РЕЦЕНЗІЯ

на дисертаційну роботу

Северіна Андрія Івановича

на тему «Алгоритмічне та програмне забезпечення захисту приватних наборів даних у задачах класифікації»,

представлену на здобуття ступеня доктора філософії

в галузі знань **12 Інформаційні технології**

за спеціальністю **121 Інженерія програмного забезпечення**

Актуальність теми дисертації.

Системи аналізу даних і штучного інтелекту використовуються у різних сферах людського життя. Зокрема, такі системи застосовуються в електронній комерції, соціальній сфері, а також як персональні інструменти для вирішення типових задач (наприклад, чатботи). Одним з важливих елементів для побудови таких систем є дані, які використовуються для навчання й тестування систем аналізу даних і штучного інтелекту. Більші обсяги різнопланових вхідних даних дозволяють розробляти більш точні програмні системи. Значна частина даних береться з реального світу, проте їх також можна генерувати програмним шляхом. Кількість створюваних та оброблюваних даних постійно зростає, однак, досить часто дані містять принаймні частину інформації, яка є приватною (конфіденційною, чутливою або секретною). Наявність приватної інформації обмежує використання наборів даних для розроблення систем з використанням штучного інтелекту, адже втрата приватності може призвести до негативних наслідків. Зважаючи на це, збереження приватності даних є важливим.

Отже, науково-технічна задача вдосконалення алгоритмічного та програмного забезпечення захисту приватних наборів даних у системах з використанням штучного інтелекту, яка вирішується у даній дисертаційній роботі для задачі класифікації, є актуальною.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Наукова новизна результатів дисертації Андрія СЕВЕРІНА полягає в наступному:

1. Уперше запропоновано архітектуру програмної системи для вирішення задачі класифікації на основі приватних даних, характерною особливістю якої є захист приватних наборів даних, шляхом функціонального шифрування, що відбувається на стороні клієнта, і

дозволяє збільшити кількість наборів даних для навчання загальнодоступних систем аналізу даних і штучного інтелекту.

2. Уперше запропоновано модифікацію програмної моделі шифрування даних, яка відрізняється від існуючої використанням двовимірних згорткових нейронних мереж, замість одновимірних, і дозволяє застосовувати модель шифрування з використанням нейронних мереж до даних, що представлені набором пікселів, з яких складається зображення.
3. Уперше розроблено алгоритмічно-програмний метод функціонального шифрування наборів даних, особливістю якого є можливість використання приватних наборів даних в загальнодоступних системах аналізу даних та штучного інтелекту шляхом зменшення їх розмірності й функціонального шифрування отриманих даних з використанням приватного ключа.
4. Уперше розроблено алгоритмічно-програмний метод пошуку нормальних поліномів серед незвідних, який відрізняється від існуючого використанням простих чисел у десятковому представленні замість поліномів, що дозволяє зменшити обчислювальні витрати алгоритму пошуку незвідних многочленів з $O(n^3)$ до $O(n \log(\log n))$ і, як наслідок, спростити міжбазисні перетворення у бінарних скінченних полях з метою пришвидшення виконання операцій над елементами поля у методах гомоморфного шифрування даних.
5. Уперше розроблено модифікований спосіб побудови матриці переходу між поліноміальним та нормальним базисами скінченного поля, який полягає у використанні рекурентної формули $\alpha_{i+1} = t^{p^{i+1}} = t^{p^i \cdot p} = (\alpha_i)^p$ замість обчислення остачі від ділення елемента $t^{p^{i+1}}$ на незвідний поліном, що дозволяє зменшити кількість використовуваної пам'яті з n^{p^i} до $n \cdot p$, а також обчислювальну складність з $O(m^{p^i})$ до $O(m^p)$.

Наукові положення та висновки дисертаційної роботи є достатньо обґрунтованими та достовірними. Це забезпечується правильністю застосування математичного апарату при викладенні наукових результатів, чіткістю формулювання задач дослідження, успішною реалізацією розроблених методів, а також підкріплюється експериментальними результатами виконаних досліджень.

Наукові дослідження були виконані здобувачем на кафедрі програмного забезпечення комп'ютерних систем КПІ ім. Ігоря Сікорського в рамках ініціативної НДР (керівник к.т.н., доцент Онай М.В., № державної реєстрації 0121U109925).

Поставлене в дисертаційній роботі наукове завдання вдосконалення алгоритмічного та програмного забезпечення захисту приватних наборів даних у системах з використанням штучного інтелекту для задачі класифікації в дисертаційній роботі виконано повністю, здобувач повною мірою оволодів методологією наукової діяльності.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.

За своїм змістом дисертаційна робота здобувача Северіна А. І. повністю відповідає Стандарту вищої освіти зі спеціальності 121 Інженерія програмного забезпечення та напрямкам досліджень відповідно до освітньої програми Інженерія програмного забезпечення.

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у науковий напрям інженерії програмного забезпечення.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Северіна Андрія Івановича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

Мова та стиль викладення результатів.

Дисертаційна робота написана українською мовою. Для тексту дисертації притаманна послідовність та логічність викладу матеріалу. Автор використовує загальноприйняту термінологію, фахову лексику та науковий стиль мовлення. Матеріал дисертаційної роботи містить достатню кількість ілюстративного матеріалу (таблиць, діаграм, графіків та фрагментів коду), завдяки чому є доступним для сприйняття, аналізу та розуміння.

