НАУКОВІ ВІСТІ КПІ

Міжнародний науково-технічний журнал

№ 2 (139) 2025

Започаткований у вересні 1997 року

У номері:

Прикладна математика

Системний аналіз
та наука про дані

Матеріалознавство

Автоматизація,
комп'ютерно-інтегровані
технології та робототехніка

Адреса редакції:
КПІ ім. Ігоря Сікорського
просп. Берестейський, 37
Київ, 03056, Україна

Тел. (+38 044) 204-94-53
E-mail: n.visti@kpi.ua
https://scinews.kpi.ua

Засновник – Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”

Внесений до реєстру суб’єктів у сфері медіа з присвоєнням ідентифікатора медіа R30-02405
(рішення Національної ради з питань телебачення і радіомовлення № 1794 від 21.12.2023)

Згідно з наказами МОН України № 1643 від 28.12.2019, № 409 від 17.03.2020 та № 392 від 05.04.2023 журнал включено до категорії «Б» Переліку наукових фахових видань України з технічних наук (спеціальності – F1 Прикладна математика, F2 Інженерія програмного забезпечення, F3 Комп’ютерні науки, F7 Комп’ютерна інженерія, F4 Системний аналіз та наука про дані, G9 Прикладна механіка, G8 Матеріалознавство, G11 машинобудування (за спеціалізаціями), G12 Авіаційна та ракетно-космічна техніка, G3 Електрична інженерія, G4 Енерговиробництво (за спеціалізацією), G1 Хімічні технології та інженерія, G5 Електроніка, електронні комунікації, приладобудування та радіотехніка, G7 Автоматизація, комп’ютерно-інтегровані технології та робототехніка)

Рекомендовано Вченою радою Національного технічного університету України
“Київський політехнічний інститут імені Ігоря Сікорського”, протокол № 7 від 30.06.2025 р.

Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”
Свідоцтво про державну реєстрацію: серія ДК № 5354 від 25.05.2017 р., просп. Берестейський, 37, Київ, 03056

Підп. до друку 18.07.2025. Формат $60\times84^{1}/_{8}$. Папір офс. Гарнітура UkrainianTimesET.
Спосіб друку – електрографічний. Ум. друк. арк. 6,77. Обл.-вид. арк. 5,4. Наклад 30 пр. Зам. № 25-051.

Видавництво “Політехніка” КПІ ім. Ігоря Сікорського
вул. Політехнічна, 14, корп. 15, Київ, 03056
тел. (044) 204-81-78

KPI SCIENCE NEWS

International research journal

№ 2 (139) 2025

Founded in September, 1997

In the issue:

Applied Mathematics

System Analysis
and Data Science

Materials Science

Automation,
Computer-Integrated
Technologies and Robotics

Editorial office:
Igor Sikorsky Kyiv Polytechnic Institute,
ave. Beresteyskyi, 37,
Kyiv, 03056, Ukraine

E-mail: n.visti@kpi.ua
https://scinews.kpi.ua

Founder – National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Entered into the register of subjects in the field of media with the assignment of media identifier R30-02405 (decision of the National Council on Television and Radio Broadcasting of Ukraine No. 1794 dated 21.12.2023).

According to the orders of MES of Ukraine from 12.28.2019 no. 1643, from 03.17.2020 no. 409, and from 04.05.2023 no. 392 the journal is included in the «B» category of the List of scientific publications of Ukraine on technical sciences (specialities – F1 Applied Mathematics, F2 Software Engineering, F3 Computer Science, F7 Computer Engineering, F4 System Analysis and Data Science, G9 Applied Mechanics, G8 Materials Science, G11 Machinery Engineering (by specializations), G12 Aviation and Aerospace Technologies, G3 Electric Engineering, G4 Energy Production (by specialization), G1 Chemical Technologies and Engineering, G5 Electronics, Electronic Communications, Instrument Making and Radio Engineering, G7 Automation, Computer-Integrated Technologies and Robotics)

Advised by the Academic Council of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, protocol No 7 on 30.06.2025.

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”
Registration Certificate – ДК № 5354 on 25.05.2017, Ave. Beresteiskyi, 37, Kyiv, 03056

Signed for printing on 18.07.2025. Format 60×84 1/8. Text-weight paper. Font UkrainianTimesET.
Print. tech. – electrographic. Convent. printed sheets 6,77. Published sheets 5,4. Edition of 30 copies. Order No 25-051.

Publishing House “Politehnika”, Igor Sikorsky Kyiv Polytechnic Institute
Politekhnichna Str., 14, building 15, Kyiv, 03056
Tel.: (044) 204-81-78

ЗМІСТ

Прикладна математика

Маслянко П.П., Мірко С.С. Розмірково-пошуковий алгоритм пошуку правової позиції на множині судових рішень судочинства України із застосуванням великих мовних моделей 7

Романюк В.В. Ефективність методу Tall Array у зниженні розмірності наборів даних на основі методу головних компонентів..... 18

Системний аналіз та наука про дані

Данилов В.Я., Зарицький О.О. Модель класифікації ракових захворювань шкіри на основі інтеграції дискримінатора в архітектуру сходових нейронних мереж..... 30

Матеріалознавство

Варченко В.Т., Костюнік Р.Є., Кущев О.В., Терентьєв О.Є., Уманський О.П., Чевичелова Т.М. Метод підвищення зносостійкості композиційного плазмового покриття на основі нікель-графіту шляхом додавання мідних сплавів 39

Автоматизація, комп'ютерно-інтегровані технології та робототехніка

Матвієнко С.М., Тимчик Г.С. Аналіз схемотехнічних методів лінеаризації температурних характеристик NTC-термісторів і характеристик вимірювального каналу 45

Автори номера..... 57

CONTENTS

Applied Mathematics

Maslianko P.P., Mirko S.S. Rational search algorithm for searching for a legal position in a set of judicial decisions of the judiciary of Ukraine using large language models..... 7

Romanuke V.V. Tall Array method efficiency in dataset dimensionality reduction by principal component analysis..... 18

System Analysis and Data Science

Danilov V.Y., Zarytskyi O.O. A skin cancer classification model incorporating a discriminator into the ladder neural network architecture..... 30

Materials Science

Varchenko V.T., Kostyunik R.Ev., Kushchev O.V., Terentjev O.Ev., Umanskyi O.P., Chevichelova T.M. Method for increasing the nickel-graphite-based composite plasma coating wear resistance by adding copper alloys.... 39

Automation, Computer-Integrated Technologies and Robotics

Matvienko S.M., Tymchik G.S. Analysis of circuit design methods for linearizing the temperature characteristics of NTC-thermistors and the measurement channel characteristics..... 45

Contributors to the issue..... 57

ПРИКЛАДНА МАТЕМАТИКА

DOI: 10.20535/kpisn.2025.2.331288
УДК 519.2+004

П.П. Маслянко[1*], С.С. Мірко[1]
[1]КПІ ім. Ігоря Сікорського, Київ, Україна
[*]corresponding author: masliankop@gmail.com

РОЗМІРКОВО-ПОШУКОВИЙ АЛГОРИТМ ПОШУКУ ПРАВОВОЇ ПОЗИЦІЇ НА МНОЖИНІ СУДОВИХ РІШЕНЬ СУДОЧИНСТВА УКРАЇНИ ІЗ ЗАСТОСУВАННЯМ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ

Проблематика. Необхідність розробити розмірково-пошуковий алгоритм, який здатний автоматизувати пошук і генерацію контекстно-обґрунтованих правових позицій та алгоритм генерації правових позицій судових рішень з наявних судових рішень в судочинстві України, що є підготовчим етапом для розмірково-пошукового алгоритму.
Мета дослідження. Розробити концепт розмірково-пошукового алгоритму, який здатний автоматизувати пошук і генерацію контекстно-обґрунтованих правових позицій. Розробити й реалізувати алгоритм генерації правових позицій судових рішень, наявних у судовій системі України.
Методика реалізації. Розроблення концепту моделі розмірково-пошукового алгоритму із трьома основними компонентами: пошуку, формування альтернатив і генерації відповіді. Розроблення та реалізація алгоритму генерації правових позицій судових рішень, який здійснює пошук використаних правових норм, декомпозицію судових рішень у послідовність дій та генерацію правових позицій.
Результати дослідження. Розроблено математичну модель алгоритму генерації правових позицій судових рішень, який забезпечує автоматичне формування бази структурованих правових позицій. Здійснено валідацію алгоритму генерації правових позицій судових рішень методом порівняння результатів аналізу алгоритму та наявних правових позицій у Базі правових позицій Верховного Суду України (ВСУ).
Висновки. Обґрунтовано необхідність автоматизації пошуку і генерації контекстно-обґрунтованих правових позицій. Запропоновано концепт розмірково-пошукового алгоритму, який здатний автоматизувати пошук і генерацію контекстно-обґрунтованих правових позицій. Реалізовано алгоритм генерації правових позицій із застосуванням великих мовних моделей і методів Data Science, валідація якого засвідчила відповідність результатів наявним позиціям ВСУ.
Ключові слова: правова позиція; судове рішення; розмірково-пошуковий алгоритм; великі мовні моделі; NLP; Data Science; автоматизація судочинства.

Вступ

Проблема пошуку правових позицій у сфері судочинства України має комплексний міждисциплінарний характер і поєднує аспекти права, філософії, системного аналізу та науки про дані (Data Science). Одним із перспективних напрямів вирішення завдань генерації правових позицій у судочинстві України є системна інженерія таких систем і окремих застосунків генерації правових позицій на основі методів і моделей штучного інтелекту (ШІ). Щодо застосування ШІ у судочинстві, то тут ми спираємось на дослідження авторитетних науковців і практиків у сфері права, що систематизовані та проаналізовані у роботі судді Касаційного адміністративного суду у складі ВСУ, доктора юридичних наук, професора Яна Берназюка [1].

Можливість генерації правових позицій на основі ШІ для усіх учасників судочинства у конкретних справах на всіх стадіях судочинства суттєво прискорює процес підготовки судового рішення.

Пропозиція для цитування цієї статті: П.П. Маслянко, С.С. Мірко, “Розмірково-пошуковий алгоритм пошуку правової позиції на множині судових рішень судочинства України із застосуванням великих мовних моделей”, *Наукові вісті КПІ*, № 2, с. 7–17, 2025. doi: 10.20535/kpisn.2025.2.331288

Offer a citation for this article: P.P. Maslianko, S.S. Mirko, “Rational search algorithm for searching for a legal position in a set of judicial decisions of the judiciary of Ukraine using large language models”, *KPI Science News*, no. 2, pp. 7–17, 2025. doi: 10.20535/kpisn.2025.2.331288

Судочинство в Україні де-юре не ґрунтується на прецедентному праві, тим не менше на практиці правові позиції ВСУ активно застосовують нижчі інстанції [2]. Утім, такі правові позиції не завжди містять чітке обґрунтування або алгоритм, на основі якого були ухвалені, що ускладнює їхнє правильне застосування іншими судами.

Однією з особливостей функціонування ВСУ є можливість відступу від раніше ухвалених правових позицій – суд може або повністю змінити свій висновок, або уточнити його, використовуючи різні методи тлумачення правових норм [1]. Це створює додаткові труднощі для суддів, які мають орієнтуватися на актуальні правові позиції під час ухвалення рішень, оскільки до моменту розгляду справи позиція ВСУ може змінитися [3].

Інформатизація судових процесів ускладнюється необхідністю ефективного пошуку актуальних правових позицій та їх контекстного обґрунтування. Наявні системи пошуку часто не можуть оперативно і точно надати потрібну інформацію через обмежену функціональність, що значно ускладнює роботу адвокатів, суддів та інших учасників судового процесу [1, 2].

З погляду наукової та інженерної складових наявні системи для пошуку правових позицій побудовані переважно на основі алгоритмів «мішка слів» і метаданих, які надають обмежений функціонал для пошуку документів. Хоча такі підходи дозволяють систематизувати і структурувати інформацію, вони не забезпечують глибокого розуміння юридичних текстів, оскільки не можуть ефективно розпізнавати семантично предметні зв'язки між правовими позиціями. Під терміном «семантично предметні зв'язки» ми маємо на увазі семантичні конструкції правничої сфери й конкретні відношення між юридичними категоріями цієї предметної області, які застосовують для формалізації тих чи інших правових позицій. Відсутність релевантних алгоритмів для допомоги у формуванні висновку щодо застосування існуючих правових позицій або допомоги у формуванні нової правової позиції на основі існуючих обмежує здатність наявних систем адаптуватися під потреби користувачів і вимагає значних затрат часу й ресурсів для оброблення та аналізу судових рішень [2]. Отже, підходи й моделі обробки природної мови і ШІ є важливими для створення нових систем і застосунків, які здатні більш точно генерувати правові позиції та швидко адаптуватися до змін у правовому полі України.

Ця стаття спрямована на розробку алгоритму пошуку контекстно-обґрунтованої правової позиції, який буде ґрунтуватися на актуальних даних правових позицій ВСУ, законодавства і кодексів. Використання сучасних технологій Data Science та великих мовних моделей (Large Language Model (LLM)) дозволить розробити науково-обґрунтований алгоритм, який автоматизує пошук релевантних правових позицій на основі відкритих баз даних судових рішень України. Алгоритм стане важливим інструментом для прискорення та оптимізації процесів судочинства в Україні за рахунок більш прозорого прийняття контекстно-обґрунтованих судових рішень на всіх рівнях судової системи.

Постановка задачі

Метою цього дослідження є розроблення моделі розмірково-пошукового алгоритму пошуку контекстно-обґрунтованої правової позиції для законодавчо встановлених видів результатів судочинства України і розроблення моделі та алгоритму підготовки даних для роботи розмірково-пошукового алгоритму. Розмірково-пошуковий алгоритм та алгоритм підготовки даних мають враховувати актуальні правові позиції ВСУ, закони і кодекси, а також забезпечити систематичний чіткий процес пошуку релевантних правових позицій, застосовуючи сучасні технології Data Science і LLM.

Предметом цього дослідження є розроблення моделі розмірково-пошукового алгоритму пошуку контекстно-обґрунтованої правової позиції для законодавчо встановлених видів результатів судочинства України і реалізація алгоритму підготовки даних для розмірково-пошукового алгоритму, який генерує правові позиції на основі аналізу судових рішень, прийнятих судами України.

У цій статті під терміном «розмірково-пошуковий алгоритм пошуку правової позиції» ми розумітимемо науково обґрунтовану модель автоматизації процесу пошуку релевантних правових позицій на множині судових рішень, що охоплює всі потрібні класи сутностей процесу пошуку і відношення між ними. Призначення розмірково-пошукового алгоритму пошуку правової позиції – імплементація системи пошуку контекстно-обґрунтованих правових позицій на множині судових рішень судочинства України [2].

Модель розмірково-пошукового алгоритму NLP-системи пошуку контекстно-обґрунтованої правової позиції на множині судових рішень судочинства України

Розмірково-пошуковий алгоритм концептуальної моделі NLP-системи пошуку контекстно-обґрунтованої правової позиції на множині судових рішень судочинства України, що ґрунтується на модифікованому бізнес-профілі Еріксона–Пенкера [2], складається із трьох компонентів для формування і формалізації правової позиції (рис. 1).

Компонент пошуку дозволяє агенту розмірково-пошукового алгоритму знаходити правові позиції у зовнішній базі знань. Компонент формування альтернатив дозволяє агенту створити дерево альтернатив, за допомогою якого можна вирішити поставлену задачу користувача. Компонент відповіді створює відповідь з наявного дерева альтернатив і відсікає з нього невалідні альтернативи. Ці три компоненти використовують чотири незалежні LLM:

1) LLM альтернатив створює дерево альтернатив вирішення поставленої задачі на основі поставленої задачі та знайдених правових позицій;

2) LLM валідації перевіряє дерево альтернатив на відповідність і релевантність альтернатив до поставленої задачі;

3) LLM відповіді генерує відповідь на основі поставленої задачі та провалідованого дерева альтернатив;

4) LLM агрегації вибирає найкращу альтернативу із згенерованого дерева альтернатив, щоб повернути як результат роботи Компонента формування альтернатив.

Кожна з описаних LLM має доступ до бази структурованих правових норм для забезпечення їх коректного розуміння та застування.

Цикл зі створення, обрізки та нарощення дерева альтернатив створює ефект саморефлекторного мислення алгоритму.

Компонент «Пошук». Для пошуку агент передає у компонент пошуку метадані пошуку правових позицій і запит пошуку, який може складатися із ключових слів і фраз. Цей компонент повертає статус пошуку і кількість релевантних правових позицій до пошукового запиту, на основі чого агент ухвалює рішення щодо своїх наступних дій. Знайдені правові позиції передаються у компонент формування альтернатив після його виклику. Цей компонент

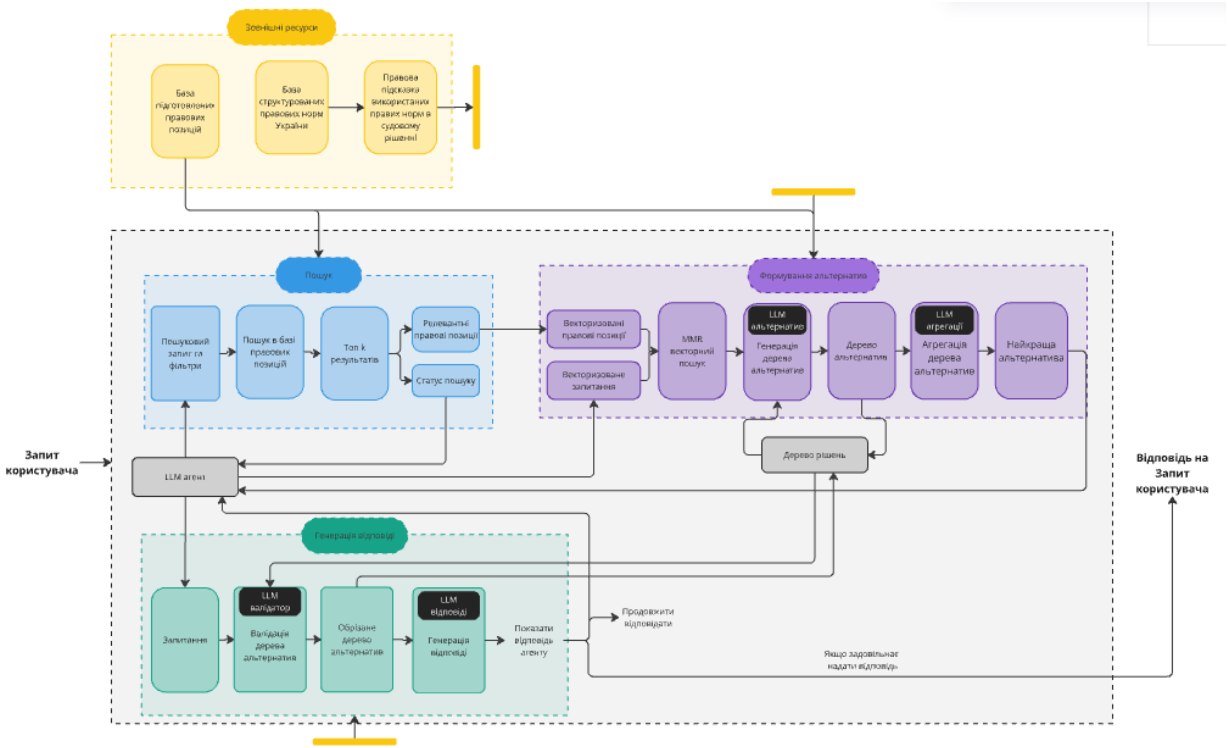


Рис. 1. Модель розмірково-пошукового алгоритму

викликається на початку роботи розмірково-пошукового алгоритму для його ініціалізації.

Компонент «Формування альтернатив». На першому етапі роботи цей компонент переранжировує список релевантних правових позицій для їх використання як LLM-альтернатив. Переранжування на основі Maximal Marginal Relevance (MMR) дозволяє реалізувати дві складові: 1) знайти найбільш релевантні правові позиції до поставленої задачі; 2) прибрати схожі правові позиції для створення якомога більш розгалуженого дерева альтернатив. LLM альтернатив на основі переранжированих правових позицій генерує дерево альтернатив або нарощує альтернативи на наявне дерево. Дерево альтернатив зберігається після кожної ітерації алгоритму для його наступного використання під час генерації остаточної відповіді. Компонент виконує агрегацію дерева альтернатив за допомогою LLM-агрегації, щоб вибрати найкращу альтернативу з дерева альтернатив. Отримана альтернатива повертається як результат роботи компонента, на основі поверхневої альтернативи агент приймає рішення щодо своїх наступних дій.

Компонент «Генерація відповіді». На початку роботи цей компонент валідує дерево альтернатив, прибираючи зайві нерелевантні гілки за допомогою LLM-валідації. Це дозволяє LLM-відповіді не використовувати невалідні й помилкові альтернативи під час формування остаточної відповіді. Провалідоване дерево альтернатив зберігається в алгоритмі для його подальшого нарощування в компоненті формування альтернатив. Далі LLM-відповіді генерують остаточну відповідь на поставлене завдання користувача. Остаточна відповідь повертається як результат роботи компонента до агента, який приймає рішення про завершення роботи й повернення відповіді користувачу, або про продовження генерації відповіді.

Підготовка даних для розмірково-пошукового алгоритму

Підготовка даних для роботи розмірково-пошукового алгоритму полягає в генерації правових позицій на основі раніше ухвалених судових рішень і збереженні згенерованих правових позицій у базі даних правових позицій судових рішень. Базу даних правових позицій судових рішень розмірково-пошуковий алгоритм використовує для пошуку релевантних правових позицій для генерації правової позиції у конкретній справі користувача.

Розмірково-пошуковий алгоритм оперує правовими позиціями замість повних текстів судових рішень. Це зроблено для того, щоб зменшити вплив нерелевантного та загального контенту судових рішень на прийняття рішень LLM, та для того, щоб збільшити кількість розглядуваних правових ситуацій. Генерація правових позицій судового рішення реалізується через послідовність етапів, зображених на рис. 2.

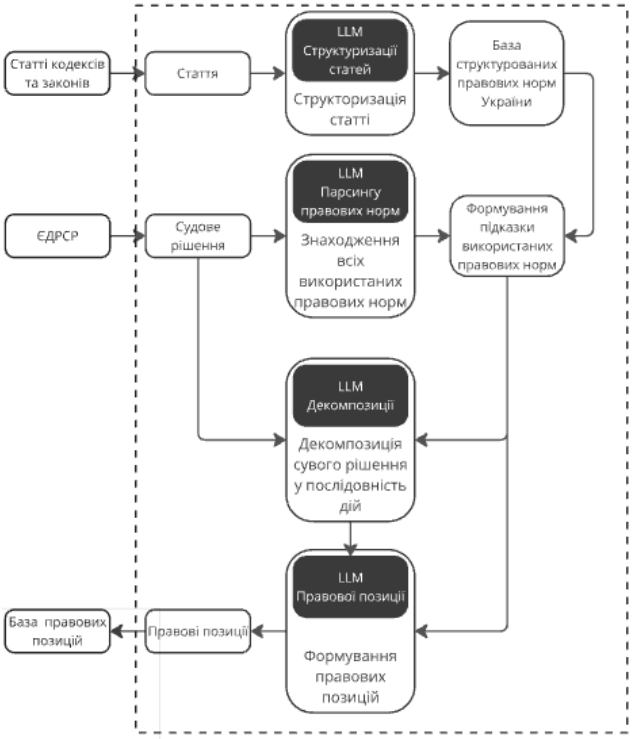


Рис. 2. Алгоритм генерації правових позицій судових рішень

На першому етапі формування правової позиції з кожного судового рішення знаходяться усі посилання на правові норми, які були в ньому використані, за допомогою LLM-парсингу правових норм. Посилання на правову норму складається з таких елементів:

– назви кодексу чи закону, в якому закріплено правову норму;

– номера статті, в якій описано правову норму;

– номера частини статті;

– номера пункту статті;

– номера підпункту статті.

Цей етап потрібний для формування підказки для LLM, оскільки у внутрішніх знаннях LLM не має достатніх знань щодо інтерпретації, наприклад, «ст. 234 ч. 2 КПК» або «стаття 234 частина 2 КПК». Через це треба передавати такі

знання разом із судовим рішенням, щоб LLM мала змогу коректно інтерпретувати й використовувати посилання на правові норми.

Щоб забезпечити змістові підказки, потрібно заздалегідь створити базу структурованих правових норм, де кожна частина статті правової норми може бути знайдена окремо від інших (тобто якщо в судовому рішенні вказано «ст. 234 ч. 2 КПК», то LLM отримає в підказці текст, який стосується 2-ї частини 234 статті Кримінально процесуального кодексу, а не увесь текст 234-ї статті).

На другому етапі сформована підказка буде використовуватися для декомпозиції судового рішення у послідовність дій ключових сутностей судового рішення. Кожний елемент послідовності включає:

– назву сутності, яка виконує дію – вказується суб'єкт, який є джерелом чи виконавцем конкретної дії у рішенні;

– опис дії – коротко формулюється, яку саме дію було вчинено;

– роз'яснення дії – деталізується, чому ця дія відбулася і які її наслідки;

– підстави для дії – зазначаються правові, фактичні або процесуальні підстави, що зумовили цю дію;

– результат дії – описується наслідок дії, який має юридичне або практичне значення;

– роз'яснення підстав для результату – обґрунтовуються правові, фактичні або процесуальні підстави, що зумовили такий результат.

Після знаходження послідовності дій судового рішення відбувається наступний, третій етап. Послідовність дій потрапляє у LLM правової позиції, де формується єдиний висновок про коректне застосування правових норм та їх взаємодію, – правова позиція. Правова позиція судового рішення складається з таких елементів:

– опису правової позиції – текст або концепція, що виражає конкретне правове тлумачення, думку або судове рішення щодо певної правової проблеми чи питання. Правова позиція показує, як конкретні правові норми застосовувалися або тлумачилися в судовому рішенні, а також визначає, яке рішення було прийняте щодо спору, і дає підґрунтя для подальшого застосування норм у подібних випадках;

– посилань на використані правові норми – конкретні посилання на статті законодавства (кодекси, закони), на основі яких сформовано правову позицію. Посилання на статті кодексів є важливими для того, щоб показати, які конкретні правові норми використовувалися для обґрунтування і прийняття судового рішення. Це забезпечує правову легітимність позиції та створює зв'язок між загальними правовими нормами та їх практичним застосуванням.

Судове рішення здебільшого містить декілька правових позицій – це пов'язано переважно з необхідністю комплексного вирішення правового спору, який охоплює різні аспекти фактичних обставин і застосування норм матеріального та процесуального права.

Правові позиції зберігаються у базі даних для їх пошуку й використання розмірково-пошуковим алгоритмом під час розгляду і вирішення правової ситуації користувача.

Математична модель алгоритму генерації правових позицій судових рішень

Алгоритм підготовки даних для розмірково-пошукового алгоритму можна розбити на кілька кроків, кожний з яких формалізує процес вилучення релевантних правових норм і створення правових позицій. Завдання алгоритму – забезпечити ефективний компактний набір даних для аналізу й використання LLM, зберігаючи всі релевантні аспекти правової інформації.

Алгоритм починає роботу зі знаходження у судовому рішенні J усіх використаних посилань на правові норми. Набір таких посилань утворює сутність ULNL (Used Legal Norms Links):

$$\text{ULNL} = \left\{ulnl_1, ulnl_2, \ldots ulnl_i \ldots ulnl_{m_{ln}}\right\}, \qquad (1)$$

де $ulnl_i$ – посилання на правову норму у форматі з визначенням усіх потрібних атрибутів для її ідентифікації;

m_{ln} – кількість знайдених посилань у судовому рішенні.

Отримання сутності ULNL можна записати таким чином:

$$\text{ULNL} = \text{LLM}\left(P_{used\ legal\ norms}(J)\right), \qquad (2)$$

де $P_{used\ legal\ norms}$ – інструкції щодо знаходження усіх посилань на правові норми у судовому рішенні; LLM – велика мовна модель, яка обробляє промт. Вона отримує цей промт і генерує відповідь, що ґрунтується на вказівках із промту.

Отримана сутність ULNL використовується для формування підказки T для правильної інтерпретації посилань на правові норми під час створення правої позиції. Формування підказки T можна записати так:

$$T = \sum_{ulnl_i \in ULNL} Search\left(ulnl_i, B_{legal\ norms}\right), \quad (3)$$

де $Search\left(ulnl_i, B_{legal\ norms}\right)$ позначає процес пошуку правової норми за посиланням $ulnl_i$ у базі правових норм $B_{legal\ norms}$.

Отже, T є об'єднанням усіх правових норм, знайдених за ULNL.

Маючи сформовану підказку T, алгоритм може розпочати аналіз тексту судового рішення. На цьому етапі алгоритм розкладає судове рішення на послідовність дій учасників процесу. Послідовність дій судового рішення позначають як CACD (Chain of Actions of a Court Decision).

