Голові разової спеціалізованої вченої ради за спеціальністю 122 «Комп'ютерні науки» в галузі знань 12 «Інформаційні технології» Національного технічного університету «Дніпровська політехніка», м. Дніпро, д.т.н., професору, професору кафедри програмного забезпечення комп'ютерних систем
Борису МОРОЗУ

офіційного опонента, д.т.н., професора, професора кафедри інформаційної та кібернетичної безпеки ім. професора Володимира Бурячка, Київського столичного університету імені Бориса Грінченка, м. Київ
Геннадія ГУЛАКА

ВІДГУК

офіційного опонента д.т.н., професора, професора кафедри інформаційної та кібернетичної безпеки ім. професора Володимира Бурячка, Київського столичного університету імені Бориса Грінченка, м. Київ
ГУЛАКА Геннадія Миколайовича
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»

1. Актуальність теми дисертаційної роботи

Дисертаційна робота Мєшкова Вадима Ігоровича присвячена розв'язанню актуальної науково-прикладної задачі підвищення ефективності виявлення атак і розпізнавання їх типів у комп'ютерних мережах шляхом розроблення інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак.

Актуальність обраної теми зумовлена інтенсивним зростанням кількості, складності та різноманітності кіберзагроз, постійним збільшенням обсягів мережевого трафіку, динамічністю його характеристик, а також необхідністю

своєчасного виявлення атакувальної активності в комп'ютерних мережах. У сучасних умовах традиційні сигнатурні підходи до виявлення вторгнень не завжди забезпечують належну ефективність, особливо під час виявлення нових, модифікованих, комбінованих або малопоширених атак. Це обумовлює потребу в розробленні інтелектуальних методів і технологій аналізу мережевого трафіку, здатних виявляти приховані закономірності у великих обсягах даних та формувати обґрунтовані рішення щодо наявності атакувальної активності.

Важливим напрямом підвищення ефективності систем виявлення атак є застосування методів машинного навчання, які дають змогу виконувати автоматизовану класифікацію мережевого трафіку, відокремлювати нормальну активність від атакувальної, а також здійснювати подальше розпізнавання типів виявлених атак. Особливої значущості набуває поєднання процедур попередньої обробки даних, формування ознакового простору, балансування вибірок, навчання моделей машинного навчання та комплексного оцінювання результатів у межах єдиної інформаційної технології.

З огляду на це, тема дисертаційної роботи є своєчасною, науково обґрунтованою та практично значущою для галузі комп'ютерних наук і кібербезпеки. Розроблення інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак відповідає сучасним потребам підвищення захищеності інформаційно-комунікаційних систем, автоматизації аналізу мережевого трафіку та підтримки прийняття рішень у процесі виявлення кіберінцидентів.

Обрана тема дисертаційного дослідження є актуальною, має важливе наукове і прикладне значення, а її розв'язання сприяє розвитку методів, моделей та інформаційних технологій інтелектуального аналізу мережевого трафіку в системах виявлення атак.

2. Зв'язок дисертаційної роботи з науковими програмами, планами і темами

Дисертаційна робота Мєшкова Вадима Ігоровича виконана в Національному технічному університеті «Дніпровська політехніка» відповідно до плану науково-дослідних робіт кафедри безпеки інформації та телекомунікацій у межах науково-дослідної роботи «Інтелектуальні методи моделювання процесів в інформаційно-телекомунікаційних системах критичної інфраструктури» із державним обліковим номером 0225U004731, термін виконання 2024-2025 рр.

Тематика дисертаційного дослідження безпосередньо відповідає зазначеному науковому напряму, оскільки пов'язана з розробленням інформаційної технології інтелектуального моніторингу трафіку комп'ютерної

мережі для систем виявлення атак, підготовкою та аналізом даних мережевого трафіку, побудовою моделей машинного навчання, а також експериментальним оцінюванням їх ефективності. У межах зазначеної науково-дослідної роботи автором виконано дослідження, спрямовані на підвищення ефективності виявлення атак і розпізнавання їх типів у комп'ютерних мережах, що узгоджується з актуальними завданнями забезпечення кібербезпеки інформаційно-телекомунікаційних систем критичної інфраструктури.

