

РЕЦЕНЗІЯ

на дисертаційну роботу **Мітікова Миколи Юрійовича**
на тему: **«Ідентифікація аномалій в роботі програмного забезпечення на
основі аналізу знімків пам'яті»,**
представлену на здобуття ступеня доктора філософії
за спеціальністю 113 «Прикладна математика»

Актуальність теми дисертації.

Предметом дисертаційного дослідження Мітікова М. Ю. є методи ідентифікації надмірного використання оперативної пам'яті програмного процесу за даними її знімка.

Для сучасних програмних систем характерним є формування значних за обсягом і складних за структурою станів оперативної пам'яті. Знімок такого стану містить відомості про множину об'єктів, їх типи, значення, розміри та зв'язки, однак самі ці дані ще не визначають, які витрати пам'яті є обґрунтованими, а які слід розглядати як прояв структурної надмірності.

Діагностування ускладнюється тим, що надмірне використання пам'яті не обов'язково супроводжується порушенням досяжності об'єктів або необмеженим зростанням керованої купи. Дублювання незмінних значень, невідповідність фактичного діапазону значень обраному типу, надлишкова ємність колекцій і втрати внаслідок вирівнювання можуть бути допустимими з погляду виконання програми, але невиправданими з погляду використання ресурсу. Тому задача потребує не лише збору даних, а й формалізованої інтерпретації виявлених станів.

У дисертації Мітікова М. Ю. цю проблему розглянуто як послідовність взаємопов'язаних процедур: побудову математичного подання знімка пам'яті, обчислення діагностичних характеристик, перевірку умов виявлення окремих класів надмірності та формування рекомендацій із кількісною оцінкою очікуваного результату. Така постановка переводить аналіз знімка пам'яті з рівня окремих евристичних спостережень до рівня системи підтримки прийняття рішень.

Отже, створення узгодженої сукупності моделей, алгоритмів і правил, за якими стан пам'яті відноситься до певного діагностичного класу, є актуальним завданням прикладної математики та відповідає потребі у відтворюваних методах аналізу складних інформаційних систем.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційне 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).

Формулювання наукової задачі, нове вирішення якої одержане в дисертації.

Наукова задача, розв'язана в роботі, полягає у встановленні відповідності між структурою зафіксованого стану керованої пам'яті та множиною діагностичних висновків щодо надмірного подання інформації. Вхідними даними є скінченна типізована сукупність об'єктів і зв'язків між ними, а результатом аналізу стає виявлений сценарій, множина об'єктів, що його утворюють, рекомендоване перетворення подання даних для поліпшення ситуації та оцінка можливого зменшення обсягу пам'яті за результатами втілення рекомендацій.

Для розв'язання цієї задачі автором послідовно формалізовано структуру знімка пам'яті, визначено ознаки та предикати для розпізнавання типових ситуацій, побудовано процедури групування, фільтрації та порівняння об'єктів, введено кількісні характеристики надмірності, сформовано продукційну модель, у якій обчислені факти є підставою для діагностичного висновку та надання рекомендацій.

Мета дослідження узгоджена з поставленою задачею і передбачає розроблення формальної моделі знімка пам'яті, критеріїв та алгоритмів ідентифікації сценаріїв надмірного використання ресурсу, а також продукційного механізму застосування отриманих знань. Перевірка результатів здійснюється шляхом зіставлення прогнозованих оцінок із фактичними змінами після реалізації рекомендацій.

Об'єктом дослідження є зафіксований стан оперативної пам'яті програмного процесу. Предмет дослідження становлять математичні моделі цього стану, методи обчислення діагностичних ознак, критерії віднесення до класів надмірності, алгоритми аналізу та продукційні правила формування висновків.

Наукова новизна отриманих автором результатів.

Наукова новизна дисертаційної роботи полягає у розробленні формалізованого підходу до ідентифікації та діагностування надмірного використання пам'яті, зокрема:

- знімок пам'яті подано як скінченну типізовану структуру, що поєднує об'єкти, їх адреси, типи, розміри, значення полів і зв'язки; це забезпечило єдину формальну основу для подальших діагностичних перетворень;

- класи надмірного використання пам'яті визначено через перевірювані ознаки та кількісні критерії, завдяки чому діагностичний висновок пов'язується з властивостями конкретної підмножини об'єктів, а не лише із загальним обсягом керованої купи;

- розвинено алгоритмічні процедури виявлення еквівалентних незмінюваних об'єктів, аналізу фактичного використання колекцій, діапазонів значень, вирівнювання та стискання масивів байтів із визначенням потенційного ресурсного ефекту;

- побудовано продукційну модель представлення знань, у якій умови правил задаються на основі формалізованих характеристик знімка, а результат

спрацювання містить діагностичне повідомлення, рекомендацію та кількісну оцінку;

– запропоновано оцінювати результативність діагностики не лише за фактом виявлення сценарію, а й за точністю, повнотою та похибкою прогнозу зміни використання пам'яті.

Практичне значення одержаних результатів.

Практична цінність результатів визначається можливістю використати розроблену систему правил як засіб підтримки прийняття рішень під час аналізу знімків пам'яті. Для кожного виявленого сценарію користувач одержує не тільки перелік проблемних ситуацій, а й пояснення причини їхнього виявлення, спосіб усунення надмірності та оцінку очікуваної економії ресурсу.

Розроблений підхід може використовуватися у промислових програмних системах, що підтверджено актом впровадження.

Обґрунтованість і достовірність наукових положень, висновків і рекомендацій, які захищаються.