CACD можна визначити як

$$\text{CACD} = \{a_i i = 1, 2, \ldots, m_a\}, \quad (4)$$

де m_a – кількість знайдених дій у судовому рішенні.

CACD використовують для структурованого подання судового рішення. Це знижує обсяг несуттєвої інформації й фокусує модель на основних правових аспектах, які забезпечують релевантне тлумачення. Процес знаходження правових норм із судового рішення можна записати формулою

$$\text{CACD} = LLM\left(P_{decompose}\left(J, T\right)\right), \quad (5)$$

де $P_{decompose}$ – інструкції щодо декомпозиції судового рішення у послідовність дій.

Після декомпозиції судового рішення у CACD формується правова позиція Legal position (LP), яка узагальнює висновок, що випливає з аналізу правових норм і підстав їх застосування у контексті конкретного судового рішення.

Модель правової позиції судового рішення

Правова позиція LP має включати:

– опис правової позиції pd – текстове або концептуальне формулювання, яке пояснює, як правові норми застосовувалися для вирішення конкретного питання;

– посилання на правові норми pl – список посилань на правові норми, що забезпечують юридичну підставу правової позиції.

Судове рішення може складатися з декількох правових позицій, тобто його можна записати формулою

$$J = \left\{LP_i \mid i = 1, 2, \ldots, m_{lp}\right\}, \quad (6)$$

де m_{lp} – кількість правових позицій у судовому рішенні J.

Формально правову позицію LP_i для судового рішення J визначають за формулою

$$LP_i = \left(pd_i, pl_i,\right), \quad (7)$$

де $pd_i = \sum_{j=1}^{m} \{a_j \mid a_j \in LP_i\}$ генерується агрегуванням та узагальненням дій a_j із CACD, які належать до правової позиції LP_i; $pl_i = \{l_j \mid l_j \in pd_i\}$ – множина всіх унікальних посилань на правові норми, які використовувалися для генерації pd_i.

Генерацію правових позицій можна описати формулою

$$\left\{LP_i \mid \text{i} = 1, 2, \ldots, m_{lp}\right\} = \\ = LLM\left(P_{legal\ position}\left(CACD, T\right)\right), \quad (8)$$

де $P_{legal\ position}$ – інструкції щодо генерації правових позицій судового рішення з послідовністю дій учасників судового рішення.

Таким чином, алгоритм формує базу судових рішень та їх правових позицій, що визначається формулою

$$B = \left\{\left(J_k, \left\{LP_1^k, LP_2^k, \ldots LP_{mk}^k\right\}\right.\right. \\ \left.\left.\text{CACD}_k, \text{ULNL}_k, meta\left(J_k\right)\right), \mid k = 1, 2, \ldots t\right\}, \quad (9)$$

де J_k – унікальне судове рішення; $\left\{LP_1^k, LP_2^k, \ldots LP_{mk}^k\right\}$ – множина правових позицій, які містить J_k; CACD_k – послідовність дій судового рішення J_k; ULNL_k – посилання на правові норми із судового рішення J_k; $meta\left(J_k\right)$, – метадані рішення J_k; t – загальна кількість унікальних судових рішень у корпусі.

Базу судових рішень та їх правових позицій використовує розмірково-пошуковий алгоритм, який може швидко ідентифікувати релевантні правові позиції для правових ситуацій користувачів, оцінюючи їх значущість і корисність у кожному конкретному випадку.

Верифікація алгоритму

У табл. 1 наведено перелік компонентів системи та відповідних великих мовних моделей LLM, які застосовуються для реалізації функцій компонентів. Основне призначення цієї таблиці – забезпечити прозорість та обґрунтованість вибору моделей LLM для кожного компонента

***Таблиця* 1.** Вибір LLM для компонентів алгоритму підготовки правових позицій судових рішень

Компонент	LLM структуризації статей	LLM структуризації статей	LLM декомпозиції	LLM Правової позиції
Модель LLM	GPT/4o-mini	GPT/4o-mini	GPT/o3	GPT/o3-mini
Посилання на модель	https://platform.openai.com/docs/models/o4-mini	https://platform.openai.com/docs/models/o4-mini	https://platform.openai.com/docs/models/o3	https://platform.openai.com/docs/models/o3-mini
Аргументація	Структуризація статей кодексів: основним завданням є розподіл тексту на окремі частини. Цей процес не потребує глибокого аналізу змісту або розширеного контексту, тому пріоритетом стає оптимізація витрат та ефективність обробки тексту	Пошук посилань на правові норми: метою є визначення та структурування правових норм за шаблоном. Оскільки процес не потребує глибокого розуміння змісту, акцент робиться на зниженні витрат і швидкому обробленні інформації	Декомпозиція судового рішення: це завдання має високий рівень складності, адже вимагає точного розуміння тексту й логічного аналізу послідовності дій. Найкраще підходять моделі, які спеціалізуються на аргументації та наукових задачах, що забезпечують високу точність результатів	Перетворення судового рішення у правову позицію: завдання потребує чіткого розуміння тексту і вміння формувати аргументовані відповіді. Утім, оскільки вхідні дані вже структуровані, можна використовувати менш складну модель без втрати якості результату

системи. Зокрема, таблиця містить інформацію про найменування компонентів, використані моделі LLM, посилання на технічну документацію цих моделей, а також аргументацію вибору кожної моделі залежно від специфіки завдання.

Вибір моделей для кожного компонента системи грунтується на характеристиках моделей і вимогах до завдань. Для завдань структуризації статей і пошуку правових норм було обрано моделі GPT/4o-mini, що характеризуються високою швидкістю обробки та оптимізованими витратами ресурсів. Це обумовлено тим, що зазначені завдання не вимагають глибокого семантичного аналізу тексту, а натомість потребують ефективної обробки та структурування інформації.

Для компонентів, що передбачають виконання більш складних завдань, таких як декомпозиція судового рішення і формування правової позиції, використовують моделі GPT/o3 та GPT/o3-mini. Ці моделі забезпечують вищий рівень розуміння тексту і здатність до аргументованих відповідей, що є критичним для зазначених завдань. Вибір моделі GPT/o3 для декомпозиції обумовлений необхідністю глибокого аналізу тексту і формування послідовності дій, а GPT/o3-mini використовують для формування правової позиції, оскільки оброблені дані дозволяють застосовувати спрощену модель без втрати якості результату.

Валідація алгоритму

Для валідації алгоритму підготовки правових позицій ми згенерували правову позицію для судового рішення із Бази правових позицій ВСУ (https://lpd.court.gov.ua/document/2065). Правова позиція у зазначеній базі має вигляд, показаний на рис. 3.

Ця правова позиція не відповідає висновку справи, оскільки виглядає не як чітке правило, яке пояснює правильне використання правових норм і їх взаємодію для схожих судових рішень, а як загальна порада без чітко визначеного формулювання, що ускладнює однозначну інтерпретацію й використання цієї правової позиції.

На першому етапі формування правової позиції із судового рішення ми отримали список використаних правових норм:

КЗпП ст. 36, КЗпП п. 2 ч. 1 ст. 36, КЗпП ст. 40, КЗпП ст. 235, ЦПК ч. 2 ст. 389, ЦПК ч. 3 ст. 3, ЦПК ч. 1 ст. 402, ЦПК ч. 1 ст. 263, ЦПК ст. 4, ЦПК ст. 12, ЦПК ст. 13, ЦПК ст. 175, КЗпП ч. 2 ст. 23, КЗпП ч. 1 ст. 39-1, КЗпП ч. 2 ст. 39-1, ЦПК п. 2 ч. 1 ст. 409, ЦПК п. 1 ч. 3 ст. 411, ЦПК ст. 400, ЦПК ст. 409, ЦПК ст. 411, ЦПК ст. 416

Рис. 3. Правова позиція з Бази правових позицій Верховного Суду

На основі цього списку було сформовано підказку щодо використаних правових норм. Оскільки ця підказка завелика, щоб її викласти повністю у цій статті, наведемо лише її частину:

частина 2 статті 23 кодексу КЗпП

Строковий трудовий договір укладається у випадках, коли трудові відносини не можуть бути встановлені на невизначений строк з урахуванням характеру наступної роботи, або умов її виконання, або інтересів працівника та в інших випадках, передбачених законодавчими актами.

частина 1 статті 39-1 кодексу КЗпП

Якщо після закінчення строку трудового договору (пункти 2 і 3 статті 23) трудові відносини фактично тривають і жодна із сторін не вимагає їх припинення, дія цього договору вважається продовженою на невизначений строк.

частина 2 статті 39-1 кодексу КЗпП

Трудові договори, що були переукладені один чи декілька разів, за винятком випадків, передбачених частиною другою статті 23, вважаються такими, що укладені на невизначений строк.

На другому етапі відбувається декомпозиція судового рішення у послідовність дій учасників судового рішення. Послідовність дій учасників судового процесу є ключовими фактами зі справи, що є достатніми для швидкого розуміння процесу прийняття судового рішення. Оскільки послідовність дій завелика, щоб її повністю відтворити у цій статті, наведемо лише її частину:

Сутність: Позивач ОСОБА_3

Дія: Подання позову до Приморського районного суду м. Одеси (травень 2017 року):

Роз'яснення дії:

Підстави: Позивач працював електромонтером ботанічного саду ОНУ ім. І. І. Мечникова з 01.04.2001 р.; наказом № 3377-18 від 10.12.2015 р. був звільнений 31.12.2015 р. за п. 2 ч. 1 ст. 36 КЗпП (закінчення строку договору). Вважає, що у 2007 р. між сторонами укладений безстроковий договір, тому підстав для звільнення не було.

Результат: Позовні вимоги: 1) поновити на посаді електромонтера; 2) стягнути середній заробіток за час вимушеного прогулу.

Аргументація результату: Посилання на ст. 36, 40, 235 КЗпП: відсутність законної підстави для звільнення, оскільки трудовий договір безстроковий, а наказ не містить підстав зі ст. 40 КЗпП.

Сутність: Позивач ОСОБА_3

Дія: Подання касаційної скарги до Верховного Суду:

Роз'яснення дії:

Підстави: Стверджує неправильне застосування матеріального права (ст. 23, 36, 39-1, 40, 235 КЗпП) та порушення процесуальних норм (ст. 389 ЦПК). Вказує, що: 1) заяв про строковий договір у 2007 р. не подавав; 2) характер його роботи постійний; 3) суди не дослідили підстави для строкового договору.

Результат: Просить скасувати постанову апеляційного суду й направити справу на новий розгляд.

Аргументація результату: Наявні підстави касаційного оскарження: неправильне застосування норм матеріального права та неповне дослідження доказів, що завадило встановити істотні факти.

Сутність: Верховний Суд

Дія: Правовий аналіз та оцінка доводів касаційної скарги:

Роз'яснення дії:

Підстави: Досліджено норми ст. 23 ч. 2, 36, 39-1, 235 КЗпП; ст. 389, 400, 402, 409, 411, 416 ЦПК; матеріали справи: заяви позивача, накази про прийняття та звільнення.

Результат: Констатовано істотні порушення норм матеріального й процесуального права судами попередніх інстанцій.

Аргументація результату: 1) Суди не встановили, чи дозволяв характер і умови роботи (електромонтер ботанічного саду) укладати саме строковий договір, чим порушили ст. 23 ч. 2 КЗпП. 2) Повторне річне переукладання договорів 2007-2015 рр. не досліджено щодо перетворення на безстроковий (ст. 39-1 ч. 2 КЗпП). 3) Предмет і підстава позову не вимагали визнання договору недійсним; суди помилково поклали на позивача обов'язок оскаржувати договір. 4) Невиконання судом обов'язку повно дослідити докази є

підставою, передбаченою п. 1 ч. 3 ст. 411 ЦПК, для скасування рішень.

На останньому, третьому етапі з послідовності дій учасників судового рішення ми формуємо правові позиції судового рішення. У результаті формування правових позицій алгоритм згенерував дві правові позиції.

1. При визначенні правомірності звільнення необхідно встановити, чи фактично було укладення строкового трудового договору, чи трудові відносини після закінчення встановленого строку фактично продовжувалися без вимоги однієї зі сторін про припинення. Якщо факт подальшої роботи свідчить про те, що договір перетворився на безстроковий (відповідно до частини 1 та частини 2 статті 39-1 Кодексу КЗпП), звільнення за п. 2 частини 1 статті 36 Кодексу КЗпП не має законних підстав.

2. Судовий процес має здійснюватися з обов'язковим і всебічним дослідженням усіх доказів, що мають значення для правильного вирішення справи. Недостатнє дослідження підстав для укладення строкового трудового договору, а також неналежна оцінка доказів щодо характеру виконуваної роботи порушують норми процесуального законодавства (зокрема, п. 1 частини 3 статті 411, статті 400, 402, 409 та 416 ЦПК) і є підставою для скасування попередніх рішень і направлення справи на новий розгляд.

Перша правова позиція стосується вирішення справи щодо незаконного звільнення працівника, оскільки строковий трудовий договір укладався кілька разів, у результаті чого він набув статусу безстрокового, а отже, звільнення є незаконним.

Друга правова позиція відповідає процесуальній частині судового рішення, яка визначає, що апеляційний суд не розглянув усіх доказів і таке рішення не може бути законим.

Отримані правові позиції відповідають висновку судової справи із вказуванням всіх потрібних правових норм і результатом їх застосування, що дає чітке розуміння, як ці правові норми можуть бути використані у схожих випадках у майбутньому.

Не менш важливими характеристиками є швидкість роботи алгоритму та його ціна.

Для обробки судового рішення і перетворення його на правові позиції алгоритм в середньому витрачає 1 хвилину 32 секунди. Середня швидкість читання людини варіюється від 200 до 400 слів за хвилину, для розрахунків взято середнє значення – 300 слів за хвилину. Оброблене судове рішення має 1873 слова, таким чином час, потрібний для прочитання потрібного документа, становить 6 хвилин 15 секунд, що суттєво перевищує час роботи алгоритму.

Ціна за обробку наведеного судового рішення склала $0,244 або 10 гривень 12 копійок.

Наведемо ще один приклад роботи алгоритму підготовки правових позицій для судового рішення з Бази правових позицій ВСУ (https://lpd.court.gov.ua/document/44869).

Правова позиція у Базі правових позицій ВСУ має вигляд, показаний на рис. 4.

Ця правова позиція вичерпно передає суть судового рішення і поставленого в ньому правового конфлікту. Утім, у цьому судовому рішенні були й інші порушення, окрім нерозгляду доказів, які не зазначені в Базі правових позицій ВСУ.

Наведемо ключові моменти з послідовності дій цього судового рішення:

Сутність: Засуджений ОСОБА_7 та захисник ОСОБА_6

Дія: Подання касаційних скарг:

Роз'яснення дії:

Підстави: Посилання на: недопустимість доказів через порушення ст. 477 КПК; відсутність складу ч. 5 ст. 191 КК та ч. 1 ст. 190 КК; відсутність обов'язкової економічної експертизи (ч. 2 п. 6 ст. 242 КПК); помилки у визначенні розміру шкоди й відповідача

Відновлення з'ясування обставин після повідомлення обвинуваченим нових обставин під час судових дебатів чи в останньому слові

КПК передбачає, що в разі якщо обвинувачений навіть в останньому слові повідомить про нові обставини, які мають істотне значення для кримінального провадження, суд за своєю ініціативою або за клопотанням учасників судового провадження відновлює з'ясування обставин, установлених під час кримінального провадження, та перевірку їх доказами, після завершення яких відкриває судові дебати щодо додатково досліджених обставин і надає останнє слово обвинуваченому. Подібне положення передбачене в КПК й у випадку повідомлення такої інформації під час судових дебатів

Рис. 4. Правова позиція з Бази правових позицій Верховного Суду

за цивільними позовами; незаконний склад апеляційної колегії.

Результат: Просили скасувати вирок і ухвалу та закрити провадження за п. 3 ч. 1 ст. 284 КПК як недоведене.

Аргументація результату: Захист стверджував, що кошти не були ввірені ОСОБА_7 особисто, правочини укладалися від імені ТОВ, а потерпілі фактично отримали квартири, тому шкоди немає.

Сутність: Верховний Суд

Дія: Оцінка вирішення цивільних позовів:

Роз'яснення дії:

Підстави: Захист надав документи про часткове відшкодування шкоди та набуття потерпілими права власності на квартири, але суди їх не оцінили.

Результат: Визнано істотне порушення вимог ст. 419 КПК; питання розміру шкоди потребує нового розгляду.

Аргументація результату: Навіть пізнє подання доказів зобов'язує суд відновити дослідження (ст. 364, 366 КПК); апеляційний суд мав можливість перевірити нові обставини, але проігнорував їх, порушивши принцип повноти судового розгляду.

Сутність: Верховний Суд

Дія: Підсумкова оцінка істотності порушень та ухвалення постанови:

Роз'яснення дії:

Підстави: Виявлено істотні порушення процесуального закону (ст. 419, п. 1 ч. 1 ст. 438 КПК), що вплинули на законність рішення апеляційного суду.

Результат: Касаційні скарги задоволено частково; ухвалу Сумського апеляційного суду від 18.03.2024 скасовано; справу направлено на новий апеляційний розгляд.

Аргументація результату: Недоліки кваліфікації та невирішеність цивільно-правових вимог не можуть бути усунуті в касаційній інстанції, тому необхідний новий апеляційний розгляд з урахуванням зауважень Суду.

У результаті формування правових позицій із судового рішення https://lpd.court.gov.ua/document/44869 алгоритм згенерував дві правові позиції.

1. Під час кваліфікації кримінального правопорушення у справах про шахрайські дії з участю службових осіб необхідно чітко розмежовувати категорії заволодіння та привласнення. Неповнота мотивувальної частини судових рішень та неточне визначення форми об'єктивної сторони злочину порушує вимоги щодо правильного застосування норм матеріального та процесуального права. Застосування норм має відбуватись із врахуванням як чіткого юридичного визначення складу злочину, так і дотримання принципів змагальності та об'єктивності судочинства, що передбачено нормами КПК (ст. 419, п. 1 ч. 1 ст. 438 КПК) та нормами касаційного провадження (ст. 433 КПК, ст. 400 КПК).

2. Повномасштабне відновлення процесуального дослідження доказів з метою встановлення належного розміру матеріальної шкоди та перевірки належності цивільно-правових рішень є необхідним для забезпечення реалізації прав потерпілих. Судовий розгляд цивільних позовів має ґрунтуватись на принципах повноти та всебічності судового розгляду (відповідно до ст. 364, 366 КПК), що включає аналіз нововиявлених доказів, навіть якщо вони подані пізніше, аби уникнути порушення права на справедливий судовий розгляд, встановленого ЦПК.

Друга правова позиція відповідає тій, що зазначена у Базі правових позицій ВСУ. Водночас алгоритм виявив ще одну правову позицію, яка стосується неправильної кваліфікації судом нижчої інстанції складу злочину обвинуваченого, що також відіграє ключову роль у скасуванні судового рішення.

Висновки

1. Запропоновано модель розмірково-пошукового алгоритму генерації контекстно-обґрунтованої правової позиції, що ґрунтується на актуальних даних правових позицій ВСУ, законодавстві та кодексах України, на основі сучасних технологій Data Science і LLM.

2. Розмірково-пошуковий алгоритм призначений для аналізу будь-яких судових рішень незалежно від виду судочинства, типу судового рішення і стадії розгляду справи та для генерації нових структурованих, контекстно-обґрунтованих правових позицій.

References

[1] Jan Bernaziuk, "Shtuchnyi intelekt ta systema pravosuddia Ukrainy: rezultaty spivpratsi u rotsi, shcho mynuv", *Portal suchasnoho prava Supreme Observer*. https://so.supreme.court.gov.ua/authors/934/shtuchnyi-intelekt-ta-systema-pravosuddia-ukrainy-rezultaty-spivpratsi-u-rotsi-sh%D1%81ho-mynuv

[2] P.P. Maslianko, S.S. Mirko, "Kontseptualna model NLP-systemy poshuku relevantnoi pravovoi pozytsii na mnozhyni sudovykh rishen sudochynstva Ukrainy", *Наукові вісті КПІ*, 2023, no. 1–4, pp. 71–85. doi: 10.20535/kpisn.2023.1-4.310302

[3] A.R. Kurbanova, “Pidstavy Vidstupu Verkhovnoho Sudu Vid Isnuiuchoi Pravovoi Pozytsii: Praktychni Pidkhody”. *Juridical scientific and electronic journal*, 2024, no. 3, pp. 508–512. URL: doi: 10.32782/2524-0374/2024-3/123 (date of access: 11.03.2025).

P.P. Maslianko, S.S. Mirko

RATIONAL SEARCH ALGORITHM FOR SEARCHING FOR A LEGAL POSITION IN A SET OF JUDICIAL DECISIONS OF THE JUDICIARY OF UKRAINE USING LARGE LANGUAGE MODELS

Background. The need to develop a reasoning search algorithm that is capable of automating the search and generation of context-based legal positions and an algorithm for generating legal positions of judicial decisions from existing judicial decisions in the judicial system of Ukraine, as a preparatory stage for a reasoning search algorithm.

Objective. Development of the concept of a reasoning search algorithm that is capable of automating the search and generation of context-based legal positions. Development and implementation of an algorithm for generating legal positions of judicial decisions available in the judicial system of Ukraine.

Methods. Development of the concept of a model of a reasoning search algorithm with three main components: search, formation of alternatives and generation of an answer. Development and implementation of an algorithm for generating legal positions of court decisions that searches for used legal norms, decomposes court decisions into a sequence of actions and generates legal positions.

Results. Mathematical model of the algorithm for generating legal positions of court decisions, which ensures automatic formation of a database of structured legal positions. Validation of the algorithm for generating legal positions of court decisions by comparing the results of the algorithm analysis and existing legal positions in the Database of Legal Positions of the Supreme Court.

Conclusions. The need for automating the search and generation of context-based legal positions is substantiated. The concept of a reasoning search algorithm is proposed, which is capable of automating the search and generation of context-based legal positions. An algorithm for generating legal positions using large language models and Data Science methods has been implemented, the validation of which has shown that the results correspond to the existing positions of the Supreme Court.

Keywords: legal position; court decision; reasoning search algorithm; large language models; NLP; Data Science; automation of judicial proceedings.

Рекомендована Радою
факультету прикладної математики
КПІ ім. Ігоря Сікорського

Надійшла до редакції
20 січня 2025 року

Прийнята до публікації
30 червня 2025 року

DOI: 10.20535/kpisn.2025.2.331279
UDC 519.7

Vadim V. Romanuke*
Vinnytsia Institute of Trade and Economics of State University of Trade and Economics, Vinnytsia, Ukraine
*corresponding author: romanukevadimv@gmail.com

TALL ARRAY METHOD EFFICIENCY IN DATASET DIMENSIONALITY REDUCTION BY PRINCIPAL COMPONENT ANALYSIS

Background. Exploratory data analysis has been extensively growing since the early 2000s. As of 2025, most real-practice datasets are classified as Big Data. The Big Data analytics workflow includes the data preprocessing step, which is the starting point of Big Data computational handling. At this step, the data are tried to get simplified as much as possible. The main paradigm is dimensionality reduction allowing simplifying and visualizing high-dimensional datasets. Principal component analysis (PCA) is a linear dimensionality reduction technique. The PCA can be sped up by applying Tall Arrays, if the data are stored on disk. The Tall Array PCA (TAPCA) computes principal components incrementally using a divide-and-conquer strategy.
Objective. The paper aims to determine when the TAPCA is factually efficient for dimensionality reduction. There are two numeric types to be studied: double and single precision.
Methods. To achieve the said objective, random large datasets are generated as matrices of a specified numeric type. Then computational time of the ordinary MATLAB PCA applied to generated matrices is measured. Next, computational time of converting in-memory arrays (generated matrices) into tall arrays is measured. Computational time of the TAPCA applied to those generated matrices, to which the PCA is applied before, is measured as well.
Results. A comparative analysis of the averaged computational times reveals that computational time complexity of both the PCA and TAPCA is rather polynomial than strictly quadratic or cubic. There is a nearly-hyperbolic margin, which alternatively could be called the TAPCA efficiency threshold, in a plane of the number of dataset observations and the number of dataset features, by which the TAPCA and the ordinary PCA take approximately the same time to compute principal components.
Conclusions. In computing principal components for dimensionality reduction of large datasets stored on disk, the Tall Array method becomes efficient by two parallel processor workers if a dataset has at least 5 to 6 million entries. The Tall Array method is more efficient on datasets with double precision whose efficiency threshold is nearly 6 million entries, whereas the efficiency threshold for datasets with single precision is between 5 to 15.2 million entries.
Keywords: dimensionality reduction; principal component analysis (PCA); Tall Arrays; efficiency threshold; double precision; single precision.

Introduction

Exploratory data analysis has been raised into likely the vastest and deepest domain encompassing various scientific fields of information technology, computer science, and applied mathematics for engineering, econometrics, sociology, bioinformatics, etc. [1, 2]. Meanwhile, datasets to be explored have been extensively or even exponentially growing since early 2000s. As of 2025, most real-practice datasets are classified as Big Data [3, 4]. The Big Data analytics workflow includes the data preprocessing step, which is the starting point of Big Data computational handling [5, 6]. At this point, before the most valuable and decisive Big Data statistics are computed, like mean and standard deviation values, the data are tried to get simplified as much as possible [7, 8]. The main paradigm is dimensionality reduction which allows simplifying and visualizing high-dimensional datasets.

Principal component analysis (PCA) is a linear dimensionality reduction technique, by which the dataset is linearly transformed into a new coordinate system such that the directions (principal components) sorted in descending order capture the largest variation in the data [3, 9]. The principal compo-

Пропозиція для цитування цієї статті: В.В. Романюк, "Ефективність методу Tall Array у зниженні розмірності наборів даних на основі методу головних компонентів", *Наукові вісті КПІ*, № 2, с. 18–29, 2025. doi: 10.20535/kpisn.2025.2.331279

Offer a citation for this article: Vadim Romanuke, "Tall Array method efficiency in dataset dimensionality reduction by principal component analysis", *KPI Science News*, no. 2, pp. 18–29, 2025. doi: 10.20535/kpisn.2025.2.331279

nents constitute an orthonormal basis in which different individual dimensions of the data are linearly uncorrelated [10, 11]. If the dataset is an $M \times N$ matrix $\mathbf{X} = [x_{mn}]_{M \times N}$ representing M observations of N-dimensional objects, where usually $M > N$, the PCA returns N principal components that are unit vectors [9]. The n-th vector, $n = \overline{2, N}$, is the direction of a line that best fits the data while being orthogonal to the first $n - 1$ vectors [11]. A best-fitting line minimizes the average squared perpendicular distance from the points to the line [3, 12].

The first principal component of a dataset with N variables represented by matrix $\mathbf{X} = [x_{mn}]_{M \times N}$ is constructed as a linear combination of the original variables, and it explains the most variance of the data. The first principal component is equivalently defined as a direction that maximizes the variance of the projected data. The second principal component explains the most variance in what is left once the effect of the first principal component is removed. Thus every next principal component explains less variance. Together the N principal components explain all the variance. The n-th principal component, $n = \overline{2, N}$, can be taken as a direction orthogonal to the first $n - 1$ principal components that maximizes the variance of the projected data. Alternatively, the PCA is defined as an orthogonal linear transformation on a real inner product space that transforms the data to a new coordinate system such that the greatest variance by some scalar projection of the data comes to lie on the first coordinate (being the first principal component), the second greatest variance lies on the second coordinate, and so on.

In practice, if the first N^* principal components explain sufficiently high amount of variance (say, typically, 90 % or above), $N^* \in \{\overline{1, N-1}\}$, the remaining $N - N^*$ principal components are ignored. Thus the initial dataset $\mathbf{X}$ is simplified to a dataset represented as an $M \times N^*$ matrix whose m-th row is a vector of values of the first N^* principal components calculated by plugging the entries of the m-th row of matrix $\mathbf{X}$ into the respective linear combinations of the N original variables. Clearly, the simplified dataset is easily visualized if $N^* = 2$ or $N^* = 3$.

If data are frequently updated (e. g., by adding new observations), it is a challenge to compute principal components timely, without delays or excessive memory occupation. The reason is the dataset is either constantly enlarged or updated, and thus the PCA slows down. The PCA can be sped up by applying Tall Arrays, if the data are stored on disk. Tall Arrays (operators and functions using the Tall Array approach) are a feature used in MATLAB for working with datasets that are too large to fit in memory [13, 14]. They allow performing computations on large datasets using familiar MATLAB functions and syntax without needing to manage low-level data processing tasks like chunking, paging, or out-of-core processing [14]. Tall Arrays handle large datasets that exceed the available memory by keeping only a relatively small portion of the data in memory at a time [13, 15]. This is called deferred evaluation, by which operations on an array are not computed immediately but are successively recorded and only executed when the data is explicitly requested [14, 16].

