

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 розробленню формальної моделі знімка пам'яті програмного процесу, критеріїв та алгоритмічних процедур ідентифікації надмірного використання оперативної пам'яті, а також продукційної моделі формування діагностичних рекомендацій.

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

У керованих середовищах виконання збільшення обсягу зайнятої пам'яті виникає не лише внаслідок витоків. Воно може бути зумовлене дублюванням незмінюваних об'єктів, нераціональним використанням колекцій, надлишковим діапазоном типів даних, вирівнюванням об'єктів або способом подання масивів байтів. Такі причини не завжди встановлюються за часовими показниками споживання ресурсу й потребують аналізу структури знімка пам'яті.

Розроблення формальних засобів подання знімка пам'яті, виявлення класів надлишкового подання інформації та кількісного оцінювання очікуваного ефекту оптимізації є актуальною науковою задачею. Тема роботи має теоретичне значення для аналізу станів програмних систем і практичне значення для високонавантажених застосунків.

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 ідентифікації надмірного використання пам'яті, побудові формальної моделі знімка пам'яті, розробленні критеріїв і алгоритмічних процедур виявлення окремих класів надлишковості, а також у створенні продукційної моделі та її програмної реалізації для автоматизованого аналізу знімків пам'яті. Результати, винесені на захист, одержані автором особисто; внесок у спільні публікації визначено в дисертації.

3. Ступінь обґрунтованості наукових положень, висновків та рекомендацій

Наукові положення, висновки та рекомендації ґрунтуються на методах математичного моделювання, теорії множин, дискретної математики, структурного й алгоритмічного аналізу даних, а також на продукційному поданні знань. У роботі послідовно пов'язано формальну структуру знімка, критерії виявлення надлишковості, алгоритмічні процедури та правила формування рекомендацій.

Достовірність результатів забезпечується аналізом промислових знімків пам'яті, обчислювальними експериментами та зіставленням прогнозованого зменшення обсягу пам'яті з фактичним результатом після внесення допустимих змін до програмного коду. Висновок продукційної системи не ототожнюється з автоматичною зміною програми: рекомендація потребує перевірки її семантичної допустимості.

4. Ступінь новизни результатів, їх теоретичне та практичне значення

У дисертації вперше запропоновано формальну модель знімка пам'яті як скінченної типізованої структури об'єктів, типів, адрес, розмірів, значень полів і зв'язків. Уперше сформульовано задачу ідентифікації надмірного використання пам'яті як виявлення класів надлишкового подання інформації

з використанням відношення еквівалентності за інформаційним вмістом і коефіцієнта надмірності.

Дістали подальшого розвитку:

- алгоритмічний підхід до пошуку дублікатів незмінюваних об'єктів на основі попереднього групування та перевірки еквівалентності в межах груп;
- метод кількісного оцінювання можливого зменшення обсягу пам'яті для дублювання даних, вибору типів полів, використання хеш-колекцій, вирівнювання та стискання масивів байтів;
- продукційна модель, що пов'язує факти про структуру знімка, предикати правил, сценарії надмірності, рекомендації та кількісні оцінки;
- система метрик валідації рекомендацій, за допомогою якої здійснюється порівняння прогнозованого і фактичного зменшення обсягу використання пам'яті.

Теоретичне значення одержаних результатів полягає у формалізації задачі аналізу структурної надмірності у пам'яті програмного процесу. Практичне значення визначається можливістю автоматизованого аналізу знімків, формування рекомендацій щодо раціонального використання пам'яті і оцінювання очікуваного результату їх реалізації. Результати апробовано на промислових програмних системах; факт упровадження підтверджено актом, наведеним у додатку до дисертації.

5. Висновок про повноту опублікування основних положень дисертації

Основні результати дисертації опубліковано у 18 наукових працях, серед яких є публікація у виданні, що індексується наукометричною базою Scopus, статті у наукових фахових виданнях України та матеріали міжнародних конференцій. Публікації відображають постановку задачі ідентифікації, побудову моделі знімка пам'яті та продукційної моделі, алгоритмічні процедури, експериментальну перевірку одержаних результатів. Результати дослідження доповідалися на наукових заходах.

6. Аналіз основного змісту роботи

Дисертація складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел зі 104 найменувань та додатків. Загальний обсяг роботи становить 170 сторінок, з них 147 сторінок основного тексту; робота містить 23 рисунки та 7 таблиць.

У вступі обґрунтовано актуальність теми, сформульовано мету дослідження, завдання, об'єкт, предмет і методи дослідження, наведено положення наукової новизни та практичне значення результатів.

У першому розділі проаналізовано підходи до діагностики проблем продуктивності й аналізу знімків пам'яті. Розмежовано витoki пам'яті та стабільне надмірне використання ресурсу, обґрунтовано необхідність формального аналізу об'єктної структури.

