

УДК 004.056

О.Безруков, Стадник Марія, к.т.н., доц.

(Тернопільський національний технічний університет імені Івана Пулюя)

ВИЯВЛЕННЯ ШАХРАЙСЬКИХ ТРАНЗАКЦІЙ З ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ

O. Bezrukov, Stadnyk Mariia, Ph.D., Assoc. Prof.

DETECTION OF FRAUD TRANSACTIONS USING MACHINE LEARNING METHODS

Виявлення шахрайства – це набір процесів і методів, які дозволяють підприємствам виявляти та запобігати несанкціонованій фінансовій діяльності, до якої, зокрема, належать шахрайські операції з кредитними картками, крадіжки, кіберзломи, шахрайство зі страхуванням тощо. Виявлення шахрайських транзакцій є однією з центральних питань безпеки таких фінансових установ, як банки, та стосується процесу моніторингу транзакцій і поведінки клієнтів, щоб точно визначити шахрайську діяльність і боротися з нею.

Шахрайство з транзакцією — це купівля товарів і послуг за допомогою викрадених платіжних даних. Кіберзлочинці або викрадають платіжну інформацію зі слабо захищених корпоративних баз даних, або ж отримують потрібні дані методами соціальної інженерії, або купують її на темних форумах. Жертви часто не знають про це, поки не перевірять свої виписки та не помітять одну або кілька несанкціонованих транзакцій, лише після того можуть подати запит до банку. Ознаками таких транзакцій часто є нетипові для людини замовлення, поспіх з доставкою або ж інша адреса доставки.

Часто виявлення шахрайських транзакцій належить до обов'язків дата аналітиків, основними завданнями яких є аналіз історичних даних та створення моделей та методів, які дозволять швидко визначати підозрілі транзакції та блокувати їх. Машинне навчання є одним з ключових методів, що дозволяє виявляти злочинні перекази. Виявлення шахрайських транзакцій можна трактувати як задачу класифікації, коли фінансова установа має історичні промітковані дані, на основі яких проводить аналіз нових даних. Або ж як задачу кластеризації, коли серед існуючого набору транзакцій необхідно виділити підозрілі та ті, які не викликають підозр.

Основними проблемами, які виникають при аналізі є те, що процес виявлення підозрілих транзакцій можна віднести до задачі виявлення аномалій (Anomaly detection). Класичні методи машинного навчання чудово працюють зі збалансованими наборами даних, коли кількість екземплярів кожного класу орієнтовно однакова. Натомість виявлення аномалій, за звичай, передбачає, що нестандартних випадків є значно менше, ніж стандартних, а отже вимагає особливих методів для роботи з такими датасетами. Ще однією специфікою таких задач є необхідність виявлення підозрілих фінансових операцій майже в реальному часі, адже сучасна екосистема транзакцій створена для швидкості та зручності. Навіть такий відносно складний процес, як заявка на кредит, можна здійснити за допомогою смартфона, тоді як більш рутинні покупки завершуються кількома натисканнями клавіш. Це високошвидкісне середовище може полегшити шахраям завершити свої злочини та зникнути до того, як їх вдасться виявити.

Ще однією проблемою може стати надмірно старанна система виявлення шахрайства, адже це може призвести до більшої кількості помилкових спрацьовувань. Це незручно для клієнтів, які в результаті можуть стати менш лояльними, і дорого для компаній, які повинні витратити час і ресурси на обробку попередження. В роботі буде продовжено дослідження виявлення аномальних для клієнтів транзакцій з метою усунення вище наведених проблем.