The Tall Array PCA (TAPCA) computes principal components incrementally using a divide-and-conquer strategy [17, 18]. First, the data are standardized by processing it in chunks. Then the covariance matrix is computed and normalized. This is done incrementally by summing outer products of chunks of data. The principal components are finally obtained by performing eigenvalue decomposition of the covariance matrix.

For standardizing each column of matrix $\mathbf{X}$, the mean and standard deviation of the matrix $\mathbf{X}$ column are computed as follows. The tall array (capitalization is used when the Tall Array approach is meant itself) representing the matrix $\mathbf{X}$ column is divided into smaller chunks that can fit into memory. The local sum, count of elements, and local sum of squares (sum of the squares of each element in the chunk) are calculated and stored temporarily for each chunk. And, once all chunks have been processed, the local sums and counts are aggregated to compute the global mean of the column as the ratio of the total sum of the local sums to the total number of elements. Then, in the second pass, for each chunk the local sum of squared differences from the global mean is calculated. These local sums are added up and the resulting global sum of squared differences is divided by the column length (i. e., the total sum of local counts) decreased by 1. The result is the global variance whose square root is the standard deviation. The column mean is subtracted from each entry of the column, whereupon every entry of the centered column is divided by its standard deviation.

Owing to the covariance matrix is smaller (its size is $N \times N$), the eigenvalue decomposition can be performed in memory. As for the rest, the TAPCA processes chunks sequentially, keeping memory usage relatively low while still being able to compute the correct coefficients of linear combinations of the original variables to return the principal components. Nevertheless, the sequential data chunking and processing may slow down computation of

principal components. Parallelization partially reduces the slowdown, but it depends on the dataset size and how it is chunked (although Tall Arrays chunk data automatically) [18, 19]. The number of parallel processor workers positively influences the TAPCA efficiency as well – Tall Arrays are generally expected to be more efficient at more parallel processor workers.

Problem Statement

The TAPCA is seemingly a very efficient method of dimensionality reduction, but there are two key aspects that must be taken into account. First, converting an in-memory (ordinary) array into a tall array takes some computational time, which ought to be added to the computational (operation) time taken by the TAPCA itself. Second, it is unclear when an in-memory array is regarded as large enough to apply the TAPCA rather than the (ordinary) PCA [6, 8, 12, 18]. This means that it is unclear when the TAPCA computes principal components faster than the (ordinary MATLAB) PCA does.

Therefore, the objective is to determine when the TAPCA is factually efficient for dimensionality reduction. The two numeric types to be studied are double and single precision. To achieve the objective, the following tasks are to be fulfilled:

1. To generate random large datasets as matrices of a specified numeric type.
2. To measure computational time of the ordinary MATLAB PCA applied to generated matrices.
3. To measure computational time of converting in-memory arrays (generated matrices) into tall arrays (which could be conditionally called tall matrices).
4. To measure computational time of the TAPCA applied to those generated matrices, to which the PCA is applied before.
5. To carry out a comparative analysis of the averaged computational times.
6. To discuss obtained results and findings from the comparative analysis.
7. To conclude on the TAPCA factual efficiency for dimensionality reduction of large datasets stored on disk.
8. To outline open questions and perspectives of further research.

Random Matrices

A random matrix $\mathbf{X} = \left[x_{mn}\right]_{M\times N}$ of a specified numeric type is generated by using the standard normal distribution having zero mean and unit variance [11, 20]. Hence, each entry x_{mn} in the matrix modeling a large dataset is a value of the normally distributed random variable with zero mean and unit variance. The numeric types to be studied are double and single precision for real numbers, so there will be two series of random datasets. For these two types the number of observations M is sequentially set to an element of set

$$\left\{\left\{10^2 \cdot k\right\}_{k=1}^{9}, \left\{10^3 \cdot k\right\}_{k=1}^{9}, \left\{10^4 \cdot k\right\}_{k=1}^{9}, \left\{10^5 \cdot k\right\}_{k=1}^{10}\right\} \quad (1)$$

and the number of initial variables (i. e., original features) N is sequentially set to an element of set

$$\left\{\overline{2, 100}\right\}. \quad (2)$$

At a given couple $\{M, N\}$ 100 random matrices $\mathbf{X} = \left[x_{mn}\right]_{M\times N}$ are generated at 100 different pseudo-random number generator seeds [12, 16, 18]. This is done to obtain statistically consistent and stable operation speed results upon averaging over those 100 generations.

Computational time

The computational time is measured on the dual-core processor Intel Core i5-7200U@2.50GHz in MATLAB R2018a. Firstly, computation of principal components is performed on the random matrix whose size is determined by the couple of integers from (1) and (2). The ordinary PCA computational time is denoted by $t_{\text{PCA}}\left(M, N, i\right)$, where i is the generation number at given $\{M, N\}$. Then, secondly, the random matrix is converted to a tall array, which can be called the random tall array (tall matrix). The time of the conversion is denoted by $\tau_{\text{TA}}\left(M, N, i\right)$. Thirdly, computation of principal components is performed on the random tall array. The TAPCA computational time is denoted by $t_{\text{TAPCA}}\left(M, N, i\right)$.

The averaged computational time of the PCA is

$$\overline{t}_{\text{PCA}}\left(M, N\right) = \frac{1}{100} \cdot \sum_{i=1}^{100} t_{\text{PCA}}\left(M, N, i\right) \quad (3)$$

at given $\{M, N\}$. The averaged computational time of the TAPCA without taking into account the array-to-tall-array conversion is

$$\overline{t}_{\text{TAPCA}}\left(M, N\right) = \frac{1}{100} \cdot \sum_{i=1}^{100} t_{\text{TAPCA}}\left(M, N, i\right). \quad (4)$$

If the conversion is regarded, then the TAPCA total time is calculated as

$$\overline{\theta}_{\mathrm{TAPCA}}\left(M,\,N\right)=\overline{\tau}_{\mathrm{TA}}\left(M,\,N\right)+\overline{t}_{\mathrm{TAPCA}}\left(M,\,N\right)$$
$$=\frac{1}{100}\cdot\sum_{i=1}^{100}\left[\tau_{\mathrm{TA}}\left(M,\,N,\,i\right)+t_{\mathrm{TAPCA}}\left(M,\,N,\,i\right)\right]. \quad (5)$$

Estimations (3)–(5) are fulfilled separately for double precision and single precision.

Analysis

The averaged computational time (3) of the PCA for double precision is shown as a mesh in Fig. 1, where some computational artifacts at

$$M\in\left\{7\cdot10^5,\,8\cdot10^5\right\}\ \text{and}\ N>30 \quad (6)$$

due to aliasing can be spotted. Obviously, the ravine at $M=9\cdot10^5$ cannot be a computational artifact. Computational time (3) is an increasing surface along each of its variables – both the numbers of observations and original variables. It roughly seems that surface (3) increases linearly. Nevertheless, the growth of computational time (3) is nonlinear, being close to quadratic or cubic. Averaged time $\overline{\tau}_{\mathrm{TA}}\left(M,\,N\right)$ of the array-to-tall-array conversion for double precision is shown in Fig. 2, where no significant trends are seen at all. Application of anti-aliasing techniques does not work to see any trend. Despite this, as the number of observations increases, the array-to-tall-array conversion time slowly grows.

Due to the array-to-tall-array conversion time for double precision does not exceed 27 milliseconds even for a million observations, averaged time (4) as a surface looks very resembling to TAPCA averaged time (5) that includes the array-to-tall-array conversion time (Fig. 3). Surface (5) in Fig. 3 has the same computational artifacts at (6) and the ravine at $M=9\cdot10^5$, although they appear to be a little bit softer. Some additional computational artifacts are visible at $M<10^5$ instead. At fewer original variables ($N=2$) and fewer observations ($M=7000$), a huge surge is seen. It is not caused by the array-to-tall-array conversion, though. Whereas the ordinary PCA handles a million-observation double-precision dataset with 100 variables within 9 seconds, the TAPCA takes no longer than just 1.9 seconds.

The averaged computational time (3) of the PCA for single precision shown as a mesh in Fig. 4 does not have any major computational artifacts. Compared to Fig. 1, this surface is much smoother. The surface increases similarly to that in Fig. 1. The

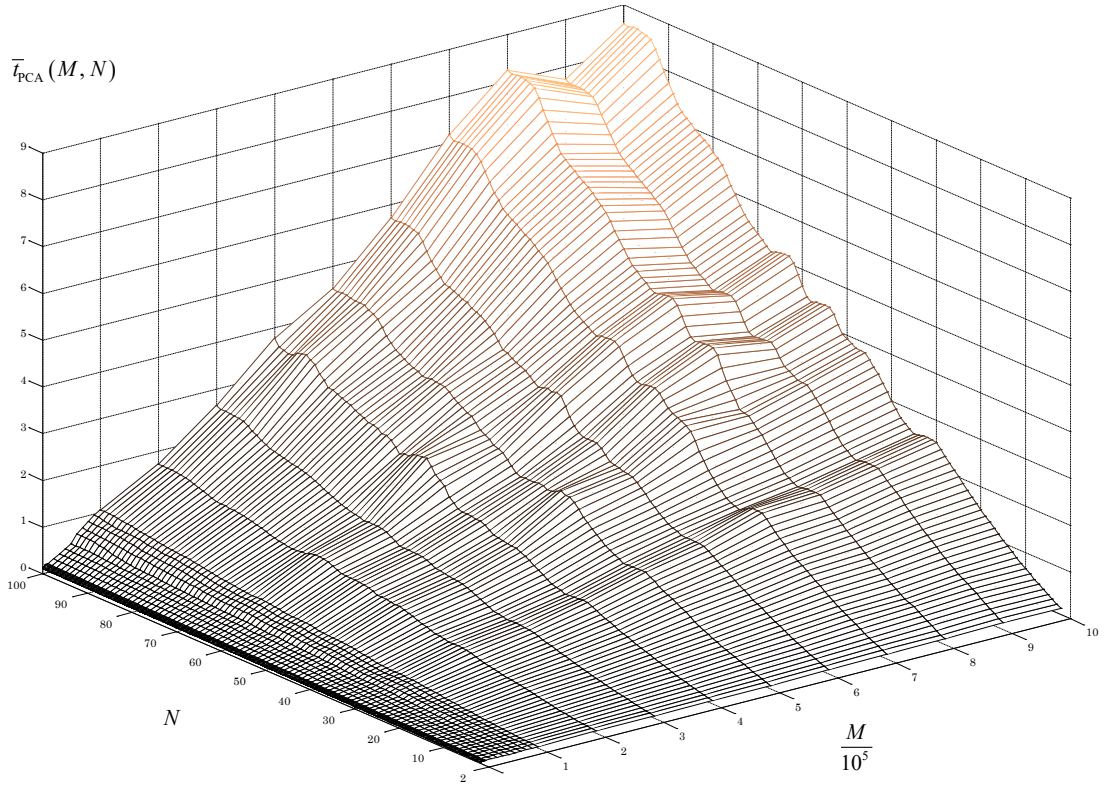


Fig. 1. The averaged computational time (3) of the PCA for double precision

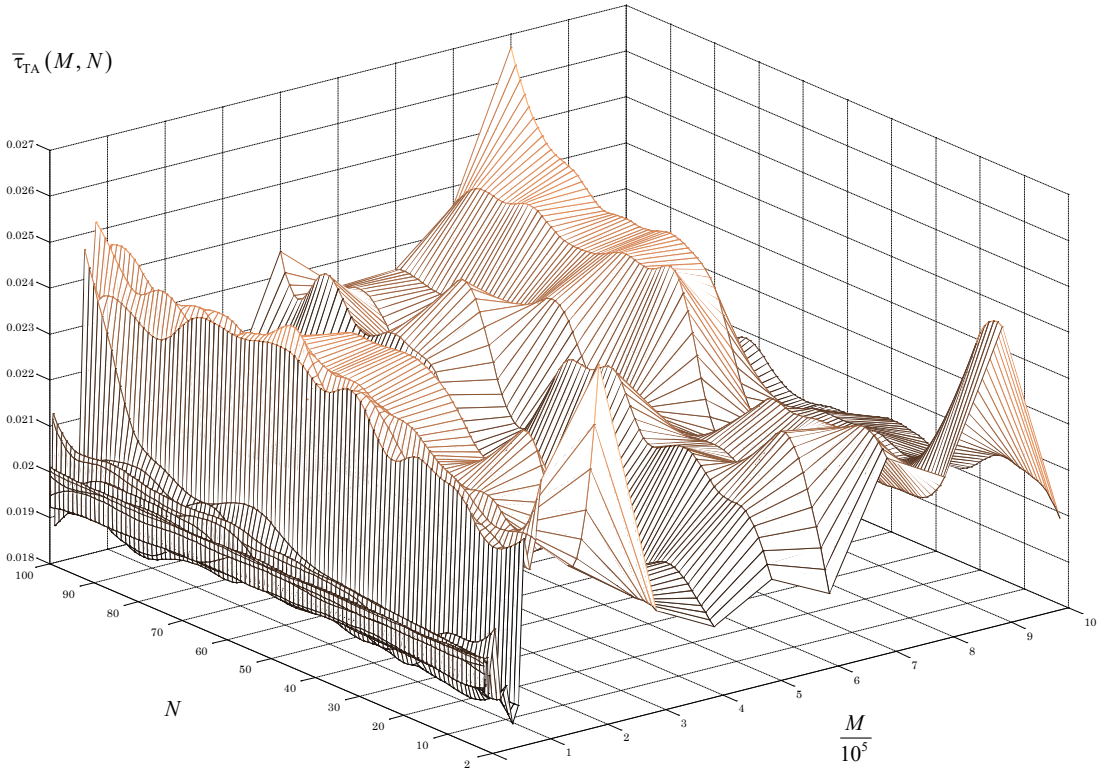


Fig. 2. The averaged time of the array-to-tall-array conversion for double precision

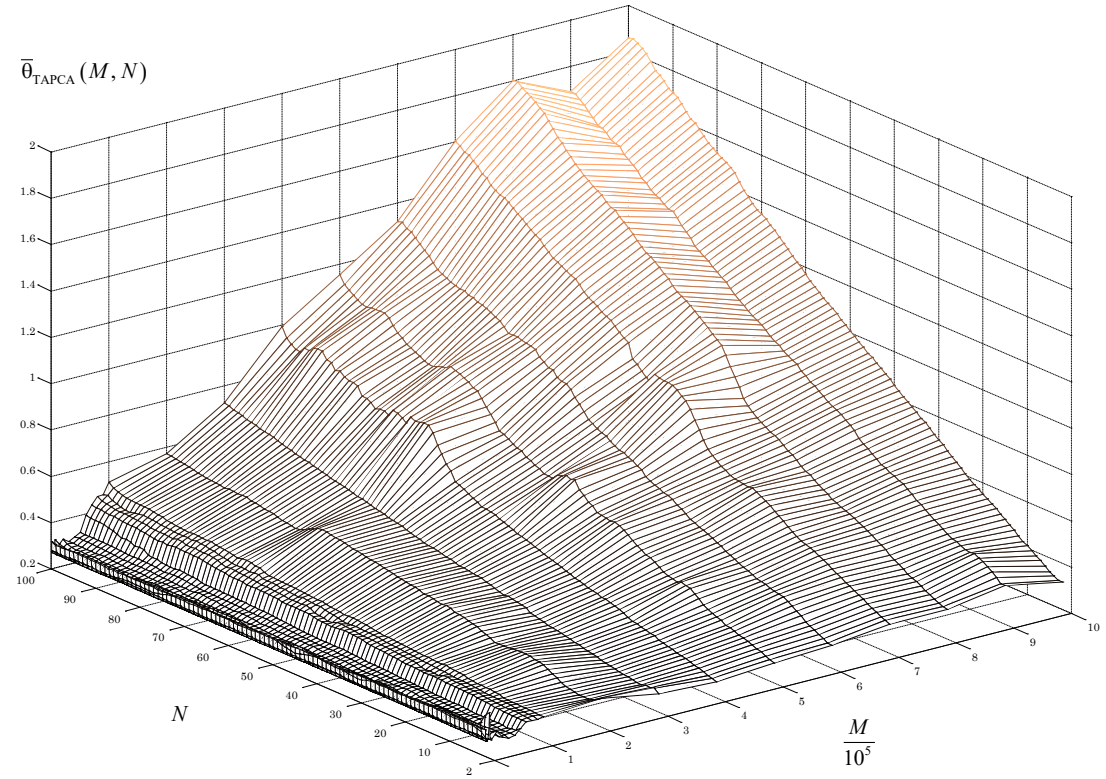


Fig. 3. The averaged computational time (5) of the TAPCA for double precision

ordinary PCA handles a million-observation double-precision dataset with 100 variables within 5.2 seconds. Overall, single-precision PCA is 1.5 to 2.1 times faster than double-precision PCA on average. Averaged time $\overline{\tau}_{\text{TA}}(M, N)$ of the array-to-tall-array conversion for single precision shown in Fig. 5 is faster as well, but the speedup is averagely 5 to 7 %. A single-precision matrix is converted into a tall array within 18 to 25 milliseconds. As the number of observations increases, the array-to-tall-array conversion time in Fig. 5 slowly grows, but the growth is more apparent than that in Fig. 2.

The averaged computational time (5) of the TAPCA for single precision is shown in Fig. 6. Compared to Fig. 3, the mesh is smoother at 10^5 observations and above, but it appears to have more non-linearities. Some computational artifacts are visible at fewer than 10^5 observations. The TAPCA takes no longer than just 1.3 seconds to compute 100 principal components of a million-observation single-precision dataset with 100 variables. Overall, single-precision TAPCA is 1.09 to 1.27 times faster than double-precision TAPCA on average.

Whichever precision or numeric type is, the TAPCA factual efficiency can be explored via ratio

$$\rho_{\text{TAPCA}}(M, N) = \frac{\overline{t}_{\text{PCA}}(M, N)}{\overline{\theta}_{\text{TAPCA}}(M, N)}. \qquad (7)$$

Clearly, the TAPCA is efficient if

$$\rho_{\text{TAPCA}}(M, N) > 1. \qquad (8)$$

The TAPCA efficiency for double precision is visualized in Fig. 7 presented as a plane view on the Cartesian product of sets (1) and (2) for abscissa and ordinate axes, respectively, where light color corresponds to (8), when the TAPCA is efficient; dark color is when the TAPCA is inefficient, i. e. inequality (8) is false. The TAPCA efficiency for single precision in Fig. 8 resembling that in Fig. 7 has a more dark color, i. e. single precision TAPCA has vaster area of inefficiency.

The TAPCA efficiency area exists beyond a nearly-hyperbolic margin (inside the hyperbola), both for double and single precision. Generally speaking, the TAPCA is inefficient at either fewer observations or fewer original variables (features). In more particular terms, it is inefficient at datasets having no more than about 50 000 observations. On the other side, the dataset with fewer than 10

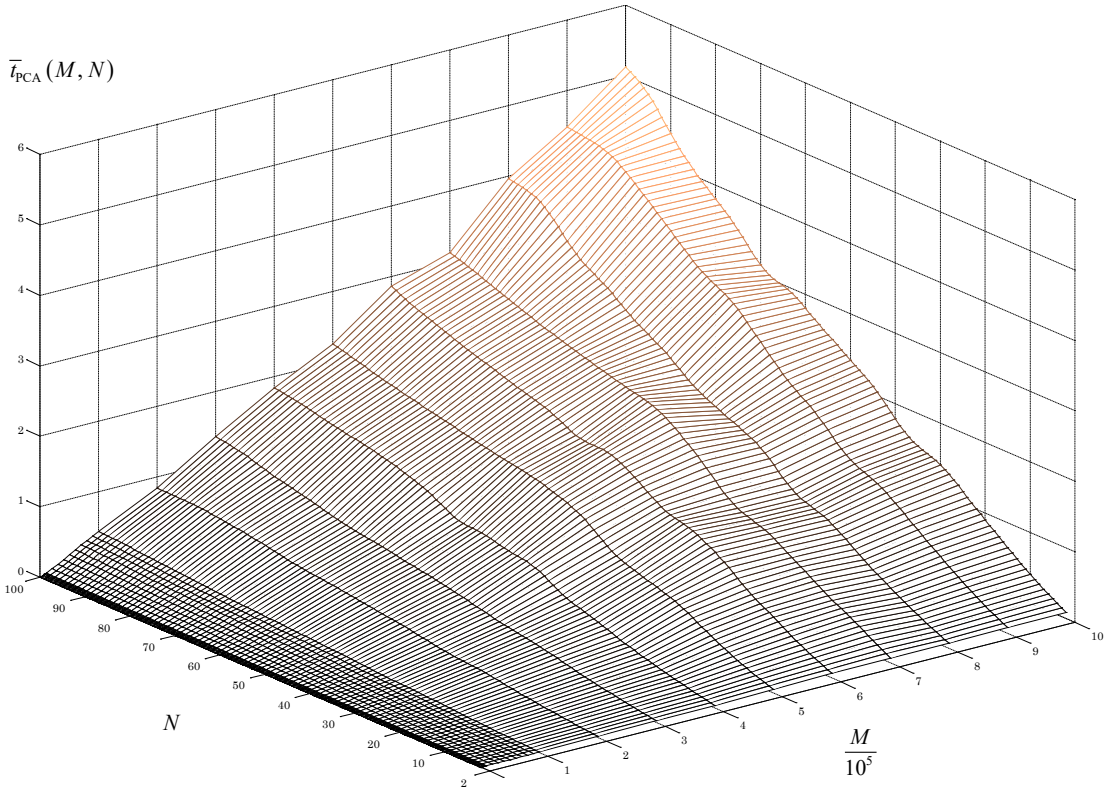


Fig. 4. The averaged computational time (3) of the PCA for single precision

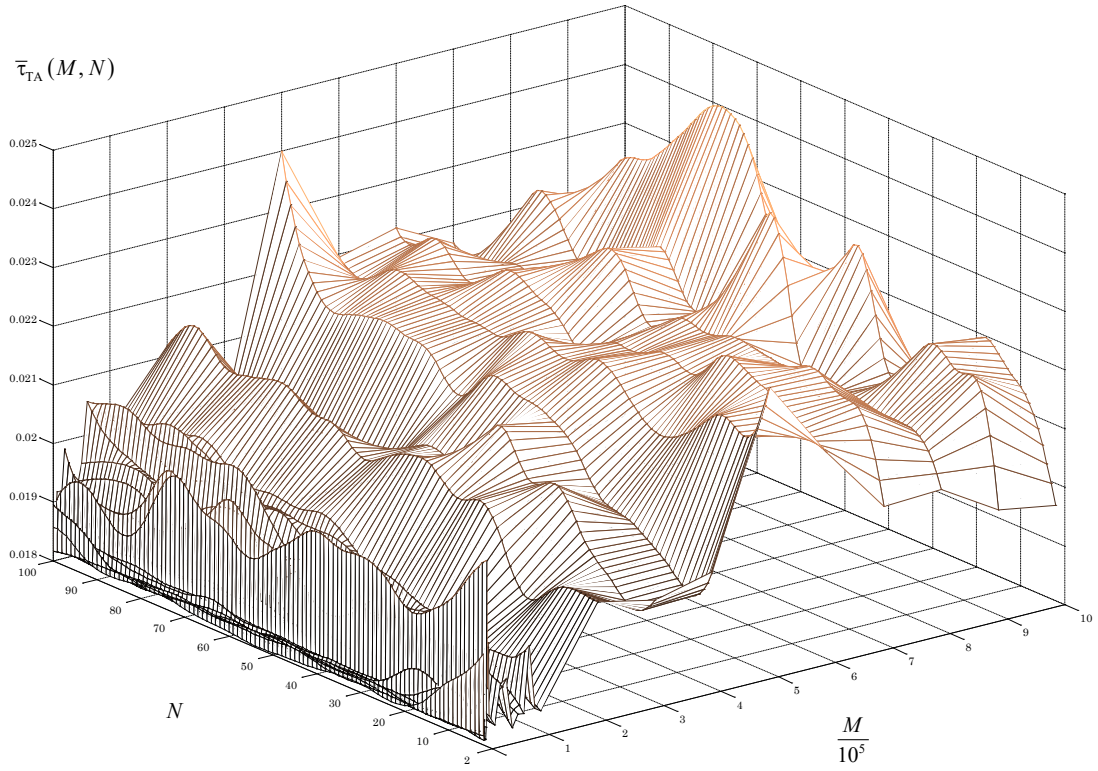


Fig. 5. The averaged time of the array-to-tall-array conversion for single precision

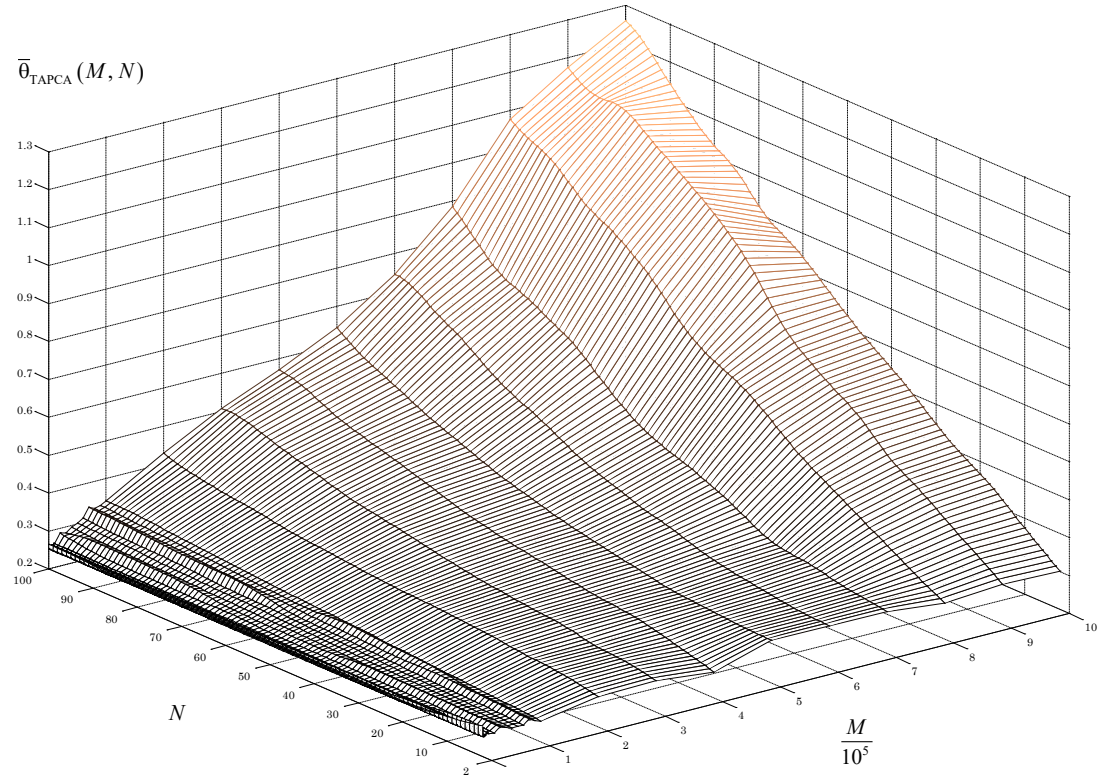


Fig. 6. The averaged computational time (5) of the TAPCA for single precision

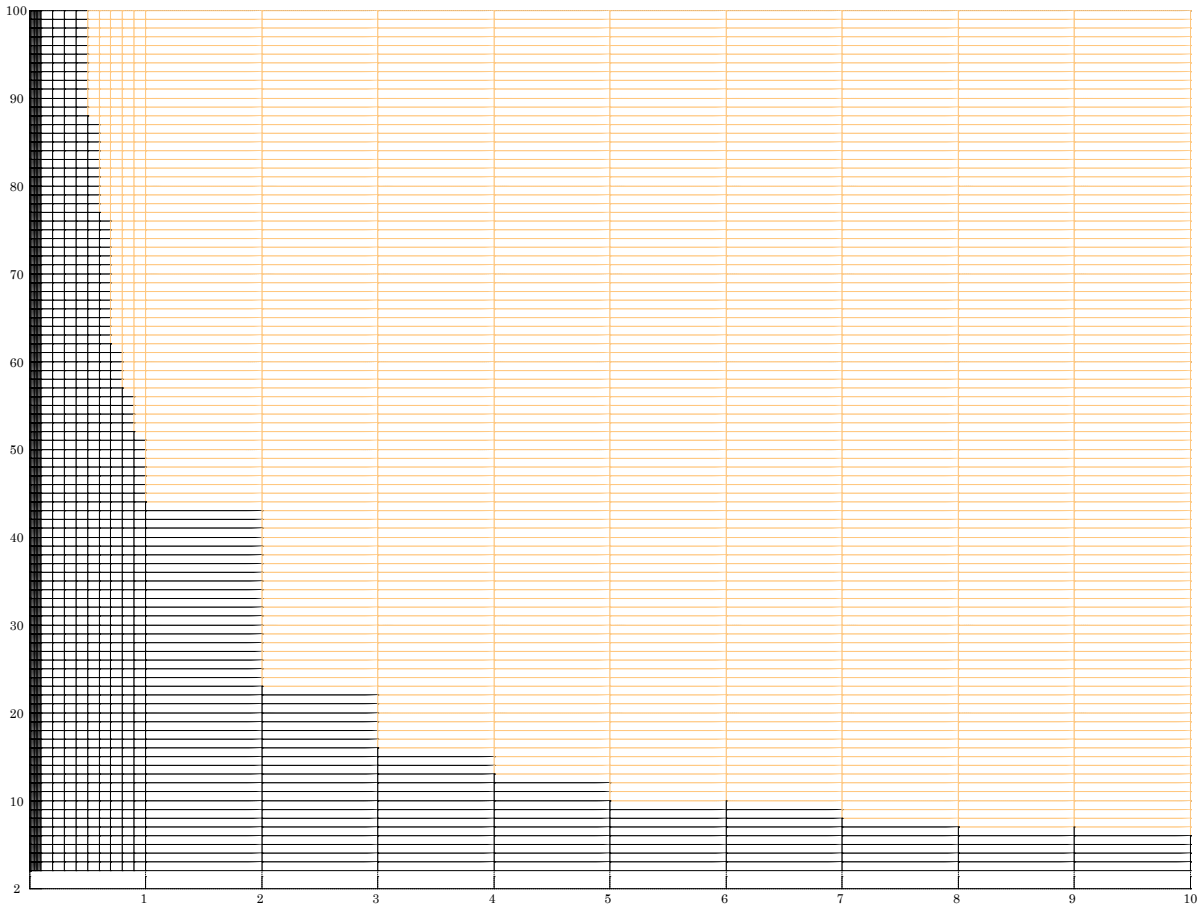


Fig. 7. Efficiency of the TAPCA for double precision on the Cartesian product of (1) and (2)

features is more efficiently handled by the ordinary PCA. The double precision TAPCA efficiency is approximately margined with hyperbola

$$N = \frac{4178585.7041}{M} + 1.661$$

$$\text{for } M \in \left[4 \cdot 10^4; 10^6\right]. \quad (9)$$

The single precision TAPCA efficiency is approximately margined with hyperbola

$$N = \frac{4919317.7503}{M} + 10.2581$$

$$\text{for } M \in \left[4 \cdot 10^4; 10^6\right]. \quad (10)$$

It is worth noting that the margin by (9) is about six times as more accurate than the margin by (10). It is also visible from Fig. 7 whose nearly-hyperbolic staircase margin is far less contorted than that in Fig. 8. An efficiency threshold can be deduced from (9) as the double-precision dataset size (the number of its entries, i. e. $M \cdot N$), at which applying the TAPCA is equivalent to applying the PCA, whereas principal components for larger datasets are better to compute by the Tall Array method. Thus, the double-precision dataset threshold is nearly 4.5 to 5.9 million entries. There is a similar finding for single-precision datasets whose threshold is nearly 5 to 15.2 million entries. The nearly-hyperbolic margins in Fig. 7 and 8 also allow concluding that the thresholds are likely to become lower for too “thin” datasets (having no more than 10 features or so) and datasets having 10^5 observations or so.

