення нейронних мереж Хеммінга для класифікації даних у форматах нечітких величин і certainty factor CF(A). Також були визначені особливості удосконаленої математичної моделі завдань нечіткої класифікації на основі набору шаблонів ознак. Представлено структуру програмного комплексу інформаційної технології управління призначенням/відбором виконавців на основі класифікації наборів шаблонів із певних нечітких ознак, який використовує процедури редукції і каппа статистики.

Бібл. 10, іл. 8.

УДК 330.3

Басько А.В., Єршова Н.М. **Методологія поетапного проектування портфеля інвестиційних проектів** // Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 2(151). – Дніпро, 2024. – С.172 – 185.

Розглянуто методику поетапного проєктування портфеля інвестиційних проєктів. Виконана оцінка ефективності трьох проектів методом аналізу ієрархій. В результаті приведених розрахунків визначені ефективності кожного проекту з альтернатив задля розподілу ресурсів. Доведено, що проєктування портфеля інвестиційних проєктів є складним процесом, і виконувати його необхідно поетапно, використовуючи на кожному з них сучасні математичні методи прийняття рішень та технології.

Бібл. 8, іл. 9.

UDC 004.42:519.85

Moroz B.I., Kruhlyk A.S., Moroz D.M., Martynenko A.A. **Mathematical model of the rational organization of the information flows processing in aircraft delivery system** // System technologies. N 2(151) Dnipro, 2024. P.3 – 12.

Using of controlled discipline for the rational organization of information processing, that based on characteristics of value and aging, may have a good prospects for dispatching information requests/flows in delivery systems by aircraft vehicles. Especially if such a model is considered as one where the cost of delivery varies depending on the terms of such delivery. And the same time keeping a fleet of aircraft to cover all peak loads is economically impractical. The proposed mathematical model should allow to process of incoming messages in such a way as to ensure the execution of all requests within the time limit that defined for such a request.

Bibl. 6, ill. 3.

UDC 004.67 , 004.042, 004.622

Rudenko K., Bozhukha L. **Integration algorithms of recommendation in the mobile trade system** // System technologies. N 2(151) Dnipro, 2024. P.13 – 20.

The process of development, implementation and optimization of recommendation systems in the field of mobile commerce with the use of cloud services is considered. Machine learning methods and recommendation algorithms are used in the work in order to increase the personalization and accuracy of offers. The results of the work are recommended for increasing the personalization of services in business (especially in retail trade and online platforms) for increasing customer satisfaction and loyalty, as well as for improving marketing strategies and analyzing market trends.

Bibl. 11, ill. 0, table 0.

UDC 621.385.6.

Zymoglyad A.Yu., Guda A.I., Klishch S. **Hardware complex for measuring the power of UHF signals** // System technologies. N 2(151) Dnipro, 2024. P.21 – 32.

A hardware complex has been developed, which consists of 2 units, the UHF signal amplifier on the QPL9547 microcircuit and the demodulating logarithmic amplifier on the AD8319. The results of the research, which were presented in the table. 1 show how the UHF signal strength meter responded to a series of input signals of varying strength.

This complex can be used in laboratory studies of the output power of transmitting devices, for qualitative evaluation of antennas or antenna comparisons. Due to the amplifier at the input, it is possible to study the signal power up to –80 dBm. The described complex also has a fairly moderate price, compared to industrial analogues.

Bible 2, fig. 8.

УДК 004.42:519.25

Zemlianyi O., Baibuz O. **Methods for imputing missing data on coronary heart disease** // System technologies. N 2(151) Dnipro, 2024. P.33 – 49.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

We suppose several modifications of algorithms for iterative multiple imputation of mixed data represented by quantitative and qualitative features. Among the quantitative ones there are both continuous and discrete ones. Qualitative features include ordinal and binary. Two types of tests are used to analyse the approaches. In the first test, a dataset is artificially filled with gaps at random positions, imputed using different methods, and the mean square error and execution time of the algorithms are estimated. In the second test, bi nary classification models are trained on datasets imputed with different methods and the classification accuracy is compared. To convert qualitative features into numerical ones, we propose our own algorithms that work with missing data and allow for the conversion back to qualitative features. Two well known datasets on observations of coronary heart disease are considered.

