

ВІДГУК

офіційного опонента на дисертаційну роботу Польгуль Тетяни Дмитрівни на тему «Інформаційна технологія виявлення шахрайства при інсталюванні мобільних додатків з використанням інтелектуального аналізу даних», представлену на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Актуальність теми дисертації. Використання мобільних додатків різних функціональних призначень стало повсякденною нормою сучасного життя. Проведення компаніями-розробниками маркетингових акцій з метою залучення користувачів до інсталювань та використання розроблених додатків слугує причиною шахрайських способів інсталяції мобільних додатків, в яких шахраї, за відповідну грошову винагороду, приводять до компаній-розробників необхідну кількість фейкових (несправжніх) «користувачів».

Існуючі методи та системи виявлення шахрайства при інсталяції мобільних додатків використовують існуючі бази з шахрайськими даними (наприклад, IP-адресами), обчислення рейтингів користувачів на основі вибіркового вхідних даних, та ін. Основним недоліком існуючих систем є те, що вони розпізнають лише відомі види шахрайства і не можуть розпізнавати нові шахрайські шаблони.

Це зумовлює необхідність розробки нових методів та систем, які на основі самонавчання відстежують та визначають шаблони шахраїв, які непомітні людині. Саме така задача вирішується в дисертаційній роботі Польгуль Т. Д., яка присвячена розробці методу, моделі та інформаційної технології виявлення шахрайства при інсталюванні мобільних додатків з використанням інтелектуального аналізу даних.

Актуальність теми дисертації підтверджується також і тим, що вона виконувалась у рамках науково-дослідних робіт Вінницького національного технічного університету.

Наукова новизна отриманих результатів полягає в тому, що:

1. Вперше запропоновано метод подолання різномірності вхідних даних, що являє собою сукупність процедур вибору ознак, зниження розмірності та нормалізації даних, відмінність якого полягає у новій моделі процесу подолання різномірності даних шляхом шкалювання за інформативністю, що дозволяє всю множину різномірних даних про користувачів звести до вектору уніфікованих ознак без зменшення діагностичної цінності інформації.

2. Удосконалено модель класифікації користувачів на основі глибинних нейронних мереж у частині зниження розмірності та нормалізації даних згідно запропонованого методу подолання різномірності даних, яка є основою для створення узагальненого портрету шахрая з метою спрощення процесів їх виявлення.

3. Вперше розроблено узагальнений метод виявлення шахрайства при інсталюванні мобільних додатків, відмінність якого полягає у використанні

запропонованої моделі класифікації користувачів та методу подолання різномірності вхідних даних, що дозволяє визначити класи користувачів та підвищити достовірність виявлення шахрайства при інсталюванні мобільних додатків.

Ступінь обґрунтованості та достовірність наукових положень, висновків, рекомендацій, сформульованих у дисертації. Обґрунтованість та достовірність наукових положень, висновків і рекомендацій забезпечується аргументованою постановкою мети й задач дослідження, повнотою формулювання умов, в яких вони розв'язуються та необхідними припущеннями і обмеженнями щодо застосування результатів, використанням сучасного математичного апарату та програмного забезпечення. Теоретичні дослідження виконано з використанням сучасних методів інтелектуального аналізу даних, статистичних методів кластеризації та класифікації, методів машинного навчання та нейромережових технологій.

Достовірність отриманих результатів підтверджується їх узгодженням із теоретичними висновками, експериментами та чисельними розрахунками, а також впровадженням розроблених моделей і методів у запроповану інформаційну технологію.

Значимість отриманих результатів для науки і практичного використання. Цінність наукових результатів роботи полягає в тому, що в ній розроблено та досліджено методи та моделі виявлення шахрайства при інсталяції мобільних додатків, що дозволило підвищити достовірність та швидкодію процесу виявлення шахраїв.

Практична корисність роботи обумовлена тим, що здобувачем розроблено алгоритмічне та програмне забезпечення інформаційної технології виявлення шахрайства при інсталяції мобільних додатків.

