Маснюк В.В.

# ОБЧИСЛЮВАЛЬНИЙ КЛАСТЕР ДЛЯ ЗАПОБІГАННЯ ФРОДУ В ПЛАТІЖНІЙ СИСТЕМІ

Вінницький національний технічний університет

**Анотація**

*Дана робота присвячена темі обчислювального кластеру для запобігання фроду в платіжній системі. В роботі проаналізовано існуючі технології і методики для реалізації автоматичної обробки динамічних незгрупованих даних. В якості підходу до створення системи пропонується використання машинного навчання та розробка програмного рішення, що використовує модель аналізу для запобігання фроду у платіжній системі.*

**Ключові слова**

Обчислювальний кластер, платіжна система, фрод, машинне навчання.

**Abstract**

*This work is dedicated to the computing cluster for avoiding of pay system fraud. In this work was done analysis of existed technologies and methods for implementing automated processing of dynamic ungrouped data. As an approach of a system is offered to use machine learning and developing a software solution that uses built analysis model to prevent fraud in a payment system.*

**Keywords**

Computing cluster, pay system, fraud, machine learning.

**Актуальність теми дослідження.** Безперервний ріст даних і збільшення швидкості їх генерації породжують проблему їх обробки та зберігання. Не дивно, що тема «великих даних» (Big Data) є однією з найбільш обговорюваних в сучасному IT-суспільстві. Все більше галузей залучають технології для роботи з великими об’ємами даних. Однією з найперших, стала галузь електронної комерції, більш відомої як e-commerce. Основою цієї галузі є платіжні системи, які стали зручним і необхідним способом обробки грошових потоків, як для звичайних користувачів, так і для великого бізнесу. Рішення в галузі платіжних систем, залежно від їх популярності, щосекундно обробляють запити від одного до тисяч користувачів, запускають автоматичні процеси для ведення статистки та збору інформації, проводять транзакції та постійно звертаються до бази даних. Зважаючи на постійний ріст кількості користувачів, розробники таких систем постійно посилюють заходи безпеки.

У платіжних системах безпека застосовується до різних аспектів їхньої роботи, якими можуть бути: безпека даних користувачів, конфіденційність різних рівнів, захист від шахрайства. Зазвичай, найбільше збитків установам, в межах чиїх платіжних систем проводяться транзакції, завдає шахрайство. Шахрайство, яке проявляється в несанкціонованих діях та неправомірному використанні ресурсів платіжних систем, компаній та банків, називається фродом. Дане поняття виникло ще до розвитку систем електронної комерції, коли шахраї використовували обмеженні ресурси обчислювальних машин фінансових установ з метою безоплатного отримання певних послуг, за допомогою SMS-спаму. В наш час фрод набув більших масштабів та складніших методів здійснення, що обумовило собою створення низки заходів, що допомагають попередити або виявити шахрайство.

Для попередження шахрайства використовують засоби, якими користувачі мають підтвердити впевненість у своїх діях та захистити систему від спроб шахрайства, які використовують вразливі місця, що відкриваються під великими навантаженнями. Виявлення шахрайства здійснюється за допомогою систем моніторингу, які аналізують звіти про роботу платіжної системи, але існуючі рішення не можуть вирішити чи є користувач шахраєм або потенційним шахраєм, а також не мають можливості тимчасово чи довічно заблокувати доступ до системи користувачу. Такі рішення приймають працівники підтримки

платіжних систем, але в їх аналізі отриманих даних і відповідних діях існує систематичність. Саме тому існує можливість і потреба створити систему для запобігання фроду, яка здатна «навчитись» розпізнавати і попереджати шахрайство [1].

Для реалізації вирішення цієї проблеми необхідні відповідні ресурси, яких буде достатньо для аналізу такого великого об'єму даних. У наш час існує достатньо технологій і сервісів, які дозволяють орендувати такі ресурси, а не купувати комплекс обладнання і в майбутньому обслуговувати його.

Одним з найвідоміших і обговорюваних проектів в області розподілених обчислень є Hadoop – розроблюваний фондом Apache Software Foundation, вільно поширюваний набір з утиліт, бібліотек і фреймворк для розробки та виконання програм розподілених обчислень. Hadoop включає у себе модулі, що призначені для обслуговування кластеру з різних сторін його використання. Він підтримується більшістю сервісів, які надають в оренду обчислювальні ресурси [2].

Аналізуючи сучасну популярність платіжних систем і відповідно швидкість розвитку методів шахрайства у них, з'явилась необхідність у розробці системи запобігання фроду, яка буде гнучкою для конфігурації, щоб адаптуватись до майбутніх умов роботи платіжної системи та нових видів фроду.

**Метою дослідження** є підвищення продуктивності роботи служби підтримки платіжної системи за рахунок розробки комп'ютерної системи, яка б визначала й попереджала ймовірні випадки шахрайства.

**Задачі дослідження**: 1) аналіз та визначення необхідних технологій і методик для організації роботи кластеру; 2) формування позитивної моделі поведінки користувачів у межах платіжної системи і вибірка характерних дій, що свідчать про можливе шахрайство; 3) проектування та реалізація системи для пошуку підозрілої активності користувачів; 4) пошук і вибір підходу навчання створеної системи, який дав би найкращі результати; 5) тестування та аналіз отриманих результатів.

**Об'єкт дослідження** – попередження та виявлення фроду у платіжних системах за допомогою аналізу даних моніторингу системи.

**Предметом дослідження** є модель поведінки користувачів, виділення ознак або дій, які можна вважати підозрілими, та машинне навчання обчислювального кластеру, для запобігання фроду у платіжній системі.

**Висновок.** Сфера застосувань машинного навчання постійно розширюється. Всеохоплююча інформатизація призводить до накопичення величезних об'ємів даних в науці, виробництві, бізнесі, транспорті, охороні здоров'я. Виникаючі при цьому задачі прогнозування, управління та прийняття рішень часто зводяться до навчання за прецедентів. Раніше, коли таких даних не було, ці завдання або взагалі не ставилися, або вирішувалися зовсім іншими методами.

На основі отриманих та проаналізованих даних можна зробити висновок, що інтеграційне машинне навчання з використанням методів лінійної регресії дозволить сформувати необхідну навчальну вибірку та буде найбільше відповідати завданню класифікації та вибірці об'єктів за ознаками.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ


1. High Performance Computing for Computational Science – VECPAR 2006: 7th International Conference / Michel Dayde – Springer Science & Business Media, 2007 – 721 p.
2. The history of Hadoop: From 4 nodes to the future of data [Електронний ресурс]: Gigaom – Режим доступу: https://gigaom.com/2013/03/04/the-history-of-hadoop-from-4-nodes-to-the-future-of-data/2/.



Маснюк Володимир Васильович, 2КІ-16м, комп'ютерна інженерія.

Науковий керівник: Ткаченко Олександр Миколайович, к.т.н., доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, alextk1960@gmail.com

Masniuk Volodymyr Vasylovych, 2KI-16m, computer engineering.