Дисертація складається з вступу, 4 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації 254 сторінки.

Вступ присвячено обґрунтуванню актуальності дисертаційного дослідження, визначенню мети і завдань дослідження, формулюванню наукової новизни й практичного значення отриманих результатів, а також характеристиці публікацій автора.

У першому розділі дисертаційної роботи наведено основні етичні аспекти, які варто враховувати під час розроблення та впровадження систем з використанням штучного інтелекту. Розглянуто загрози приватності даних й проведено комплексний порівняльний аналіз методів збереження приватності в машинному навчанні. Розроблено вимоги до програмного забезпечення захисту приватних наборів даних.

У другому розділі розроблено алгоритмічні методи міжбазисних перетворень елементів полів Галуа. Розглянуто особливості використання скінченних полів у гомоморфних методах збереження приватності. Проведено аналіз обчислювальної складності методів виконання операцій над елементами полів Галуа в поліноміальному й нормальному базисах. Запропоновано метод пошуку нормальних поліномів. Розроблено модифікований спосіб побудови матриці міжбазисних перетворень.

Третій розділ присвячено розробленню алгоритмічно-програмного методу захисту приватних наборів даних. Проаналізовано математичне підґрунтя побудови методів шифруванням з використанням нейронних мереж. Модифіковано модель шифрування даних, внаслідок чого її можна застосовувати для шифрування даних, що представлені набором пікселів, з яких складається зображення. Запропоновано метод функціонального шифрування даних, що дозволяє, шляхом зменшення розмірності даних й функціонального шифрування з використанням приватного ключа, використовувати приватні набори даних в машинному навчанні. Проведено аналіз метрик оцінки методів захисту наборів даних.

У четвертому розділі розроблено програмне забезпечення, що реалізує розроблені методи захисту приватних наборів даних та проведено експериментальні дослідження. Запропоновано архітектуру програмної системи для вирішення задачі класифікації на основі приватних даних. Розроблено програмну систему виконання обчислень над елементами поля $GF(p^m)$ в процесі гомоморфного шифрування з використанням різних базисів. Проведено експериментальні дослідження розроблених методів міжбазисних перетворень. Розроблено програмну систему вирішення задачі класифікації на приватних наборах даних, що реалізує метод функціонального шифрування для захисту приватних наборів даних. Проведено експериментальні дослідження розробленого методу функціонального шифрування. Розглянуто способи інтеграції розроблених програмних систем.

У висновках наведено наукові та практичні результати, отримані в дисертаційній роботі.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

Оприлюднення результатів дисертаційної роботи.

Наукові результати дисертації висвітлені у 4 наукових публікаціях здобувача, серед яких: 2 статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України; 2 статті у періодичних наукових виданнях, проіндексованих у базах даних Web of Science

Core Collection та Scopus, з яких 1 статтю у виданні, віднесеному до третього квартилю (Q3) відповідно до класифікації SCImago Journal and Country Rank.

Також результати дисертації були апробовані на 3 наукових фахових конференціях.

Наукові публікації здобувача повністю висвітлюють ключові результати дисертаційної роботи. Науковий рівень публікацій здобувача є високим. У всіх опублікованих працях автор дотримується принципів академічної доброчесності. Особистий внесок здобувача до всіх наукових публікацій, опублікованих із співавторами, є значним та переконливим.

Таким чином, наукові результати описані в дисертаційній роботі повністю висвітлені у наукових публікаціях здобувача.

Недоліки та зауваження до дисертаційної роботи.

1. У другому розділі дослідження проаналізовані найбільш відомі криптографічні схеми гомоморфного шифрування, в основі яких лежать скінченні поля. Було б доцільніше навести приклади операцій над елементами полів Галуа, що використовуються в цих схемах.
2. У підрозділі 4.2 при аналізі експериментальних досліджень, згадується додатковий параметр k, залежно від значень якого проводиться дослідження операції піднесення до степеня й операції Фробеніуса. Однак, цей параметр по різному визначається для цих операцій, тож автору краще було б навести визначення для цього параметру й у цьому підрозділі, а не лише в підрозділі 2.2, де операції піднесення до степеня й операції Фробеніуса розглядаються докладно.
3. У підрозділі 4.1 бажано було б детальніше описати складові запропонованої архітектури. Зокрема, на схемі в генератора ключів шифрування немає вхідних параметрів, що ймовірно зроблено для уникнення перенасиченості схеми, однак, це доцільно було б зазначити в тексті дисертації.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

Висновок про дисертаційну роботу.

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Северіна Андрія Івановича на тему «Алгоритмічне та програмне забезпечення захисту приватних наборів даних у задачах класифікації» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі

знань 12 Інформаційні технології. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в 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», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Северін Андрій Іванович заслуговує на присудження ступеня доктора філософії в галузі знань 12 Інформаційні технології за спеціальністю 121 Інженерія програмного забезпечення.

Рецензент:

доцент кафедри системного
програмування і спеціалізованих
комп'ютерних систем
КПІ ім. Ігоря Сікорського,
кандидат технічних наук, доцент Ярослав КЛЯТЧЕНКО

М.П.

«21» травня 2024 року