РЕЦЕНЗІЯ

доктора технічних наук, професора Кучука Георгія Анатолійовича на дисертаційну роботу **Челака Віктора Володимировича "Методи та засоби захисту інформації в комп'ютерних системах та мережах"**, подану на здобуття наукового ступеня доктора філософії з галузі знань 12 – інформаційні технології за спеціальністю 123 – комп'ютерна інженерія

Ступінь актуальності теми дисертаційної роботи. У сучасному світі з кожним днем зростає кількість та цінність інформації, з'являються нові компанії, підприємства, пристрої та відповідно до цього зростання з'являються й нові загрози, мета яких отримати конфіденційну інформацію кримінальним шляхом. Одним з перспективних напрямків розвитку захисту інформації в комп'ютерних системах та мережах є розробка та удосконалення систем виявлення вторгнень, які дозволяють вчасно ідентифікувати аномальний стан, знайти джерело або причину такого функціонування системи та видалити загрозу. Тому тема дисертаційної роботи Челака Віктора Володимировича, "Методи та засоби захисту інформації в комп'ютерних системах та мережах", яка спрямована на вирішення завдання підвищення точності та швидкості ідентифікації стану комп'ютерних систем та мереж за рахунок розробки нових методів з використанням технології машинного навчання, є актуальною з наукової та практичної точок зору та має важливу технічну значущість.

Зв'язок теми дисертаційної роботи з науковими планами, програмами, фундаментальними та прикладними дослідженнями. Тематика дисертаційної роботи відповідає пріоритетним напрямам розвитку науки і техніки в Україні з розділу «Інформаційні та комунікаційні технології». Дисертаційна робота була виконана у межах науково-дослідних робіт кафедри «Комп'ютерна інженерія та програмування» в рамках двох науково-дослідних тем: «Моделі і методи обробки та захисту інформації в комп'ютерних системах» (ДР №0122U200526, ТОВ «Передові цифрові рішення», м. Харків), в якій здобувач був керівником дослідження, «Моделі і методи обробки даних і розподілу мережних ресурсів в

комп'ютерних системах» (ДР №0122U200527, компанія «LineUp», м. Харків), в якій здобувач брав участь у якості виконавця.

Ступінь обґрунтованості та достовірності наукових положень, висновків та рекомендацій, сформульованих у дисертаційній роботі. Ґрунтовно проаналізувавши дисертаційну роботу можна відмітити, що наукові положення, висновки та рекомендації, що висвітлені в роботі, є достатніми, повними, а також належними чином повністю обґрунтованими. Для їх отримання та підтвердження автором було проведено як теоретичні, так і емпіричні, експериментальні дослідження, при цьому використовувалися вітчизняні та міжнародні вузькопрофільні та актуальні джерела. Достовірність положень і висновків зроблених автором підтверджується використанням класичних і сучасних методів досліджень, зокрема глибоким логічним аналізом літературних джерел, коректністю поставлених актуальних завдань, що потребують розв’язання та вирішення. Результати експериментальних та теоретичних досліджень доповідались та обговорювались на міжнародних науково-технічних конференціях, а також опубліковані в наукових фахових виданнях. Крім того, про достовірність отриманих результатів свідчить їх взаємоузгодженість, відповідність літературним даним і позитивні результати впровадження. У результаті проведення дисертаційного дослідження дисертанту вдалось розкрити та вирішити в повному обсязі мету та завдання, що були сформовані на початку. До кожного пункту роботи приведені логічні висновки, які дозволяють коротко та повно зрозуміти суть кожного етапу дослідження та практичну значущість отриманих результатів. Також достовірність заявлених положень обґрунтовується комплексним підходом у вивченні визначеного об’єкта.

Вищевикладене свідчить про обґрунтованість та достовірність наукових положень, висновків і рекомендацій, що викладено у дисертаційній роботі Челака Віктора Володимировича.

Наукова новизна положень, висновків та рекомендацій, сформульованих у дисертації. Наукова новизна отриманих результатів обумовлена теоретичним узагальненням і новим рішенням важливого наукового

завдання, сутність якого полягає в підвищенні точності та швидкості ідентифікації стану комп'ютерних систем та мереж за рахунок розробки нових методів з використанням технології машинного навчання. У дисертаційній роботі отримані такі основні науково обґрунтовані результати:

вперше запропоновано метод побудови дерева з багатовимірними вузлами рішень для формування деревоподібних моделей з урахуванням кореляційних зв'язків між показниками функціонування комп'ютерної системи;

вперше запропоновано метод побудови нечіткого дерева рішень, який відрізняється від відомих наявністю спеціальної автоматизованої процедури формування нечітких множин та їх функцій належності;

удосконалено метод побудови дерева рішень, за рахунок використання у якості критерію прийняття рішень мінімальної помилки класифікації, використання направленого вибору ознак та застосування алгоритму бінарного пошуку для визначення оптимального значення порогу розщеплення вузла ДР;

удосконалено ансамблевий метод класифікації на основі мета-алгоритму бустінгу за допомогою використання у якості базових моделей розроблених дерев рішень та процедури попередньої обробки даних.

