

ВІДГУК

офіційного опонента на дисертаційну роботу
Сінька Дмитра Павловича
на тему «Комп'ютерна модель прогнозування стану розщеплення
кластеру розподіленої інформаційної системи з мікросервісною
архітектурою»,

представлену на здобуття ступеня доктора філософії
в галузі знань 12 - Інформаційні технології
за спеціальністю 122 - Комп'ютерні науки

Актуальність теми дисертації

Сучасні інформаційні системи дедалі частіше будуються за мікросервісною архітектурою та розгортаються як розподілені кластери, що обслуговують критично важливі галузі, зокрема фінансові сервіси, телекомунікації та енергетику. Перехід від монолітних рішень до розподілених був зумовлений потребою масштабування й неперервної доступності, проте докорінно змінив природу відмов: у розподіленому середовищі збій поширюється каскадом і часто проявляється не там, де виник.

Найнебезпечнішим наслідком часткової втрати мережевої зв'язності є стан розщеплення кластера, за якого система розпадається на кілька частин, кожна з яких незалежно вважає себе повноцінною та обслуговує запити. Наявні механізми протидії, такі як кворумні політики та огороження вузлів, є за своєю природою реактивними: вони спрацьовують у момент, коли поділ уже стався, і в найкращому разі обмежують його шкоду. Водночас кількісна оцінка того, наскільки кластер близький до цього стану, у відомих джерелах залишається поза увагою.

Для України ця проблематика має додаткову гостроту, оскільки вітчизняна критична інфраструктура працює в умовах підвищених ризиків, спричинених нестабільним енергопостачанням, пошкодженнями каналів зв'язку та цілеспрямованими кібератаками на телекомунікаційні й енергетичні вузли. Кожен із цих чинників безпосередньо породжує ті часткові втрати зв'язності, що ведуть до розщеплення кластера.

З огляду на зазначене, наукове завдання, вирішене у роботі Сінька Д.П., яке полягало у розробленні та обґрунтуванні комп'ютерної моделі прогнозування відстані до розщеплення кластера на основі регресійних моделей машинного навчання та графово-інваріантного представлення стану кластера, є безумовно актуальним.

Зв'язок роботи з науковими програмами, планами, темами та грантами

Дисертаційне дослідження виконувалось у межах планових науково-дослідних робіт Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, зокрема:

- НДР «Розвиток наукових засад алгебраїчної теорії сильного штучного інтелекту стосовно кібернетичної безпеки об'єктів критичної інфраструктури в галузі енергетики» (№ ДР 0123U100913, 2023–2027 рр.)

• НДР «Розвиток розподіленої енергетики в умовах ринку електричної енергії України з використанням технологій та систем цифровізації, Розділ 1. Організаційні та математичні моделі взаємодії учасників децентралізованого ринку електроенергії» (№ ДР 0125U000237, 2025–2026 рр.)

Оцінка обґрунтованості наукових результатів дисертації, їх новизни

Обґрунтованість отриманих наукових положень підтверджується коректністю математичних побудов, повним звіренням розробленого алгоритму обчислення цільової величини з еталонним перебором усіх підмножин ребер, за якого розбіжностей не виявлено, а також результатами експериментальних досліджень, виконаних за єдиним протоколом оцінювання на незалежній контрольній вибірці. У дисертації отримано такі наукові результати, що визначають її новизну:

1. Вперше поставлено та розв'язано задачу регресійного прогнозування відстані до стану розщеплення кластера. На відміну від наявних підходів, які зводяться до виявлення факту розщеплення після його настання, запропонована постановка зберігає інформацію про ступінь близькості конфігурації до критичного стану.

2. Вперше формалізовано саме поняття відстані до розщеплення як мінімальної кількості видалень ефективних ребер, що переводять кластер у стан із щонайменше двома повнотипними функціональними островами, чим уведено кількісну міру структурної крихкості.

3. Вперше побудовано представлення стану кластера, інваріантне водночас до перестановки вузлів і до перейменування типів сервісів, що досягається канонічним впорядкуванням вузлів та канонічним маркуванням типів за структурним рангом.

4. Вперше розроблено комп'ютерну модель, яка поєднує точний алгоритм обчислення цільової величини, конвеєр формування інваріантних ознак, регресійну модель на основі градієнтного бустингу та механізми калібрування прогнозу.

5. Удосконалено спосіб ознакового опису стану кластера завдяки поєднанню канонізованого матричного представлення з набором графово-інваріантних структурних ознак.