Discussion

The obtained results visualized in Fig. 1, 3, 4, 6–8 reveal what Tall Arrays are factually capable of and when they are efficient in computing principal components for dimensionality reduction of large datasets stored on disk. Computational time complexity of both the PCA and TAPCA is rather polynomial than strictly quadratic or cubic, although it

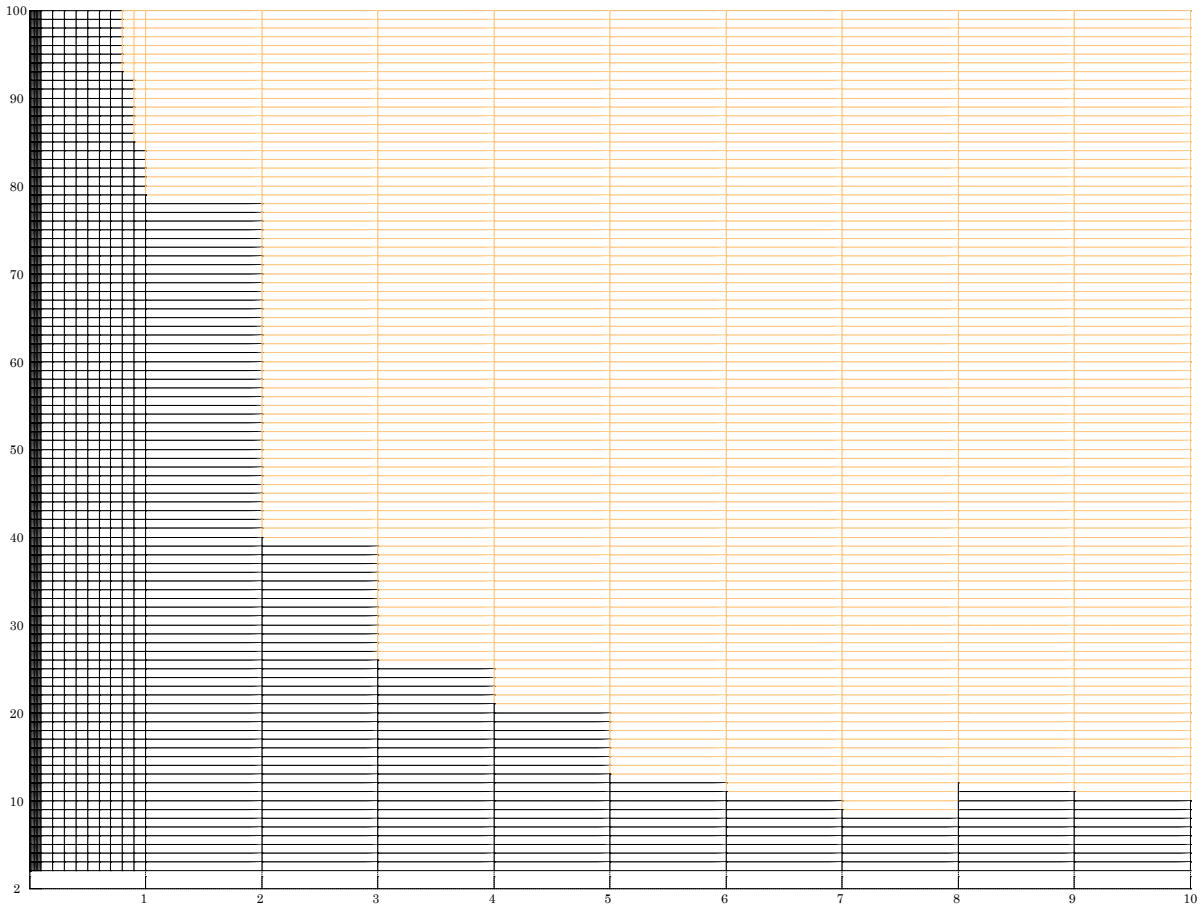


Fig. 8. Efficiency of the TAPCA for single precision on the Cartesian product of (1) and (2)

seems to be quasilinear. Nevertheless, computational time spans have been registered for the TAPCA by two parallel processor workers, so it is naturally expected that the TAPCA will be more efficient by more parallel processor workers. However, as the dataset size increases, whether in its number of observations or features, or both, the growth of ratio (7) reaches its saturation (see it in Fig. 9 for double precision) scarcely exceeding 5. Ratio (7) for the TAPCA factual efficiency for single precision looks similarly.

The nearly-hyperbolic margin, which alternatively could be called the TAPCA efficiency threshold, implies the size of a dataset (or the size of a Big Data instance) as matrix $\mathbf{X} = [x_{mn}]_{M \times N}$, by which the TAPCA and the ordinary PCA take approximately the same time to compute N principal components. For datasets whose size is above the threshold, the TAPCA is faster. Here, it is necessary to remember that starting parallel processor workers takes some time, so applying the TAPCA to smaller datasets, even if their size is above the threshold (but the size is close to the margin inside the hyperbola), is reasonable only for multiple times. Applying the TAPCA just once is reasonable if the dataset (say, of 10 features at most) cannot be loaded into the workspace. For instance, 10 principal components of a dataset as matrix $\mathbf{X} = [x_{mn}]_{550000 \times 10}$ of 550 000 observations and 10 features are computed by the PCA within 370.4 milliseconds, whereas the TAPCA takes about 375.5 milliseconds (i. e., it is 1.3789 % slower) if to count time spent on converting this matrix into a tall array. Without taking into account the array-to-tall-array conversion, the TAPCA takes about 353.1 milliseconds (which is 4.6733 % faster). So, the conversion taking here 22.4 milliseconds does not ruin the TAPCA efficiency (and thus it may be conditionally neglected) only if matrix $\mathbf{X}$ is converted into a tall array once and then the TAPCA is applied at least twice – to the tall array and its modification (such a modification is expected to be a subdataset of the initial dataset, rather than an expansion of the dataset).

A large dataset (a Big Data instance) is usually stored on disk by some privacy and security reasons.

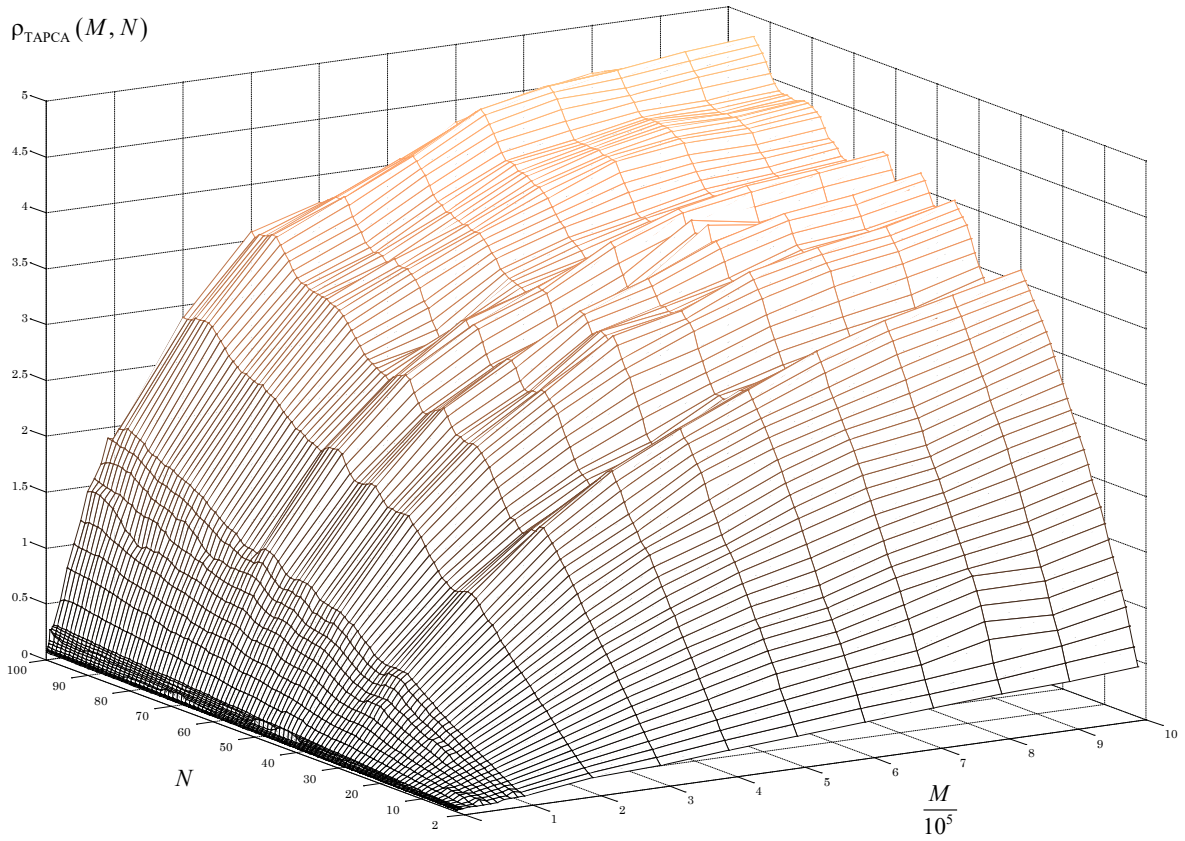


Fig. 9. Ratio (7) showing the TAPCA factual efficiency for double precision

This is when Tall Arrays become useful as they do not disclose factual data. In this comprehension, the array-to-tall-array conversion can be considered as a tradeoff (payment) for privacy and security along with simplified interoperability and manageability owing to dimensionality reduction by principal components.

Conclusions

In computing principal components for dimensionality reduction of large datasets stored on disk, the Tall Array method becomes efficient by two parallel processor workers if a dataset has at least 5 to 6 million entries. The Tall Array method is more efficient on datasets with double precision whose efficiency threshold is nearly 6 million entries, whereas the efficiency threshold for datasets with single precision is between 5 to 15.2 million entries. Both the thresholds may become lower as the number of dataset features is dropped below 10 or the number of observations does not exceed 10^5.

The presented research is nearly the worst-case scenario, in which only two parallel processor workers are deployed. As the number of workers increases, the TAPCA factual efficiency threshold is expected to drop further. The computational time complexity of both the PCA and TAPCA is polynomial, so the drop will be significant even for four parallel processor workers (it is also a commonly widespread computational architecture), let alone batch computation of principal components on computer clusters by using the Tall Array approach.

An open question is about the reasonability or efficiency of storing a large dataset on disk (prior to dimensionality reduction), when Big Data clusters like Hadoop or cloud-based solutions are available [21, 22]. The second open question, lying nearly in parallel to the mentioned one, is about efficiently using the MapReduce technique for dimensionality reduction by principal component analysis. Another open question is how to further speed up the PCA by implementing it (or TAPCA) on graphic processing units (GPUs). These questions are the perspective for further research. Furthermore, a general methodology of efficient computation of principal components for dimensionality reduction of large datasets should be formulated and justified.

References

[1] Ü. Demirbaga *et al.*, Big Data Analytics. Theory, Techniques, Platforms, and Applications, Springer, Cham, 2024. Retrieved from: doi: 10.1007/978-3-031-55639-5.

[2] A. Jamarani *et al.*, "Big data and predictive analytics: A systematic review of applications", *Artificial Intelligence Review*, 2024, Vol. 57, no. 176. Retrieved from: doi: 10.1007/s10462-024-10811-5

[3] I. Si-ahmed *et al.*, "Principal component analysis of multivariate spatial functional data", in *Big Data Research*, 2025, vol. 39, Art. no. 100504. Retrieved from: doi: 10.1016/j.bdr.2024.100504

[4] A. Meepaganithage *et al.*, "Enhanced Maritime Safety Through Deep Learning and Feature Selection", *Advances in Visual Computing. ISVC 2024. Lecture Notes in Computer Science*, Vol. 15047, Springer, Cham, 2025, pp. 309–321. Retrieved from: doi: 10.1007/978-3-031-77389-1_24

[5] W.J. Ewens and K. Brumberg, Introductory Statistics for Data Analysis, Springer, Cham, 2023. Retrieved from: doi: 10.1007/978-3-031-28189-1

[6] S. Akter and S.F. Wamba, Handbook of Big Data Research Methods, Edward Elgar Publishing, 2023. Retrieved from: doi: 10.4337/9781800888555

[7] V.V. Romanuke, "Fast Kemeny consensus by searching over standard matrices distanced to the averaged expert ranking by minimal difference", *Research Bulletin of NTUU "Kyiv Polytechnic Institute"*, 2016, no. 1, pp. 58–65. Retrieved from: doi: 10.20535/1810-0546.2016.1.59784

[8] J. Cao, "Data Collection in the Era of Big Data", *E-Commerce Big Data Mining and Analytics. Advanced Studies in E-Commerce*, Springer, Singapore, 2023, pp. 19–28. Retrieved from: doi: 10.1007/978-981-99-3588-8_2

[9] W.K. Ндrdle *et al.*, "Principal Component Analysis", *Applied Multivariate Statistical Analysis*, Springer, Cham, 2024, pp. 309–345. Retrieved from: doi: 10.1007/978-3-031-63833-6_11

[10] I.T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2016, Vol. 374, Iss. 2065, no. 20150202. Retrieved from: doi: 10.1098/rsta.2015.0202

[11] A.C. Olivieri, *Principal Component Analysis*, *Introduction to Multivariate Calibration*, Springer, Cham, 2024, pp. 71–87. Retrieved from: doi: 110.1007/978-3-031-64144-2_4

[12] V.V. Romanuke, "Speedup of the *k*-means algorithm for partitioning large datasets of flat points by a preliminary partition and selecting initial centroids", in *Applied Computer Systems*, 2023, Vol. 28, no. 1, pp. 1–12. Retrieved from: doi: 110.2478/acss-2023-0001

[13] S. Ekici *et al.*, "Electricity Consumption Analysis with Matlab Tall Arrays", 1*st International Engineering and Technology Symposium (1st IETS)*, May 2018, Batman University, Batman, Turkey, 2018. Retrieved from: https://www.researchgate.net/publication/327987720_Electricity_Consumption_Analysis_with_Matlab_Tall_Arrays

[14] M. Paluszek and S. Thomas, "Data for Machine Learning in MATLAB", *MATLAB Machine Learning Recipes*, Apress, Berkeley, CA, 2024, pp. 21–48. Retrieved from: doi: 10.1007/978-1-4842-9846-6_2

[15] V.V. Romanuke, "Limitation of effectiveness in using MATLAB gpuArray method for calculating products of transpose-symmetrically sized matrices", *Herald of Khmelnytskyi national university. Technical sciences*, 2015, no. 5, pp. 243–248. Retrieved from: https://elar.khmnu.edu.ua/handle/123456789/4611

[16] V.V. Romanuke, "Maximum-versus-mean absolute error in selecting criteria of time series forecasting quality", *Bionics of intelligence*, 2021, no. 1, pp. 3–9. Retrieved from: doi: 10.30837/bi.2021.1(96).01

[17] C.L. Valenzuela and A. J. Jones, "Evolutionary divide and conquer (I): A novel genetic approach to the TSP", *Evolutionary Computation*, 1993, Vol. 1, Iss. 4, pp. 313–333. Retrieved from: doi: 10.1162/evco.1993.1.4.313

[18] V.V. Romanuke, "Deep clustering of the traveling salesman problem to parallelize its solution", in *Computers & Operations Research*, 2024, Vol. 165, no. 106548. Retrieved from: doi: 10.1016/j.cor.2024.106548

[19] T. Weinzierl, Principles of Parallel Scientific Computing: A First Guide to Numerical Concepts and Programming Methods, Springer, Cham, 2021. Retrieved from: doi: 10.1007/978-3-030-76194-3

[20] V.V. Romanuke, "Optimal construction of the pattern matrix for probabilistic neural networks in technical diagnostics based on expert estimations", *Information, Computing and Intelligent Systems*, 2021, no. 2, pp. 19–25. Retrieved from: doi: 10.20535/2708-4930.2.2021.244186

[21] R. Han and Y. Wang, "The Advance and Performance Analysis of MapReduce", *Proceedings of 2nd International Conference on Artificial Intelligence, Robotics, and Communication. ICAIRC 2022. Lecture Notes in Electrical Engineering*, Vol. 1063, Springer, Singapore, 2023, pp. 205–213. Retrieved from: doi: 10.1007/978-981-99-4554-2_20

[22] S. Hedayati *et al.*, "MapReduce scheduling algorithms in Hadoop: a systematic study", *Journal of Cloud Computing*, 2023, Vol. 12, no. 143. Retrieved from: doi: 10.1186/s13677-023-00520-9.

В.В. Романюк

ЕФЕКТИВНІСТЬ МЕТОДУ TALL ARRAY У ЗНИЖЕННІ РОЗМІРНОСТІ НАБОРІВ ДАНИХ НА ОСНОВІ МЕТОДУ ГОЛОВНИХ КОМПОНЕНТІВ

Проблематика. Розвідковий аналіз даних швидко розвивається ще з початку 2000-х років. На 2025 рік більшість реальних наборів даних класифікують як великі дані. Робочий процес аналітики великих даних включає етап попередньої обробки даних, який є початковою точкою обробки великих даних. На цьому кроці дані намагаються максимально спростити. Основною парадигмою є зниження розмірності, що дозволяє спростити та візуалізувати масиви даних великої розмірності. Метод головних компонентів (PCA) є лінійним методом зниження розмірності. PCA можна прискорити, застосувавши Tall Arrays, якщо дані зберігаються на диску. Tall Array PCA (TAPCA) обчислює головні компоненти поступово, використовуючи стратегію divide-and-conquer.

Мета дослідження. Мета полягає у тому, щоб визначити, коли TAPCA фактично ефективний для зниження розмірності. Вивчаються два типи чисел – з подвійною та одинарною точністю.

Методика реалізації. Для досягнення зазначеної мети генеруються випадкові великі набори даних у вигляді матриць певного числового типу. Потім вимірюється час обчислень звичайного PCA у середовищі MATLAB, застосованого до згенерованих матриць. Далі вимірюється обчислювальний час перетворення масивів (згенерованих матриць) у пам'яті у tall-масиви. Також вимірюється час обчислення TAPCA, застосованого до тих згенерованих матриць, до яких PCA застосовувався раніше.

Результати дослідження. Порівняльний аналіз усередненого часу обчислень показує, що часова складність обчислень як PCA, так і TAPCA є радше поліноміальною, ніж строго квадратичною чи кубічною. Існує майже гіперболічна границя, яку альтернативно можна назвати порогом ефективності TAPCA, у площині кількості спостережень набору даних і кількості ознак набору даних, за якою TAPCA та звичайний PCA потребують приблизно однакового часу для обчислення головних компонентів.

Висновки. В обчисленні головних компонентів для зниження розмірності великих наборів даних, що зберігаються на диску, метод Tall Array стає ефективним за двох паралельних процесорів, якщо набір даних містить принаймні 5-6 мільйонів записів. Метод Tall Array більш ефективний для наборів даних з подвійною точністю, де його поріг ефективності становить майже 6 мільйонів записів, тоді як поріг ефективності для наборів даних з одинарною точністю становить від 5 до 15,2 мільйонів записів.

Ключові слова: зниження розмірності; метод головних компонентів (PCA); Tall Arrays; поріг ефективності; подвійна точність; одинарна точність.

Рекомендована Радою
факультету прикладної математики
КПІ ім. Ігоря Сікорського

Надійшла до редакції
30 січня 2025 року

Прийнята до публікації
30 червня 2025 року

СИСТЕМНИЙ АНАЛІЗ ТА НАУКА ПРО ДАНІ

DOI: 10.20535/kpisn.2025.2.320063
УДК 004.85.032.26:004.93’1:616.5-006](045)

В.Я. Данилов, О.О. Зарицький*
КПІ ім. Ігоря Сікорського, Київ, Україна
*Відповідальний автор: zaritskiy.alexey@gmail.com

МОДЕЛЬ КЛАСИФІКАЦІЇ РАКОВИХ ЗАХВОРЮВАНЬ ШКІРИ НА ОСНОВІ ІНТЕГРАЦІЇ ДИСКРИМІНАТОРА В АРХІТЕКТУРУ СХОДОВИХ НЕЙРОННИХ МЕРЕЖ

Проблематика. Напівкероване навчання (НН) є одним із перспективних напрямів глибокого навчання, особливо для задач, де отримання міток є складним і дорогим процесом, як у випадку класифікації ракових захворювань шкіри. Наявні підходи, зокрема Adversarial Autoencoders (AAE) та Ladder Networks (LN), ефективно використовують нерозмічені дані, але мають обмеження у точності реконструкції та регуляризації.
Мета дослідження. Розробка та дослідження моделі класифікації ракових захворювань шкіри на основі інтеграції дискримінатора в архітектуру сходових нейронних мереж.
Методика реалізації. Розроблена модель поєднує регуляризуючі властивості сходових нейронних мереж із використанням функції втрат на основі дискримінатора, що оцінює якість реконструкції зображень, орієнтуючись на їх структуру, форму та ключові візуальні ознаки. Експерименти проводилися на датасеті HAM10000 з різними співвідношеннями розмічених і нерозмічених даних (30 %, 10 %, 5 %).
Результати дослідження. Експерименти показали, що запропонована модель підвищила F_1-score для класу злоякісних утворень на 4 % порівняно з базовою сходовою мережею за умов 5 % розмічених даних. За 30 % маркованих зразків F_1-score досяг 74,8 %, що всього на 1 % менше за повністю керовану модель. Відносний показник R_{F_1} за 5 % розмічених даних становив 0,939, перевищуючи аналогічний коефіцієнт для STFL (0,877), що підтверджує ефективність використання нерозмічених даних у запропонованій моделі.
Висновки. Запропонована модель вдосконалює наявні методи напівкерованого навчання (LN, AAE), забезпечуючи високу ефективність використання нерозмічених даних для регуляризації латентного простору енкодера у задачі класифікації ракових захворювань шкіри. Перспективи подальших досліджень включають використання дискримінатора для порівняння латентних представлень і вдосконалення функції Reconstruction Cost для розширення її застосування в інших задачах аналізу зображень.
Ключові слова: класифікація ракових захворювань шкіри; HAM10000; напівкероване навчання; сходові нейронні мережі; дискримінатор.

Вступ

Напівкероване навчання є одним із найбільш актуальних напрямів дослідження глибокого навчання. Особливо корисним воно є для задач, де отримання міток є дуже складним або дорогим процесом, як, наприклад, у сфері медичних даних. У контексті задач розпізнавання та класифікації ракових захворювань шкіри, що є однією з найпоширеніших і небезпечних форм онкології, напівкероване навчання дає можливість ефективно використовувати велику кількість нерозмічених даних для навчання моделі. Складність отримання розмічених даних для навчання моделі у класифікації ракових захворювань шкіри також проявляється у тому, що для точного діагнозу недостатньо візуального вивчення утворення, а необхідно проводити складні дослідження із забором тканин утворення. Таким чином, дослідження методів, що можуть використовувати дуже малу кількість розмічених даних у поєднанні з великою кількістю нерозмічених, є перспективним напрямом у цій сфері.

Автоенкодери є поширеним інструментом у напівкерованому навчанні, оскільки їх структура дозволяє використовувати нерозмічені дані для регуляризації латентного простору енкоде-

Пропозиція для цитування цієї статті: В.Я. Данилов, О.О. Зарицький, “Модель класифікації ракових захворювань шкіри на основі інтеграції дискримінатора в архітектуру сходових нейронних мереж”, *Наукові вісті КПІ*, № 2, с. 30–38, 2025. doi: 10.20535/kpisn.2025.2.320063

Offer a citation for this article: V.Y. Danilov, O.O. Zarytskyi, “A skin cancer classification model incorporating a discriminator into the ladder neural network architecture”, *KPI Science News*, no. 2, pp. 30–38, 2025. doi: 10.20535/kpisn.2025.2.320063

ра, таким чином покращуючи узагальнюючу здатність моделі. Серед найбільш інноваційних підходів до використання автоенкодерів у напівкерованому навчанні є AAE (змагальні автоенкодери), що використовують дискримінатор для покращення регуляризації через порівняння латентного представлення з певним апріорним заданим розподілом, та LN (сходові нейронні мережі), що використовують більш складну знешумлювальну структуру у вигляді «драбини» і дозволяють зберігати важливу інформацію на різних рівнях абстракції. Хоча LN і пропонують регуляризацію за рахунок драбинної структури та додавання шуму до вхідних шарів декодера, для порівняння латентних шарів і, відповідно, обрахунку reconstruction cost здебільшого використовують прості методи типу MSE або інші аналогічні грубі методи порівняння.

У цій статті пропонується об'єднати регуляризуючу здатність обох цих підходів, використавши дискримінатори для визначення reconstruction cost у порівнянні латентних представлень енкодера та відповідних реконструйованих представлень декодера у сходовій нейронній мережі.

Для експериментів було вибрано відомий датасет HAM10000, що містить вибірку ракових захворювань шкіри різних класів, а також багато прикладів доброякісних утворень, таких як родимки. Вибір датасету пояснюється тим, що автори статей, які є основою цього дослідження, демонстрували результати експериментів тільки на «неприкладних» датасетах, таких як MNIST, де припущення напівкерованого навчання виконуються занадто очевидно й ефективність цих підходів для задачі класифікації складніших даних (а саме фотографій ракових захворювань шкіри) є недослідженою.

Постановка задачі

Метою роботи є дослідження і розробка моделі до напівкерованого навчання для задачі класифікації ракових захворювань шкіри за допомогою поєднання регуляризаційних властивостей сходових нейронних мереж із використанням дискримінатора для оцінювання якості реконструкції.

Огляд наявних досліджень

Нині у відкритому доступі не виявлено публікацій, у яких сходові нейронні мережі або інші denoising-autoencoder архітектури порівнювали саме на датасеті HAM10000. Єдиним винятком є робота [1], результати якої і будуть використані для прямого порівняння з новою моделлю.