Bibl. 15, pict. 0, tabl. 2.

UDC 004.942:

Poliakov O., Zhurakovskyi B. **Prototyping of control units for systems with industrial controllers** // System technologies. N 2(151) Dnipro, 2024. P.50 – 61.

A method of prototyping system control devices with software implementation of con trol algorithms in SFC, FBD, LD languages of the IEC 61131 3 standard are proposed. In the process of prototyping, the decomposition of project models of operating and controlling au tomata of the control device is performed, their implementation in the OpenPLC program envi ronment in the form of typical components of the organization of the control system pro grams, which are executed in the Arduino board. An example of the application of the pro posed methodology is presented, which can be useful for teaching programming of industrial controllers.

Bibl. 10, ill. 8.

UDC 621.771.068

Hrechanyi O.M., Vasilchenko T.O., Vlasov A.O., Vypryzhkin P.O., Yakymchuk D.I. **Simula tion modeling in the research of metallurgical equipment operation** // System technolo gies. N 2(151) Dnipro, 2024. P.62 – 75.

The technological regimes of winding hot rolled rods of different thicknesses are ana lyzed. The dependence of the roll diameter on the winding time has been established. The form of free oscillations occurring in the drive of the hot rolled strip winding is determined. The relationship between the elastic deformations due to the resistance forces of the electric motor and the rotating parts of the winder drum has been established.

Ref. 13, fig. 4, table 3

UDC 519.8

Karavaiev K.D. **On the necessary conditions for the existence of dense sequencing in the classical parallel sequencing problem** // System technologies. N 2(151) Dnipro, 2024. P.76 – 91.

We consider the classical problem of minimizing the length of a sequencing at a given width, in which the target sequencing is dense. The necessary conditions for the existence of a dense sequencing when using the branch and bound method, related to the limited capaci

ty of places and the possibility of filling them, are investigated. The obtained conditions were reduced to a single one, and efficient algorithms to test it in general and for graphs with all vertices on critical paths were proposed. In addition, the study also resulted in new improved lower bound estimates of the sequencing length and generalization of special sequencings in which the vertices occupy the leftmost and rightmost possible places, that take into account the sequencing width.

Bible. 15, ill. 2, tab. 0.

UDC 519.85:004.42

Koshel E. **Neural network assisted continuous embedding of univariate data streams for real time time series analysis** // System technologies. N 2(151) Dnipro, 2024. P.92 – 101.

The problem of embedding the univariate time series data into higher dimensional space is a ubiquitous problem that is relevant in many fields of study. The variety of methods that exists to solve this problem usually rely on some form of delay embedding that attempts to reconstructs the phase space of the underlying system by creating high dimensional vec tors from the time series data values that are far away from each other in time. We propose a more flexible method that uses some rules to create a linear combination of the time series values to arrive at a representation of the time series that more accurately represents the de pendence between the historical values of the series.

Bible 17, ill. 2, tab. 2.

UDC 669.15 194.53

Babachenko O, Podolskyi R., Kononenko G., Merkulov O., Safronova O., Dudchenko S. **Analysis of the influence of the cooling rate on the hardness of steel for railway rails of the pearlite and bainetic classes** // System technologies. N 2(151) Dnipro, 2024. P.102 – 112.

The process of operating vehicles determines the interaction of the wheel and rail. Traf fic safety and the main technical and economic indicators of track management and rolling stock largely depend on these parameters. The result is the effect arising from the rolling fric tion and especially from the friction of the wheel sliding on the rail during braking, relative to these changes there is a significant increase in the intensity of wear of the wheels of the roll ing stock, which, in turn, can lead to catastrophic results for the locomotive industry. Also, in the process of operation of the rail in most cases, defects are formed that have the character of a complicated state: its head is subject to wear, crumpling, cracking and buckling, contact fatigue damage can develop in the metal.