6. Дістали подальшого розвитку методи калібрування регресійного виходу та оцінювання невизначеності прогнозу в застосуванні до задачі оцінювання структурної крихкості.

Практичне значення отриманих результатів

Практичне значення полягає у розробленні програмного комплексу, який прогнозує відстань до розщеплення та надає неперервну оцінку ризику з інтервалом надійності через вебсервіс. Комплекс легко інтегрується у системи моніторингу розподілених систем для завчасного виявлення вразливих конфігурацій. Окремо слід відзначити інженерну якість реалізації: єдина точка формування ознак гарантує однаковість ознакового опису під час навчання та обслуговування запитів, чим усувається поширене й важко локалізоване

джерело помилок промислових систем машинного навчання. Реалізацію виконано з дотриманням вимог відтворюваності, а вихідний код доступний у відкритому репозиторії, що дозволяє незалежну перевірку наведених результатів.

Теоретичне значення дисертаційної роботи

Теоретичне значення роботи полягає у розвитку наукових основ оцінювання стійкості розподілених інформаційних систем шляхом переведення задачі аналізу стану розщеплення кластера з площини бінарної констатації факту в площину неперервного вимірювання. Ключовим теоретичним внеском є формалізація відстані до розщеплення як міри структурної крихкості, що дорівнює мінімальній кількості видалень ефективних ребер до утворення щонайменше двох повнотипних островів. Таке означення, на відміну від наявних підходів, зберігає інформацію про ступінь близькості конфігурації до критичного стану та задає природне впорядкування конфігурацій за рівнем ризику, чого принципово не дає класифікаційна постановка.

Самостійне теоретичне значення має обґрунтоване в роботі подвійно інваріантне представлення стану кластера, що є інваріантним водночас до перестановки вузлів і до перейменування типів сервісів. Автор коректно показує, що обидві ці симетрії є несуттєвими для самої задачі, оскільки і стан кластера, і відстань до розщеплення залежать виключно від розбиття вузлів на типи, а не від конкретних позначень чи порядку запису елементів системи. Побудова представлення, що відповідає справжній симетрії задачі, усуває штучну надлишковість опису й дозволяє моделі навчатися саме структурним закономірностям, а не артефактам кодування. Окремо слід відзначити встановлений зв'язок запропонованих структурних ознак із цільовою величиною через її межі та релаксації, зокрема через глобальний мінімальний розріз, алгебраїчну зв'язність, мости та точки зчленування, що надає ознаковому опису змістовне теоретичне підґрунтя, а не суто емпіричний характер.

Оцінка змісту, структури та завершеності роботи

Загальний обсяг роботи становить 204 сторінки, з яких 158 сторінок припадає на основний текст. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел, що налічує 119 найменувань, та додатка. Зміст роботи ілюстровано 19 таблицями.

У вступі обґрунтовано актуальність теми, сформульовано мету та завдання дослідження, визначено об'єкт і предмет, викладено положення наукової новизни та практичного значення одержаних результатів, зазначено особистий внесок здобувача та наведено відомості про апробацію роботи.

У першому розділі розглянуто еволюцію архітектурних підходів до побудови програмних систем — від монолітної та сервіс-орієнтованої до мікросервісної й гексагональної архітектур. Проаналізовано природу стану розщеплення кластера, моделі реплікації та алгоритми узгодження, а також

фундаментальні обмеження, що описуються теоремою CAP і результатом про неможливість детермінованого консенсусу в асинхронному середовищі. Показано, що наявні рішення трактують розщеплення як дискретну подію та реагують на неї постфактум.

У другому розділі побудовано математичну модель кластера у вигляді орієнтованого графа з типізованими вузлами, у якій ефективним вважається лише взаємне ребро. Уведено поняття функціонального острова та відстані до розщеплення, обґрунтовано практично значущий діапазон цільової величини. Запропоновано інваріантне представлення фіксованої розмірності та набір із 17 структурних ознак, а також виконано порівняльний аналіз методів машинного навчання, за результатами якого обрано ансамблі дерев рішень.

Третій розділ присвячено програмній реалізації. Описано стек технологій та модульну архітектуру комплексу, наведено й детально розібрано точний алгоритм обчислення відстані з двома оптимізаціями, а також процедуру його валідації повним перебором. Викладено конвеєр формування ознак, конфігурацію та навчання базових моделей, застосування ізотонічного калібрування й конформних інтервалів, інтеграцію у вебсервіс на Django та автоматизовані тести інваріантності.