Поза межами тематики дослідження є багато методів НН, що демонструють результати на HAM10000 (STFL [2], MixMatch [3], Mean-Teacher [4]), тому для зовнішнього порівняння було обрано одну з таких архітектур НН – Self-feedback Threshold Focal Learning (STFL), що показала стабільні результати за невеликої кількості розмічених даних (500 розмічених зразків) і використовує ResNet-50 як базову мережу. Головна ідея цієї моделі ґрунтується на автоматичному коригуванні порогів впевненості псевдоміток, а також використанні focal loss для боротьби з дисбалансом вибірки. STFL було вибрано для порівняння тому, що:

1. Наскільки нам відомо, ця стаття є найостаннішою публікацією з методів НН для класифікації ракових захворювань шкіри.

2. Стаття містить відкрито опубліковані метрики F_1, accuracy тощо саме на датасеті HAM10000, що дає змогу їх порівняти з моделлю, запропонованою у цій роботі. Подальші деталі про те, як будуть зіставлятися метрики, подано в розділі «Результати експериментів».

Таким чином, далі розглянемо теоретичні засади двох ліній досліджень – AAE та LN – щоб показати, як саме запропонована модель поєднує їхні регуляризуючі здібності та розширює можливості НН для задач класифікації ракових захворювань шкіри.

Adversarial Autoencoders

Adversarial Autoencoders [5] – це ймовірнісні автоенкодери, що використовують генеративно змагальну мережу [6] для накладання заданого апріорного розподілу на латентний простір автоенкодера. Нехай виходом енкодера в автоенкодерній архітектурі є деякий апостеріорний розподіл $q(z)$. Автори статті [5] визначають його таким чином:

$$q(z) = \int_x q(z|x)\, p_{data}(x) dx,$$

де $q(z \mid x)$ – це латентний розподіл, що генерує енкодер; $p_{data}(x)$ – це розподіл вхідних даних.

Формально навчання змагального автоенкодера можна описати у два етапи:

1. Фаза реконструкції, де автоенкодер мінімізує помилку реконструкції відновленого зображення відносно оригіналу.

2. Фаза регуляризації, де дискримінатор порівнює апостеріорний розподіл $q(z)$ з деяким

апріорним розподілом $p_{prior}(z)$ і намагається визначити, який із них є оригіналом, а який є результатом енкодера.

Архітектуру AAE зображено на рис. 1. У цій взаємодії роль генератора виконує енкодер, а дискримінатор додається окремо для порівняння апріорного та апостеріорного розподілу. Оцінювання дискримінатора у цій системі дає змогу енкодеру (тобто генератору) виділити ознаки таким чином, щоб апостеріорний розподіл був дуже схожим на апріорний, з яким його порівнює дискримінатор. Автори показують, що такий спосіб регуляризації може суттєво покращити здатність моделей виділяти важливі ознаки, ефективно «підказуючи» їм, як приблизно повинен виглядати цільовий розподіл.

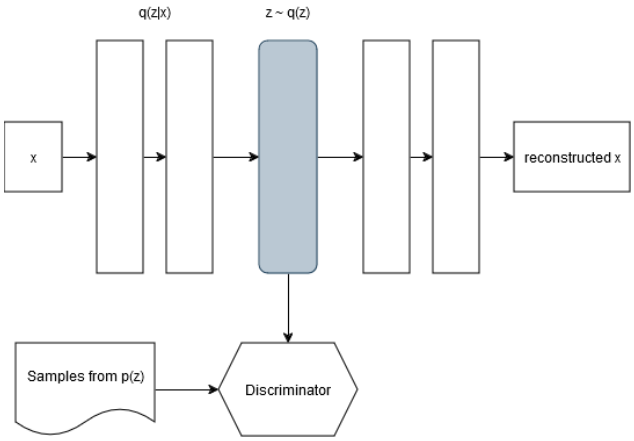


Рис. 1. Архітектура змагального автоенкодера. Схему створено за описом у [3]

Хоча заздалегідь невідомий точний розподіл даних, проте можна припустити, який це тип розподілу, і підказати за допомогою дискримінатора, як краще енкодеру виділяти ознаки, щоб апостеріорний розподіл відповідав гіпотезі.

Adversarial Autoencoders in semi-supervised learning

Якщо у змагальних автоенкодерах використати апостеріорний розподіл як вхідний вектор для повнозв'язного шару, то отримаємо модель напівкерованого навчання для класифікації, де розмічені дані використовуються для керованої частини, а reconstruction cost та regularization cost – для некерованої частини. Такий підхід був запропонований у статті про використання AAE для напівкерованого навчання [7], що розвиває ідеї, описані у [8] та [5].

Автори пропонують таку функцію втрат:

$$L = \lambda_{NLL} L_{NLL} + \lambda_{rec} L_{rec} + \lambda_{reg} L_{reg},$$

де L_{NLL} – втрати класифікації між міткою моделі та справжньою міткою; L_{rec} – вартість реконструкції між реконструйованим зображенням та оригіналом; L_{reg} – втрати регуляризації, які визначає дискримінатор.

Сходові нейронні мережі

Сходові нейронні мережі (LN) [9], [10], [11] – це сучасний архітектурний підхід до напівкерованого навчання, що ґрунтується на ідеї знешумлюючих (denoising) автоенкодерів. У загальному розумінні структура сходової нейронної мережі для використання у напівкерованому навчанні нагадує будову автоенкодера [8], проте на вході у декодер латентне представлення зашумлюється і завдання декодера полягає у відновленні оригінального зображення.

Сходова нейронна мережа складається з трьох основних компонент:

1. Зашумлений енкодер (noisy encoder), що генерує латентні представлення за допомогою додавання гаусівського шуму.
2. Чистий енкодер створює еталонні латентні представлення без шуму.
3. Декодер, що реконструює латентні представлення на кожному кроці енкодера із зашумлених версій цих представлень.

На кожному шарі енкодера та декодера використовується батч-нормалізація (batch normalization).

Енкодер і декодер працюють у двох потоках: у чистому та зашумленому. Опис архітектури формалізується таким чином [9]:

$$\tilde{x}, \tilde{z}^1, \ldots, \tilde{z}^L, \tilde{y} = Encoder_{noisy}(x);$$

$$x, z^1, \ldots, z^L, y = Encoder_{clean}(x);$$

$$\hat{x}, \hat{z}^1, \ldots, \hat{z}^L, \tilde{y} = Decoder\left(\tilde{z}^1, \ldots, \tilde{z}^L\right),$$

де x, y та $\tilde{y}$ – вхідні дані, знешумлені та зашумлені виходи; z^1, $\tilde{z}^1$, $\hat{z}^1$ – приховане представлення, його зашумлена версія та його реконструйоване представлення.

Кожен шар декодера реконструює зашумлене латентне представлення енкодера. Для цього він використовує реконструйоване попередніми шарами декодера латентне представлення у комбінації із зашумленим представленням енкодера із поточного рівня:

$$\hat{z}^l = g\left(\tilde{z}^l, u^{l+1}\right),$$

де $\hat{z}^l$ – реконструйоване представлення; u^{l+1} – вертикальний сигнал із наступного шару після застосування батч-нормалізації; g – функція комбінування.

Цільова функція складається із зваженої суми втрат за класифікацію та за реконструкцію:

$$L = \lambda_{sup} L_{sup} + \lambda_{rec} L_{rec}.$$

Для задачі класифікації як L_{sup} використовується Cross-Entropy Loss, а як L_{rec} використовується MSE.

Спрощену схему архітектури сходової нейронної мережі зображено на рис. 2, де CE – це Cross-Entropy Loss, RC – це функції втрат реконструкції, сума яких формує загальну функцію втрат L_{rec}. На виході незашумленого енкодера зазвичай використовують повнозв'язний шар, що використовується для класифікації.

Ця архітектура забезпечує можливість одночасного навчання моделі як на мічених, так і на немічених даних. Для немічених даних використовується шлях «$Encoder_{noisy} \rightarrow Decoder$», де зворотне поширення помилки дозволяє коригувати ваги модуля $Encoder_{clean}$, покращуючи здатність моделі до виділення важливих ознак.

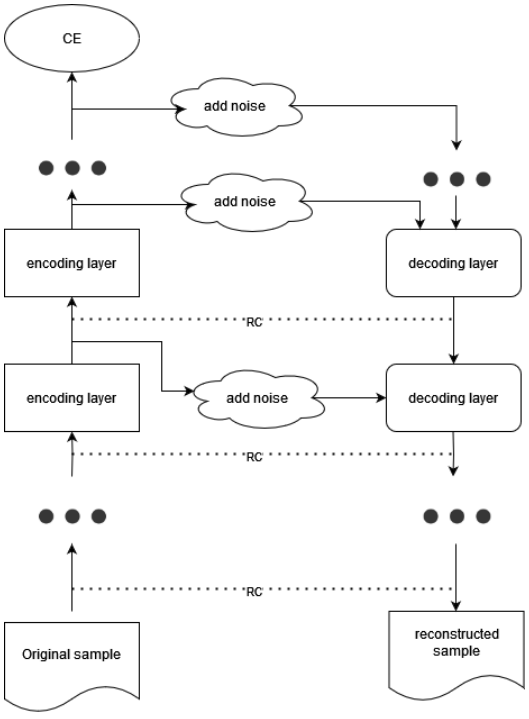


Рис. 2. Спрощена схема архітектури сходової нейронної мережі. Схему створено на основі опису з [9]

Для мічених даних додається стандартний енкодер із повнозв'язним шаром на виході, який навчається паралельно. У результаті модель має два окремі потоки: «$Encoder_{noisy} \rightarrow Decoder$» для обробки немічених даних і «$Encoder_{clean} \rightarrow$ повнозв'язний шар» для роботи з міченими зразками.

Запропонована модель

Після огляду наявних підходів було встановлено ключовий недолік: більшість сучасних методів НН, розроблених для класифікації ракових захворювань шкіри, покладаються на евристичні методи псевдомаркування, при цьому структура латентного простору залишається поза увагою. Ці спостереження підказують, що перспективними є дослідження моделей з регуляризацією, таких як сходові нейронні мережі та змагальні автоенкодери.

У [7] автори розглядають використання дискримінатора виключно для формування L_{reg}, тобто втрат за регуляризацію латентного простору на виході енкодера. Якщо використати дискримінатор для оцінювання втрат звичайної реконструкції L_{rec}, то це не буде мати сенсу, оскільки оригінальний автоенкодер має повну інформацію про оригінальне зображення, а отже, цінність дискримінатора в цій архітектурі невелика. Утім сходові нейронні мережі використовують зашумлення на етапі передачі латентного представлення до декодера, тому суттєва частина інформації втрачається. Як наслідок, використання дискримінатора в цій моделі набуває сенсу.

Модель класифікації ракових захворювань шкіри на основі інтеграції дискримінатора в архітектуру сходових нейронних мереж

У нашому дослідженні пропонується використовувати дискримінатор як більш м'який (у сенсі soft computing) спосіб оцінювання вартості реконструкції сходової нейронної мережі.

Запропонована функція реконструкції у сходовій нейронній мережі може бути виражена таким чином:

$$L_{rec} = E_{x \sim p_{data}}(x)\left[logD(\tilde{x})\right], \qquad (1)$$

де $\tilde{x}$ – частково реконструйоване зображення, яке декодер намагається відновити із зашумленого латентного представлення.

В (1) D – це дискримінатор, що навчається розрізняти відновлені зображення від справжніх,

що дозволяє йому «м'яко» контролювати якість реконструкції, орієнтуючись на глобальні ознаки, такі як текстура і форма, замість грубих порівнянь, які використовуються в оригінальній архітектурі сходових нейронних мереж.

Використання зашумлених даних у сходових нейронних мережах створює ефективну навчальну задачу для дискримінатора, оскільки він оцінює відновлення зображення у складніших умовах, ніж у звичайному автоенкодері. Водночас дискримінатор своїм оцінюванням спонукає генератор (тобто декодер) відновити важливі ознаки замість того, щоб грубо намагатися мінімізувати загальну похибку між відновленим зображенням та оригіналом.

Пропонується така функція втрат для цієї моделі:

$$L = \lambda_{NLL} L_{NLL} + \lambda_{rec} L_{rec}, \quad (2)$$

де L_{rec} можна використовувати як подання (1), так і комбінований варіант, що також частково враховує грубі методи оцінювання реконструкції:

$$L_{rec} = E_{x \sim p_{data}}(x)\left[logD(\tilde{x})\right] + \lambda_s E_{x \sim p_{data}}(x) R\left(x, \hat{x}\right), \quad (3)$$

де $R\left(x, \hat{x}\right)$ – це функція грубого оцінювання реконструкції (наприклад, MSE); λ_s – коефіцієнт, що набуває значень [0,1].

Фрагмент моделі запропонованої мережі схематично зображено на рис. 3.

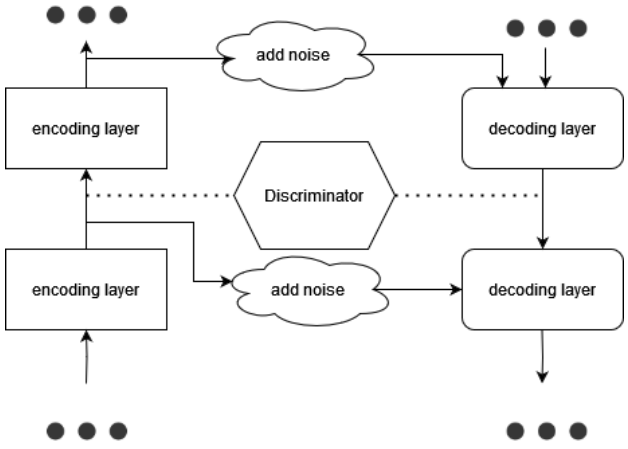


Рис. 3. Архітектура фрагмента сходової мережі із дискримінатором для оцінювання реконструкції

Тут RC замінена дискримінатором, який виконує функцію оцінювання якості реконструкції. Заміна функції втрат реконструкції на дискримінатор можна реалізувати як на кожному рівні сходової нейронної мережі, так і вибірково. Наприклад, можна використати класичний спосіб оцінювання якості реконструкції для латентних представлень, а спосіб із дискримінатором використати для порівняння реконструйованого зображення з оригіналом.

Результати експериментів

Експерименти проводилися на датасеті HAM10000, що містить зображення ракових захворювань шкіри (6 типів), а також здорових утворень (родимок) і відповідних міток цих зображень. Усі експерименти проводилися для різних співвідношень розмічених і нерозмічених даних у вибірці: 30 %, 10 % та 5 % розмічених даних. Оскільки для дослідження моделей напівкерованого навчання необхідно припустити, що більшість зображень є нерозміченими, дані було поділено на тип «злоякісні» та «доброякісні», тобто всі шість типів злоякісних утворень було згруповано в один клас (рис. 4).

Для порівняння розглянутих моделей НН створена спрощена згорткова нейронна мережа із чотирьох згорткових шарів, а також одного лінійного шару. Було використано всього один лінійний шар для наочності демонстрації регуляризуючих здібностей автоенкодера. Ця базова модель використовувалася як еталонна для порівняння якості запропонованих моделей НН. Еталонна «погана» модель – це модель, навчена тільки на розміченій частині даних. Еталонна «хороша» модель – це модель, навчена так, ніби всі 100 % даних є розміченими. Якщо метричні показники автоенкодера будуть наближатися до показників еталонної «поганої» моделі, то це означатиме, що модель НН є неефективною, оскільки використання нерозмічених даних не покращує результат. Аналогічно, якщо якісні показники автоенкодера близькі до еталонної «хорошої» моделі, то можна зробити висновок, що запропонована модель НН є ефективною.

Базова модель використовувалася як основа для енкодера у всіх моделях автоенкодера із мінімально можливими модифікаціями, які вимагаються для реалізації. У побудові автоенкодерів для розрахування Reconstruction Cost використано звичайний MSE для всіх прихованих шарів і дискримінатор для порівняння зображення з оригіналом. На рис. 5 зображено схематичний опис моделі згорткового автоенкодера та сходової нейронної мережі з використанням дискримінатора.

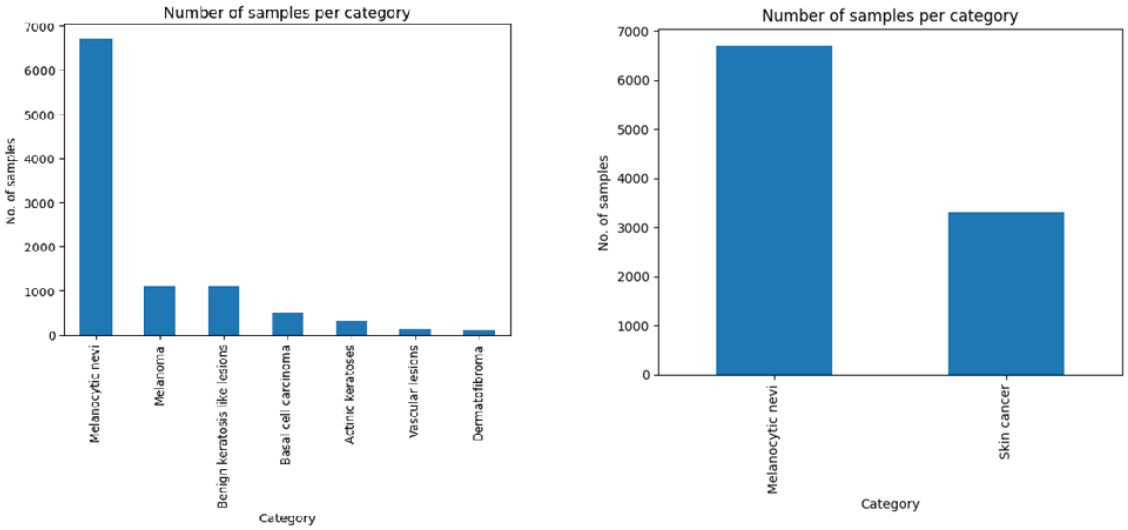


Рис. 4. Розподіл даних у датасеті до та після ребалансування вибірки

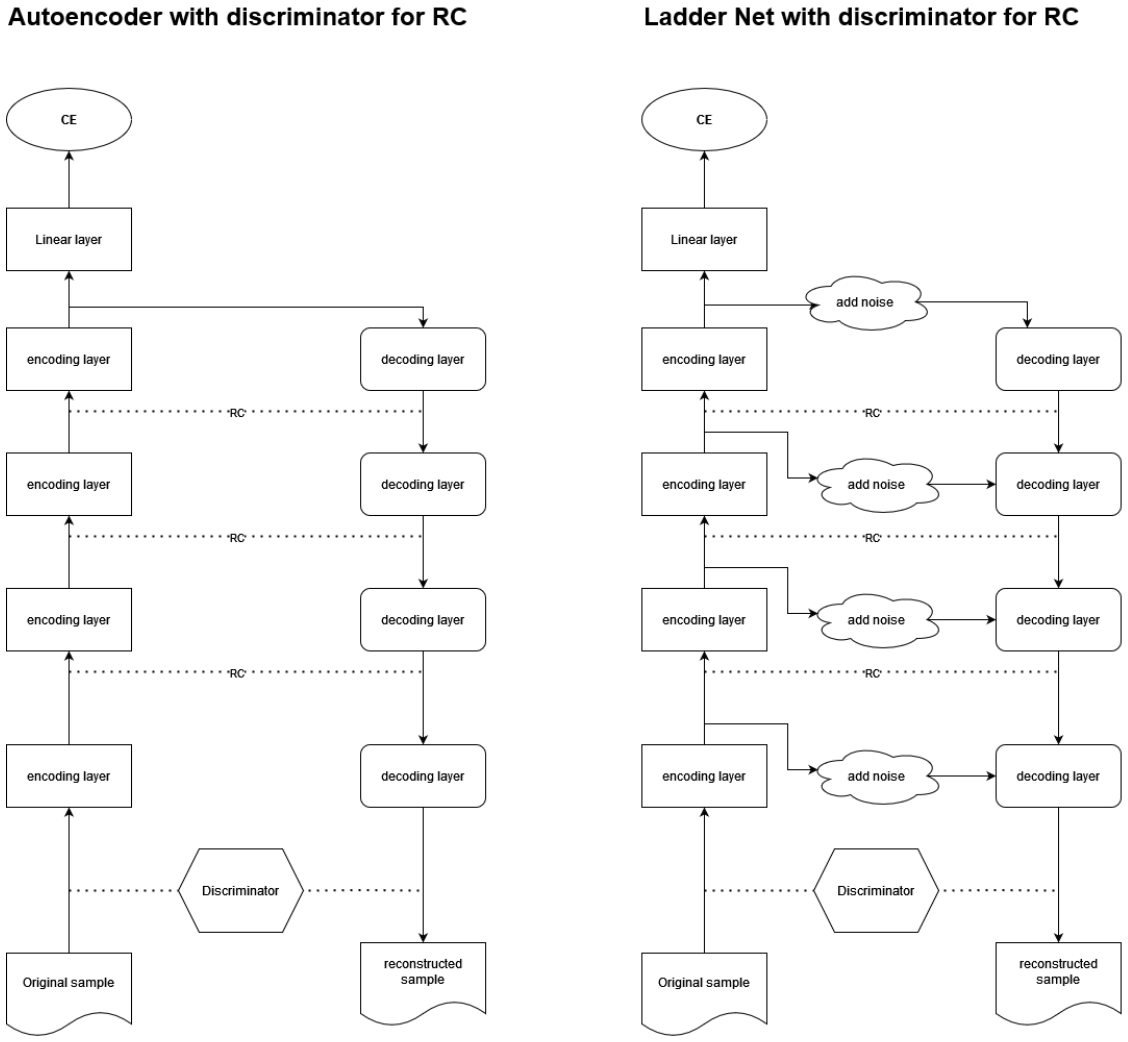


Рис. 5. Автоенкодер та сходова нейронна мережа з дискримінатором

У табл. 1 відображено метричні показники моделей для різних співвідношень розмічених даних у вибірці. Тут CA – це звичайний згортковий автоенкодер, LN – це сходова нейронна мережа, simple означає, що використовується оригінальна архітектура без дискримінатора, discriminator RC – це модель із дискримінатором для оцінювання reconstruction cost (як зображено на рис. 5), combined RC – це така сама модель, але функція втрат визначається як зважена сума оцінювання дискримінатора та MSE (тобто за формулою (3)).

З результатів експериментів випливає, що використання дискримінатора виявилося ефективним і збільшило F_1-score для класу злоякісних захворювань як для звичайного автоенкодера, так і для сходової нейронної мережі. Також із результатів експерименту видно, що використання комбінації сходових нейронних мереж із дискримінатором у більшості випадків є ефективнішим порівняно з використанням того самого підходу для звичайного автоенкодера (F_1-score становив 74,8 % проти 73,5 % для 30 % розмічених даних, 74,.3 % проти 74,6 % для 10 %, 71 % проти 68,6 % для 5 %).

Особливо відчутна перевага такого підходу на 5 % розмічених даних, де приріст F_1-score метрики склав 4 % порівняно зі звичайною LN та 2,4 % порівняно з аналогічною моделлю звичайного згорткового автоенкодера. Також помітно, що використання комбінованого підходу обчислення функції втрат за формулою (3) демонструє кращі результати на вибірках з малою кількістю розмічених даних. Для 30 % розмічених даних використання виключно дискримінатора без MSE (формула (2)) показало себе краще як для сходової нейронної мережі, так і для згорткового автоенкодера (73,5 % проти 71,8 % для CA, а також 74,8 % проти 72,5 % для LN), при цьому показник F_1-score досяг 74,8 %, що всього на 1 % менше за повністю керовану модель.

***Таблиця* 1.** Метричні показники моделей, навчених різними підходами

Модель НН	Accuracy	Precision (melanocytic nevi)	Precision (skin cancer)	Recall (melanocytic nevi)	Recall (skin cancer)	F1 (melanocytic nevi)	F1 (skin cancer)
30 % розмічених даних							
CA (simple)	0,801	0,828	0,739	0,876	0,660	0,851	0,698
CA (discriminator RC)	0,827	0,844	0,789	0,902	0,687	0,872	0,735
CA (combined RC)	0,818	0,834	0,780	0,900	0,664	0,866	0,718
LN (simple)	0,807	0,824	0,767	0,896	0,641	0,858	0,699
LN (discriminator RC)	0,818	0,875	0,722	0,841	0,775	0,857	0,748
LN (combined RC)	0,807	0,854	0,721	0,849	0,729	0,852	0,725
Etalon bad model	0,773	0,779	0,751	0,908	0,519	0,839	0,614
Etalon good model	0,847	0,852	0,796	0,903	0,712	0,883	0,756
10 % розмічених даних							
CA (simple)	0,797	0,853	0,698	0,831	0,733	0,842	0,715
CA (discriminator RC)	0,816	0,874	0,720	0,839	0,775	0,856	0,746
CA (combined RC)	0,820	0,839	0,777	0,896	0,679	0,867	0,725
LN (simple)	0,794	0,844	0,702	0,839	0,710	0,841	0,706
LN (discriminator RC)	0,809	0,855	0,723	0,851	0,729	0,853	0,726
LN (combined RC)	0,805	0,889	0,685	0,800	0,813	0,842	0,743
Etalon bad model	0,743	0,786	0,654	0,844	0,552	0,819	0,591
Etalon good model	0,847	0,852	0,796	0,903	0,712	0,883	0,756
5 % розмічених даних							
CA (simple)	0,778	0,797	0,727	0,884	0,580	0,838	0,645
CA (discriminator RC)	0,786	0,813	0,722	0,871	0,626	0,841	0,671
CA (combined RC)	0,795	0,821	0,737	0,878	0,641	0,848	0,686
LN (simple)	0,802	0,803	0,799	0,922	0,576	0,858	0,670
LN (discriminator RC)	0,803	0,835	0,735	0,869	0,769	0,852	0,706
LN (combined RC)	0,794	0,850	0,696	0,831	0,725	0,840	0,710
Etalon bad model	0,691	0,913	0,534	0,582	0,897	0,711	0,670
Etalon good model	0,847	0,852	0,796	0,903	0,712	0,883	0,756

Для зовнішнього порівняння із запропонованою моделлю було обрано модель STFL [2]. Оскільки у статті як базову модель було використано ResNet-50, що має суттєво більшу кількість параметрів, ніж базова модель у нашому дослідженні, результати STFL коректно порівнювати лише відносно. У [2] наведено метричні показники моделі STFL для двох режимів: 5 % маркованих зразків (0,7462) та 100 % (0,8507). Пропонується оцінити відносну якість моделі як співвідношення показника F_1 для напівкерованої моделі ($F_{1_{SSL}}$) до відповідного показника для повністю керованої моделі ($F_{1_{SUP}}$):

$$R_{F_1} = \frac{F_{1_{SSL}}}{F_{1_{SUP}}} = \frac{0.7462}{0.8507} = 0.877.$$

Запропонована модель має відчутно вищий показник R_{F_1} для 5 % розмічених даних (0,939), що може свідчити про високу ефективність використання нерозмічених даних для покращення якості моделі.

Висновки

Досліджено та встановлено відсутність у відкритих джерелах експериментів, де denoising-autoencoder архітектури, зокрема LN, оцінювалися на HAM10000. Виявлено, що наявні SSL-методи для класифікації ракових захворювань шкіри (FixMatch, MixMatch, Mean-Teacher, STFL) не враховують регуляризації латентного простору.

Розроблено напівкеровану модель, яка інтегрує дискримінатор AAE у декодер сходової нейронної мережі для комбінованого оцінювання Reconstruction Cost (MSE + дискримінатор).

Модель забезпечує суттєве покращення якості. Наприклад, за 5 % розмічених даних F_1-score для злоякісних утворень зріс на 4 % відносно класичної LN і на 2,4 % відносно аналогічної моделі звичайного автоенкодера з дискримінатором. Відносний коефіцієнт $R_{F_1} = 0,929$ перевищив аналогічний показник STFL (0,877), що підтверджує ефективність використання нерозмічених даних у процесі напівкерованого навчання моделі.

Перспективи подальших досліджень включають вивчення використання дискримінатора для порівняння латентних представлень та аналіз впливу різних параметрів функції Reconstruction Cost.

References

[1] О.О. Зарицький та В.Я. Данилов, «Порівняння ефективності методів напівкерованого навчання на основі автоенкодерів для задач класифікації фотографій ракових захворювань шкіри», *Вісник Вінницького політехнічного інституту*, 2024. № 2. с. 71–77. doi: 10.31649/1997-9266-2024-173-2-71-77.

[2] W. Yuan, Z. Du, and S. Han, “Semi-supervised skin cancer diagnosis based on self-feedback threshold focal learning”, *Discover Oncology*, 2024, vol. 15. no. 1, 180 p. doi: 10.1007/s12672-024-01043-8.

