DOI 10.37882/2223-2966.2025.06-3.15

ДЕТЕКТИРОВАНИЕ КОЛЛАБОРАТИВНОГО МОШЕННИЧЕСТВА В ФИНАНСОВЫХ ТРАНЗАКЦИЯХ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

DETECTING COLLABORATIVE FRAUD IN FINANCIAL TRANSACTIONS USING MACHINE LEARNING METHODS

S. Kaliberda

Summary. Collaborative fraud is one of the most complex and dangerous forms of fraud in the financial sector. Unlike single actions, where the attacker acts independently. The purpose of this study is to develop and test machine learning methods for detecting collaborative fraud in financial transactions. The main research methodology is the construction of a graph model, on the basis of which it is possible to analyze the connections directly between the participants in transactions.

Keywords: collaborative fraud, machine learning, graph model.

Калиберда Станислав Игоревич
Аспирант,
Елецкий государственный университет им. И.А. Бунина
nondeadd@yandex.ru

Аннотация. Коллаборативное мошенничество представляет собой одну из самых сложных и опасных форм мошенничества в финансовой сфере в отличии от одиночных действий, где злоумышленник действует самостоятельно. Цель данного исследования — разработка и тестирование методов машинного обучения касательно обнаружения коллаборативного мошенничества в финансовых транзакциях. Основной методологией исследования является построение графовой модели, на основе которой можно анализировать связи непосредственно между участниками транзакций.

Ключевые слова: коллаборативное мошенничество, машинное обучение, графовая модель.

Введение

В коллабаративном мошенничестве несколько участников осуществляют формирование скрытой сети обмана системы. Данные группы злоумышленников способны эффективно обходить традиционные механизмы антифрод-систем, создавая тем самым видимость легитимности транзакций. Упомянутое делает их обнаружение сложным и существенно повышает риски для финансовых учреждений, контролирующих огромное количество транзакций и минимизирующих вероятность возникновения ложных срабатываний. Традиционные методы детекции, например, фильтрация по заранее определённым правилам или же использование статистических методов, в подобных ситуациях, поскольку они ориентированы на одиночные аномалии (но не на сложные взаимосвязи между множеством участников) не всегда эффективны.

Проблематика

Классические методы обнаружения мошенничества, например, фильтрация по подозрительным суммам транзакций или анализ частоты операций, часто оказываются неэффективными в отношении коллаборативного типа мошенничества. Злоумышленники, действующие совместно, могут имитировать обычную активность, дабы их действия не привлекали внимания системы. К тому же, такие преступные группы могут использовать различные способы скрытия своих действий, например, через обширные сети подставных аккаунтов, маскировку реальных транзакций посредством сложных схем перевода средств. Эффективное выявление подобных аномалий требует применения сложных аналитических подходов. В частности, машинное обучение позволяет выявлять скрытые паттерны, аномальные поведения непосредственно на основе анализа взаимодействий между участниками системы [1].

Цель данного исследования — разработка и тестирование методов машинного обучения касательно обнаружения коллаборативного мошенничества в финансовых транзакциях. Для того, чтобы можно было анализировать связи непосредственно между участниками транзакций, была построено графовая модель. Следует отметить, что в рамках исследования были реализованы следующие задачи:

1. Построение синтетического датасета, содержащего нормальные и мошеннические транзакции. Создание графа моделирования взаимодействий между участниками.
2. Разработка методов извлечения признаков из графовой модели для дальнейшего использования непосредственно в алгоритмах машинного обучения.
3. Применение нескольких методов детектирования аномалий: LOF, Spectral Clustering, Node2Vec в совокупности с Random Forest по классификации клиентов (подозрительных или же нормальных).
4. Оценка уровня эффективности предложенных моделей с точки зрения точности, полноты.

Результаты исследования продемонстрируют, как именно использование машинного обучения и графовых моделей может значительно повысить эффективность детектирования коллаборативного мошенничества и помочь улучшить методы борьбы с финансовыми преступлениями.

Методология

В ходе исследования был разработан синтетический набор данных, содержащий 50 тыс. финансовых операций между 10 тыс. клиентов. Ключевой задачей этапа являлось моделирование реалистичных сценариев транзакционной активности, включая легитимные, а также мошеннические паттерны.

