ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему: **ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ У МОБІЛЬНИХ ДОДАТКАХ**

Виконав: студент 2 курсу
другого (магістерського) рівня вищої освіти
спеціальності 126 «Інформаційні системи та технології»
ОПП «Інформаційні системи та технології»
Генс О.С.

Керівник: Райко Г.О.

Рецензент: Огнєва О.Є., к.т.н., доцент
кафедри ПЗ і Т
(прізвище та ініціали)

Херсон – 2020 року

РЕФЕРАТ

Кваліфікаційна робота магістра містить 87 сторінок, 33 рисунки, 1 таблиця, список використаних джерел із 43 найменування.

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ У МОБІЛЬНИХ ДОДАТКАХ

У першому розділі роботи представлено теоретичні основи дослідження аномалій при інсталюванні мобільних додатків, характеристику та аналіз методів пошуку аномалій в даних.

У другому розділі розкриті особливості застосування інтелектуального аналізу даних в методології виявлення аномалій при інсталюванні мобільних додатків, представлений аналіз та класифікація різнорідних даних при інсталюванні мобільних додатків.

У третьому розділі представлено дослідження інформаційної технології виявлення аномалій при інсталюванні мобільних додатків, характеристика алгоритмів виявлення шахрая, послідовність формування алгоритму створення узагальненого портрету шахрая.

КЛЮЧОВІ СЛОВА: ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, УПРАВЛІННЯ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ, МОБІЛЬНИЙ ДОДАТОК.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ACID - Atomicity, consistency, isolation, durability

ANN - Artificial neural networks

API - Application programming interface

CNN - Convolutional neural network

DNN - Deep neural network

DQN - Deep Q Learning

EM - Expectation-maximization

GAN - Generative Adversarial Nets

LB - Load Balancer

LOF - Local outlier factor

MTTI - Mean-time-to-install

NN - Neural networks

RNN - Recurrent neural network

SVM - Support Vector Machine

TPR - True positive rate

TTI - Time-to-install Outliers

ІАД - Інтелектуальний аналіз даних

ІТ - Інформаційна технологія

КН - Комп’ютерні науки

ШІ - Штучний інтелект

ВСТУП

Актуальність теми дослідження. На сьогоднішній день в умовах зростання кількості користувачів мобільних додатків, компанії-розробники вимушені звернутися за послугами до маркетингових кампаній, з метою залучення інсталювань саме до їхнього додатку. Саме така потреба у маркетингових кампаніях стала однією з причин появи шахраїв та їх шахрайських способів інсталювання мобільних додатків. Шахраї, у свою чергу, приводять до компаній-розробників необхідну кількість фейкових (несправжніх) «користувачів» та отримують за це відповідну грошову винагороду. Проте такі «користувачі» ніколи не повертаються у мобільний додаток, оскільки є фейковими, ми ж їх називатимемо шахрайськими.

Очевидно, що причиною недоліків систем є відсутність єдиного підходу до виявлення шахрайства на основі всіх наявних даних. Також, недоліком існуючих систем є те, що вони розпізнають лише відомі види шахрайства і не можуть розпізнавати нові шахрайські шаблони. А в сучасному світі важливою є можливість системи адаптуватись, тому необхідним є створення відповідних інформаційних технологій, що матимуть змогу самонавчатися.

Об’єкт дослідження – процеси виявлення шахрайства як аномалій в даних при інсталюванні мобільних додатків.

Предмет дослідження – моделі, методи та інформаційні технології виявлення шахрайства при інсталюванні мобільних додатків з використанням інтелектуального аналізу даних.

Метою роботи є дослідження процесів виявлення шахрайства при інсталюванні мобільних додатків.

Для досягнення вказаної мети в роботі розв’язуються такі основні *завдання:*

– аналіз методів виявлення шахрайства при інсталюванні мобільних додатків;

– формалізація процесу виявлення шахрайства як аномалії в даних;

– аналіз та класифікація різнорідних даних при інсталюванні мобільних додатків;

– дослідження методів виявлення шахрайства при інсталюванні мобільних додатків;

– характеристика методів подолання різнорідності вхідних даних;

– характеристика інформаційної технології виявлення шахрайства при інсталюванні мобільних додатків.

Методи дослідження, що використані в роботі: методи шкалювання під час вирішення задач аналізу та класифікації різнорідних даних при інсталюванні мобільних додатків та методу подолання різнорідності вхідних даних; теорія множин для вирішення задачі формалізації процесу виявлення шахрайства як аномалії в даних, а також методи класифікації, статистичні методи, методи машинного навчання, інтелектуальний аналіз даних, методи кластеризації, нейромережеві методи для вирішення задач розробки узагальненого методу виявлення шахрайства при інсталюванні мобільних додатків.

Результати проведених в роботі досліджень опубліковано у матеріалах III Всеукраїнської науково-практичної інтернет - конференції молодих вчених та студентів «Сучасні інформаційні системи та технології», з назвою: «Інтелектуальний аналіз даних в методах виявлення аномалій даних», що проходила 30 листопада 2020 року у Херсонському національному технічному університеті.