Bibl. 17, ill. 4, table. 2.

UDC 004.4:004.8

Hnatushenko V.V., Pavlenko I.V. **The use of generative artificial intelligence in soft ware testing** // System technologies. N 2(151) Dnipro, 2024. P.113 – 123.

The article discusses the possibilities of generative artificial intelligence when applied in software testing. Recognizing the rapid advancement of pre trained generative AI models

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

with transformers, such as OpenAI's ChatGPT, Google's Gemini, and others, the authors interro gate the current and future roles of these technologies in improving software testing meth odologies. The focus is on the benefits, challenges, and practical implications of utilizing generative AI in test generation, maintenance, and documentation.

Bibl. 16, Fig. 2.

UDC 51 74

Huk K.G., Sheveleva A.E. **A neural network approach to the identification of room occupationalness according to air parameters** // System technologies. N 2(151) Dnipro, 2024. P.124 – 132.

Different approaches to predicting air parameters and the number of people in the premises based on the results of real time air condition monitoring are considered. An ap proach to forecasting using an artificial neural network, which allows taking into account var ious factors, the nature and quantitative values of which cannot be accounted for and cannot be added to the model parameters, is proposed. The issue of choosing the architecture and parameters of the neural network for solving the forecasting problem has been studied. The possibility of using the proposed approach to determine the number of people in office prem ises is shown. Recommendations are given to improve the performance of the model.

Bibl. 4, il. 2, tabl. 1.

UDC 004.021, 004.042, 519.688

Zhushman V.V., Zaytseva T.A. **A complex approach to solving the problem of interac tion between a rigid double connected punch and an elastic half space** // System tech nologies. N 2(151) Dnipro, 2024. P.133 – 143.

A complex approach is proposed to solve the problems of creating and analyzing math ematical and computer models of contact interaction of a rigid cylindrical punch with an elas tic half space. Finite element models of interaction between a punch and an elastic half space are constructed in the ANSYS software package. An expert system was built using the CLIPS software system to identify the shape of the punch for cases where the dimensions of the contact zone were changed randomly or according to a certain law. Additionally, software was developed using the C++ language.

Bib. 15, Fig. 1.

UDC 378.14:629.331

Rudyk O. Yu., Dykha O. V., Golenko K. E. **Innovative approaches in teaching automo tive disciplines** // System technologies. N 2(151) Dnipro, 2024. P.144 – 154.

For the educational discipline "Modeling of technological processes of road transport enterprises" in the SolidWorks Simulation and Ansys Workbench environments, the basic prin ciples and provisions of automated design in the field of computer modeling of units, assem blies and parts of vehicles, as well as devices for their repair (lifts, jacks, stands, filmers, etc.). The main attention is paid to the theory and practical use of finite element methods and the acquisition of skills in the design and calculations of road transport details. Mandatory ele ments of research in SolidWorks and practical skills of modeling various load modes of road and special vehicles in Ansys Workbench are defined. In order to extend the service life of

structural elements and parts of road transport, methods of their restoration and increase in wear resistance are defined.

Bibl. 12, ill. 10.

UDC 004.6: 007.5: 004.94

Skalozub V.V., Horiachkin V.M., Klymenko I., Terlitskyi I.A. **Research of intellectual management models based on classification procedures of uncertain data with estab lished requirements of result reliability** // System technologies. N 2(151) Dnipro, 2024. P.155 – 171.

The work explored the possibilities of improving Hamming neural networks for data classification in the formats of fuzzy quantities and certainty factor CF(A). The features of the improved mathematical model of fuzzy classification tasks based on a set of feature templates were also determined. The structure of the information technology software complex for man aging the appointment/selection of executors is presented based on the classification of sets of templates from certain fuzzy features, which uses reduction procedures and kappa statis tics.

Bibl. 10, ill. 8.

