

ВІДГУК

офіційного опонента на дисертаційну роботу

Мітікова Миколи Юрійовича

**«Ідентифікація аномалій в роботі програмного забезпечення на основі
аналізу знімків пам'яті»,**

подану на здобуття ступеня доктора філософії
за спеціальністю 113 «Прикладна математика»

Дисертаційну роботу Мітікова Миколи Юрійовича присвячено формалізації й автоматизації аналізу надмірного використання оперативної пам'яті у програмних застосунках керованого середовища виконання.

1. Актуальність обраної теми дисертації, її зв'язок з науковими програмами

Збільшення обсягів даних і кількості об'єктів у корпоративних та розподілених програмних системах підвищує вимоги до раціонального використання оперативної пам'яті. Невдалий вибір типів даних, колекцій і способів подання інформації створює ресурсні витрати навіть за відсутності класичного витоку пам'яті. Наслідками можуть бути зростання частоти збирання сміття, збільшення затримок обробки запитів і зниження стійкості системи за навантаження.

Знімок пам'яті фіксує фактичний стан програмного процесу, але його інтерпретація часто залежить від ручного аналізу. Розроблення формальної моделі, критеріїв, алгоритмічних процедур і продукційних правил для виявлення структурної надмірності є актуальною задачею. Тематика роботи відповідає напряму прикладної математики, пов'язаному з моделюванням та алгоритмічним аналізом складних інформаційних систем.

Дисертаційні дослідження виконано відповідно до планів наукової роботи кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара в межах тем «Детерміновані та стохастичні алгоритми комп'ютерного моделювання об'єктів та процесів різної природи» (номер державної реєстрації 0122U001467) і «Розробка високоефективних

комп'ютерних алгоритмів для аналізу та ідентифікації параметрів математичних моделей» (номер державної реєстрації 0125U002277).

2. Особистий внесок здобувача в отримання наукових результатів

Особистий внесок здобувача полягає у формулюванні задачі дослідження, побудові формальної моделі знімка пам'яті, дослідженні типових ситуацій надмірного використання та створенні на основі такого аналізу продукційної моделі представлення знань. Автором розроблено критерії для оцінювання ефективності запропонованих змін, які застосовано під час експериментальної перевірки на промислових знімках пам'яті. Положення, що виносяться на захист, одержані автором особисто; внесок у спільних публікаціях чітко визначено в тексті дисертації.

3. Ступінь обґрунтованості наукових положень, висновків та рекомендацій

Наукові положення й висновки базуються на методах математичного моделювання, теорії множин, дискретної математики, структурного та алгоритмічного аналізу даних, продукційного подання знань і експериментальної перевірки. У дисертації від визначення структури знімка пам'яті здійснено перехід до критеріїв надлишковості, обчислення кількісних характеристик, формування рекомендацій та їхньої валідації.

Обґрунтованість результатів підтверджується результатами аналізу промислових знімків, оцінюванням дублювання, структури хеш-колекцій, типів даних, вирівнювання та масивів байтів. Для сформульованих рекомендацій виконано порівняння прогнозованого й фактичного результатів після реалізації змін. Така перевірка є важливою, оскільки сама по собі економія пам'яті не гарантує допустимості зміни структури даних для конкретного застосунку.

4. Ступінь новизни результатів, їх теоретичне та практичне значення

У дисертаційній роботі вперше запропоновано формальну модель знімка пам'яті програмного процесу як скінченної типізованої структури, що

містить об'єкти, типи, адреси, розміри, значення полів і зв'язки між об'єктами. Вперше сформульовано задачу ідентифікації надмірного використання пам'яті через виявлення класів надлишкового подання інформації з урахуванням еквівалентності за інформаційним вмістом і коефіцієнта надмірності.

Дістали подальшого розвитку:

- алгоритмічний підхід до пошуку дублікатів незмінюваних об'єктів шляхом попереднього їхнього групування та перевірки еквівалентності у сформованих підмножинах;
- метод кількісного оцінювання можливого зменшення пам'яті в разі дублювання даних, надлишкового діапазону типів, вирівнювання, неефективного використання хеш-колекцій і стискання масивів байтів;
- продукційна модель представлення знань, яку побудовано за результатами аналізу типових сценаріїв надмірності;
- система метрик валідації рекомендацій, за допомогою якої здійснюється порівняння прогнозованих та фактичних змін використання пам'яті після застосування рекомендацій.

Теоретичне значення роботи полягає у формалізації задачі виявлення структурної надмірності використання пам'яті. Практичне значення результатів полягає у можливості автоматизованого аналізу знімків пам'яті, формування рекомендацій щодо подання даних і прогнозування результату від реалізації рекомендацій. Результати впроваджено у промислові системи, що підтверджується актом у додатку до дисертації.

5. Висновок про повноту опублікування основних положень дисертації

Основні результати відображено у 18 наукових працях: 10 публікацій у фахових виданнях категорії Б; 1 тези доповіді на міжнародній конференції, матеріали якої індексуються базою Scopus; 7 тез доповідей на міжнародних наукових конференціях.

Публікації висвітлюють формулювання задачі ідентифікації, розроблення моделей, методів та алгоритмів виявлення сценаріїв надмірного використання пам'яті, аналіз структур даних, експериментальну перевірку підходу до розв'язання задачі дослідження. Результати дослідження доповідалися на міжнародних наукових конференціях.

6. Аналіз основного змісту роботи

Дисертація складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел зі 104 найменувань та додатків. Загальний обсяг роботи становить 169 сторінок, з них 147 сторінок основного тексту; робота містить 23 рисунки та 7 таблиць.

У вступі обґрунтовано актуальність теми, сформульовано мету і завдання, визначено об'єкт, предмет та методи дослідження, наведено наукову новизну і практичне значення результатів.

