

ВІДГУК

рецензента 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 критичного призначення будуються як розподілені кластери, працездатність яких безпосередньо залежить від цілісності мережевої взаємодії між вузлами. Втрата зв'язності призводить до стану розщеплення кластера, за якого система розпадається на кілька незалежно функціонуючих частин, що спричиняє розходження станів і порушення цілісності даних. У фінансовій сфері подібний стан здатен спричинити розбіжність балансів, у телекомунікаціях — втрату маршрутної інформації, а в державних реєстрах руйнує довіру до самого джерела істини.

Особливої гостроти ця проблематика набуває для об'єктів критичної інфраструктури України, які функціонують в умовах систематичних перебоїв зі зв'язком та електропостачанням, а також цілеспрямованих кібератак на телекомунікаційні й енергетичні вузли. У цьому контексті здатність завчасно оцінювати структурну крихкість кластера перестає бути суто академічною задачею й набуває значення елемента забезпечення стійкості критичної інфраструктури.

Наявні механізми протидії реагують на вже наявне розщеплення, а не передбачають наближення до нього. Тому наукове завдання, вирішене у роботі Сінька Д.П., що полягало у розробленні комп'ютерної моделі проактивного прогнозування відстані до розщеплення кластера, є актуальним і має практичну цінність.

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

Робота виконувалась відповідно до плану наукових досліджень Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України:

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

2. 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 pp.)

Оцінка обґрунтованості наукових результатів дисертації, їх новизни

У дисертаційній роботі повністю виконано поставлене наукове завдання: розроблено та обґрунтовано комп'ютерну модель прогнозування відстані до розщеплення кластера на основі регресійних моделей машинного навчання та графово-інваріантного представлення стану кластера. Запропонована модель дає змогу завчасно виявляти вразливі конфігурації та переходити від реактивного до превентивного управління кластерними системами.

В рамках дисертаційної роботи одержані такі наукові результати:

1. Вперше формалізовано відстань до стану розщеплення кластера як неперервну міру структурної крихкості та поставлено задачу її регресійного прогнозування, що, на відміну від бінарної класифікації факту розщеплення, зберігає інформацію про ступінь близькості до критичного стану.

2. Вперше запропоновано подвійно інваріантне представлення стану кластера, інваріантне водночас до перестановки вузлів і до перейменування типів.

3. Вперше розроблено комп'ютерну модель прогнозування відстані до розщеплення, яка поєднує точний алгоритм обчислення цільової величини, конвеєр формування інваріантних ознак, регресійну модель CatBoost та механізми калібрування.

4. Удосконалено спосіб представлення стану кластера завдяки поєднанню канонізованого матричного опису з графово-інваріантними структурними ознаками.

5. Дістали подальшого розвитку методи калібрування та оцінювання невизначеності прогнозу, зокрема ізотонічне калібрування та конформні інтервали, застосовані до задачі регресії структурної крихкості.

Обґрунтованість отриманих наукових положень дисертації підтверджується використанням загальноприйнятих методів досліджень, підтвердженням коректності розробленого алгоритму повним звіренням з еталонним перебором, за якого розбіжностей не виявлено, а також результатами експериментальних досліджень на незалежній контрольній вибірці.

Практичне значення результатів роботи

1. Розроблений програмний комплекс прогнозує відстань до розщеплення та надає неперервну оцінку ризику з інтервалом надійності у зручній для користувача формі.

2. Комплекс може бути інтегрований у системи моніторингу розподілених інформаційних систем для завчасного виявлення вразливих конфігурацій та переходу від реактивного до превентивного управління кластерними системами.

3. Запропонований підхід до інваріантного представлення структури може бути адаптований для інших задач, у яких оцінюється структурна стійкість мережевих об'єктів.

Теоретичне значення роботи

Теоретичне значення дисертаційної роботи полягає у розвитку наукових основ забезпечення стійкості розподілених інформаційних систем через формалізацію поняття структурної крихкості кластера. Автором уперше введено відстань до стану розщеплення як неперервну міру, що дорівнює мінімальній кількості відмов ефективних зв'язків, після настання яких кластер розпадається на щонайменше два повнотипні функціональні острови. На відміну від наявних підходів, які трактують розщеплення як дискретну подію, що або відбулася, або ні, така постановка зберігає інформацію про ступінь близькості системи до критичного стану й дозволяє впорядковувати конфігурації за рівнем ризику. Тим самим стан розщеплення розглядається не як аварійна подія, а як процес, розвиток якого може бути виявлений та кількісно оцінений заздалегідь, що є методологічно важливим переходом від реактивної до превентивної парадигми забезпечення стійкості.

Окремий теоретичний внесок становить обґрунтоване в роботі подвійно інваріантне представлення стану кластера, що є інваріантним водночас до перестановки вузлів і до перейменування типів сервісів. У дисертації показано, що обидві ці симетрії є несуттєвими для самої задачі, оскільки і стан кластера, і відстань до розщеплення визначаються виключно розбиттям вузлів на типи, а не конкретними позначеннями чи порядком запису елементів системи. Побудова представлення, яке відповідає справжній симетрії задачі, усуває штучну надлишковість опису та створює коректне підґрунтя для застосування методів машинного навчання до аналізу топологічних структур. Отримані результати закладають основу для подальших досліджень у сфері кількісного оцінювання вразливості мережевих структур критичного призначення.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності

Структура дисертації побудована відповідно до мети та задач дослідження. Дисертація Сінька Д.П. викладена на 204 сторінках та складається зі вступу, 4-х розділів, висновків, списку використаних джерел та додатка. Обсяг основного тексту дисертації складає 158 сторінок друкованого тексту. Робота ілюстрована 19 таблицями. Список використаних джерел містить 119 найменувань.