[3] C.B. Hansen, V. Nath, R. Gao, C. Bermudez, Y. Huo, K.L. Sandler, P.P. Massion, J.D. Blume, T.A. Lasko, and B.A. Landman, “Semi-supervised machine learning with MixMatch and equivalence classes”, *Interpretable and Annotation-Efficient Learning for Medical Image Computing* (iMIMIC, MIL3ID & LABELS Workshops, MICCAI 2020, Lima, Peru), Lecture Notes in Computer Science, 2020, vol. 12446, pp. 112–121. doi: 10.1007/978-3-030-61166-8_12.

[4] S. Kitada and H. Iyatomi, “Skin lesion classification with ensemble of squeeze-and-excitation networks and semi-supervised learning”, arXiv preprint arXiv:1809.02568, 2018. [Online]. Retrieved from: doi: 10.48550/arXiv.1809.02568.

[5] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, “Adversarial autoencoders”, arXiv preprint arXiv:1511.05644, 2015. [Online]. Retrieved from: https://arxiv.org/abs/1511.05644. doi: 10.48550/arXiv.1511.05644.

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets”, *Advances in Neural Information Processing Systems*, 2014, vol. 27, pp. 2672–2680. doi: 10.1145/3422622.

[7] R. Tachibana, T. Matsubara, and K. Uehara, “Semi-supervised learning using adversarial networks”, *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, Okayama, Japan, 2016, pp. 1–6. doi: 10.1109/icis.2016.7550881.

[8] A. Gogna and A. Majumdar, “Semi supervised autoencoder”, *Neural Information Processing, 23rd International Conference, ICONIP 2016*, Kyoto, Japan, October 16–21, 2016, Proceedings, Part II 23, Springer International Publishing, 2016. doi: 10.1007/978-3-319-46672-9_10

[9] M. Pezeshki, L. Fan, P. Brakel, A. Courville, and Y. Bengio, “Deconstructing the Ladder Network Architecture”, *Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA, 2016, pp. 2368–2376. doi: 10.48550/arXiv.1511.06430.

[10] A. Rasmus, H. Valpola, M. Honkala, M. Berglund and T. Raiko “Semi-Supervised Learning with Ladder Networks”, *NIPS’15: Proceedings of the 29th International Conference on Neural Information Processing Systems*, 2015, vol. 2, pp. 3546–3554. doi: 10.48550/arXiv.1507.02672.

[11] D.P. Kingma, D.J. Rezende, S. Mohamed and M. Welling, “Semi-supervised learning with deep generative models”, *Advances in Neural Information Processing Systems*, 2014, vol. 27, pp. 3581–3589. doi: 10.48550/arXiv.1406.5298.

V.Y. Danilov, O.O. Zarytskyi

A SKIN CANCER CLASSIFICATION MODEL INCORPORATING A DISCRIMINATOR INTO THE LADDER NEURAL NETWORK ARCHITECTURE

Background. Semi-supervised learning (SSL) is one of the most promising areas of deep learning, especially for tasks where label acquisition is a complex and expensive process, such as skin cancer classification. Existing approaches, such as Adversarial Autoencoders (AAE) and Ladder Networks (LN), effectively use unlabelled data but have limitations in reconstruction and regularization accuracy.

Objective. Development and study of a model for classifying skin cancers based on the integration of a discriminator into the architecture of ladder neural networks.

Methods. The developed model combines the regularizing properties of ladder neural networks with the use of a discriminator-based loss function that evaluates the quality of image reconstruction based on their structure, shape, and key visual features. The experiments were conducted on the HAM10000 dataset with different ratios of labelled and unlabelled data (30 %, 10 %, 5 %).

Results. The experiments showed that the proposed model improved the F_1-score for the malignancy class by 4 % compared to the baseline ladder network with 5 % labelled data. With 30 % labeled samples, the F_1-score reached 74.8 %, which is only 1 % less than the fully supervised model. The relative indicator R_{F_1} at 5 % labelled data was 0.939, exceeding the similar coefficient for STFL (0.877), which confirms the effectiveness of using unlabelled data in the proposed model.

Conclusions. The proposed model improves on existing semi-supervised learning methods (LN, AAE) by providing high efficiency of using unlabelled data to regularize the encoder latent space in the task of skin cancer classification. Prospects for further research include using a discriminator to compare latent spaces and improving the Reconstruction Cost function to expand its application to other medical image analysis tasks.

Keywords. skin cancer classification; HAM10000; semi-supervised learning; ladder neural networks; discriminator.

Рекомендована Радою
Навчально-наукового інституту прикладного системного аналізу
КПІ ім. Ігоря Сікорського

Надійшла до редакції
18 січня 2025 року

Прийнята до публікації
30 червня 2025 року

МАТЕРІАЛОЗНАВСТВО

DOI: 10.20535/kpisn.2025.2.331048
УДК 621.793

О.В. Кущев[1*], О.Є. Терентьєв[1], В.Т. Варченко[1], Т.М. Чевичелова[1],
Р.Є. Костюнік[2], О.П. Уманський[1]

[1] Інститут проблем матеріалознавства ім. І.М. Францевича НАН України,
[2] Державне некомерційне підприємство «Державний університет
«Київський авіаційний інститут», м. Київ, Україна

[*]Відповідальний автор: aleks.tmu@gmail.com

МЕТОД ПІДВИЩЕННЯ ЗНОСОСТІЙКОСТІ КОМПОЗИЦІЙНОГО ПЛАЗМОВОГО ПОКРИТТЯ НА ОСНОВІ НІКЕЛЬ-ГРАФІТУ ШЛЯХОМ ДОДАВАННЯ МІДНИХ СПЛАВІВ

Проблематика. У статті розглянуто можливість розширення сфер застосування плазмових покриттів на основі системи нікель-графіт завдяки додаванню до складу безолов'яних мідних сплавів, що мають спорідненість до нікелевої оболонки порошку системи нікель-графіт і, окрім збереження триботехнічних властивостей нікель-графіту, можуть забезпечити потрібний рівень когезії покриття.
Мета дослідження. Метою роботи є дослідження впливу вмісту мідного сплаву БрАЖ10-4 після його додавання у нікель-графіт (НПГ-75) на триботехнічні характеристики композиційного плазмового покриття.
Методика реалізації. У межах роботи для визначення впливу вмісту мідного сплаву в нікель-графіті на триботехнічні властивості плазмового композитного покриття проведено випробування на тертя і зношування зразків із масовою часткою мідного сплаву БрАЖ10-4 35, 50, 75 та 90 % у плакований нікелем графіт НПГ-75 (масова частка нікелю становить 75 %). Як еталон використовували подвійний сплав мідь-цинк Л90.
Результати дослідження. Експериментально встановлено, що плазмові покриття на основі НПГ-75 з додаванням масової частки БрАЖ10-4 50 % мають зносостійкість, більшу у 10 разів щодо чистого НПГ-75 за значно меншого (0,36) за еталон Л90 (0,52) коефіцієнта тертя. Намазування на контртіло зі сталі 45 (твердість 45-48 HRC) при цьому найменше серед трибопар, що були досліджені в межах цієї роботи. Зменшення вмісту мідного сплаву до масової частки 35 % недоцільне, оскільки призводить до інтенсифікації зношування матеріалу покриття і його намазування на контртіло. Подальше збільшення вмісту порошку БрАЖ10-4 у складі спричиняє схоплення у зоні контакту, про що свідчить різке (у 10 разів) підвищення зносу зразків покриття, зростання температури у зоні контакту і збільшення приблизно у 2 рази коефіцієнта тертя.
Висновки. Оптимальна масова частка БрАЖ10-4 у плазмовому покритті на основі НПГ-75 становить близько 50 %. Матеріал із таким складом дозволяє зменшити з 0,52 (Л90) до 0,36 коефіцієнт тертя у парі зі сталлю 45. Масове зношення сталевого контртіла також значно знижується через, ймовірно, зміну механізму зношування з адгезійного (схоплення) на окисне зношування. Це дозволяє суттєво розширити сфери застосування матеріалу Ni-C + БРАЖ10-4 в авіаційній і космічній техніці.
Ключові слова: нікель-графіт; композиційний порошок; плазмове покриття; хімічний склад; структура, тертя та зношення.

Вступ

Композитні плазмові покриття, виконані на основі порошку системи нікель-графіт, характеризуються досить високою термостійкістю і гарними антифрикційними властивостями. Такі покриття застосовують переважно як ущільнювальні матеріали газотурбінних двигунів. Враховуючи унікальні властивості графіту утворювати на поверхнях тертя шар твердого термостійкого змащувального матеріалу, нікель-графітові композити є привабливим для використання в багатьох трибопарах, утім, їх використання ускладнене невеликою твердістю таких матеріалів.

Пропозиція для цитування цієї статті: О.В. Кущев, О.Є. Терентьєв, В.Т. Варченко, Т.М. Чевичелова, Р.Є. Костюнік, О.П. Уманський, "Метод підвищення зносостійкості композиційного плазмового покриття на основі нікель-графіту шляхом додавання мідних сплавів", *Наукові вісті КПІ*, № 2, с. 39–44, 2025. doi: 10.20535/kpisn.2025.2.331048

Offer a citation for this article: O.V. Kushchev, O.E. Terentjev, V.T. Varchenko, T.M. Chevichelova, R.E. Kostyunik, O.P. Umanskyi, "Method for increasing the nickel-graphite-based composite plasma coating wear resistance by adding copper alloys", *KPI Science News*, no. 2, pp. 39–44, 2025. doi: 10.20535/kpisn.2025.2.331048

Підвищення вмісту нікелю у складі порошку є найпростішим способом збільшення твердості й когезії нікель-графітових плазмових покриттів, але це спричиняє суттєве погіршення їх антифрикційних характеристик, оскільки нікель не є антифрикційним матеріалом.

Широко відоме використання як антифрикційних матеріалів сплавів на мідній основі, зокрема бронз. Олов'яні бронзи мають достатню твердість (близько 75 НВ), але їх термостійкість зазвичай обмежена властивостями певних легуючих елементів, що утворюють окрему фазу в гетерогенній структурі матеріалу, і становить близько 250–300 °С [1, 2]. Значно більш термостійкими є безолов'янисті бронзи (наприклад, мідний сплав БрАЖ10-4, DIN 17665 CuAl10Ni5Fe4), що здатні витримувати близько 400–500 °С [3, 4]. Утім, твердість таких сплавів також значно більша (220 НВ та вище) за олов'яні бронзи, що висуває ряд вимог до відповідного матеріалу трибопари.

Відоме використання у вузлах тертя матеріалів, утворених спеченим порошком мідного сплаву та графіту [5, 6, 7]. Вони характеризуються значною (75–85 %) пористістю, що не може не впливати негативно на когезію, швидкість зношування та коефіцієнт тертя [8].

Автори запропонували розширити сфери застосування плазмових покриттів на основі системи нікель-графіт завдяки додаванню до складу безолов'яних мідних сплавів, що мають спорідненість до нікелевої оболонки, порошку системи нікель-графіт і, окрім збереження триботехнічних властивостей нікель-графіту, можуть забезпечити потрібний рівень когезії покриття.

У межах дослідницької роботи виготовлено й досліджено на триботехнічні властивості зразки, в яких масова частка мідного сплаву в нікель-графіті становить 35, 50, 75 та 90 %.

Постановка задачі

Триботехнічні характеристики пар тертя можуть бути підвищенні завдяки використанню композитів на основі металів і твердого змащувального матеріалу. Автори статті для підвищення зносостійкості вузлів тертя запропонували використати метод створення композиційного плазмового покриття на основі плакованого нікелем графіту (НПГ-75) з додаванням мідного сплаву з масовою часткою БрАЖ10-4 від 35 до 90 %.

Методи дослідження

Під час проведення дослідження використано порошки плакованого графіту НПГ-75 виробництва ТОВ «Композиційні системи» (м. Запоріжжя). За основу композиційних порошків обрано графіт марки ГАК-1 (99 % С, ТОВ «Заваллівський графіт», Україна), а також порошок мідного сплаву БрАЖ10-4 (DIN 17665 CuAl10Ni5Fe4).

Товщину плакувального шару порошку НПГ-75 обирали так, щоб вона забезпечувала потрібну масову частку компонентів (75 % нікелю та 25 % графіту) і становила від 6,4 до 12,3 мкм.

Щоб збільшити адгезійний зв'язок матеріалу покриття і сталевої основи, а також вирівняти коефіцієнти термічного розширення, на підкладку наносили проміжний шар з термореагуючого матеріалу ПГ-Ю5Н (склад: 95 % Ni + 5 % Al (ISO 9001: 2008)) зернистістю від –100 до +40 мкм, товщина підшару становить близько 100 мкм.

Для змішування порошків було використано обладнання, що реалізує принцип «п'яної бочки».

Щоб підвищити адгезію, перед нанесенням покриттів поверхню сталевих зразків піддавали струминно-абразивній обробці порошком електрокорунду зернистістю F22–F24 (ISO 8486-86) з метою її очищення, активації та надання шорсткості (R_z = 63–80). Відстань, на якій здійснювалась обробка поверхні, становить 90–150 мм, кут – від 60° до 90°. Тиск повітря становив 0,5–0,7 МПа.

Плазмові покриття наносили у відкритій атмосфері (APS) на установці УПУ-3Д у захисній камері з маніпулятором 15ВБ і плазмотроном F4-MB фірми Metco на зразки зі сталі 45 (зарубіжні аналоги – AISI 1045, JIS S45C, DIN C45, за міжнародним стандартом EN 10277-2 – сталь 1.1191). Як плазмоутворювальний газ використовували суміш аргону з воднем.

Дослідження морфології композиційного порошку і структури отриманих плазмових покриттів проведено на шліфах поперечного перерізу на растровому електронному мікроскопі РЕМ-106И.

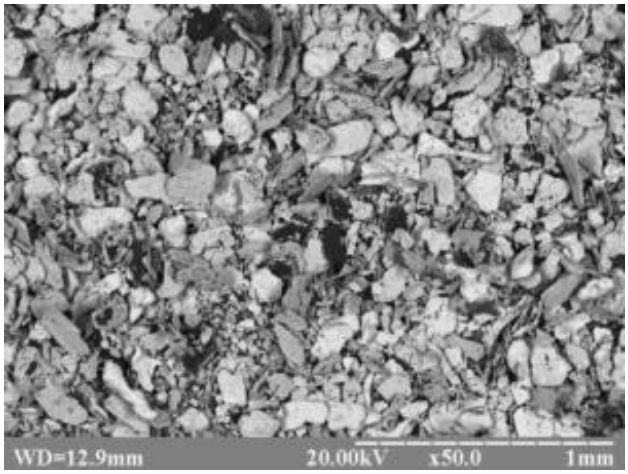
Дослідження на тертя і зношування проводили на машині для випробовування матеріалів на тертя і зношування М-22М відповідно до вимог ГОСТ 26614-85 «Матеріали антифрикційні порошкові. Метод визначення триботехнічних властивостей». Для проведення порівняльних трибологічних випробувань як базову трибопару було використано таку пару: зразок Л90 (зарубіжні аналоги – ASTM CW501L, DIN CuZn10) і контрзразок сталь 45.

Як пару тертя досліджуваного матеріалу було використано сталевий контрзразок зі ста-

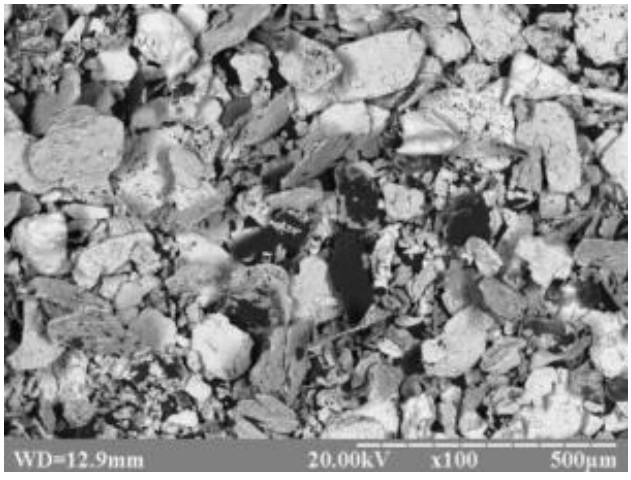
лі 45 термооброблений до твердості 45-48HRC. Швидкість тертя становила 4 м/с, питоме навантаження на контакті – 15 кг/см^2, шлях тертя – 1000 м.

Аналіз отриманих результатів

Щоб отримати вихідний матеріал, придатний для формування композиційних покриттів плазмовим методом, було підготовлено суміш, яка складалася з порошків плакованого нікель-графіту з масовою часткою нікелю 75 % (НПГ-75) та мідного сплаву. Морфологію отриманої суміші порошків зображено на рис. 1.



а



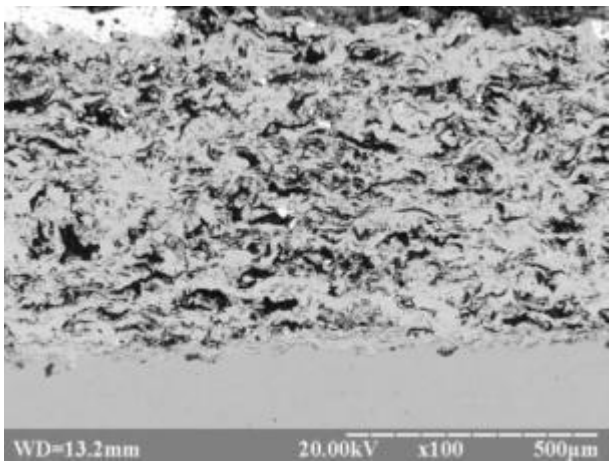
б

Рис. 1. Морфологія суміші порошків НПГ-75 та БрАЖ10-4 зі збільшенням х50 (*а*) та х100 (*б*)

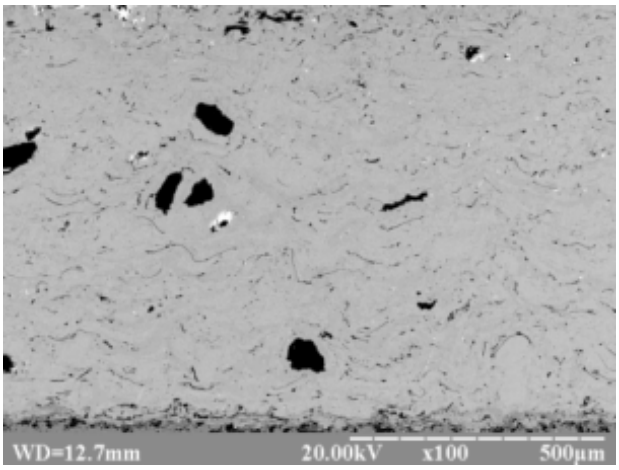
Розмір частинок у сплаві БрАЖ10-4 перебуває в діапазоні від 20 до 150 мкм, у НПГ-75 від 100 до 160 мкм.

З метою визначення складу, який зможе забезпечити в обраній системі оптимальні триботехнічні властивості (коефіцієнт тертя і зносостійкість), підготовлено чотири типи сумішей, що характеризуються різним співвідношенням компонентів покриття, а саме НПГ-75 та мідного сплаву БрАЖ10-4. Масова частка мідного сплаву у запропонованих зразках становила 35, 50, 75 та 90 %.

З отриманих порошків плазмовим методом було нанесено чотири типи покриттів, приклади структур поперечних шліфів яких показано на рис. 2. Покриття наносили на підкладку зі сталі.



а



б

Рис. 2. Структура композиційного плазмового покриття на основі НПГ-75 (збільшення х100), що містить у складі: *а* – масову частку порошку БрАЖ10-4 50 %; *б* – масову частку порошку БрАЖ10-4 90 %

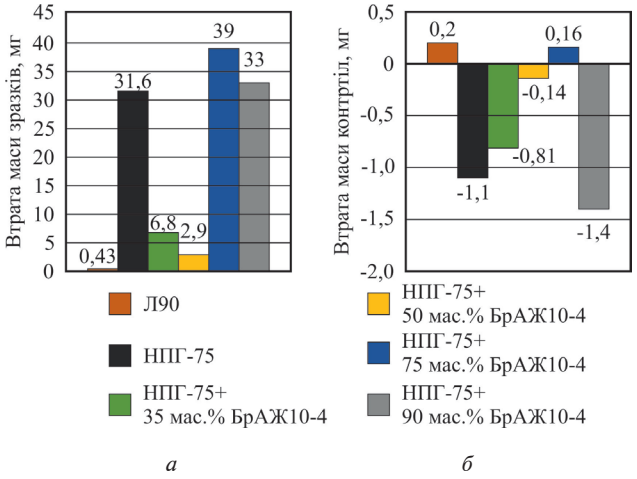
Як видно з рис. 2, структура покриття двофазна, складається з металевої матриці із включенням графіту. Нікелева плакувальна оболонка

порошку НПГ-75 не виражена, що може свідчити про високий рівень когезії отриманого покриття.

Варто зазначити, що зі збільшенням масової частки мідного сплаву у покритті змінюється структура, а ламелі стають менш вираженими.

Виходячи зі структури покриття можна передбачити, що зі збільшенням масової частки нікель-графіту має зростати ефективність змащування трибоконтакту графітовою фазою. Наслідком цього буде суттєво менше зношування сталевого контртіла або масоперенесення на нього матеріалу покриття за помірного зношування зразка зі значним (масова частка 25 % і більше) вмістом НПГ-75.

Щоб визначити протизношувані властивості запропонованих покриттів, було проведено випробування шести типів зразків, з-поміж яких були латунь Л90, плазмове покриття із НПГ75 й чотири типи композитних плазмових покриттів на основі нікель-графіту з додаванням мідного сплаву у співвідношенні (масова частка, %) НПГ-75/БрАЖ10-4: 65/35, 50/50, 25/75 та 10/90. Отримані результати триботехнічних випробувань показано на рис. 3, 4 й 5.



а *б*

Рис. 3. Масове зношування зразків (*а*) і контрзразків (*б*) у результаті досліджень на зношування

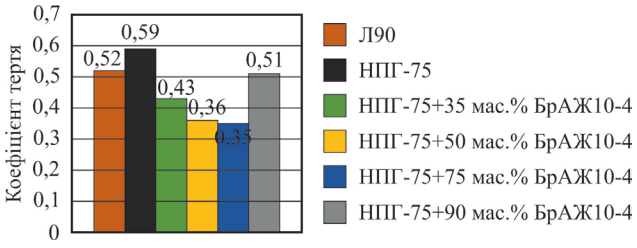


Рис. 4. Залежність коефіцієнта тертя зразків по сталі 45 від вмісту БрАЖ10-4 у покритті на основі НПГ-75

Як можна бачити з рис. 4, плазмове покриття на основі НПГ-75 дозволяє зменшити зношення сталевого контрзразка. Спостерігалося активне намазування матеріалу покриття на контрзразок, що призводило до збільшення його маси після тертя на 1,1 мг порівняно з вихідною. Утім, руйнування зразка покриття дуже інтенсивне. Коефіцієнт тертя більший, ніж у вихідної трибопари «сталь 45 – Л90», що може свідчити про адгезійний характер зношування через схоплення нікелевої фази покриття зі сталевим контрзразком.

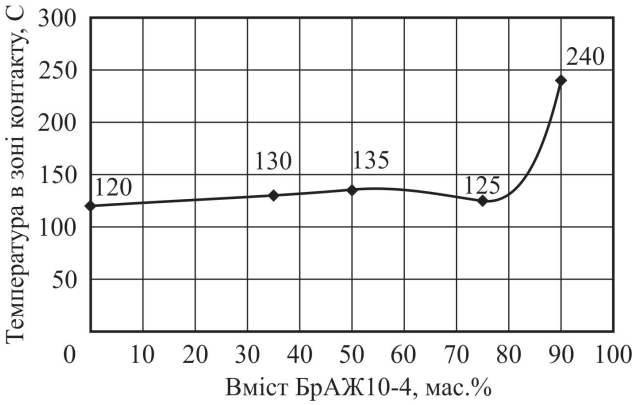


Рис. 5. Температура в зоні трибоконтакту залежної від вмісту БрАЖ10-4 у покритті на основі НПГ-75

Введення порошку БрАЖ10-4 у склад покриття масовою часткою 35 % дозволяє більш ніж у 4,5 рази зменшити зношування покриття і приблизно на 27 % знизити коефіцієнт тертя. При цьому спостерігається все ще значне намазування продуктів зношування на сталеве контртіло – приріст маси становив 0,81 мг.

Збільшення масової частки мідного сплаву у покритті до 50 % призвело до подальшого підвищення зносостійкості (приблизно у 10 разів) покриття, при цьому намазування продуктів тертя на контрзразок зменшилось приблизно у 5 разів (приріст маси контртіла зменшився з 0,81 мг до 0,14 мг), а коефіцієнт тертя найнижчий з-поміж досліджених у роботі трибопар, що може свідчити про зміну характеру зношування з адгезійного на окисний.

Подальше підвищення вмісту мідного сплаву у складі покриття масовою часткою до 75 та 90 % спричинило значне збільшення зношування зразка покриття та інтенсифікацію намазування на контріло (приріст маси найбільший серед усіх трибопар), й підвищення температури у зоні контакту і коефіцієнта тертя, що є свідченням схоплення у процесі тертя.

Висновки

У роботі з метою підвищення зносостійкості вузлів тертя автори запропонували метод формування композиційного плазмового покриття на основі плакованого нікелем графіту (НПГ-75) з додаванням мідного сплаву БрАЖ10-4 масовою часткою від 35 до 90 %.

Додавання БрАЖ10-4 у склад покриття на основі НПГ-75 масовою часткою 35 % дозволяє зменшити приблизно на 27 % коефіцієнт тертя щодо покриття з чистого НПГ-75. При цьому спостерігається активне зношування матеріалу покриття і його намазування на контртіло.

Зі збільшенням масової частки мідного сплаву до 50 % у складі порошку на основі Ni-C спостерігається тенденція поліпшення основних триботехнічних характеристик покриттів: зносостійкість збільшується у 10 разів, а значення коефіцієнта тертя зменшується з 0,59 до 0,36. Намазування на контртіло при цьому найменше серед трибопар, що були досліджені у межах цієї роботи. Це дозволяє суттєво розширити сфери застосування матеріалу Ni-C + БРАЖ10-4 в авіаційній та космічній техніці.

Подальше збільшення масової частки порошку БрАЖ10-4 у складі недоцільне, оскільки призводить до виникнення схоплення у зоні контакту, про що свідчить суттєве (у 10 разів) підвищення зношуваня зразків покриття, зростання температури у зоні контакту та зростання приблизно у 2 рази коефіцієнта тертя.

Такі принципові відмінності триботехнічних характеристик запропонованих трибопар швидше за все виникають через те, що змінюється механізм зношування з адгезійного (схоплення) на окисне зношування. У наступних роботах доцільно продовжити дослідження впливу на триботехнічні характеристики покриттів системи нікель-графіт інших антифрикційних мідних сплавів і дослідити хімічний склад продуктів, що утворюються у зоні тертя.

Щоб визначити механізми формування плазмового покриття системи плакований нікелем графіт-мідний сплав у наступних роботах має сенс дослідити змочування компонентів, що входять до складу плазмового покриття.

Подяка

Публікація містить результати досліджень, проведених за фінансової підтримки Національного фонду досліджень України, проєкт № 2023.04/0066.

References

[1] K.M. Zohdy, M.M. Sadawy, M. Ghanem, “Corrosion behavior of leaded-bronze alloys in sea water”, *Materials Chemistry and Physics*, Vol. 147, Iss. 3, 15 October 2014, pp. 878–883. doi.org: 10.1016/j.matchemphys.2014.06.033

[2] Yang Li, Kang He, Chengwei Liao, Chunxu Pan, “Measurements of mechanical properties of α-phase in Cu–Sn alloys by using instrumented nanoindentation”, *Journal of Materials Research*, January 2012, no. 27 (01), pp. 192–196. doi.org: 10.1557/jmr.2011.299

[3] J. S. Carlton. FREng, in Marine Propellers and Propulsion (Fourth Edition), 2018. doi.org: 10.1016/C2014-0-01177-X

[4] J. Freudenberger, L. Tikana, W.F. Hosford. Alloys: Copper. Encyclopedia of Condensed Matter Physics (Second Edition), 2024, Vol. 5, pp. 601–634. doi.org: 10.1016/B978-0-323-90800-9.00116-5.

[5] ASTM B438 Standard Specification for Sintered Bronze Bearings (Oil-Impregnated), (2004 EDITION), 10 p.