У *першому розділі* наведено аналітичний огляд підходів до виявлення проблем продуктивності та аналізу знімків пам'яті. Розмежовано витoki пам'яті та стабільне надмірне використання ресурсу.

У *другому розділі* сформовано математичну основу дослідження: побудовано формальну модель знімка пам'яті, введено критерії та класи надмірного використання, подано процедури аналізу дублювання, хеш-колекцій, діапазонів типів, вирівнювання й стискання масивів байтів.

У *третьому розділі* наведено результати експериментальної перевірки. Виконано кількісний аналіз промислових знімків, досліджено прискорений пошук дублікатів та перевірено вплив запропонованих змін на використання пам'яті.

У *четвертому розділі* подано продукційну експертну систему. Визначено базу фактів, базу правил, механізм логічного виведення та структуру результатів застосування правил.

Роботу завершено висновками.

7. Оцінка структури дисертації, мови та стилю викладення

Робота складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел і додатків. Матеріал викладено послідовно з використанням сучасної наукової термінології, а структура дисертації відповідає поставленим меті та завданням. Це відповідає вимогам до наукових праць на здобуття ступеня доктора філософії, а зміст роботи висвітлює основні результати наукових досліджень.

8. Зауваження щодо змісту дисертації

1. Формальна модель $M(u) = (B, O, T, TH)$, наведена у (2.10), орієнтована передусім на об'єкти керованої купи, їх типи та потоки. Водночас пікове використання пам'яті процесом може включати некеровані виділення, стеки потоків, ЛІТ-код, завантажені модулі, службові структури збирача сміття, зарезервовані, але не зафіксовані області. Тому твердження про зменшення пам'яті процесу потребують чіткого відмежування від оцінок керованої купи або розширення моделі додатковими компонентами резидентної пам'яті.
2. У підрозділі 2.5.1 послідовність знімків використано для виявлення кандидатів на витік пам'яті, тоді як інші класи надмірності переважно оцінюються за одним станом процесу. Потребує додаткового обґрунтування, за яких умов дублювання, надлишкова ємність колекцій або надлишкове подання полів, виявлені в одному знімку, вважаються стійкими, а не наслідком тимчасового профілю навантаження. Серійний аналіз доцільно поширити на валідацію продукційних правил для структурної надмірності з урахуванням порівнюваності режимів роботи системи.
3. У розділі 3 ефект рекомендацій оцінюється переважно за зменшенням зайнятого обсягу пам'яті. Водночас компактніше подання даних може змінювати частоту й тривалість збирання сміття, локальність доступу, ефективність кешування, пропускну здатність і час виконання. У роботі слід було чіткіше визначити, чи є часова ефективність обмеженням

задачі оптимізації, і навести процедуру відхилення рекомендації, якщо економія пам'яті супроводжується неприйнятним погіршенням затримки або процесорних витрат.

4. У підрозділах 3.4–3.5 та правилах 4.3.2–4.3.5 розмірні оцінки для полів, вирівнювання і *Dictionary* $\langle K, V \rangle$ отримано для конкретного 64-бітного середовища CLR. Внутрішнє компонування об'єктів, розмір заголовка, крок вирівнювання, послідовність зміни місткості колекцій і службові структури залежать від версії середовища виконання та архітектури. Доцільно було б явно параметризувати відповідні моделі цими характеристиками й навести межі похибки перенесення оцінок на інші конфігурації.

Наведені зауваження не зменшують наукової та практичної цінності дисертаційної роботи і не можуть істотно вплинути на загальну позитивну оцінку дисертації. Але їх доцільно врахувати під час подальшого розвитку тематики дослідження.

9. Відсутність порушень академічної доброчесності

Використані у роботі літературні джерела супроводжуються бібліографічними посиланнями, особистий внесок здобувача у спільні публікації визначено в дисертації. За результатами ознайомлення з текстом дисертації підстави щодо порушення принципів академічної доброчесності відсутні.

10. Відповідність дисертації спеціальності, за якою вона подається до захисту

Зміст дисертації відповідає спеціальності 113 «Прикладна математика». Робота містить формальну математичну модель об'єкта дослідження, постановку задачі ідентифікації, дослідження та аналіз типових ситуацій надмірного використання пам'яті та побудовану на результатах аналізу продукційні модель. Розроблено критерії визначення надмірності та алгоритмічні процедури, систему кількісних оцінок; виконано експериментальну перевірку запропонованого підходу.

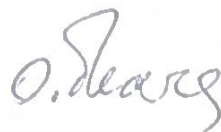
11. Загальні висновки

Дисертаційна робота Мітікова Миколи Юрійовича є завершеним самостійним науковим дослідженням, у якому розв'язано актуальну задачу ідентифікації надмірного використання оперативної пам'яті на основі аналізу знімків пам'яті.

За актуальністю, науковою новизною, обґрунтованістю та достовірністю отриманих результатів, їх теоретичним і практичним значенням, повнотою опублікування та рівнем виконання дисертаційна робота «Ідентифікація аномалій в роботі програмного забезпечення на основі аналізу знімків пам'яті» відповідає вимогам Порядку присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України № 44 від 12 січня 2022 року зі змінами, а її автор Мітіков Микола Юрійович заслуговує на присудження ступеня доктора філософії за спеціальністю 113 «Прикладна математика».

Офіційний опонент

доктор технічних наук, професор,
професор кафедри програмних систем
і технологій Київського національного
університету імені Тараса Шевченка



Ольга ТКАЧЕНКО



В І Д Ч У Ю

ПРОРЕКТОР З НАУКОВОЇ РОБОТИ

ЖИЛІНСЬКА