Дисертаційна робота має чіткий зв'язок із науковими програмами, планами і темами Національного технічного університету «Дніпровська політехніка», а отримані результати відповідають змісту виконуваної науково-дослідної роботи та спрямовані на розвиток методів, моделей і інформаційних технологій інтелектуального аналізу мережевого трафіку в системах виявлення атак.

3. Ступінь обґрунтованості та достовірності наукових положень, висновків і рекомендацій

Наукові положення, висновки та рекомендації, сформульовані в дисертаційній роботі Мєшкова Вадима Ігоровича, є достатньо обґрунтованими та достовірними. Вони базуються на коректному використанні методів системного аналізу, інтелектуального аналізу даних, машинного навчання, статистичного опрацювання результатів, методів попередньої обробки мережевого трафіку, формування ознакового простору, зниження розмірності даних, балансування вибірок та експериментального оцінювання якості класифікації.

Достовірність отриманих результатів забезпечується логічною послідовністю проведеного дослідження: від аналізу сучасних методів і засобів моніторингу мережевого трафіку та систем виявлення атак до розроблення методу підготовки даних, побудови інформаційної технології, її програмної реалізації та експериментальної перевірки ефективності. У роботі послідовність обробки трафіку реалізовано як впорядкований програмний конвеєр: завантаження та інтеграція даних, первинний аналіз, попередня обробка, формування ознакового простору, побудова бінарної або багатокласової вибірки, балансування даних, навчання моделей машинного навчання, оцінювання та візуалізація результатів.

Обґрунтованість висновків підтверджується тим, що експериментальні дослідження проведено для двох взаємопов'язаних постановок задачі: бінарного виявлення атакувального трафіку та багатокласового розпізнавання типів атак. Для бінарного режиму використано моделі логістичної регресії та методу опорних векторів, а для багатокласового режиму – моделі випадкового лісу,

дерева рішень і методу k-найближчих сусідів. Такий добір моделей дозволив дослідити різні алгоритмічні принципи класифікації мережевого трафіку та порівняти їх ефективність за однакових умов підготовки даних .

Важливим чинником достовірності результатів є застосування комплексної системи оцінювання якості моделей. У дисертації використано сукупність взаємодоповнювальних критеріїв: accuracy, результати 5-fold cross-validation, матриці помилок, ROC-криві, Precision–Recall-криві, а також покласові показники precision, recall і F1-score. Проведення порівняльного аналізу за єдиним підготовленим ознаковим простором, після очищення даних, стандартизації та зниження розмірності методом головних компонент, забезпечує коректність зіставлення отриманих результатів і дає підстави для обґрунтованих висновків щодо практичної придатності розробленої інформаційної технології.

Слід відзначити, що сформульовані в роботі висновки узгоджуються з результатами експериментального дослідження та відображають фактичну ефективність запропонованої інформаційної технології. Застосування репрезентативного набору даних мережевого трафіку, єдиної методики підготовки вибірок, уніфікованих метрик оцінювання та порівняння кількох моделей машинного навчання підвищує надійність отриманих результатів.

Наукові положення, висновки та рекомендації дисертаційної роботи є належним чином обґрунтованими, методично послідовними та достовірними. Отримані результати підтверджують коректність запропонованого підходу до побудови інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак.

4. Наукова новизна отриманих результатів

Наукова новизна дисертаційної роботи Мєшкова Вадима Ігоровича полягає в розвитку методів, моделей і засобів інтелектуального аналізу мережевого трафіку для систем виявлення атак, а також у побудові цілісної інформаційної технології, що забезпечує автоматизоване виконання повного циклу обробки даних, навчання моделей машинного навчання та оцінювання результатів класифікації.

До найбільш суттєвих наукових результатів дисертаційної роботи слід віднести такі.