Результати роботи впроваджені у: Garuda AI В. V. (м. Схіпгол, Нідерланди) ТОВ «ВІН ІНТЕРАКТИВ» (м. Вінниця, Україна), ТОВ «4ХайТек» (м. Вінниця, Україна), ПП «Літсофт» (м. Київ, Україна) та у навчальний процес кафедри комп'ютерних наук Вінницького національного технічного університету (ВНТУ) та кафедри інформатики, програмної інженерії та економічної кібернетики Херсонського державного університету.

Повнота викладення в публікаціях та апробація роботи. Основні наукові положення, висновки і рекомендації, які сформульовані в дисертаційній роботі, достатньо повно відображені в публікаціях здобувача і пройшли апробацію на міжнародних науково-технічних конференціях.

За темою дисертації опубліковано 20 наукових праць, в тому числі: 5 статей у наукових фахових виданнях України (одна з яких також входить до наукометричної бази даних Scopus); 6 статей в міжнародних наукових виданнях, п'ять з яких входять до міжнародної наукометричної бази Scopus, 7 тез доповідей у збірках матеріалів конференцій (три з яких міжнародні), отримано 2 свідоцтва про реєстрацію авторського права на твір.

Структура дисертації цілком відповідає логіці й послідовності рішення поставлених задач. Дисертація складається зі вступу, 4-х розділів, висновків, списку використаних джерел та додатків.

У вступі обґрунтовано актуальність теми дисертаційної роботи, зазначено зв'язок з науковими програмами, планами, темами, сформульовано мету та задачі, об'єкт та предмет дослідження, визначено наукову новизну та практичне значення одержаних результатів, наведено відомості про апробацію роботи, впровадження результатів та публікації.

У першому розділі, на підставі порівняльного аналізу існуючих методів, моделей, інформаційних технологій виявлення шахрайства, запропоновано розглядати задачу виявлення шахрайства при інсталяції мобільних додатків як задачу пошуку аномалій в даних. Показано, що існуючі методи не в повній використовують важливі вхідні дані, та не можуть розпізнавати нові шахрайські шаблони. Обґрунтована необхідність розробки інформаційної технології виявлення шахрайства, яка має змогу самонавчатися та відстежувати і визначати нові шаблони шахраїв. Визначена мета та задачі дослідження.

Другий розділ присвячено формалізації процесу виявлення шахрайства як аномалії в даних з використанням теорії множин, що дозволило визначити властивості аномальних і неаномальних даних, виділити ті інформативні характеристики даних, які дозволяють визначити клас користувача – шахрай чи органічний (не шахрай).

Розроблено узагальнений метод виявлення шахрайства та формалізовано його етапи. Запропоновано метод подолання різномірності вхідних даних шляхом їх шкалювання. Удосконалена модель класифікації вхідних даних на основі глибинної нейронної мережі.

У третьому розділі розроблено інформаційну технологію виявлення шахрайства. Розроблено алгоритм ініціалізації / корекції моделі виявлення шахрайства і бази знань та алгоритм виявлення шахрайства при наявності ініціалізованої моделі і бази знань.

У четвертому розділі апробовано розроблену інформаційну технологію. Розроблено методіку проведення експериментальних досліджень, виділені основні етапи. Визначені критерії ефективності та здійснено аналіз результатів тестування. Виконано порівняльний аналіз ефективності розробленого узагальненого методу виявлення шахрайства та існуючих методів.

Автореферат дисертації ідентичний за змістом з основними положеннями дисертації і достатньо повно відображає основні наукові положення, практичну значимість і висновки. Дисертаційна робота та автореферат оформлені у відповідності з встановленими вимогами.

Недоліки та зауваження щодо змісту дисертації та автореферату:

1. В підрозділі 1.2 «Варіантний аналіз методів пошуку аномалій в даних» варто було приділити більше уваги математичному формулюванню методів, які аналізуються.

2. В підрозділі 1.2.3 «Методи класифікації» наводиться метод найближчих сусідів, а ідеологічно близькі методи порівняння з прототипом (еталоном) чи потенційних функцій відсутні.

3. Наведена на рис. 2 автореферату та відповідного рис. 2.11 дисертації класифікація різнорідних даних є не зовсім коректною: дихотомічні дані не завжди є підмножиною номінальних даних – частіше вони є граничним випадком ординальних (рангових) даних; крім того, наведені приклади ординальних даних (мобільна платформа, тип події та ін.) слід віднести до номінальних даних, тому що над ними неможливо виконувати операції порівняння.