Наукова та практична цінність одержаних результатів. Робота має чітку послідовність постановки задач та отриманих рішень, достатню доказову базу та аргументованість результатів. Використано сучасний математичний апарат для реалізації сформованої мети. Порівняльні оцінки запропонованих автором нових рішень щодо результатів, які отримані провідними вченими та дослідниками в галузі, достатньо аргументовані та відповідають списку приведених першоджерел. Висновки та рекомендації, які сформульовані в дисертаційній роботі, враховують сутність та актуальність наукового завдання роботи та її мету, вони є придатними для практичного використання.

Усі теоретичні розробки дисертації здобувач довів до конкретних методик та алгоритмів. Зокрема, отримані такі практичні результати:

– метод та програмне забезпечення побудови дерев з багатовимірними вузлами рішень дозволяє зменшити кількість розгалужень, підвищуючи таким чином оперативність ідентифікації стану КС до 50% та точність до 12%;

– метод та програмне забезпечення формування нечітких множин та їх функцій належності для побудови нечітких дерев рішень дозволяє підвищити точність класифікації до 30% при великої кількості даних, які знаходяться на межі розмежування класів та швидкість до 23%, порівнюючи з класичними деревами рішень;

– удосконалений метод побудови дерева рішень дозволяє зменшити час навчання дерев з одновимірними вузлами рішень до 4,5 раз;

– у досконалений ансамблевий метод класифікації на основі мета-алгоритму бустінгу дозволяє підвищити точність класифікації до 32%;

Результати дисертації впровадженні та використані в діяльності компаній SoftInWay, Inc. та ТОВ «ФТ ГРУП», а також використовуються в навчальному процесі Національного технічного університету «ХПІ».

Повнота викладення наукових і прикладних результатів дисертації в опублікованих працях. Основні ідеї здобувача та результати дослідження викладено у 9 фахових статтях, 2 закордонних статтях, 2 статтях у інших наукових українських виданнях, 1 розділі колективної монографії у співавторстві. Також здобувач активно приймав участь в багатьох міжнародних конференціях та симпозіумах, де була проведена апробація наукових результатів дисертаційного дослідження. Основні результати дисертаційної роботи у цих публікаціях відображено достатньо повно.

Оцінка змісту дисертації, її завершеності й оформлення. Побудова дисертації відповідає прийнятим для наукового дослідження нормам. Усі положення, винесені на захист, висвітлені в тексті дисертації. Зміст дисертаційної роботи відповідає її назві.

Дисертація написана грамотною науковою мовою та оформлена відповідно до існуючих нормативних документів, текст і графічний матеріал виконані акуратно з використанням комп'ютерної техніки.

Дисертація складається з анотацій, вступу, 5 розділів, висновків, списку використаних джерел і 6 додатків. Обсяг основного тексту дисертації (без анотацій, змісту, списку використаних джерел і додатків) становить 145 сторінок, що відповідає встановленим вимогам.

У **вступі** автором представлена загальна характеристика роботи, обґрунтована актуальність наукової теми, сформульовані мета і задачі дослідження, відображено наукову новизну та практичну цінність отриманих результатів і висновків, наведено дані щодо їх апробації та впровадження.

У першому розділі дисертації проведений аналіз науково-технічної проблеми захисту даних в комп'ютерних системах та мережах, на основі отриманих результатів сформульовані мета, задачі, об'єкт та предмет дослідження.

У другому розділі здобувач, використовуючи статистичні методи, виконав оцінку кореляційних зв'язків між показниками функціонування комп'ютерних систем та визначив набір атрибутів, які будуть використані в якості ознак тестової та навчальної вибірки та забезпечуватимуть високу точність ідентифікації стану комп'ютерних систем. Опираючись на теорію графів здобувачем запропоновано метод побудови дерева рішень, який використовує бінарний пошуку для знаходження порогу розщеплення вузла дерева рішень.

У третьому розділі, опираючись на методи математичної статистики та кластеризації розроблено процедуру формування нечітких множин з функціями належності, що дозволяє підвищити точність та оперативність ідентифікації стану КС за рахунок зменшення вплив експертів на результат через ініціалізацію вихідних даних.

В четвертому розділі здобувач, використовуючи ансамблеві методи машинного навчання, розробив метод ідентифікації функціонування комп'ютерних систем та мереж зі спеціальною процедурою попередньої обробки даних та використання у якості базових класифікаторів удосконалених дерев рішень з бінарним пошуком оптимального значення порогу розщеплення вузла дерева та використанням у якості критерія прийняття рішень мінімальної похибки класифікації, що надало можливість обробляти шуми в даних.

В п'ятому розділі, на основі методів оцінки якості моделей машинного навчання (матриця невідповідності та ROC-аналіз) виконано оцінку експериментальних результатів, що отримані в ході дослідження. Вибір методів досліджень дозволяє забезпечити достовірність отриманих результатів.

