Голові разової спеціалізованої вченої ради
Національного технічного університету
«Дніпровська політехніка»
д.т.н., професору, професору кафедри
програмного забезпечення комп'ютерних
систем **Борису МОРОЗУ**

РЕЦЕНЗІЯ

к.т.н., доцента, доцента кафедри безпеки інформації та телекомунікацій,
Національного технічного університету «Дніпровська політехніка»
ГЕРАСІНОЇ Олександри Володимирівни
на дисертаційну роботу **МЄШКОВА Вадима Ігоровича**
«Інформаційна технологія інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак»,
подану на здобуття наукового ступеня доктора філософії
за спеціальністю 122 «Комп'ютерні науки»,
галузі знань 12 «Інформаційні технології»

Актуальність теми дисертації. Наразі ефективність функціонування будь-якої організації, підприємства та держави залежить не тільки від надійності функціонування інформаційно-комунікаційних систем і мереж (ІКМ), а й у значній мірі від захищеності їх інформаційних ресурсів та інформації взагалі.

Сучасні системи кіберзахисту повинні забезпечувати не лише фіксацію відомих шаблонів атак, а й виявлення аномальної або нетипової активності, яка може свідчити про нові чи модифіковані загрози. У таких умовах мережевий трафік є важливим джерелом інформації про стан ІКМ, оскільки в його характеристиках відображаються як нормальні режими функціонування, так і ознаки атакувальної поведінки.

Використання методів машинного навчання для аналізу мережевого трафіку є доцільним напрямом розвитку систем виявлення атак, оскільки такі методи дають змогу працювати з багатовимірними наборами ознак, виявляти статистичні залежності між параметрами трафіку та формувати рішення щодо належності мережевої активності до нормальної або атакувальної. У дисертації обґрунтовано, що ефективність інтелектуального виявлення атак залежить не тільки від вибору моделі машинного навчання, а й від якості підготовки даних,

формування ознакового простору, урахування дисбалансу класів і коректності оцінювання результатів.

Важливо, що робота орієнтована на розв'язання двох взаємопов'язаних задач: бінарного виявлення атакувальної активності та багатокласового розпізнавання типів атак. Такий підхід відповідає практичним потребам систем кібербезпеки, для яких важливо не лише встановити факт наявності загрози, а й отримати деталізовану інформацію про її характер.

Тема дисертаційної роботи є актуальною, має науково-прикладне значення та відповідає сучасним напрямам розвитку комп'ютерних наук, інтелектуального аналізу даних і технологій захисту комп'ютерних мереж.

Зв'язок дисертаційної роботи з науковими програмами, планами і темами. Дисертаційна робота відповідає пріоритетним напрямам розвитку інформаційних технологій в Україні. Робота є складовою частиною досліджень, проведених в Національному технічному університеті «Дніпровська політехніка», зокрема науково-дослідної роботи «Інтелектуальні методи моделювання процесів в інформаційно-телекомунікаційних системах критичної інфраструктури» (державний обліковий номер 0225U004731, термін виконання 2024-2025 рр.).

Особистий внесок здобувача у виконання зазначених науково-дослідних робіт полягає у розробці інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі, підготовці даних мережевого трафіку, побудові моделей машинного навчання та експериментальному оцінюванні їх ефективності для задач виявлення атак.

Обґрунтованість і достовірність наукових положень, висновків та рекомендацій. Сформульовані в дисертації наукові результати та висновки повною мірою відповідають установленим вимогам до досліджень відповідного рівня. Автор чітко визначив мету, об'єкт та предмет дослідження, що дозволило сформувати чітку логіку наукового пошуку. Наукові положення, висновки та рекомендації, сформульовані в роботі, є достатньо аргументованими та підтвердженими результатами проведеного дослідження. Обґрунтованість роботи забезпечується поєднанням теоретичного аналізу сучасних підходів до виявлення атак, методичного опрацювання задачі класифікації мережевого трафіку та експериментальної перевірки запропонованих рішень.

Високий ступінь обґрунтованості, достовірності та об'єктивності наукових результатів і висновків дисертації забезпечується також використанням сучасного апарату системного та статистичного аналізу, методів попередньої обробки даних, а також методів машинного навчання.

Достовірність отриманих результатів підтверджується використанням відкритого набору даних, який містить приклади нормального та атакувального мережевого трафіку і дає змогу досліджувати задачі як бінарної, так і багатокласової класифікації.