4. В Розділі 2 наведено узагальнену модель класифікації на основі глибинної нейронної мережі, а в Розділі 3 – база знань та алгоритми класифікації на основі нечіткої логіки. Варто було показати взаємозв'язок вказаних моделей класифікації.

5. Формалізація нечіткої моделі формування портрету шахряя варто було перенести з Розділу 3 в Розділ 2.

6. В «алгоритмі автоматичного виявлення шахраїв» (Розділ 3) варто було більш чітко виділити етапи: навчання (створення узагальненого шаблону шахрая); класифікація користувачів; налаштування моделі класифікації та коригування узагальненого шаблону шахрая. Слід також уточнити умови виконання кожного етапу, тому що в наведеному алгоритмі всі етапи обов'язково виконуються для **кожного** користувача, що не є ефективним.

7. З тексту дисертації не зрозуміло, яким чином та які параметри моделі класифікації налаштовуються згідно формули (3) автореферату та відповідної формули (3.11) дисертації.

8. Зміст табл. 3 «Порівняльний аналіз ефективності...» автореферату не відповідає аналогічній табл. 4.3 дисертації: дані табл. 3 помилково подані в %, відсутні дані відомих методів.

9. Наведені у Розділі 3 приклади запитів до БД та відповідей, а у Розділі 4 – фрагменти коду (Рис.4.13, 4.20 – 4.25) є мало інформативними і їх варто було винести в Додаток.

10. По тексту автореферату та дисертації вживається словосполучення «точність класифікації», хоча мова йде про «достовірність класифікації».

11. Оформлення дисертаційної роботи та автореферату виконано акуратно, послідовно, логічно, але у тексті дисертації зустрічаються незначні технічні помилки, зокрема:

- синтаксичні помилки;

- текстові позначення на рис. 4, 5, 7, 8 автореферату та відповідних рис. дисертації (а також інших рисунків дисертації, які є копіями екрану) менші мінімально-допустимих.

Однак зазначені зауваження не носять принциповий характер і не знижують цінності проведеного здобувачем дослідження, актуальності, новизни та практичної значущості дисертаційної роботи.

Висновки, щодо відповідності дисертації встановленим вимогам.

Дисертація є завершеною науковою роботою, в якій отримано нові науково-обґрунтовані теоретичні та експериментальні результати в галузі інформаційних технологій, що в сукупності вирішують актуальну науково-технічну задачу підвищення точності та швидкодії процесу виявлення шахраїв шляхом розробки моделей, методів та алгоритмів створення узагальненого портрету шахрая, подолання різномірності вхідних даних та пошуку аномалій в даних, які покладені в основу запропонованої інформаційної технології.

Тематика та зміст дисертації відповідають паспорту спеціальності і профілю спеціалізованої вченої ради.

Вважаю, що представлена дисертаційна робота «Інформаційна технологія виявлення шахрайства при інсталюванні мобільних додатків з використанням інтелектуального аналізу даних» за актуальністю вибраної теми, обсягом та рівнем виконаних теоретичних і експериментальних досліджень, достовірністю та обґрунтованістю висновків, новизною досліджень, значенням для науки і практики відповідає вимогам пп. 9, 11, 12 «Порядку присудження наукових ступенів», затвердженого постановою Кабінету Міністрів України від 24 липня 2013р. № 567, щодо кандидатських дисертацій, а здобувач Польгуль Тетяна Дмитрівна заслуговує присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент,
професор кафедри обчислювальної
техніки та програмування Національного
технічного університету «Харківський
політехнічний інститут»,
доктор технічних наук, професор



А. І. Поворознюк



Підпис Поворознюк А.І.
ЗАСВІДЧУЮ:
ПРИСУДЖЕННИЙ СЕКРЕТАР
НАЦІОНАЛЬНОГО-ТЕХНІЧНОГО УНІВЕРСИТЕТУ
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»
Заковоротний О.Ю.
" 03 " 20 20 р.