[6] R. Selver, Remzi Varol. Some thermal and physical characteristics of sintered tin bronze bearings, January 2002, Metall 57(1-2):28-32. Retrieved from: https://www.researchgate.net/publication/288103212_Some_thermal_and_physical_characteristics_of_sintered_tin_bronze_bearings

[7] K. Ramasamy, K. Subramanian, G. Sozhan, Su. Venkatesan, “Studies on mechanical properties of the sintered bronze-graphite composites”, *International Journal of Rapid Manufacturing*, January 2019, 8(1/2):16, Published Online: December 18, 2018, pp. 16–33. doi.org: 10.1504/IJRAPIDM.2019.097028

[8] Rasha Hussein, Haydar Al-Ethari, Talib A. Jasim. Microstructure, hardness, and wear rate of hot squeezed Cu-10Sn-graphite composite prepared by stir casting. Microstructure, Hardness, and Wear Rate of Hot Squeezed Cu-10Sn-Graphite Composite Prepared by Stir Casting, July 2023, Conference: 4th International Scientific Conference of Engineering Sciences and Advances Technologies At: University of Babylon, doi.org: 10.1063/5.0163884

O.V. Kushchev, O.E. Terentjev, V.T. Varchenko, T.M. Chevichelova, R.E. Kostyunik, O.P. Umanskyi

METHOD FOR INCREASING THE NICKEL-GRAPHITE-BASED COMPOSITE PLASMA COATING WEAR RESISTANCE BY ADDING COPPER ALLOYS

Background. The article considers the possibility of expanding the areas of application of plasma coatings based on the nickel-graphite system by adding lead-free copper alloys to the composition, which have an affinity for the nickel shell of the nickel-graphite powder and, in addition to preserving the tribotechnical properties of nickel-graphite, can provide the necessary level of the coating cohesion.

Objective. The paper aims to study the effect of the content of the copper alloy CuAl10Ni5Fe4 when added to nickel-graphite (NPG-75) on the tribotechnical characteristics of the composite plasma coating.

Methods. As part of the work to determine the influence of the copper alloy content in nickel-graphite on the tribotechnical properties of the plasma composite coating, friction and wear tests were conducted on samples with 35, 50, 75 and 90 wt. % copper alloy CuAl10Ni5Fe4 in nickel-clad graphite NPG-75 (75 wt. % nickel). The double copper-zinc alloy CuZn10 was used as a standard.

Results. It has been experimentally established that plasma coatings based on NPG-75 with the addition of 50 wt. % CuAl10Ni5Fe4 have wear resistance increased by 10 times compared to pure NPG-75 with a friction coefficient significantly lower (0.36) than the CuZn10 standard (0.52). The coating on the counterbody of DIN C45 (hardness 45-48 HRC) is the smallest among the friction pair that were investigated in this work. Reducing the copper alloy content to 35 wt. % is impractical, since this results in active wear of the coating material and its spreading onto the counter body. Further increase in the content of CuAl10Ni5Fe4 powder in the composition leads to the occurrence of seizure in the contact zone, as evidenced by a sharp (10-fold) increase in the wear of the coating samples, an increase in the temperature in the contact zone, and a ~2-fold increase in the friction coefficient.

Conclusions. The optimal content of CuAl10Ni5Fe4 in a plasma coating based on NPG-75 is about 50 wt. %. This material allows you to reduce the friction coefficient from 0.52 (CuZn10) to 0.36 when paired with steel DIN C45. The mass wear of the steel counterbody is also significantly reduced, probably due to a change in the wear mechanism from adhesive (seizing) to oxidative wear. This allows us to significantly expand the scope of application of the Ni-C + CuAl10Ni5Fe4 material in aviation and space technology.

Keywords: nickel-graphite; composite powder; plasma coating; chemical composition; structure; friction and wear.

Рекомендована Радою
Навчально-наукового інституту матеріалознавства
та зварювання імені Є.О. Патона КПІ ім. Ігоря Сікорського

Надійшла до редакції
11 листопада 2024 року

Прийнята до публікації
30 червня 2025 року

АВТОМАТИЗАЦІЯ, КОМП'ЮТЕРНО-ІНТЕГРОВАНІ ТЕХНОЛОГІЇ ТА РОБОТОТЕХНІКА

DOI: 10.20535/kpisn.2025.2.331284
UDC 536.531

S.M. Matvienko[1*], G.S. Tymchik[1]
Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
[*]Corresponding author: s.matvienko@kpi.ua

ANALYSIS OF CIRCUIT DESIGN METHODS FOR LINEARIZING THE TEMPERATURE CHARACTERISTICS OF NTC-THERMISTORS AND THE MEASUREMENT CHANNEL CHARACTERISTICS

Background. Temperature measurement devices play a crucial role in the automation of various systems and in monitoring diverse physicochemical parameters. thermoresistors, thermistors, or thermocouples are commonly used as temperature sensors. However, the nonlinear temperature-resistance characteristic of thermistors presents challenges for accurate temperature-to-electric signal conversion. Therefore, the issue of linearizing the output characteristic of the temperature measurement channel remains highly relevant. The task of linearizing the temperature dependence of the measurement channel can be addressed using hardware (circuit-based), software, or combined hardware-software approaches.
Objective. The purpose of the research is to analyse the effectiveness of circuit-based (schematic) methods for linearizing the temperature characteristics of NTC-type thermistors and to evaluate their impact on the overall performance of the temperature measurement channel.
Methods. A structural model of a temperature measurement device utilizing an NTC-thermistor as a sensor was developed in the MATLAB Simulink environment. Five different circuit configurations for connecting the thermistor to the measurement channel were analysed to assess the effectiveness of linearization within the temperature ranges of ±15 °C and ±25 °C. Based on the proposed connection schemes, a simulation of the temperature characteristics of NTC-thermistors was carried out.
Results. The results of the study for five circuit-based methods of linearizing the temperature characteristics of NTC-thermistors provide an opportunity to evaluate the effectiveness of applying different circuit-based methods for linearizing the temperature characteristics of NTC-thermistors and the measurement channel characteristics in two temperature ranges: ±15 °C and ±25 °C. Based on the presented simulations, it can be stated that by applying these circuit-based methods for linearizing the temperature characteristics of NTC-thermistors and the measurement channel characteristics, the measurement error can be reduced by 40 % in a limited temperature range compared to using a balanced measurement bridge, and corrections for compensating the self-heating effect of the thermistor can be determined.
Conclusions. By modelling the circuit-based methods for linearizing the temperature characteristics of NTC-thermistors, it becomes possible to evaluate the effectiveness of linearizing the temperature dependence of the measurement channel while investigating various options for connecting the thermistor to the measurement channel. Additionally, the parameters of the measurement channel elements of the temperature-measuring device can be determined.
Keywords: temperature measurement; NTC-thermistor; MATLAB Simulink; linearization.

Problem statement

Regardless of where temperature measurement is performed using different devices, the primary requirement for the results of such measurements is their accuracy. The overall technical and economic parameters of the temperature measurement system are determined by the sensor, the type of which depends on the requirements needed from the system as a whole. Sensors, particularly commonly used thermistors with a TCR – negative temperature coefficient (or NTC – Negative Temperature Coefficient), have a fairly wide working temperature range, remote monitoring capabilities, can operate in strong magnetic fields, and have small dimensions. One of the most significant drawbacks of thermistors with a negative TCR is the nonlinearity of their $R(T)$ characteristic – the relationship between the

Пропозиція для цитування цієї статті: С.М. Матвієнко, Г.С. Тимчик, "Аналіз схемотехнічних методів лінеаризації температурних характеристик NTC-термісторів і характеристик вимірювального каналу", *Наукові вісті КПІ*, № 2, с. 45–56, 2025. doi: 10.20535/kpisn.2025.2.331284

Offer a citation for this article: S.M. Matvienko, G.S. Tymchik, "Analysis of circuit design methods for linearizing the temperature characteristics of NTC-thermistors and the measurement channel characteristics", *KPI Science News*, no. 2, pp. 45–56, 2025. doi: 10.20535/kpisn.2025.2.331284

resistance of the thermistor and its temperature. This relationship has an exponential form.

The linearity of the measurement channel in a temperature measurement device with a thermistor sensor is influenced not only by the nonlinearity of its $R(T)$ characteristic but also by the nonlinearity of the transfer characteristic of the interface circuit to which the thermistor is connected and the self-heating effect of the thermistor when an electric current passes through it. Thus, the task of linearizing the temperature dependence of the measurement channel remains open. Among the circuit techniques, the most common are passive correction circuits and resistive voltage dividers, which allow forming a linear section on the $R(T)$ characteristic within a specific temperature range. These methods are quite economical compared to software methods that require a microcontroller.

The problem can be solved by determining a rational design for the measuring probe, selecting the correct characteristics and operating modes of the thermistor, and implementing optimal operating parameters for the measurement channel. At the stage of developing the temperature measurement device, it is important to carry out mathematical modelling of the measurement channel to choose the optimal linearization method and define the parameters of the components. Creating mathematical models to study linearization methods allows for determining the characteristics of the measurement channel components depending on the specific application. In this work, using the developed mathematical model created in MATLAB Simulink, methods for linearizing the device characteristics were studied, measurement errors were evaluated, and recommendations were developed to improve measurement accuracy.

Analysis of recent research and publications

Nonlinearity in the characteristics of most sensors and measurement channels in devices is an important problem. Their linearized characteristics simplify the design, and calibration, and improve measurement accuracy. To compensate for the nonlinearity of the characteristics of thermistors and the measurement channel, various linearization methods are presented in the literature. In [1, 2], an analysis of such methods is conducted. These works provide an overview of different methods applied for linearizing sensor characteristics. It is noted that due to the availability of high-performance analogue devices, circuit-based (analogue) methods are still popular among many researchers, although the use of digital methods combined with software methods provides better results and ensures flexibility and efficiency. The popularity of circuit-based linearization methods is explained by the simplicity of implementation and, compared to software methods, their more economical nature.

As is known, the nonlinearity of the characteristics of measurement channels in devices is primarily influenced by: the nonlinearity of the $R(T)$ dependence of the thermistor's resistance on its temperature, the self-heating effect of the thermistor when an electric current passes through it, the nonlinearity of the voltage across the thermistor in a voltage divider or a measurement bridge arm concerning its resistance, and the nonlinearity of the signal amplifier's characteristic.

Modelling the specific implementation of devices with the required characteristics using the necessary linearization method significantly simplifies the development process and helps to optimally and quickly select the required component parameters and achieve the desired result. Manufacturers of nonlinear components also offer their own models for connecting thermistors. For example, online services like "NTC Thermistor Simulation" on the TDK website [3] and the Murata Electronics website [4]. Mitsubishi Materials offers the "Chip Thermistor Resistance Simulator" as an Excel file [5]. The models provided by the manufacturers help solve the problem of selecting a specific type of thermistor for a particular case and partially address the nonlinearity of the thermistor's $R(T)$ characteristic. Scientists, engineers, and specialists also propose their own ways of solving the problem of nonlinearity in the $R(T)$ characteristic of the sensor [6, 7]. These models and methods do not provide a full solution to the problem in the device, considering the specific requirements required from the device as a whole.

The research aims to construct (develop) an original structural-mathematical model of the measurement channel of a temperature measurement device using an NTC-thermistor in the MATLAB Simulink environment, to investigate using the created model different circuit-based methods of linearizing the temperature characteristics of NTC-thermistors and measurement channel elements, analyze the causes of measurement errors, choose the optimal method and characteristics of the measurement channel elements to improve the temperature measurement accuracy. The goal of the work is to investigate, using the built models, various ways of linearizing the temperature characteristics of NTC-thermistors and elements of the measurement channel.

The presentation of the main research material

The essence of temperature measurement using an NTC-thermistor is to use the dependence of the thermistor resistance on its temperature, that is, on the temperature of the environment surrounding the thermistor.

The dependence of the electrical resistance of an NTC-thermistor on temperature has the form [8]:

$$R_T = R_N \exp B\left(\frac{1}{T} - \frac{1}{T_N}\right), \quad (1)$$

where: R_T – NTC-thermistor resistance by temperature T, K;

R_{TN} – NTC-thermistor resistance at nominal temperature T_N, K;

T, T_N – temperature, K;

B – constant coefficient that depends on the thermistor material, K.

Fig. 1 shows the dependence of the thermistor resistance on temperature. As can be seen from Fig. 1, this dependence is exponential in nature, and therefore is nonlinear, which creates certain difficulties when creating temperature measuring devices and introduces additional error when attempting to linearize this characteristic. The value of the error will depend on the linearization method used.

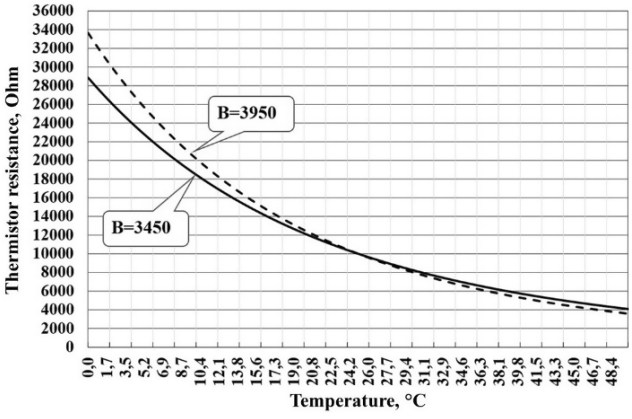


Fig. 1. Dependence of thermistor resistance on temperature at coefficient values B = 3450 K and B = 3950 K

Often, the form of the temperature characteristic of an NTC-thermistor does not meet the requirements for a specific application of the device. The use of passive correction circuits (Linear resistance networks) [6, 7] allows for modifying the $R(T)$ dependence in such a way that a maximally linear dependence segment is achieved within a specific temperature working range.

The simplest circuit-based options for such networks to correct the $R(T)$ dependence of the thermistor are the series, parallel, or series-parallel connection of resistors (Fig. 2), whose temperature coefficient of resistance (TCR) is very small compared to TCR of the thermistor. Connecting a resistor in parallel with the thermistor compensates for its nonlinearity, a series connection adjusts the sensor's sensitivity, and a series-parallel connection creates a linear characteristic with the required slope.

When a resistor is connected in series with the thermistor, the slope of the $R(T)$ characteristic decreases by an amount proportional to the ratio $\frac{R_T}{R_1 + R_T}$, which slightly reduces the non-linearity of the characteristic, but at the same time, the sensitivity of the sensor decreases proportionally. When a resistor is connected in parallel to balance the $R(T)$ characteristic, the resistance of the parallel resistor is determined by the formula (see Fig. 2, b) [9]:

$$R_1 = R_{TN} \times \frac{B - 2T_N}{B + 2T_N}, \quad (2)$$

where: R_1 – resistance of a resistor connected in parallel, *Ohm*;

R_{TN} – resistance of the NTC-thermistor at nominal temperature T_N, K;

T, T_N – temperature, K;

B – constant coefficient that depends on the thermistor material, K.

Series-parallel connection of resistors (see Fig. 2, c) is used to form a linear characteristic with a given slope. The total resistance R in series-parallel connection is determined by the formula [10] (3):

$$R = \frac{(R_1 + R_T)R_2}{R_1 + R_2 + R_T}. \quad (3)$$

If resistors R_1 and R_2 have a fixed ratio with respect to the resistance of the thermistor at its nominal temperature R_N (average temperature in a given range). So, having made the substitution $R_1 = a \cdot R_N$ and $R_2 = b \cdot R_N$, we can write formula (3) [11] as:

$$R = \frac{(aR_N + R_T)bR_N}{aR_N + bR_N + R_T}. \quad (4)$$

For the predicted linear function to pass at the inflection point of the nonlinear function T_N, and the transfer function to have a sensitivity (slope) at this point equal to the desired value of Ω/K, it's necessary to determine the values for a and b. Based on

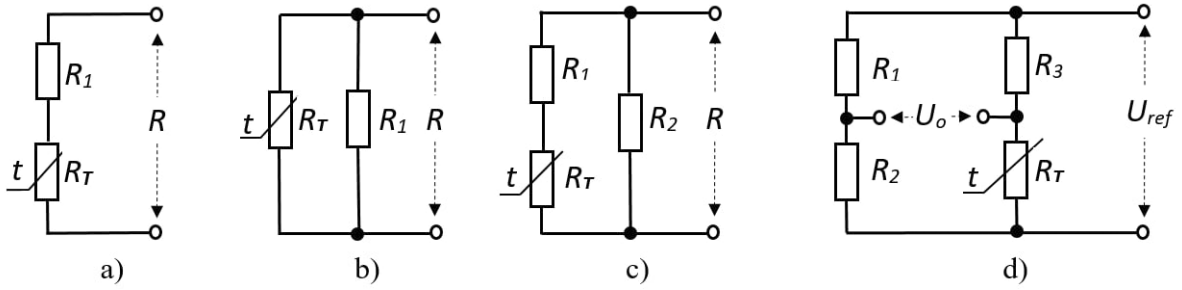


Fig. 2. Correction of the $R(T)$ thermistor dependence by sequential (a), parallel (b), series-parallel (c) resistor connections and using the non-linearity of the voltage divider in the Wheatstone bridge circuit (d)

formula (2), when resistors are connected in parallel to the thermistor $R_1 = a \cdot R_N$ and $R_2 = b \cdot R_N$ the sum of a and b is determined by the formula (5) [10]:

$$a + b = \frac{B - 2T_N}{B + 2T_N}. \quad (5)$$

The coefficients a are respectively equal to [9]:

$$a = \frac{B - 2T_N}{B + 2T_N} - b, \quad (6)$$

and coefficient b [6]:

$$b = \frac{2T_N}{B + 2T_N}\sqrt{\frac{R'_N}{R_N}} = \frac{2T_N}{B + 2T_N}\sqrt{\frac{-mB}{R_N}}, \quad (7)$$

where: R'_N is the nonstandard value and R'_N – is the standard value thermistor resistance at nominal temperature T_N, m – required slope of the linearized transfer function, Ω/K.

Temperature measurement using a thermistor can be carried out by measuring the voltage across the thermistor, connected to a DC source or through a divider to a voltage source. The voltage across the thermistor U_T is proportional to its resistance and, accordingly, proportional to the temperature of the thermistor. The diagram of such a divider is shown in Fig. 3.

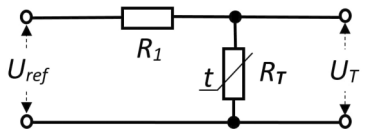


Fig. 3. Voltage divider circuit

The voltage dependence on the thermistor in the divider has the form (8):

$$U_T = U_{ref} \times \frac{R_T}{R_1 + R_T}. \quad (8)$$

In Fig. 4, the dependence of the voltage across the thermistor on its resistance is shown for different ratios $k = \frac{R_1}{R_{TN}}$, where R_{TN} is the resistance of the NTC-thermistor at the nominal temperature T_N, assuming the resistance of the thermistor changes linearly. In this case, for example, $U_{ref} = 10$ V, and the thermistor's resistance changes linearly from 30 $k\Omega$ to 4 $k\Omega$. The resistance range is based on the characteristics of the RH18 6H103 thermistor from Mitsubishi Materials, with a temperature range from 0 °C to +50 °C. At 25 °C, the resistance of RH18 is $R_{TN} = 10$ $k\Omega$ ±1 %.

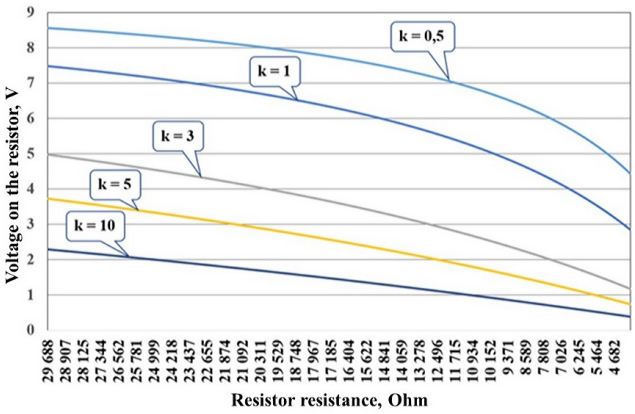


Fig. 4. Dependence of the voltage across the R_T resistance on the thermistor resistance at different ratios k, where $k = R_1/R_{TN}$

As seen from formula (8) and Fig. 3, the nonlinearity has a hyperbolic nature, and as the value of k increases, i.e., as the resistance of resistor R_1 increases, the degree of linearity – the approximation of the curve $U_T(R_T)$ to a straight line – grows. This feature is used to correct the nonlinear dependence $R(T)$ of the NTC-thermistor. By selecting the appropriate value of k (the resistance of resistor R_1) in the voltage divider, a certain degree of linearity of the characteristic can be achieved.

For accurate measurement of the electrical parameters of sensors when creating the interface circuit, Wheatstone impedance bridges are widely used. This bridge circuit consists of two legs of voltage dividers, one of which contains the thermistor. Wheatstone bridge circuits are designed to measure an unknown electrical resistance by unbalancing two sides of the bridge circuit – one of which contains an unknown component. The bridges can be symmetric, asymmetric, balanced, and unbalanced. Wheatstone bridge circuits are simple, efficient, and very useful for temperature monitoring.

The Wheatstone bridge has a single element with a variable impedance, and the further its resistance is from the balance point, the more the voltage in the bridge diagonal becomes nonlinear with respect to the resistance of the element with variable impedance. When using the Wheatstone bridge for compensating the nonlinearity of the thermistor, the character of the bridge's nonlinearity should be opposite to that of the thermistor's nonlinearity. This is difficult to achieve, but partial compensation can be obtained in a narrow temperature range.

To compensate for the nonlinearity of the Wheatstone bridge, researchers and developers propose several circuit design solutions [11], which are based on introducing feedback with a specific coefficient β, defined by the formula:

$$\beta = \frac{1+k}{G \cdot k}, \qquad (9)$$

where: β – feedback coefficient; G – bridge unbalanced amplifier gain; k – coefficient of the ratio of the resistances of the resistors in the bridge arm,

$$k = \frac{R_2}{R_{TN}}.$$

Among the simple and common circuit solutions for compensating for the nonlinearity of the Wheatstone bridge, the circuits shown in Fig. 5 [11, 12] are used.

In the electrical circuit of the temperature measurement device, an electric current flows through the NTC-thermistor, which heats it up. This phenomenon is called the self-heating of the thermistor. The self-heating temperature of the thermistor depends on the magnitude of the current passing through it, the materials and design of the sensor, and the thermal conductivity of the surrounding environment [13, 14]. This phenomenon is used in devices for measuring the thermal-physical properties of materials or the rate of substance flow [14, 15, 16]. The self-heating of the NTC-thermistor leads to an additional decrease in the thermistor's resistance, which distorts the measurement result. Therefore, in temperature measurement devices, a correction must be applied to the obtained measurement result, and the calculation of this correction is performed using the following formula [13]:

$$T_A = T - \frac{U^2}{\delta_{th} + R(T)} = T - \frac{I^2 \times R(T)}{\delta_{th}}, \qquad (10)$$

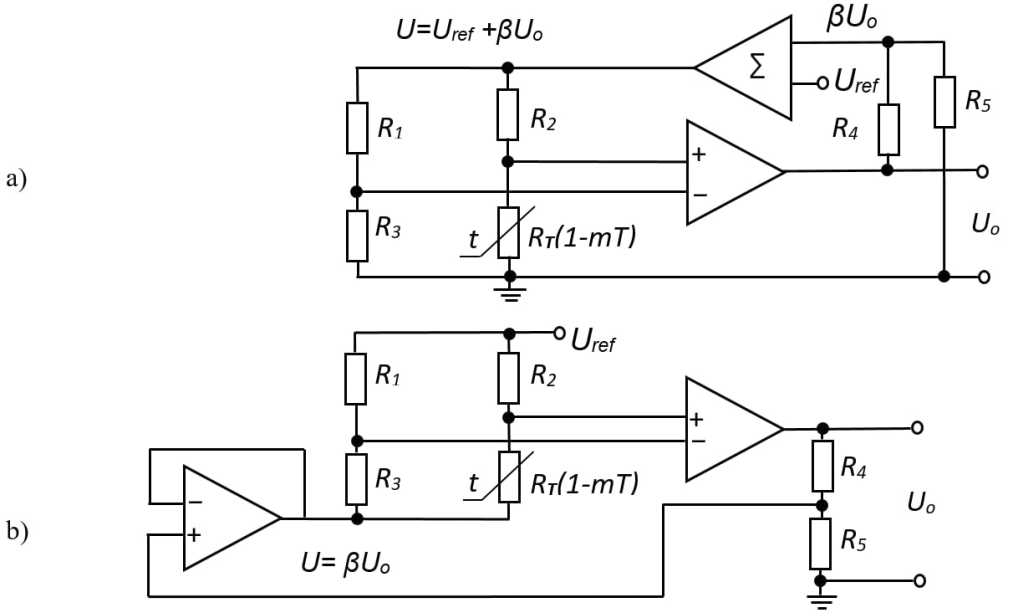


Fig. 5. Schematic solutions for compensation of Wheatstone bridge nonlinearity using feedback: a – with adjustment of the bridge supply voltage, b – with regulation of the voltage of the base (zero) pole of the bridge

where: T_A – the actual value of the controlled temperature; T – the measured temperature value; U – the instantaneous voltage across the thermistor; I – the instantaneous current flowing through the thermistor; $R(T)$ – the thermistor resistance corresponding to the temperature T; δ_{th} – the thermal dissipation coefficient in the measurement environment.

Using as a basis the circuit solutions presented in Fig. 2, 3, 5, a model of methods for compensating for the nonlinearity of the dependence of the thermistor resistance on its temperature and the influence of the thermistor self-heating effect, taking into account the nonlinearity of the transfer function of the sensor interface circuit, was designed and implemented.

Fig. 6 shows the developed measurement channel model in MATLAB Simulink, which meets these requirements.

The model consists of 3 groups of blocks for simulating the electrical circuit of the measurement channel: the "*Sensor interfacing circuit*" group, the "*Amplifier*" group, and the "*Feedback*" group.

The "*Sensor interfacing circuit*" group is the connection scheme of the thermistor model – "*Thermistor*" RT to the differential amplifier model "*Fully Differential Op-Amp*" in the "*Amplifier*" group via the appropriately configured resistor connection models "*Resistor*" R1, R2, R3, ..., R7.

The "*Thermistor*" can be connected to the Wheatstone bridge leg R1, R2, R3, RT or simply through the voltage divider R_2/R_T to the "*Fully Differential Op-Amp*". The Wheatstone bridge legs R1-R3 and R2-RT can be connected using the switch models "*Switch*" S1, S2, S3 to current source models "*Current Source*" or voltage source models "DC Voltage Source".

The electrical "*Feedback*" group consists of a resistive voltage divider R8/R9, which defines the feedback coefficient β, an operational amplifier, and switches S4, S5, which are used to configure the feedback circuit for compensating the nonlinearity of the Wheatstone bridge.

To form the current temperature value over a time series, the "*Repeating Sequence*" model is used, where a sequence of time-temperature values is input. The temperature values are then converted from Simulink format to physical signal data (temperature in K) using the "*Simulink-PS Converter*". These data are sent to the "*Controlled Temperature Source*", which is an ideal energy source in the thermal network that can maintain the controlled temperature difference. From this source, the physical temperature data are sent to the thermistor's T port.

The "*Solver Configuration*" module determines the general modelling parameters.

The "*Amplifier*" group consists of the differential amplifier model "*Fully Differential Op-Amp*" and the ideal operational amplifier model "*Op-Amp*", which provide the necessary voltage gain of the Wheatstone bridge imbalance G.

Using the "*Voltage Sensor*" models, the voltage values at the output of the "*Amplifier*" and across the

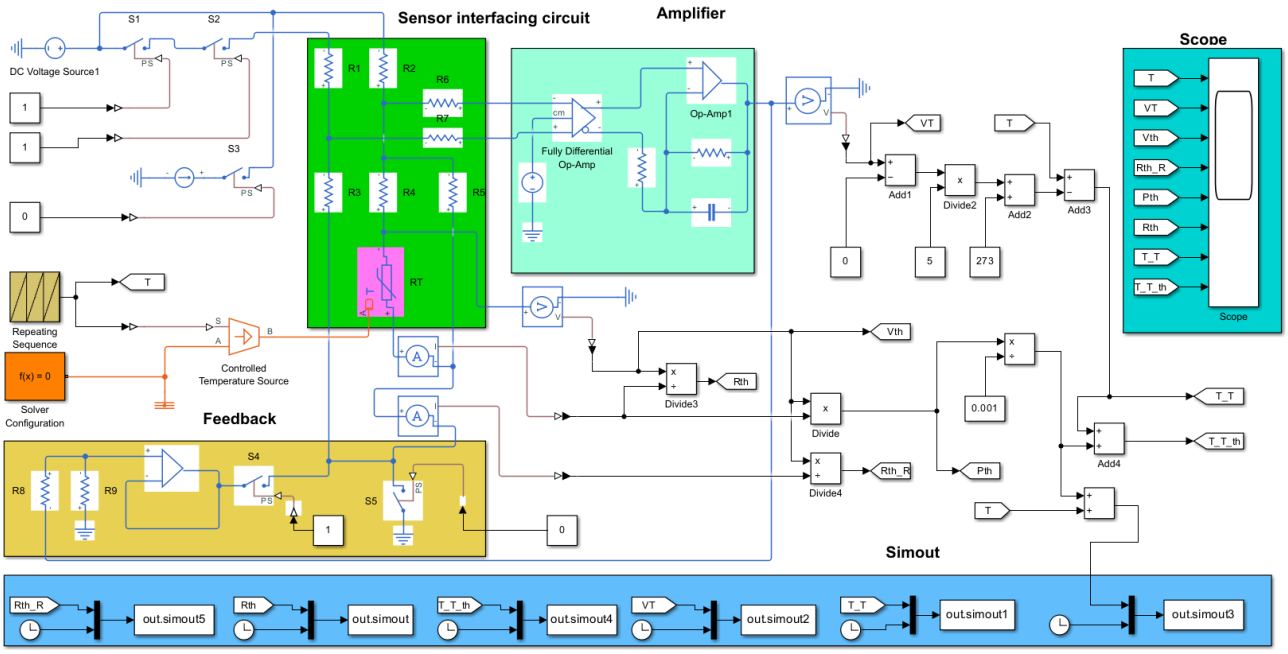


Fig. 6. Measurement channel model in MATLAB Simulink

"Thermistor" at the current time are recorded. Using the "*Current Sensor*" model, the current passing through the thermistor is recorded.