У четвертому розділі наведено результати експериментального дослідження за єдиним протоколом оцінювання: порівняння представлень, зіставлення базових моделей за якістю та обчислювальною вартістю, обґрунтування вибору функції втрат, оцінювання ефекту калібрування, а також додаткові абляції щодо обсягу даних, вагування прикладів і внеску канонізації.

Мова дисертації є академічною, посилання на джерела оформлено коректно, використання матеріалів інших авторів супроводжується належними цитуваннями. Робота є завершеною та внутрішньо цілісною, її аналіз підтверджує дотримання принципів академічної доброчесності.

Повнота публікацій матеріалів досліджень

Основні наукові положення та результати дисертаційної роботи повною мірою висвітлені у 12 наукових працях. До цього переліку входять статті у провідних фахових виданнях України категорій «А» та «Б», статті у міжнародних виданнях, проіндексованих у наукометричних базах даних Web of Science Core Collection та/або Scopus, а також 4 тези доповідей на наукових конференціях.

Особливу увагу варто приділити особистому внеску здобувача у працях, опублікованих у співавторстві. Аналіз цих внесків, чітко визначених для кожної публікації, підтверджує, що здобувач відігравав ключову роль на всіх етапах дослідження. Основні напрямки особистого внеску автора включають постановку задачі дослідження, обґрунтування поняття відстані до розщеплення як оцінки ризику розпаду кластера на острови, проведення огляду методів машинного навчання та формування гіпотези про склад ансамблю моделей, а також практичну реалізацію проактивної системи виявлення розщеплення кластера.

Зауваження щодо результатів, змісту та оформлення дисертації:

1. Означення відстані до розщеплення (формула 2.4) записано у змішаному словесно-математичному вигляді, де умова належності до стану розщеплення сформульована словами. Доцільно було б увести окремий предикат стану та подати означення повністю формально, що зробило б запис компактнішим і однорідним із рештою формул розділу.

2. Лістинг 3.1 подано мовою реалізації з англomовними ідентифікаторами всередині україномовного тексту. Наведення псевдокоду або принаймні україномовних коментарів полегшило б сприйняття, тим паче що подальший покроковий розбір алгоритму виконано ґрунтовно й докладно.

3. Лістинг 3.2, що містить функцію перетворення відстані у показник ризику, фактично складається з двох рядків і дублює формулу (2.6). Його доцільно було б подати безпосередньо в тексті без виділення в окремий лістинг.

4. Таблиця 3.3 перелічує лише назви ключових гіперпараметрів базових моделей без зазначення їхніх конкретних значень. Наведення використаних значень підсилило б підрозділ 3.9, присвячений відтворюваності результатів.

5. Конвеєр формування ознак (підрозділ 3.3) описано послідовно текстом. Блок-схема з чотирма етапами перетворення зробила б структуру конвеєра очевидною з першого погляду та полегшила б повторну реалізацію запропонованого підходу.

6. Підрозділ 3.1.1, присвячений стеку технологій, викладено доволі розлого, з детальним описом переваг кожної бібліотеки. Частину цього матеріалу можна було б згорнути у таблицю формату «компонент — призначення — версія», що водночас підсилило б вимоги до відтворюваності.

Наведені зауваження мають рекомендаційний характер, стосуються переважно форми викладення матеріалу та не ставлять під сумнів обґрунтованість основних наукових положень дисертації і не впливають на загальну позитивну оцінку роботи.

Висновок про дисертаційну роботу

Дисертаційна робота Сінька Д.П. «Комп'ютерна модель прогнозування стану розщеплення кластеру розподіленої інформаційної системи з мікросервісною архітектурою» є завершеною, самостійно виконаною працею, в якій отримано нові теоретичні та практичні результати. За актуальністю обраної теми, обсягом та рівнем виконаних досліджень, повнотою вирішення наукових і практичних задач, новизною і ступенем обґрунтованості отриманих результатів та практичних висновків дисертаційна робота відповідає вимогам, які висуваються до дисертацій на здобуття ступеня доктора філософії з галузі знань 12 «Інформаційні технології».

Робота відповідає вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії»,

затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Сінько Дмитро Павлович заслуговує на присудження ступеня доктора філософії в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Офіційний опонент:

В.о. завідувача кафедри ігрових та інтелектуальних систем
факультету інформатики та обчислювальної техніки

Національного технічного університету України «Київський політехнічний
інститут імені Ігоря Сікорського»

доктор технічних наук, професор



Андрій МУСІЄНКО

«27» серпня 2026 року