Важливим чинником надійності результатів є те, що ефективність запропонованої інформаційної технології оцінювалася не за одним окремим показником, а за системою взаємодоповнювальних метрик. У роботі використано accuracy, precision, recall, F1-score, ROC-AUC, Precision–Recall-криві, матриці помилок і п'ятикратну перехресну перевірку, що дозволяє комплексно характеризувати якість моделей машинного навчання.

Висновки дисертаційної роботи є логічним наслідком проведеного аналізу, розроблених методичних положень, програмної реалізації та експериментальної перевірки. Використані методи дослідження відповідають поставленій меті, а отримані результати мають належний рівень достовірності й обґрунтованості.

Наукова новизна одержаних результатів. Наукова новизна дисертації Мєшкова В.І. полягає у розробленні та обґрунтуванні інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі, яка забезпечує комплексне розв'язання задач підготовки даних, класифікації мережевого трафіку та оцінювання ефективності моделей машинного навчання для систем виявлення атак. Серед основних результатів варто відзначити:

- *уперше запропоновано* інформаційну технологію інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак, яка, на відміну від існуючих рішень, забезпечує в межах єдиного програмного модуля повний цикл підготовки даних мережевого трафіку, формування ознакового простору, побудови вибірок, балансування, навчання моделей машинного навчання та комплексного оцінювання результатів для двох взаємопов'язаних постановок задачі – бінарного виявлення атак і багатокласового розпізнавання їх типів;

- *удосконалено* метод підготовки даних, класифікації та оцінювання мережевого трафіку для систем виявлення атак, який відрізняється інтегрованим поєднанням процедур попередньої обробки даних, аналізу та зниження розмірності ознакового простору, формування й балансування вибірок, навчання моделей машинного навчання та оцінювання їх якості, що забезпечує підвищення ефективності підготовки даних і розв'язання задач бінарного виявлення атак та багатокласового розпізнавання їх типів;

- *набули подальшого розвитку* методичні положення комплексного оцінювання ефективності моделей машинного навчання в задачах моніторингу мережевого трафіку, які відрізняються уніфікованим поєднанням покласових

метрик якості класифікації, ROC-AUC, Precision–Recall-кривих, матриць помилок і п’ятикратної перехресної перевірки та дають змогу визначати ефективні конфігурації моделей для двох режимів функціонування інформаційної технології: первинного бінарного виявлення атак і подальшого багатокласового розпізнавання їх типів.

Оцінка змісту, структури та завершеності дисертаційної роботи. Повний обсяг дисертаційної роботи складає 230 сторінок, з яких 223 сторінки – основний текст. Список використаних джерел містить 150 найменувань.

У **вступі** обґрунтовано актуальність теми, визначено мету, завдання, об’єкт і предмет дослідження, наведено методи дослідження, сформульовано наукову новизну та практичне значення отриманих результатів. Вступна частина загалом дає повне уявлення про спрямованість роботи та очікуваний науковий результат.

Перший розділ присвячено аналізу теоретичних основ інтелектуального моніторингу мережевого трафіку та виявлення атак. У ньому розглянуто сучасні методи й засоби моніторингу, класифікацію IDS, підходи до виявлення загроз, а також особливості використання методів машинного навчання для аналізу мережевого трафіку. Спираючись на результати проведеного аналізу сформульовано задачі дисертаційного дослідження.

У **другому розділі** розроблено метод підготовки даних, класифікації та оцінювання мережевого трафіку. Позитивним є те, що автор розглядає підготовку даних як важливу складову ефективності системи виявлення атак. У розділі послідовно подано процедури очищення, уніфікації, стандартизації, аналізу ознак, зниження розмірності, формування та балансування вибірок для бінарної та багатокласової класифікації.

Третій розділ містить опис інформаційної технології інтелектуального моніторингу трафіку комп’ютерної мережі. У ньому представлено архітектуру технології, конвеєр обробки трафіку, двоетапну схему класифікації, структуру програмного модуля та основні етапи його реалізації. Цей розділ демонструє прикладну спрямованість роботи та доведення запропонованих методичних положень до рівня програмного рішення.

У **четвертому розділі** наведено результати експериментального дослідження ефективності розробленої інформаційної технології. Автором виконано оцінювання моделей машинного навчання для двох режимів роботи: бінарного виявлення атак і багатокласового розпізнавання їх типів. Використання системи взаємодоповнювальних метрик дає змогу більш повно оцінити якість класифікації та обґрунтувати отримані висновки.