Уперше розроблено інформаційну технологію інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак, яка, на відміну від наявних науково-технічних 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. У роботі підкреслено, що така технологія підтримує два режими функціонування та може бути використана як основа для побудови інтелектуальних компонентів IDS у комп'ютерних мережах.

Удосконалено метод підготовки даних, класифікації та оцінювання мережевого трафіку для систем виявлення атак. Його відмінність полягає в інтегрованому поєднанні процедур попередньої обробки даних, очищення, уніфікації, стандартизації, аналізу ознак, зниження розмірності ознакового простору, формування й балансування вибірок, навчання моделей машинного навчання та оцінювання їх якості. Такий підхід забезпечує методично узгоджену підготовку даних і підвищує коректність розв'язання задач бінарної та багатокласової класифікації мережевого трафіку.

Набули подальшого розвитку методичні положення комплексного оцінювання ефективності моделей машинного навчання в задачах інтелектуального моніторингу мережевого трафіку. Їх особливістю є уніфіковане використання покласових метрик якості класифікації, ROC-AUC, Precision–Recall-кривих, матриць помилок і п'ятикратної перехресної перевірки, що дає змогу обґрунтовано визначати ефективні конфігурації моделей для первинного виявлення атакувального трафіку та подальшого розпізнавання типів атак.

Слід зазначити, що наукова новизна роботи має не декларативний, а прикладно підтверджений характер, оскільки запропоновані положення реалізовано в межах програмного модуля та перевірено експериментально на наборі даних CIC-IDS2017 у двох постановках задачі. Проведене дослідження дало змогу встановити ефективні моделі для бінарного виявлення атак і багатокласового розпізнавання їх типів, що підтверджує практичну придатність запропонованої інформаційної технології.

Отримані в дисертаційній роботі результати мають належний рівень наукової новизни, відповідають предметній області спеціальності 122 «Комп'ютерні науки» та становлять внесок у розвиток інформаційних технологій інтелектуального моніторингу мережевого трафіку для систем виявлення атак.

5. Аналіз змісту дисертаційної роботи

Дисертаційна робота Мєшкова Вадима Ігоровича має логічну, послідовну та завершену структуру, яка відповідає поставленій меті, сформульованим завданням і загальній логіці наукового дослідження. Повний обсяг дисертації становить 230 сторінок, з яких 190 сторінок основного тексту. Робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел із 150

найменувань та 5 додатків; містить 47 рисунків, 13 таблиць, що свідчить про достатній рівень деталізації теоретичних, методичних, програмних та експериментальних результатів дослідження.

У вступі автором обґрунтовано актуальність теми дисертаційного дослідження, сформульовано мету й завдання роботи, визначено об'єкт, предмет і методи дослідження, розкрито наукову новизну та практичне значення отриманих результатів. Мета роботи спрямована на розв'язання науково-прикладної задачі підвищення ефективності виявлення атак і розпізнавання їх типів у комп'ютерних мережах шляхом розроблення інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак, що базується на методах машинного навчання та функціонує у бінарному і багатокласовому режимах аналізу.

У першому розділі дисертації розглянуто теоретичні основи інтелектуального моніторингу мережевого трафіку та виявлення атак. Здобувачем проаналізовано сучасні методи й засоби моніторингу мережевого трафіку, класифікацію систем IDS/IPS, особливості сигнатурних, аномальних і гібридних підходів до виявлення загроз. Окрему увагу приділено методам інтелектуального аналізу трафіку в системах виявлення атак, типовим кіберзагрозам та їх проявам у мережевому трафіку. За результатами аналізу автором обґрунтовано наявність недоліків у традиційних підходах і сформульовано постановку задачі дослідження. Розділ є важливою теоретичною основою роботи, оскільки демонструє обізнаність здобувача із сучасним станом предметної області та дозволяє логічно перейти до розроблення власного методу.