The "*PS-Simulink Converter*" modules convert the input physical electrical signal (voltage or current value) to the output signal in Simulink format. The output signal format in Simulink corresponds to the physical signal format. Then, in Simulink format, mathematical operations for calculating the current value of the thermistor's resistance, its power, and the thermistor's self-heating temperature are performed using the "*Add*" (addition or subtraction) or "*Divide*" (multiplication or division) modules. Constant coefficients and values are input using the "*Constant*" block.

The current voltage value at the output of the "*Amplifier*" is proportional to the set temperature at the current time. Using the "*PS-Simulink Converter*" module, this is converted into an output signal in Simulink format, then multiplied by the necessary coefficient in the "*Divide*" block and converted into the temperature value in K. The calculated self-heating temperature of the thermistor is added to this value. The obtained data of each measured value are recorded by the oscilloscope "*Scope*" and saved to the specified time series or array in the base workspace of MATLAB Simulink using the "*Simout*" module group.

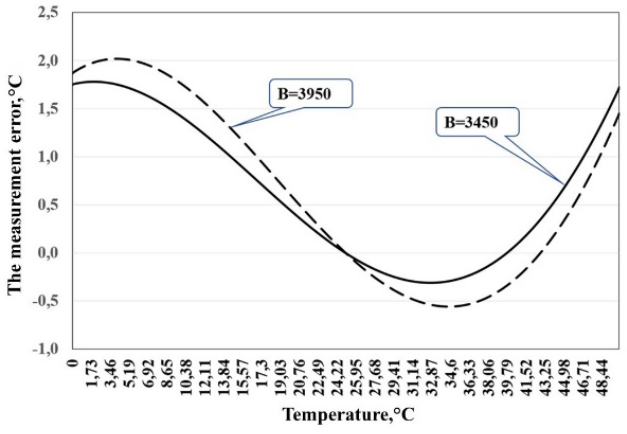
Research Results

The research was conducted in two temperature ranges for the thermistor: 298±25 K and 298±15 K, corresponding to the ranges of 0 ... +50 °C and +10 ... +40 °C, respectively. The parameters of the thermistor used were similar to those of the RH18 thermistor from Mitsubishi, with $R = 10$ $k\Omega$ (25 °C) and coefficients $B = 3450$ K and $B = 3950$ K. This thermistor has an epoxy resin casing and small dimensions (diameter 1.8 mm, length 7 mm), making it sensitive to self-heating ($\delta_{th} = 1$ mW/°C in air). This allows for the investigation of the effect on measurement errors from linearization methods in different temperature ranges with various B coefficient values, as well as determining correction values from the impact of the thermistor's self-heating effect.

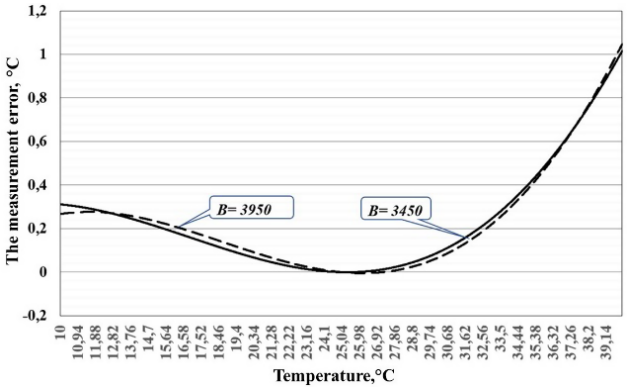
The simulation was carried out using five different interface circuit configurations for connecting the thermistor:

Variant 1. The thermistor is placed in one leg of the balanced Wheatstone bridge, where the resistance of the bridge resistors equals the thermistor's resistance at the nominal temperature of 25 °C (R1, R2, R3, $R_{TN} = 10$ $k\Omega$). The bridge is powered by a 10 V DC voltage source. Is given by S1 – switch by MATLAB Simulink model, given in Fig. 6.

Fig. 7 shows the error dependence for temperature determination in the temperature ranges from 0 °C to 50 °C and from 10 °C to 40 °C, respectively.



a



b

Fig. 7. Dependence of temperature determination error in temperature ranges from 0 °C to 50 °C (a) and from 10 °C to 40 °C (b)

The values of the root mean square (RMS) error in different temperature ranges for different values of the thermistor's B coefficient are shown in the table. These values are used for comparison with the error values when applying other linearization methods.

Variant 2. Is given by S2 – switch by MATLAB Simulink model, given in Fig. 6. The thermistor is placed in one leg of the Wheatstone bridge at different resistance ratios of the bridge $\frac{R_1}{R_3} = \frac{R_2}{R_{TN}} = k.$ The bridge is powered by a 10 V DC voltage source. It was determined that the minimum error value occurs at $k = 0.75$ in each of the temperature ranges for

both values of the thermistor's B coefficient, that is shown in Fig. 8.

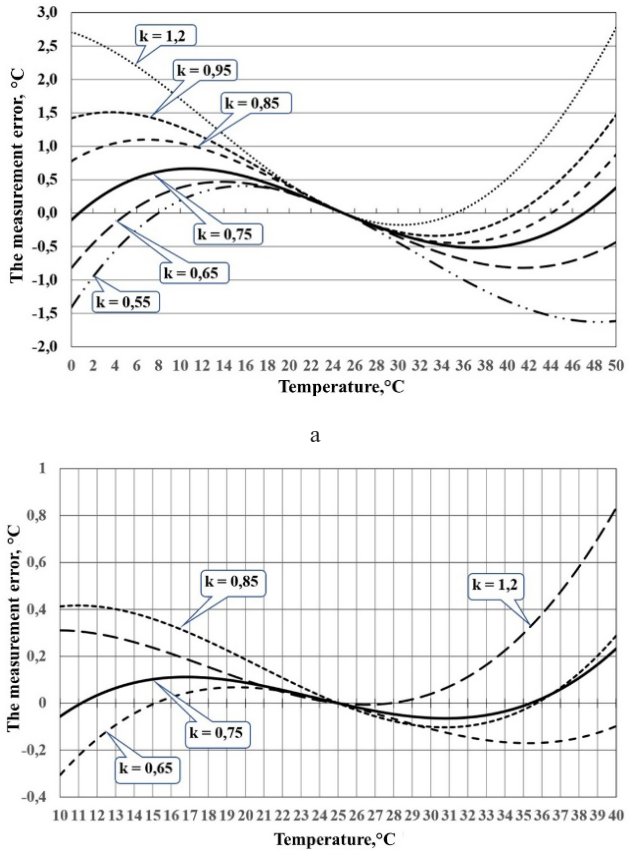


Fig. 8. Dependence of temperature measurement error in the temperature ranges from 0 °C to 50 °C (a) and from 10 °C to 40 °C (b) for the thermistor with $B = 3450$ K

The root mean square (RMS) error values in different temperature ranges for different values of the thermistor's B coefficient at $k = 0.75 \pm 0.05$ are shown in Table 1.

Variant 3. Is given by S3 – switch by MATLAB Simulink model, given in Fig. 6. The thermistor is placed in one leg of the Wheatstone bridge with the parallel connection of resistor R5 = 7054 Ω for the thermistor with B = 3450 K and R5 = 7378 Ω for the thermistor with B = 3950 K (Fig. 9 a, b), at the resistance ratio $\frac{R_1}{R_3} = \frac{R_2}{R_p} = k$, where:

$$R_p = \frac{R_{TN} \times R_5}{R_5 + R_{TN}}. \qquad (11)$$

The values of R5 are determined by formula (2). The bridge is powered from a 10 V DC voltage source.

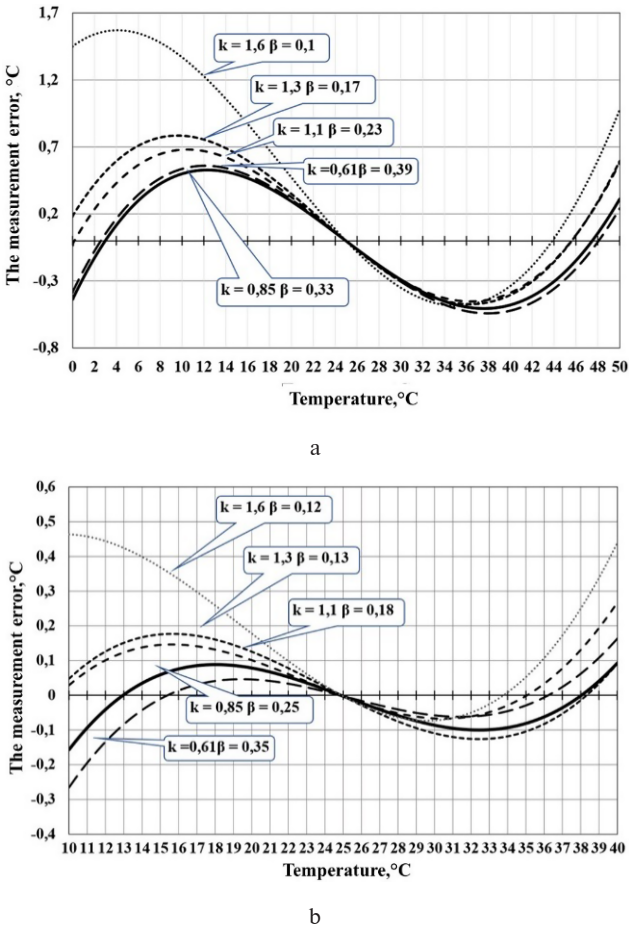


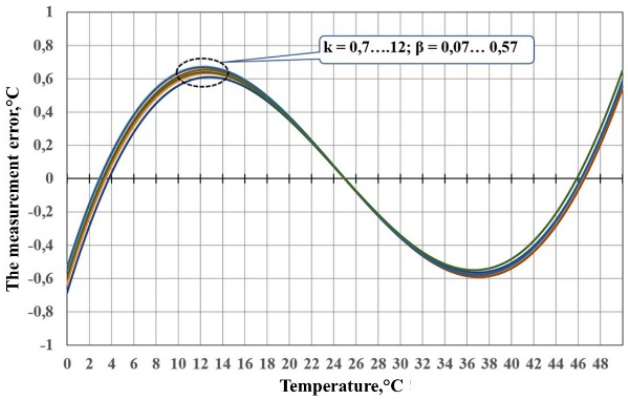
Fig. 9. Dependence of temperature measurement error in the temperature ranges from 0 °C to 50 °C (a) and from 10 °C to 40 °C (b) for the thermistor with $B = 3450$ K

In this variant, the error has a minimum value at $k = 0.85$ in each temperature range and for both values of the thermistor's B coefficient. The root mean square (RMS) error value – σ in different temperature ranges for different values of the thermistor's B coefficient at $k = 0.85 \pm 0.05$ is shown in Table 1.

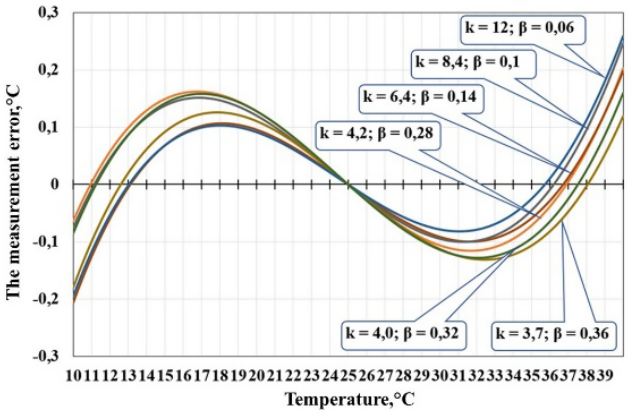
Variant 4. The thermistor is placed in one leg of the Wheatstone bridge with the parallel connection of resistor R5 = 7054 Ω for the thermistor with B = 3450 and R5 = 7378 Ω for the thermistor with B = 3950 (see formula (2)) at the resistance ratio $\frac{R_1}{R_3} = \frac{R_2}{R_p} = k$ (see formula (11)). Is given by S4 – switch by MATLAB Simulink model, given in Fig. 6. For linearizing the bridge characteristic, feedback is introduced on the operational amplifier with a feedback coefficient β, which is equal to:

$$\beta = \frac{R_9}{R_9 + R_8}. \qquad (12)$$

The bridge is powered with a 10 V DC voltage source. In this case, that's shown in Fig. 10, the error has approximately the same value for k from 0.7 to 12 in each temperature range and for both values of the thermistor's B coefficient.



a



b

Fig. 10. Dependence of temperature measurement error in the temperature ranges from 0 °C to 50 °C (a) and from 10 °C to 40 °C (b) for the thermistor with $B = 3950$ K

The average value of the root mean square (RMS) error – σ in different temperature ranges for different values of the thermistor's B coefficient is shown in the table. With an increase in k, the gain coefficient of the differential amplifier G increases, while the feedback coefficient β decreases.

Variant 5. Is given by S5 – switch by MATLAB Simulink model, given in Fig. 6. The thermistor is directly connected to the input «–» of the amplifier with the parallel connection of resistor R5 = 7054 Ω (R4 = 0 Ω) for the thermistor with $B = 3450$ K and R5 = 7378 Ω for the thermistor with $B = 3950$ K. The thermistor is powered through a resistor from a constant current source from 300 μA to 1 mA via resistor R2 = 4154 Ω for the thermistor with $B = 3450$ K and R2 = 4246 Ω for the thermistor with $B = 3950$ K. The value of R2 corresponds to the value of R_p (see formula (7)), then $k = 1$. The «+» input of the amplifier is connected to a voltage divider $\frac{R_1}{R_3} = k = 1$. The voltage source of this divider is a constant DC voltage source of 10 V.

In this variant, the thermistor is powered through a resistor from a constant current source from 300 μA to 1 mA. Regardless of the current value, the RMS error in each temperature range remains unchanged. As the current value increases, the gain coefficient of the amplifier G decreases and the correction value U – the offset voltage, or T – the temperature correction value for compensating the self-heating effect of the thermistor increases.

The generalized data of the research results are presented in Table 1. The results of the research for five circuit-technical methods of linearization of the temperature characteristics of NTC-thermistors provide an opportunity to evaluate the effectiveness of the application of different variants of circuit-technical methods of linearization of the temperature characteristics of NTC-thermistors and the characteristics of the measuring channels.

Notes: 1 – The range of k, where $k = \frac{R_2}{R_p}$ (R_p is calculated using formula (7)); 2 – The average value of RMS in the range of k; 3 – The value increases as k decreases; 4 – The value increases as k decreases; 5 – For current values of the current source from 300 μA to 1 mA; 6 – The average value of RMS for current source values from 300 μA to 1 mA; 7 – The value increases as the current source current decreases; 8 – The value decreases as the current source current decreases.

Conclusions

In the conducted study, a structural-mathematical model of the measurement channel in MATLAB Simulink was developed to study different methods of linearizing the temperature characteristics of NTC-thermistors and the measurement channel. After modelling, graphical dependencies of the temperature measurement error were obtained in the temperature ranges from 0 °C to 50 °C and from 10 °C to 40 °C for thermistors with values of $B = 3450$ K and 3950 K, using several methods for linearizing the temperature characteristics of NTC-thermistors and the measurement channel. The smallest error value was achieved by using the method of linearizing the temperature characteristic

***Table* 1.** The results of the study for five circuit-based methods of linearizing the temperature characteristics of NTC-thermistors

Variant	*k*	*B*, K	Δ*T*, K	Without self-heating thermistor			Taking into account the self-heating temperature of the thermistor in air				
				σ, K	Parameters		σ, K	Amendments			
					G	β		G	β	*U*, B	*T*, K
1	1 ±0.05	3450	±25	0.74	1.5	–	0.61	0	–	–0.53	–2.65
			±15	0.24	1.4	–	0.20	0	–	–0.53	
		3950	±25	0.92	1.35	–	0.80	0	–	–0.52	–2.6
			±15	0.24	1.35	–	0.19	0	–	–0.52	
2	0.75 ±0.05	3450	±25	0.41	1.55	–	0.41	+0.05	–	–0.7	–3.5
			±15	0.07	1.45	–	0.09	+0.02	–	–0.7	
		3950	±25	0.39	1.35	–	0.54	+0.04	–	–0.7	
			±15	0.07	1.4	–	0.09	+0.03	–	–0.7	
3	0.85 ±0.05	3450	±25	0.36	2.2	0.33	0.35	+0.16	0	–1.0	–5
			±15	0.07	2.1	0.25	0.07	0	0	–1.0	
		3950	±25	0.41	2.0	0.30	0.41	+0.16	0	–0.97	–4.85
			±15	0.08	1.5	0.17	0.09	+0.06	0	–0.96	–4.8
4	0.7 ... 12[1]	3450	±25	0.35[2]	4.3...13.1[3]	0.07...0.57	0.37[2]	0...+0.2[4]	0	0...–0.65[4]	0...–3.25[4]
			±15	0.09[2]	4.0...12.6[3]	0.06...0.36	0.08[2]	0...+0.3[4]	0	0...–0.62[4]	0...–3.1[4]
		3950	±25	0.44[2]	3.7...11.2[3]	0.07...0.41	0.46[2]	0...+0.25[4]	0	0...–0.62[4]	
			±15	0.1[2]	3.5...10.6[3]	0.07...0.41	0.1[2]	0...+0.3[4]	0	0...–0.62[4]	
5	1	3450	±25	0.31[6]	2.1...7.2[7]	–	0.33[6]	0...–0.17[8]	–	0...–0.2[8]	0...–1[8]
			±15	0.07[6]	2.0...7.0[7]	–	0.07[6]	0...–0.1[8]	–	0...–0.2[8]	
		3950	±25	0.47[6]	1.9...6.4[7]	–	0.45[6]	0...–0.1[8]	–	0...–0.2[8]	
			±15	0.09[6]	1.8...6.9[7]	–	0.08[6]	0...–0.1[8]	–	0...–0.2[8]	

of the NTC-thermistor with a parallel resistor connection and introducing feedback to the Wheatstone bridge (variant 4). Additionally, the minimum error is achieved by applying variant 5 of the thermistor connection – connecting it to a DC source with linearization of the characteristic by connecting a parallel resistor. This method has the drawback of requiring high accuracy in the current value of the current source. The measurement error when applying these methods is reduced approximately two times compared to the error obtained when connecting the thermistor in a balanced Wheatstone bridge arm (variant 1). As the *B* coefficient of the thermistor increases, the error increases for all the specified methods, while narrowing the temperature range reduces the error. Thus, in the temperature range of ±15 °C, measurement accuracy can reach less than 0.1 °C (RMS).

As the coefficient *k* decreases, i.e., when the current flowing through the thermistor increases, the self-heating temperature of the thermistor increases, introducing an error in the measurement. Except in cases where this phenomenon is used [14, 15, 16], the current value should be kept at a minimum – less than 100 μA.

As seen from the obtained modelling results, this mathematical model can be used to investigate various temperature measurement circuits, select the circuit, and determine the optimal characteristics of the measurement channel elements to achieve maximum accuracy and meet the required specifications.

References

[1] T. Islam, S.C. Mukhopadhyay, "Linearization of the sensors characteristics: a review", *International Journal on Smart Sensing and Intelligent Systems*. Vol. 12, Iss. 1, pp. 1–21, 2019. doi: 10.21307/ijssis-2019-007

[2] A.J. Lopez-Martin, A. Carlosena, "Sensor signal linearization techniques: A comparative analysis", *Proceedings of the IEEE 4th Latin American Symposium on Circuits and Systems (LASCAS)*, Cusco, Peru. February 27– March 01, 2013, pp. 1–4. doi: 10.1109/LASCAS.2013.6519013

[3] TDK Corporation, Chip NTC Thermistors (Sensor). [Online]. Retrieved from: https://product.tdk.com/en/search/sensor/ntc/chip-ntc-thermistor/simulation.

[4] Murata Manufacturing Co. NTC Thermistor Performance Simulator. [Online]. Retrieved from: https://ds.murata.co.jp/simsurfing/ntcthermistor.html?rgear=suaykx&rgearinfo=com&md5=67c837df0f254f67edb244383dec4b71

[5] Mitsubishi Materials. Chip Thermistor Resistance Simulator. [Online]. Retrieved from: https://www.mmc.co.jp/adv/en/solution/simulator.html

[6] J.G. Webster. "Measurement, Instrumentation, and Sensors Handbook", *CRC Press*, 1999. doi: 10.1201/9781003040019

[7] S.B. Stankovic, P.A. Kyriacou, "Comparison of thermistor linearization techniques for accurate temperature measurement in phase change materials", *Journal of Physics: Conference Series*. Vol. 307, iss. 1, pp. 1–6, 2011. doi: 10.1088/1742-6596/307/1/012009

[8] TDK Corporation – NTC Thermistors, General technical information: [Online]. Retrieved from: https://www.tdk-electronics.tdk.com/download/531116/19643b7ea798d7c4670141a88cd993f9/pdf-general-technical-information.pdf

[9] D.R. White, "Temperature Errors in Linearizing Resistance Networks for Thermistors". Vol. 36, pp. 3404−3420, 2015. doi: 10.1007/s10765-015-1968-2. [Online]. Retrieved from: https://link.springer.com/article/10.1007/s10765-015-1968-2

[10] Math Encounters Blog, Two-Resistor Thermistor Linearizer. [Online]. Retrieved from: https://www.mathscinotes.com/2013/12/two-resistor-thermistor-linearizer/

[11] M. Mohan, T. Geetha, P. Sankaran and J. Kumar, "Linearization of the Output of a Wheatstone bridge for a Single Active Sensor", *Proceedings of the XVIth IMEKO TC4 International Symposium on Electrical Measurements and Instrumentation*, Florence, Italy, September 22–24, 2008. Retrieved from: https://www.researchgate.net/publication/235643444_Linearisation_of_the_Output_of_a_Wheatstone_bridge_for_a_Single_Active_Sensor

[12] A.B. Narayanan, "Linearization of Wheatstone-Bridge", *Maxim Integrated*, *Application* Note 6244, 2014. [Online]. Retrieved from: www.analog.com/media/en/technical-documentation/tech-articles/how-to-linearize-wheatstonebridge-circuit-for-better-performance.pdf

[13] S. Ghosh, A. Mukherjee, K. Sahoo, S. K. Sen, and A. Sarkar, "A novel sensitivity enhancement technique employing Wheatstone's Bridge for strain and temperature measurement", *Proceedings of the 2015 Third International Conference* on Computer, Communication, Control and Information Technology (C3IT), Hooghly, India, February 07–08, 2015, pp. 1–6. doi: 10.1109/C3IT.2015.7060116

[14] S. Matvienko, V. Shevchenko, M. Tereshchenko, A. Kravchenko, and R. Ivanenko, "Determination of composition based on thermal conductivity by thermistor direct heating method", *EEJET*, February, 2020, Vol. 1, no. 5 (103), pp. 19–29. doi: 10.15587/1729-4061.2020.193429

[15] S. Matvienko, S. Vysloukh, A. Matvienko, and A. Martynchyk, "Determination thermal and physical characteristics of liquids using pulse heating thermistor method", *International Journal of Engineering Research & Science (IJOER)*. Vol. 2, Iss. 5, pp. 250–258, 2016. Retrieved from: https://scholar.google.com.ua/scholar?oi=bibs&cluster=16255719573973202880&btnI=1&hl=en

[16] G.S. Tymchik, S.M. Matvienko, I. Sikorsky, P. Kisała, K. Nurseitova, and A. Iskakova, "Improving the way of determination substances thermal physical characteristics by direct heating thermistor method", *Przegląd Elektrotechniczny*, Vol. 95, Iss. 4, pp. 121–126, 2019. doi: 10.15199/48.2019.04.21

С.М. Матвієнко, Г.С. Тимчик

АНАЛІЗ СХЕМОТЕХНІЧНИХ МЕТОДІВ ЛІНЕАРИЗАЦІЇ ТЕМПЕРАТУРНИХ ХАРАКТЕРИСТИК NTC-ТЕРМІСТОРІВ І ХАРАКТЕРИСТИК ВИМІРЮВАЛЬНОГО КАНАЛУ

Проблематика. Пристрої вимірювання температури відіграють важливу роль для автоматизації різних систем і моніторингу різних фізико-хімічних параметрів. Як сенсори використовують терморезистори, термістори або термопари. Утім, нелінійна температурна характеристика терміторів створює труднощі для точного перетворення температури в електричний сигнал, тому питання лінеаризації вихідної характеристики вимірювального каналу пристрою для вимірювання температури залишається актуальним досі. Завдання лінеаризації температурної залежності вимірювального каналу може бути вирішене апаратними (схемотехнічними), програмними та апаратно-програмними методами.

Мета дослідження. Метою роботи є аналіз ефективності схемотехнічних методів лінеаризації температурних характеристик терміторів NTC-типу, а також їх впливу на загальну характеристику вимірювального каналу.

Методика реалізації. Розроблено структурну модель пристрою для вимірювання температури із сенсором на основі термістора у середовищі MATLAB Simulink. Проведено аналіз п'яти схемотехнічних способів підключення термістора до вимірювального каналу з метою дослідження ефективності лінеаризації у межах температурних діапазонів ±15 °C та ±25 °C. На основі запропонованих способів підключення термістора здійснено моделювання температурних характеристик NTC-терміторів.

Результати дослідження. Результати досліджень для п'яти схемотехнічних способів лінеаризації температурних характеристик NTC-терміторів надають можливість оцінити ефективність застосування різних варіантів схемотехнічних методів лінеаризації температурних характеристик NTC-терміторів і характеристик вимірювального каналу у двох температурних діапазонах: ±15 °C та ±25 °C. На основі поданого моделювання можна стверджувати, що за рахунок застосування вказаних схемотехнічних способів лінеаризації температурних характеристик NTC-терміторів і характеристик вимірювального каналу можна зменшити похибку вимірювання на 40 % в обмеженому діапазоні температур порівняно з використанням збалансованого вимірювального мосту, і визначити потрібні корекційні заходи для компенсації ефекту саморозігріву термістора.

Висновки. За рахунок моделювання схемотехнічних способів лінеаризації температурних характеристик NTC-термісторів надано можливість оцінювати результативність лінеаризації температурної залежності вимірювального каналу, досліджуючи при цьому різноманітні варіанти підключення до вимірювального каналу термістора, а також визначити параметри елементів вимірювального каналу пристрою для вимірювання температури.

Ключові слова: вимірювання температури; NTC-термістор; MATLAB Simulink; лінеаризація.

Рекомендована Радою
приладобудівного факультету
КПІ ім. Ігоря Сікорського

Надійшла до редакції
25 лютого 2025 року

Прийнята до публікації
30 червня 2025 року

АВТОРИ НОМЕРА

Варченко Віктор Трифонович
ORCID: 0009-0009-8885-8434

Данилов Валерій Якович
ORCID: 0000-0003-3389-3661

Зарицький Олексій Олексійович
ORCID: 0009-0003-1243-5845

Костюнік Руслан Євгенович
ORCID: 0000-0003-0232-9208

Кущев Олексій Вікторович
ORCID: 0009-0003-9056-0123

Маслянко Павло Павлович
ORCID: 0000-0003-4001-7811

Матвієнко Сергій Миколайович
ORCID: 0000-0002-7547-4601

Мірко Сергій Сергійович
ORCID: 0009-0002-8254-7351

Романюк Вадим Васильович
ORCID: 0000-0001-9638-9572

Терентьєв Олександр Євгенович
ORCID: 0000-0002-4723-9305

Тимчик Григорій Семенович
ORCID: 0000-0003-1079-998X

Уманський Олександр Павлович
ORCID: 0000-0003-3629-7224

Чевичелова Тетяна Михайлівна
ORCID: 0009-0001-8324-2456