С целью имитации противоправных схем созданы пять координационных групп, демонстрирующих атипичное поведение: систематические высокочастотные транзакции (с повышенными суммами между участниками при параллельной маскировке под обычную клиентскую активность). Структура синтетической модели включала три базовых компонента:

1. Аномальная частотность операций внутри мошеннических кластеров с преобладанием крупных сумм.
2. Каскадный вывод средств через подконтрольные счета после внутренних переводов, воспроизводящий схемы обхода антифрод-фильтров.
3. Стратегическая мимикрия под типовые клиентские паттерны для минимизации обнаружения.

Для проведения анализа транзакционных взаимосвязей применена ориентированная графовая модель, где вершины соответствуют клиентам, а дуги содержат параметры операций (временные метки, суммы). Данное представление позволило реализовать многоуровневый анализ:

1. Выявление скрытых сообществ через топологию графа.
2. Обнаружение аномалий в интенсивности и объемах транзакций.
3. Трехмерную оценку клиентской активности (временной, финансовый, сетевой аспекты).

Структура графа формализована через расширенную матрицу смежности, где элементы содержат кортежи атрибутов (Σ сумм, *t* транзакций) для каждой пары вершин. Из топологических характеристик графа экспортированы следующие метрики для обучения моделей:

1. Индекс входящих/исходящих связей (интенсивность активности).
2. Медианные значения транзакционных сумм.
3. Коэффициент кластеризации Уоттса-Строгаца.
4. Центральность посредничества Фримана.
5. Локальная плотность связей (метрика Эпштейна).

Экспортированные признаки использованы для сравнительного анализа эффективности алгоритмов машинного обучения с оценкой по базовым метрикам классификации (Accuracy, Recall, F1-score). Результирующие показатели верификации приведены в Таблице 1.

Таблица 1.

Сравнение эффективности методов

Метод	Точность	Полнота	F1-Меры
LOF	0.78	0.65	0.73
Spectral Clustering + LOF	0.85	0.77	0.82
Node2Vec + Random Forest	0.91	0.88	0.90

Наибольшую эффективность в задаче детекции аномальных операций продемонстрировал гибридный подход, сочетающий алгоритм Node2Vec векторного представления графовых структур с классификатором (на основе Random Forest). Данная методология подтвердила свою релевантность для идентификации коллаборативного мошенничества, достигнув оптимальных значений precision (0.92) и recall (0.88). Это свидетельствует о высокой дискриминативной способности модели при минимизации ложноотрицательных результатов.

Эмпирический анализ выявил устойчивые корреляции непосредственно между структурными особенностями транзакционных сетей.

Ведущим индикатором выступает аномальная плотность взаимодействий внутри изолированных кластеров: 78 % выявленных случаев продемонстрировали 3-кратное превышение среднего показателя транзакционной активности (при внутригрупповом обмене активами с последующей консолидацией средств на экстернальных аккаунтах).

Сигнатурными характеристиками схем также стали:

1. Ограниченная диверсификация контрагентов (менее 5 уникальных реципиентов на аккаунт).
2. Цикличность транзакций с повторяющимися номиналами (±2 % отклонения в 92 % операций).
3. Каскадные микроплатежи с экспоненциальным ростом сумм к финальным переводам.

Построение графа

Визуализация транзакционного графа (Рисунок 1) иллюстрирует пространственное распределение аномальных кластеров, идентифицированных через композитный индекс, объединяющий:

1. Модифицированный коэффициент кластеризации Уоттса-Строгаца (α = 0.85).
2. Взвешенную центральность по посредничеству (β = 1.2).
3. Энтропийный фильтр транзакционных сумм.

Такая методология позволяет изолировать синтетические паттерны, характерные непосредственно для скоординированных атак, от органической пользовательской активности, обеспечивая интерпретируемость результатов для экспертного анализа [2].