У другому розділі викладено метод підготовки даних, класифікації та оцінювання мережевого трафіку для систем виявлення атак. У цьому розділі сформульовано постановку задачі інтелектуального виявлення атак у мережевому трафіку в межах двох взаємопов'язаних постановок: бінарної класифікації, спрямованої на відокремлення нормального трафіку від атакувального, та багатокласової класифікації, орієнтованої на розпізнавання конкретних типів атак. Автором охарактеризовано набір даних CIC-IDS2017, обґрунтовано його використання як експериментальної основи дослідження, а також розглянуто процедури очищення, уніфікації, стандартизації, аналізу ознак, зниження розмірності, формування та балансування вибірок. Важливо, що метод подано не як набір окремих процедур, а як послідовну технологічну схему підготовки даних і побудови класифікаційних моделей.

У третьому розділі розроблено інформаційну технологію інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак і представлено її програмну реалізацію. Автором визначено архітектуру інформаційної технології, описано конвеєр обробки мережевого трафіку та

обґрунтовано двоетапну схему класифікації, яка передбачає послідовне розв'язання задач бінарного виявлення атак і багатокласового розпізнавання їх типів. У роботі розглянуто вибір засобів програмної реалізації, структуру програмного модуля, порядок завантаження, інтеграції та первинного аналізу даних, реалізацію процедур попередньої обробки, формування ознакового простору, побудови вибірок, навчання моделей машинного навчання, оцінювання та візуалізації результатів. Такий зміст розділу свідчить про доведення запропонованих методичних положень до рівня програмної реалізації.

У четвертому розділі наведено результати експериментального дослідження ефективності розробленої інформаційної технології. Експерименти виконано на основі набору даних CIC-IDS2017 для двох режимів функціонування: бінарного виявлення атак і багатокласового розпізнавання їх типів. Для оцінювання якості моделей використано фіксований поділ вибірок на навчальну та тестову частини, п'ятикратну перехресну перевірку, accuracy, precision, recall, F1-score, ROC-AUC, Precision–Recall-криві та матриці помилок. У роботі встановлено, що в бінарній постановці найефективнішою є конфігурація методу опорних векторів, а в багатокласовій постановці найвищу точність продемонструвала модель k-найближчих сусідів, тоді як модель випадкового лісу показала кращу узагальнювальну здатність.

Загалом зміст дисертаційної роботи повністю відповідає її темі та поставленій меті. Розділи роботи логічно пов'язані між собою: від аналізу сучасного стану систем виявлення атак і методів інтелектуального аналізу мережевого трафіку – до розроблення методу, побудови інформаційної технології, її програмної реалізації та експериментальної перевірки. Виклад матеріалу є послідовним, науково аргументованим і достатньо деталізованим. Основні положення, висновки та рекомендації дисертації випливають зі змісту проведених досліджень і підтверджуються результатами експериментальної перевірки.

Дисертаційна робота Мєшкова Вадима Ігоровича за змістом, структурою та логікою викладення є завершеним науковим дослідженням, у якому послідовно розв'язано поставлені завдання та отримано результати, що мають наукове і практичне значення для розвитку інформаційних технологій інтелектуального моніторингу мережевого трафіку в системах виявлення атак.

6. Практичне значення одержаних результатів

Практичне значення дисертаційної роботи полягає в розробленні програмно реалізованої інформаційної технології інтелектуального моніторингу

трафіку комп'ютерної мережі, яка може бути використана як основа для побудови інтелектуальних компонентів сучасних систем виявлення атак.

Запропонована технологія забезпечує автоматизовану обробку мережевого трафіку, формування ознакового простору, навчання моделей машинного навчання, виявлення атакувальної активності та подальше розпізнавання типів атак. Її особливістю є підтримка двох режимів функціонування: бінарного виявлення атак і багатокласової класифікації їх типів.

Отримані результати підтверджують придатність розробленої інформаційної технології для використання в системах кібербезпеки, зокрема для підтримки прийняття рішень під час аналізу мережевих загроз. Результати роботи впроваджено в діяльність ТОВ «Центум-Д» та в освітній процес Національного технічного університету «Дніпровська політехніка».