У вступі обґрунтовано актуальність теми дисертації, розкрито сутність і стан науково-технічної проблеми, що досліджується, відображено зв'язок роботи з науковими програмами та планами НДР Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, сформульовано мету і завдання досліджень, наукову новизну і практичне значення отриманих результатів, зазначено особистий внесок здобувача, наведено апробації роботи та публікації її результатів.

У першому розділі проведено системний аналіз сучасних архітектур інформаційних систем, розглянуто природу стану розщеплення кластера та наявні механізми протидії йому. Обґрунтовано, що проактивне попередження переходу кластера до цього стану залишається відкритою науково-інженерною задачею.

Другий розділ закладає теоретичний фундамент роботи: побудовано графову модель кластера, уведено поняття функціонального острова та відстані до розщеплення, обґрунтовано інваріантне представлення стану й запропоновано набір структурних ознак.

Третій розділ присвячено практичній реалізації комп'ютерної моделі у вигляді модульного програмного комплексу, що охоплює генерацію даних, точне обчислення цільової величини, формування ознак, навчання й калібрування моделей та вебсервіс.

У четвертому розділі наведено результати експериментального дослідження, що підтверджують перевагу запропонованого представлення та обраної моделі, а також коректність процедур калібрування й оцінювання невизначеності.

Висновки по дисертаційній роботі підкреслюють новизну та практичну цінність проведених досліджень. Список використаної літератури із 119 найменувань охоплює сучасні публікації за метою дисертаційних досліджень. В роботі відсутні порушення академічної доброчесності. В цілому структура, обсяг та оформлення дисертації відповідають чинним вимогам, які ставляться до дисертацій на здобуття ступеня доктора філософії з галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Відповідність дисертаційної роботи спеціальності

Дисертація за змістом та усіма матеріалами, зокрема науковими публікаціями та виступами на науково-технічних конференціях, повністю відповідає спеціальності 122 «Комп'ютерні науки».

Повнота публікацій матеріалів досліджень

За темою роботи опубліковано 12 наукових праць, зокрема статті у фахових виданнях України категорій «А» та «Б», статті у міжнародних виданнях, що індексуються в наукометричних базах даних Web of Science Core Collection та/або Scopus, а також тези доповідей на науково-технічних конференціях. Для кожної праці, опублікованої у співавторстві, здобувачем зазначено особистий внесок, який охоплює постановку задачі, розроблення математичної моделі кластера, огляд методів машинного навчання та практичну реалізацію системи виявлення розщеплення кластера.

Таким чином, можна зробити висновок про те, що у науково-технічних виданнях є достатньо повна інформація про результати досліджень, викладені в дисертації здобувача.

Зауваження щодо результатів, змісту та оформлення дисертації:

1. У підрозділі 2.4.5 та таблиці 2.5 обґрунтовано відмову від застосування нейромережових моделей, зокрема через потребу у значних обсягах навчальних даних. Водночас навчальна вибірка в роботі формується синтетично й може бути згенерована у практично довільному обсязі, що дещо послаблює цей аргумент. Демонстрація хоча б базової нейромережової моделі як додаткової точки порівняння зробила б висновок про перевагу ансамблів дерев рішень більш переконливим, а гнучкість таких моделей у налаштуванні могла б дати додаткові висновки щодо межі досяжної якості на запропонованому представленні.

2. У дисертаційній роботі зазначається (с. 27), що практичне значення роботи полягає в можливості розробленого програмного комплексу прогнозувати відстань до проблеми розщеплення кластера та надавати неперервну оцінку ризику з інтервалом надійності в зручній для користувача формі (вебінтерфейс). Доцільно було б в основному тексті детальніше описати функціональні можливості розробленого користувацького вебінтерфейсу.

3. Результати експериментального дослідження запропонованого автором підходу наведені у розділі 4 дисертації. Разом з тим, оформлення даного дослідження з дотриманням вимог теорії планування та оформлення експерименту дало б змогу встановити зв'язок між метою експерименту, його завданням та отриманими результатами.

4. Таблиця 1.1, що містить порівняння сучасних методів запобігання розщепленню, не містить посилань на джерела безпосередньо в рядках, хоча відповідні посилання наведено в тексті поруч. Додавання посилань підвищило б довідкову цінність цієї таблиці.

5. Дисертаційне дослідження присвячене графовим структурам і топології кластерів, однак на всю роботу наведено лише один рисунок (рис. 2.1). Робота сприймалася б значно краще, якби опис архітектури системи чи конвєсра ознак було проілюстровано додатковими схемами чи діаграмами.

Зазначені зауваження стосуються оформлення та подання матеріалу, не піддають сумніву вагомість отриманих результатів та не впливають на загальну позитивну оцінку дисертаційної роботи Сінька Д.П.

Висновок про дисертаційну роботу

Детальний аналіз матеріалів дисертаційної роботи та опублікованих наукових праць дає змогу стверджувати, що дисертаційна робота Сінька Дмитра Павловича «Комп'ютерна модель прогнозування стану розщеплення кластеру розподіленої інформаційної системи з мікросервісною архітектурою» містить усі ознаки завершеної наукової роботи. В ній отримано нові науково обґрунтовані теоретичні результати, що дають змогу вирішити актуальне наукове завдання проактивного прогнозування переходу розподіленої інформаційної системи мікросервісної архітектури в стан розщеплення кластера.

Враховуючи актуальність теми дисертаційної роботи, її теоретичну та практичну цінність, отримані наукові результати та достатню повноту висвітлення основних положень в опублікованих працях, вважаю, що представлена дисертаційна 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».

Рецензент:

заступник директора з науково-технічної роботи
Інституту проблем моделювання в енергетиці
ім. Г.Є. Пухова НАН України
д. т. н., ст. досл.



Сергій ГОНЧАР