У другому розділі побудовано математичну модель знімка пам'яті, сформульовано задачу ідентифікації та введено критерії й кількісні характеристики для оцінювання. Подано процедури аналізу дублювання незмінюваних об'єктів, хеш-колекцій, діапазонів типів, вирівнювання та стискання масивів байтів.

У третьому розділі наведено результати аналізу промислових знімків. Досліджено ефективність прискореного пошуку дублікатів і зіставлене прогнозовані та фактичні обсяги використання пам'яті після реалізації рекомендацій.

У четвертому розділі подано продукційну модель експертної системи з механізмом прямого логічного виведення. На основі досліджених у другому та третьому розділах типових ситуацій надмірного використання пам'яті сформульовано базу фактів та правил, на основі яких визначається сценарій надмірності, формулюється рекомендація щодо її усунення та надається кількісна оцінка.

Кожний розділ дисертації та робота в цілому завершуються висновками, які впливають зі змісту розділів, є логічними та відображають основні результати.

7. Оцінка структури дисертації, мови та стилю викладення

Дисертація складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел і додатків. Викладення основних положень є послідовним; терміни, позначення та результати загалом відповідають предмету дослідження.

8. Зауваження щодо змісту дисертації

1. У підрозділі 3.7 (с. 113-115) рекомендація класифікується як істинно позитивна або хибно позитивна за результатами її підтвердження розробником, зміною програмного коду та повторним знімком пам'яті. Водночас не визначено процедуру побудови повної еталонної множини сценаріїв, без якої неможливо встановити кількість хибно негативних результатів і, відповідно, відтворювано обчислити заявлений показник повноти. Доцільно було б формально задати спосіб незалежного формування еталонної множини, кількість перевірених випадків та інтервальні оцінки точності, повноти і похибки прогнозування.
2. У розділі 4 визначено продукційні правила та структуру результату логічного виведення, проте не формалізовано порядок дій у разі одночасного спрацьовування кількох правил для тих самих або пов'язаних множин об'єктів. За таких умов оцінки можливої економії можуть бути неадитивними, а запропоновані зміни – несумісними. Слід було б визначити відношення конфлікту між рекомендаціями та функцію пріоритету, яка враховує прогнозований ефект, достовірність оцінки, трудомісткість реалізації й ризик порушення функціональної поведінки програми.
3. Апробацію продукційного механізму виконано на знімках обсягом від 20 до 180 ГБ, але для повного циклу формування первинних, структурних та агрегованих фактів не наведено верхніх оцінок допоміжної пам'яті. Це ускладнює встановлення граничного розміру

знімка, який може бути опрацьовано за заданого обсягу ресурсів аналітичної системи. Доцільно було б подати просторову складність основних процедур залежно від кількості об'єктів, типів і сформованих фактів, а також розглянути потокове або зовнішнє формування частини індексів.

4. У співвідношенні (2.1) знімок пам'яті названо скінченною послідовністю байтів або частковим відображенням адресного простору, але записано як множину $B_u = \{b_i\}$. Множинний запис не зберігає порядок байтів і не розрізняє однакові байтові значення, розташовані за різними адресами. Унаслідок цього потребує уточнення також зміст підмножини $B_u = \{b_i\}$ у (2.5). Для строгості формальної моделі доцільно задати байтовий стан як часткове відображення адрес у байтові значення, а представлення об'єкта – як його впорядковане обмеження на відповідний діапазон адрес.

Наведені зауваження конкретизують межі застосовності отриманих результатів і напрями подальшого розвитку, але не зменшують загальної позитивної оцінки дисертаційної роботи.

9. Відсутність порушень академічної доброчесності

В дисертаційній роботі всі результати інших авторів супроводжуються коректними посиланнями на використані джерела, особистий внесок здобувача у спільні праці чітко визначено в дисертації. За результатами ознайомлення з роботою підстав вважати, що в ній порушено принципи академічної доброчесності, не встановлено.

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

Дисертація відповідає спеціальності 113 «Прикладна математика», оскільки ґрунтується на формальній постановці задачі ідентифікації, розробленої математичної моделі для прикладного об'єкта дослідження,

побудові продукційної моделі представлення знань про досліджені типові сценарії надмірного використання пам'яті, містить розроблені алгоритмічні процедури ідентифікації типових сценаріїв та результати їх експериментальної перевірки із застосуванням кількісних оцінок.

11. Загальні висновки

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

За актуальністю, науковою новизною, обґрунтованістю та достовірністю 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 програмного
забезпечення комп'ютерних систем
Національного технічного університету
«Дніпровська політехніка»

Борис МОРОЗ

Підпис
засвідчую:
вчений секретар
Вченої ради