У **додатках** представлені акти впровадження та використання результатів дисертаційного дослідження.

Загалом зміст дисертації відповідає її темі, меті та поставленим завданням. Розділи роботи логічно пов’язані між собою, матеріал викладено послідовно, а висновки узгоджуються з результатами проведених досліджень. Загальні висновки по роботі коректні і містять основні наукові результати отримані в дисертації здобувачем.

Дисертація є завершеною кваліфікаційною науковою працею, оформленою на належному науково-технічному рівні з використанням відповідної термінології, ілюстративного матеріалу, таблиць та результатів експериментального аналізу.

Практичне значення отриманих результатів. Практичне значення дисертаційної роботи полягає у можливості використання запропонованих методичних і програмних рішень для інтелектуального аналізу мережевого трафіку в системах виявлення атак. Розроблена інформаційна технологія може слугувати основою для створення програмних компонентів, які забезпечують попереднє опрацювання даних, класифікацію мережевої активності та формування результатів, придатних для подальшого аналізу фахівцями з кібербезпеки.

Здобувачем реалізовано послідовний процес обробки даних мережевого трафіку: від завантаження й очищення даних до навчання моделей машинного навчання, оцінювання їх якості та візуалізації результатів. Зазначене дає змогу використовувати розроблену технологію не лише для експериментального порівняння моделей, а й як прототип інтелектуального модуля підтримки прийняття рішень у системах моніторингу мережевої безпеки.

Важливим практичним результатом є підтримка двох режимів аналізу: визначення факту атакувальної активності та подальше розпізнавання типу атаки, що підвищує інформативність результатів моніторингу, оскільки система може не тільки виявити відхилення від нормального трафіку, а й надати деталізовану характеристику потенційної загрози.

Результати дисертаційного дослідження можуть бути використані в діяльності підприємств і організацій, які експлуатують комп’ютерні мережі та потребують засобів автоматизованого аналізу трафіку, а також у навчальному процесі під час підготовки фахівців за спеціальністю 122 «Комп’ютерні науки» та суміжними освітніми програмами у сфері кібербезпеки та захисту інформації.

Практичне значення роботи полягає у створенні програмно реалізованої інформаційної технології, придатної для застосування в задачах виявлення атак, аналізу мережевих загроз і підтримки прийняття рішень у сфері захисту комп’ютерних мереж.

Повнота висвітлення результатів дисертації в наукових публікаціях та їх апробація. Ознайомлення з дисертаційною роботою та науковими працями дозволяє зробити висновок про дотримання здобувачем вимог щодо повноти викладення результатів дисертації у наукових виданнях. Основні результати досліджень викладено в 11 публікаціях. Серед них: 4 статті у фахових виданнях України категорії Б, а 7 праць у матеріалах конференцій і симпозіумів, у тому числі 5 міжнародних. Наведений перелік публікацій, їх зміст та обсяг у достатній мірі відображають особистий внесок автора і відповідають вимогам, що висуваються до дисертації на здобуття наукового ступеня доктора філософії.

Відповідність дисертаційної роботи спеціальності 122 «Комп'ютерні науки». Дисертаційна робота Мєшкова Вадима Ігоровича за своїм змістом, поставленою метою, використаними методами та отриманими результатами відповідає спеціальності 122 «Комп'ютерні науки».

Тематика дослідження безпосередньо пов'язана з розробленням інформаційної технології, алгоритмізацією процесів обробки даних, побудовою моделей машинного навчання, програмною реалізацією запропонованих рішень та експериментальним оцінюванням їх ефективності. У роботі розглянуто задачі підготовки й аналізу мережевого трафіку, формування ознакового простору, класифікації даних, зниження розмірності, балансування вибірок і комплексного оцінювання якості моделей.

Отримані результати мають чітку інформаційно-технологічну спрямованість і належать до сфери розроблення методів, алгоритмів та програмних засобів інтелектуального аналізу даних. Запропонована технологія орієнтована на інтелектуальний моніторинг мережевого трафіку та підтримку прийняття рішень у системах виявлення атак.

Таким чином, дисертаційна робота за тематикою, змістом і результатами відповідає галузі знань 12 «Інформаційні технології» та спеціальності 122 «Комп'ютерні науки».