Висновки, сформульовані у роботі, висвітлюють результати дослідження як вирішення висунутих в дисертації завдань. Висновки відповідають вимогам, які висуваються до результатів дисертаційного дослідження на здобуття наукового ступеня доктора філософії.

Список використаних джерел широко охоплює предметне поле дослідження, певною мірою відображає опрацювання автором значної кількості джерел пов'язаних з захистом інформації, інтелектуальними методами та метриками оцінки їх ефективності.

Додатки доповнюють дисертацію прикладами функціонування розроблених методів, аналізом існуючих метрик оцінки якості та структурними схемами розроблених засобів з детальним описанням кожного блоку представлених у додатку схем.

Зауваження до дисертаційної роботи. В процесі ознайомлення з роботою позитивне враження справило докладне обґрунтовування усіх висунутих у роботі положень, використання сучасних математичних методів.

Але при цьому виникли такі зауваження та недоліки:

1. У першому розділі здобувачем проведений детальний аналіз сучасних методів виявлення вторгнень. Але постановочна частина дисертації виглядала б краще, якби більш наглядно (у вигляді діаграм та графіків) були б наведені результати аналітичного огляду основних характеристик розглянутих методів. Це підвищило б ступінь обґрунтованості зробленого автором висновку щодо необхідності розробки нових та удосконалення існуючих методів і засобів ідентифікації стану комп'ютерних систем.

2. В першому розділі було б доцільно додатково розглянути детальну класифікацію існуючих способів розповсюдження окремих загроз, саме на детектування яких налаштоване евристичне сканування.

3. У другому розділі здобувач в якості критерію розщеплення дерева рішень обирає функцію помилки. Але в роботі відсутнє детальне обґрунтування такого вибору, бажано було б провести порівняльний аналіз з іншими розповсюдженими критеріями, наприклад, індексом Джині та ентропійним підходом, також наведеними в роботі.

4. У третьому розділі здобувач навів головний недолік системи нечіткого виведення та в якості його усунення запропонував ефективну техніку використання нечітких дерев рішень. Безсумнівно, були покращені характеристики процесу ідентифікації стану комп'ютерної системи, але все ж таки потрібно було обґрунтувати такий вибір.

5. У четвертому розділі треба було б посилити аргументацію вибору кількості базових класифікаторів в ансамблі (підрозділі 4.2). Так, при аналізі залежності точності класифікації ансамблю від кількості базових класифікаторів (рис. 4.6) зроблено висновок, що оптимальною кількістю класифікаторів є п'ятдесят базових моделей. Однак при 30 класифікаторах точність класифікації знаходиться вже на достатньо високому рівні (97%), при цьому задіяна менша кількість класифікаторів, що прискорить швидкість класифікації. Отже, бажано було б знайти баланс між цими параметрами.

6. Результати наукових досліджень здобувача отримали широке впровадження у компаніях України та Сполучених Штатах Америки. Без сумніву, це відмінні результати, але на мою думку, робота ще б виграла, якщо б були більш детально розкрити результати проведеної реалізації, наприклад, наприкінці п'ятого розділу. Це ще б підсилило висновки автора щодо практичної цінності дисертаційного дослідження.

Відповідність дисертації встановленим вимогам і загальні висновки. Зазначені недоліки не впливають на загальне позитивне враження від роботи, не зменшують її якості, а також наукової та практичної цінності. Вони не є визначальними і можуть бути враховані як напрямки подальших досліджень. Під час вивчення та аналізу дисертаційної роботи **випадків порушення академічної доброчесності** виявлено не було.

Дисертаційна робота Челака Віктора Володимировича є завершеною науково-дослідною роботою, яка містить науково-обґрунтовані результати, має наукову новизну та дає перспективи подальших досліджень. Тема дослідження відповідає спеціальності 123 – «Комп'ютерна інженерія».

За змістом, актуальністю, ступенем новизни, обґрунтованістю, науковою та практичною значимістю одержаних результатів дисертаційна робота "Методи

та засоби захисту інформації в комп'ютерних системах та мережах" відповідає вимогам п.п. 6–9 "Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії", із змінами, внесеними згідно з Постановою Кабінету Міністрів України № 341 від 21.03.2022, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44, а її автор, Челак Віктор Володимирович, заслуговує на присудження наукового ступеня доктора філософії за спеціальністю 123 – комп'ютерна інженерія.

Професор кафедри комп'ютерної інженерії та програмування
Національного технічного університету
«Харківський політехнічний інститут»
доктор технічних наук, професор Георгій КУЧУК

"23" жовтня 2023 р.

Підпис проф. Георгія Кучука
ЗАСВІДЧУЮ:
ВЧЕНИЙ СЕКРЕТАР
НАЦІОНАЛЬНОГО-ТЕХНІЧНОГО УНІВЕРСИТЕТУ
"ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"
ЗАЯЦЕВ Ю. І.
"23" 10 2023 р.