UDC 330.3

Basko A.V., Ershova N.M. **Methodology of step by step design of investment project portfolio** // System technologies. N 2(151) Dnipro, 2024. P.172 – 185.

The method of step by step design of a portfolio of investment projects is considered. Evaluation of the effectiveness of three projects was carried out using the method of hierar chy analysis. As a result of the above calculations, the efficiency of each project was deter mined from the alternatives for the distribution of resources. It has been proven that the de sign of a portfolio of investment projects is a complex process, and it must be carried out in stages, using modern mathematical decision making methods and technologies for each of them.

Bible 8, ill. 9.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Системні технології

ЗБІРНИК НАУКОВИХ ПРАЦЬ

Випуск 2 (151)

Головний редактор: к.т.н., доц. Т.В. Селівьорстова

Технічний редактор та секретар збірки: к.т.н., доц. К.Ю. Островська

Здано до набору 19.02.2024. Підписано до друку 29.03.2024.
Формат 60х84 1/16. Друк - різограф. Папір типограф.
Умов. друк арк. – 14,143. Обл.–видавн. арк. –12,375.
Тираж 300 прим. Замовл. – 02/24

Український державний університет науки і технологій,
ННІ «Інститут промислових та бізнес технологій»,
кафедра Інформаційних технологій та систем: ІВК «Системні технології»
49600, Дніпро, а/с 493
http://journals.nmetau.edu.ua/index.php/st
Свідоцтво про державну реєстрацію друкованого засобу масової інформації:
Серія КВ № 8684 від 23 квітня 2004 рік

Редакційна колегія

Селівьорстова Тетяна Віталіївна ***(головний редактор)*** доцент, кандидат технічних наук	Український державний університет науки і технологій, ННІ «Інститут промислових та бізнес технологій», Україна
Алпатов Анатолій Петрович Член-кореспондент НАН України, професор, доктор технічних наук	Інститут технічної механіки НАНУ і ДКАУ, Україна
Архипов Олександр Євгенійович професор, доктор технічних наук	Національний технічний університет України «Київський політехнічний інститут» імені Ігоря Сікорського", Україна
Бабічев Сергій Анатолійович доцент, доктор технічних наук	Jan Evangelista Purkyně University in Ústí nad Labem Університет імені Яна Євангеліста Пуркінє, Усті над Лабем, Чеська Республіка
Білозьоров Василь Євгенович професор, доктор фізико-математичних наук	Дніпровський національний університет імені Олеся Гончара, Україна
Гече Федір Елемирович професор, доктор технічних наук	Ужгородський національний університет, Україна
Гуда Антон Ігорович ***(заст. головного редактора)*** професор, доктор технічних наук	Український державний університет науки і технологій, ННІ «Інститут промислових та бізнес технологій», Україна
Гнатушенко Вікторія Володимирівна ***(вчений секретар)*** професор, доктор технічних наук	Український державний університет науки і технологій, ННІ «Інститут промислових та бізнес технологій», Україна
Гнатушенко Володимир Володимирович професор, доктор технічних наук	Національний технічний університет «Дніпровська політехніка», Україна
Гожий Олександр Петрович професор, доктор технічних наук	Чорноморський національний університет імені П.Могили, Україна
Єрьомін Олександр Олегович професор, доктор технічних наук	Український державний університет науки і технологій, ННІ «Інститут промислових та бізнес технологій», Україна
Кіріченко Людмила Олегівна професор, доктор технічних наук	Харківський національний університет радіоелектроніки, Україна
Світличний Дмитро Святозарович професор, доктор технічних наук	Akademia Górniczo-Hutnicza Краківська гірничо-металургійна академія ім. С. Сташіца, Польща
Скалозуб Владислав Васильович професор, доктор технічних наук	Український державний університет науки і технологій, ННІ «Дніпровський інститут інфраструктури і транспорту»Україна
Хандецький Володимир Сергійович професор, доктор технічних наук	Дніпровський національний університет імені Олеся Гончара, Україна