7. Повнота викладення результатів у наукових публікаціях та апробація

Основні наукові положення, результати досліджень і практичні напрацювання дисертаційної роботи Мєшкова Вадима Ігоровича достатньо повно відображені в опублікованих наукових працях.

За темою дисертації опубліковано наукові праці у фахових виданнях України, зокрема з питань машинного навчання для бінарної та багатокласової класифікації мережевого трафіку, попередньої обробки даних для інтелектуальних систем виявлення атак, підвищення кібербезпеки електронних комунікаційних систем та аналізу систем інтелектуального моніторингу трафіку.

Результати дисертаційного дослідження апробовано на міжнародних і всеукраїнських науково-практичних конференціях, де представлено матеріали щодо архітектури інформаційної технології, аналізу наборів даних мережевого трафіку, кореляційного аналізу ознак та розроблення технології інтелектуального моніторингу трафіку для систем виявлення атак.

Опубліковані праці за тематикою, змістом і спрямованістю відповідають основним положенням дисертації та достатньою мірою висвітлюють отримані наукові результати.

8. Відповідність дисертаційної роботи спеціальності 122 «Комп'ютерні науки»

Дисертаційна робота Мєшкова Вадима Ігоровича за змістом, метою, завданнями та отриманими результатами відповідає спеціальності 122 «Комп'ютерні науки».

У роботі розроблено інформаційну технологію інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак, що передбачає застосування методів машинного навчання, алгоритмів попередньої обробки даних, формування ознакового простору, класифікації та оцінювання якості моделей.

Отримані результати належать до галузі комп'ютерних наук, оскільки спрямовані на розроблення методів, алгоритмів, програмних засобів та інформаційної технології інтелектуального аналізу даних мережевого трафіку. Це підтверджує відповідність дисертації спеціальності 122 «Комп'ютерні науки».

9. Дотримання принципів академічної доброчесності

За результатами ознайомлення з дисертаційною роботою ознак порушення академічної доброчесності не виявлено. Використані в роботі положення, ідеї, результати та матеріали інших авторів супроводжуються відповідними посиланнями на джерела.

Основні наукові результати дисертації є результатом самостійної роботи здобувача. У публікаціях, виконаних у співавторстві, відображено особистий внесок Мєшкова Вадима Ігоровича у розв'язання поставленої науково-прикладної задачі.

Таким чином, дисертаційна робота відповідає вимогам академічної доброчесності.

10. Дискусійні положення та зауваження до дисертаційної роботи

Позитивно оцінюючи дисертаційну роботу Мєшкова Вадима Ігоровича, слід зазначити окремі зауваження, які мають дискусійний і рекомендаційний характер та не знижують загальної наукової й практичної цінності отриманих результатів.

1. У дисертаційній роботі основну увагу приділено моделям машинного навчання, зокрема логістичній регресії, методу опорних векторів, дереву рішень, випадковому лісу та методу k-найближчих сусідів. Разом із тим доцільно було б ширше розглянути можливість порівняння запропонованого підходу із сучасними методами глибокого навчання, зокрема нейронними мережами, автоенкодерами або рекурентними моделями, які також застосовуються для аналізу мережевого трафіку та виявлення аномалій.

2. У роботі запропоновано архітектуру інформаційної технології та описано її програмну реалізацію, однак питання інтеграції розробленого рішення в реальні IDS/IPS-системи розкрито недостатньо детально. Було б доцільно

конкретизувати можливі сценарії практичного впровадження технології, зокрема її взаємодію з існуючими засобами моніторингу, джерелами мережевого трафіку, журналами подій, SIEM-системами та іншими компонентами інфраструктури кібербезпеки.

3. Експериментальне дослідження виконано на основі набору даних CIC-IDS2017, який є поширеним і репрезентативним для задач дослідження систем виявлення атак. Водночас додаткова перевірка запропонованої інформаційної технології на інших наборах даних або на фрагментах реального мережевого трафіку дала б змогу ширше оцінити її узагальнювальну здатність, стійкість до зміни структури трафіку та придатність до використання в різних мережевих середовищах.