Дотримання вимог академічної доброчесності. За результатами аналізу дисертації Мєшкова В.І. можна зробити висновок, що дослідження виконано з дотриманням принципів академічної доброчесності. У тексті роботи використання наукових джерел, результатів інших авторів, методичних положень і довідкових матеріалів супроводжується відповідними бібліографічними посиланнями.

Дисертаційна робота має самостійний характер, а основні наукові та практичні результати узгоджуються з опублікованими працями здобувача за темою дослідження. У дисертації наведено перелік публікацій автора, у яких

висвітлено ключові результати роботи, зокрема щодо машинного навчання для класифікації мережевого трафіку, попередньої обробки даних та інтелектуального моніторингу трафіку для систем виявлення атак.

Ознак академічного плагіату, некоректного використання результатів інших дослідників, фабрикації чи фальсифікації даних під час ознайомлення з дисертаційною роботою не встановлено. Отже, дисертація відповідає вимогам академічної доброчесності, що висуваються до кваліфікаційних наукових праць на здобуття ступеня доктора філософії.

Дискусійні положення та зауваження до дисертаційної роботи. Позитивно оцінюючи дисертаційну роботу Мєшкова Вадима Ігоровича, доцільно зазначити окремі дискусійні положення та зауваження, які не повторюють раніше висловлені зауваження і можуть бути враховані в подальших дослідженнях.

1. У дисертації доцільно було б більш чітко розмежувати процедури попередньої обробки даних, які виконуються до поділу вибірки на навчальну і тестову частини, а також процедури, які мають виконуватися лише на навчальних даних. Це важливо для уникнення потенційного витоку інформації з тестової вибірки під час стандартизації, зниження розмірності або балансування даних.

2. У дисертаційній роботі значну увагу приділено порівнянню моделей машинного навчання за узагальненими метриками, однак доцільно було б детальніше проаналізувати покласову поведінку моделей для окремих типів атак. Такий аналіз дав би змогу краще визначити, для яких саме класів атак запропонована технологія демонструє найвищу ефективність, а для яких потребує додаткового налаштування або розширення навчальної вибірки.

3. При моделюванні та статистичній обробці експериментальних даних недостатню увагу приділено визначенню адекватності моделей та значущості отриманих результатів.

4. У роботі недостатньо уваги приділено питанню обмежень запропонованої технології під час аналізу зашифрованого або частково прихованого трафіку. Оскільки в сучасних мережах значна частина взаємодій відбувається із застосуванням шифрування, доцільно було б окремо зазначити, які саме ознаки залишаються доступними для аналізу та як це може впливати на якість виявлення атак.

5. В роботі мають місце деякі граматичні та стилістичні помилки; декілька раз зустрічаються введення одних й тих самих скорочень: наприклад, «системи виявлення атак (СВА)» наведено тричі – на стор. 2, 3, 20, але такого небагато по усьому тексту дисертації.

Вказані зауваження не знижують загального позитивного враження від дисертації, як з точки зору актуальності теми, так і з точки зору наукової новизни та практичної цінності отриманих результатів.

Загальний висновок. Дисертаційна робота Мєшкова Вадима Ігоровича на тему «Інформаційна технологія інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак» є самостійним, оригінальним і логічно завершеним науковим дослідженням. У роботі успішно розв'язано важливу науково-прикладну задачу розроблення інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак, що базується на методах машинного навчання та забезпечує функціонування в бінарному і багатокласовому режимах аналізу і, таким чином, дозволяє підвищити ефективність виявлення атак і розпізнавання їх типів у комп'ютерних мережах. Робота виконана на високому науковому рівні та не порушує принципів академічної доброчесності.

За актуальністю теми, обсягом і рівнем виконаних досліджень, науковою новизною та практичною цінністю одержаних результатів ця дисертаційна робота відповідає вимогам чинного законодавства України, що передбачені Порядком підготовки здобувачів вищої освіти ступеня доктора філософії та доктора наук у закладах вищої освіти (наукових установах), затвердженого постановою Кабінету Міністрів України від 23 березня 2016 р. № 261, пп. 6, 7, 8 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.

Вважаю, що здобувач МЄШКОВ Вадим Ігорович заслуговує на присудження наукового ступеня доктора філософії з галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Офіційний рецензент,
кандидат технічних наук, доцент,
доцент кафедри безпеки інформації
та телекомунікацій Національного
технічного університету
«Дніпровська політехніка» Олександра ГЕРАСІНА