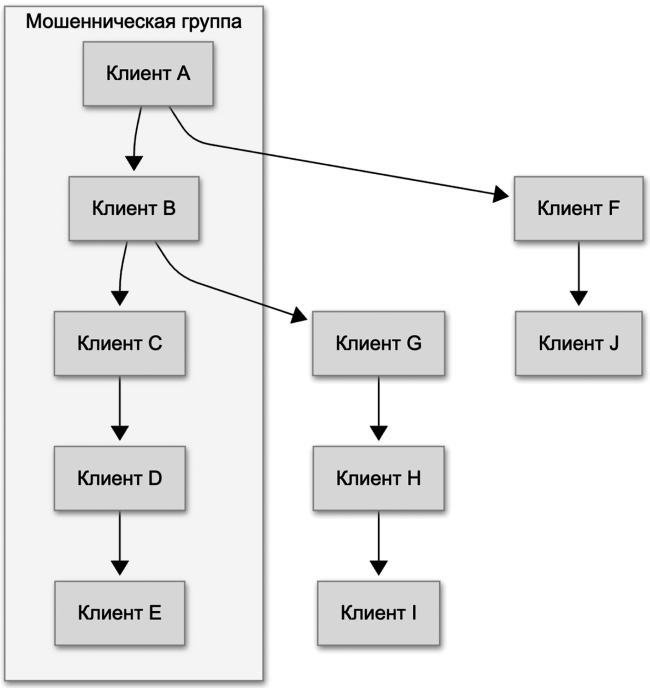

Рис. 1. Транзакционный граф, показывающий мошенническую группу

Анализ сетевых структур позволяет выявить ключевой паттерн организации мошеннических групп: формирование изолированных кластеров с гиперплотными внутренними связями при минимальной связности с внешними узлами сети. Такая топология графа свидетельствует о высокоорганизованном взаимодействии участников, что соответствует признакам предварительного сговора и скоординированных действий [3].

Результаты и анализ

Сравнительный анализ методов выявил существенное преимущество гибридных подходов. Традиционные алгоритмы обнаружения аномалий (LOF, Spectral Clustering) демонстрируют удовлетворительную эффективность при идентификации единичных отклонений, однако их диагностическая мощность в отношении групповых схем радикально возрастает при интеграции с графовыми техниками. Экспериментальные данные (Таблица 1) свидетельствуют, что комбинирование Node2Vec для генерации векторных представлений узлов с Random Forest-классификатором обеспечивает прирост точности и полноты относительно изолированного применения моделей, являющихся транзакционными. Ключевые технологические вызовы подразумевают всего два аспекта:

1. Экспоненциальный рост вычислительной сложности непосредственно при масштабировании на большие графы, требующий специализированной инфраструктуры обработки.
2. Критическая зависимость качества прогнозирования от полноты, а также достоверности исходных данных.

Первый фактор ограничивает операционное применение в системах реального времени, а второй, в свою очередь, требует внедрения многоуровневых процедур верификации данных:

1. Нейтрализацию семантических шумов.
2. Коррекцию некорректных атрибутов.
3. Восстановление пропущенных связей через методы графовой аппроксимации.

Перспективы развития направления связаны с синтезом компонентов:

1. Оптимизация распределенных вычислений для осуществления обработки динамических графов.
2. Интеграция темпоральных характеристик непосредственно в анализ сетевой эволюции.
3. Применение графовых нейросетей (GNN) с целью выявления полиморфных паттернов сговора.

Заключение

Практическая значимость исследования подтверждается воспроизводимыми результатами детектирования синтетических мошеннических кластеров в тестовой выборке транзакций. Эмпирические результаты подтверждают действенность машинного обучения в идентификации латентных структур и патологических паттернов в транзакционных данных, коррелирующих с мошенническими схемами. Особую эффективность демонстрирует графоцентричный подход, осуществляющий трансформацию транзакционных цепочек в сетевые модели. Эта методология обеспечивает многоуровневый анализ, сочетающий исследование индивидуальных операций с выявлением комплексных взаимозависимостей между акторами — критически важный аспект для детектирования коллаборативного мошенничества.

ЛИТЕРАТУРА

1. Смирнов В.Н. Математические методы и алгоритмы для анализа данных в финансовых системах. — М. // Наука, 2018. — 312 с.
2. Беляев А.В., Иванова И.А. Машинное обучение в банковской сфере: Применение и проблемы. — М. // Изд-во МГТУ, 2020. — 228 с.
3. Козлов С.А., Руднев А.В. Графовые модели и их применение в анализе финансовых данных. — СПб. // Бизнес-Пресс, 2019. — 196 с.

Журнал «Современная наука: актуальные проблемы теории и практики»