4. У дисертації наведено результати навчання та оцінювання моделей машинного навчання, однак процедуру вибору й налаштування гіперпараметрів окремих моделей можна було б подати більш докладно. Зокрема, доцільно було б показати, яким чином зміна ключових параметрів моделей впливає на точність, повноту, F1-score, кількість помилкових спрацювань і пропущених атак, оскільки це має важливе значення для практичного застосування систем виявлення атак.

5. Дискусійним залишається питання стійкості розробленої інформаційної технології до зміни характеристик мережевого трафіку в часі. У реальних комп'ютерних мережах структура трафіку, поведінка користувачів і характер атак можуть істотно змінюватися. Тому в подальших дослідженнях доцільно було б розглянути механізми періодичного перенавчання моделей, адаптації до нових типів атак і виявлення деградації якості класифікації в процесі експлуатації.

6. У роботі використано комплексну систему метрик оцінювання якості моделей, що є позитивною стороною дослідження. Разом із тим варто було б ширше проаналізувати практичні наслідки помилкових спрацювань і пропусків атак для реальних систем кібербезпеки. Наприклад, у промислових або корпоративних мережах хибнопозитивні спрацювання можуть перевантажувати аналітиків безпеки, тоді як хибнонегативні результати можуть призводити до пропуску критичних інцидентів.

7. Доцільно було б детальніше висвітлити питання обчислювальної складності та швидкодії запропонованої інформаційної технології. Для практичного використання в системах моніторингу мережевого трафіку важливими є не лише показники точності класифікації, а й час обробки даних, вимоги до оперативної пам'яті, масштабованість рішення та можливість роботи з великими потоками трафіку в умовах, наближених до реального часу.

Зазначені зауваження мають переважно дискусійний і рекомендаційний характер, не знижують достовірності основних наукових положень, висновків і практичних результатів дисертаційної роботи, а також можуть бути враховані

автором у подальших дослідженнях, спрямованих на розвиток інформаційної технології інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак.

11. Загальний висновок

Дисертаційна робота Мєшкова Вадима Ігоровича «Інформаційна технологія інтелектуального моніторингу трафіку комп'ютерної мережі для систем виявлення атак» є завершеною кваліфікаційною науковою працею, у якій розв'язано актуальну науково-прикладну задачу підвищення ефективності виявлення атак і розпізнавання їх типів у комп'ютерних мережах шляхом розроблення інформаційної технології інтелектуального моніторингу мережевого трафіку.

Отримані в дисертації результати мають наукову новизну, належний рівень обґрунтованості та достовірності, а також практичне значення для розвитку інтелектуальних компонентів систем виявлення атак. Основні положення роботи достатньо повно висвітлені в наукових публікаціях здобувача та апробовані на науково-практичних конференціях.

За змістом, структурою, науковим рівнем і практичною спрямованістю дисертаційна робота відповідає спеціальності 122 «Комп'ютерні науки» та чинному Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженому постановою Кабінету Міністрів України від 12.01.2022 № 44.

З огляду на викладене, дисертаційна робота Мєшкова Вадима Ігоровича заслуговує на позитивну оцінку, а її автор – на присудження наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки».

Офіційний опонент:
д.т.н., професор, професор кафедри
інформаційної та кібернетичної
безпеки ім. проф. Володимира Бурячка,
Київського столичного університету
імені Бориса Грінченка, м. Київ

Геннадій ГУЛАК

КИЇВСЬКИЙ СТОЛИЧНИЙ УНІВЕРСИТЕТ ІМЕНІ БОРИСА ГРІНЧЕНКА Код ЄДРПОУ 45307965
КИЇВСЬКИЙ СТОЛИЧНИЙ УНІВЕРСИТЕТ ІМЕНІ БОРИСА ГРІНЧЕНКА Код ЄДРПОУ 45307965 * УКРАЇНА *
ОСОБИСТИЙ ПІДПИС
Г. Гулака ЗАСВІДЧУЮ
(ПІБ)
начальник ВК
(посада)