Обґрунтованість положень дисертації забезпечується узгодженістю вихідного подання даних, введених критеріїв та алгоритмів, за допомогою яких ці критерії перевіряються. Умови продукційних правил спираються на характеристики, що обчислюються з формальної моделі знімка, тому між первинними даними й отриманим висновком простежується логічний зв'язок.

Достовірність кількісних результатів підтверджено обчислювальними експериментами на промислових знімках пам'яті та порівнянням прогнозованого зменшення обсягу пам'яті з фактичним результатом після внесення відповідних змін. Основні положення роботи оприлюднено у 18 наукових працях, що відображають етапи формування моделей, алгоритмів і засобів їх перевірки.

Особистий внесок здобувача.

Наведені в дисертації постановки, моделі, алгоритмічні процедури, результати експериментів і висновки отримані здобувачем особисто. Для праць у співавторстві внесок Мітікова М. Ю. конкретизовано в тексті дисертації.

Автором виконано формалізацію об'єктної структури пам'яті, сформовано діагностичні ознаки та правила їх інтерпретації, розроблено процедури оцінювання надмірності, створено експериментальну реалізацію і проведено перевірку отриманих рекомендацій.

Структура та обсяг дисертації.

Дисертація містить вступ, чотири розділи, загальні висновки, список із 104 використаних джерел і додатки. Із 170 сторінок роботи 147 сторінок становить основний текст; матеріал ілюстровано 23 рисунками та 7 таблицями.

Перший розділ визначає місце поставленої задачі серед засобів діагностики продуктивності та методів дослідження пам'яті. Принциповим для подальшого викладу є розмежування витоку пам'яті й стійкого надмірного подання інформації.

У другому розділі побудовано математичну модель подання знімка пам'яті та сформовано апарат діагностичних ознак. Розглянуто окремі класи надмірності й алгоритми, за якими для них визначається множина об'єктів-кандидатів та оцінюється можливе зменшення пам'яті.

Третій розділ присвячено обчислювальній перевірці запропонованих процедур. Подані результати характеризують як виявлений обсяг надмірності, так і витрати на виконання окремих алгоритмів та відповідність прогнозованого ефекту фактичному.

У четвертому розділі результати попередніх етапів організовано у вигляді продукційної системи: визначено базу фактів, множину правил, порядок формування діагностичних висновків і структуру рекомендацій. В такий спосіб здійснено перехід від формальної моделі до її застосування в діагностичній процедурі.

Висновки відповідають завданням дослідження та відображають одержані теоретичні й експериментальні результати.

Позитивно оцінюючи дисертаційну роботу в цілому, вважаю за необхідне висловити такі зауваження.

1. У співвідношеннях (2.7) та (2.8) поряд використовуються еквівалентність за байтовим поданням, належність об'єктів до одного типу та еквівалентність за інформаційним значенням. Ці відношення утворюють різні розбиття множини об'єктів. Для однозначності математичної моделі їх доцільно було б позначити окремо та явно визначити відповідні класи еквівалентності.
2. Повноту бази знань оцінено переважно через результати застосування окремих правил до наявних знімків пам'яті. Разом із тим у роботі не сформульовано критерію покриття множини діагностичних ситуацій. Доцільно подати матрицю «клас сценарію -> ознаки -> правило -> рекомендація» і визначити ситуації, які не охоплюються базою знань.
3. Умови правил експертної системи мають детермінований характер, хоча частина діагностичних ознак може змінюватися залежно від стану програмного процесу. Перспективним є доповнення моделі мірою впевненості у висновку або градаціями критичності. Це дозволило б відокремити встановлений сценарій від ситуації, що потребує перевірки.
4. На початку розділу 3 зазначено, що проаналізовано понад 80 знімків пам'яті, тоді як у висновках до цього розділу йдеться про вибірку з понад 60 знімків. Потребує уточнення склад кожної вибірки, кількість досліджених програмних систем, число знімків однієї системи та режими навантаження. Без такого опису складно оцінити незалежність спостережень і межі перенесення експериментальних висновків.

Надані зауваження не змінюють загальної позитивної оцінки дисертаційної роботи та можуть бути використані в якості рекомендацій для подальшої роботи.

Дисертація є завершеним науковим дослідженням, що має теоретичне й практичне значення. Викладення матеріалу відповідає науковому стилю, використані джерела належно позначені бібліографічними посиланнями. Порушень принципів академічної доброчесності не виявлено. За змістом моделей, алгоритмів, критеріїв і способів їх верифікації робота відповідає спеціальності 113 «Прикладна математика».

Загальний висновок.

За сукупністю отриманих результатів, рівнем їхнього обґрунтування, науковою новизною та практичною значущістю вважаю, що дисертаційна робота «Ідентифікація аномалій в роботі програмного забезпечення на основі аналізу знімків пам'яті» відповідає вимогам наказу Міністерства освіти і науки України № 40 від 12.01.2017 р. «Про затвердження Вимог до оформлення дисертації» та Порядку присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44 зі змінами, а її автор Мітіков Микола Юрійович заслуговує на присудження ступеня доктора філософії зі спеціальності 113 «Прикладна математика».

Рецензент

Доцент кафедри комп'ютерних технологій
факультету прикладної математики
та інформаційних технологій

Дніпровського національного університету
імені Олеся Гончара,

кандидат технічних наук, доцент



Костянтин ЗОЛОТЬКО

Підпис кандидата технічних наук, доцента Золотька К. Є. засвідчую

Проректор з наукової роботи
Дніпровського національного університету
імені Олеся Гончара



Олег МАРЕНКОВ