DOI: 10.58168/PPSPM2025_192-198
УДК 004.9

ПРИМЕНЕНИЕ ГРАФОВЫХ БАЗ ДАННЫХ ДЛЯ ВЫЯВЛЕНИЯ МОШЕННИЧЕСКИХ СХЕМ В ФИНАНСОВОЙ СФЕРЕ

А.М. Тюнина, Д.С. Кукуева, В.В. Кондусова

ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»

В статье рассматривается применение графовых баз данных Neo4j для обнаружения и анализа сложных мошеннических схем в финансовых операциях. Исследование демонстрирует преимущества графового подхода перед традиционными реляционными базами данных при анализе связанных данных. Подробно описаны ключевые алгоритмы анализа графов: PageRank для выявления влиятельных узлов, Betweenness Centrality для обнаружения посреднических элементов и алгоритмы обнаружения сообществ для идентификации организованных групп. Приведены практические примеры запросов на языке Cypher и схемы организации типичных мошеннических операций. Результаты исследования показывают, что графовый анализ позволяет перейти от реактивного к проактивному подходу в борьбе с финансовым мошенничеством, обеспечивая выявление сложных взаимосвязей и превентивное блокирование противоправных схем.

Ключевые слова: графовые базы данных, Neo4j, финансовое мошенничество, анализ социальных сетей, PageRank, центральность по посредничеству, обнаружение сообществ, Cypher, связанные данные, проактивная безопасность.

THE USE OF GRAPH DATABASES TO IDENTIFY FRAUDULENT SCHEMES IN THE FINANCIAL SECTOR

A.M. Tyunina, D.S. Kukueva, V.V. Kondusova

Voronezh State University of Forestry and Technologies named after G.F. Morozov

The article discusses the use of Neo4j graph databases to detect and analyze complex fraudulent schemes in financial transactions. The study demonstrates the advantages of the graph approach over traditional relational databases in the analysis of related data. The key graph analysis

algorithms are described in detail: PageRank to identify influential nodes, Betweenness Centrality to detect intermediary elements, and community detection algorithms to identify organized groups. Practical examples of queries in the Cypher language and schemes for organizing typical fraudulent operations are given. The results of the study show that graph analysis makes it possible to move from a reactive to a proactive approach in the fight against financial fraud, ensuring the identification of complex relationships and the preventive blocking of illegal schemes.

Keywords: graph databases, Neo4j, financial fraud, social network analysis, PageRank, mediation centrality, community discovery, Cypher, linked data, proactive security.

Современные финансовые мошенничества эволюционировали от единичных противоправных действий к сложным организованным схемам, использующим множество взаимосвязанных счетов и транзакций для сокрытия незаконной деятельности. Традиционные системы мониторинга, основанные на реляционных базах данных и анализе статистических аномалий, демонстрируют ограниченную эффективность против таких схем. Основная проблема заключается в неспособности реляционных моделей эффективно анализировать сложные сети взаимосвязей между сущностями, что приводит к пропуску скоординированных мошеннических операций. Кроме того, существующие методы часто фокусируются на реактивном подходе, выявляя инциденты после их совершения, вместо проактивного предотвращения. Это создает серьезные уязвимости в финансовых системах и приводит к значительным экономическим потерям. В связи с этим возникает насущная потребность в разработке и внедрении инновационных подходов к анализу финансовых данных, способных выявлять сложные сетевые взаимосвязи и пресекать мошеннические схемы на ранних стадиях их формирования.

Традиционные реляционные базы данных и методы анализа, основанные на проверке отдельных записей, зачастую не способны эффективно выявлять такие скрытые взаимосвязи. Они могут дать ответы на вопросы «что?» и «сколько?», но не могут ответить на вопросы «почему?» и «с кем связаны эти транзакции?».

Однако, в настоящее время активно используются графовые базы данных, такие как Neo4j. Эти системы позволяют моделировать и анализировать финансовые экосистемы, включая взаимосвязи между клиентами, счетами, компаниями и транзакциями. Это делает их мощным инструментом для обнаружения мошеннических схем, которые остаются незамеченными традиционными системами.

Таблица 1 – Сравнение реляционный vs графовый подход

Критерий	Реляционные БД (SQL)	Графовые БД (Neo4j)
Модель данных	Таблицы с записями	Узлы и связи
Запросы связей	Сложные JOIN-ы	Прямой обход связей
Производительность	Замедляется при глубине связей	Постоянная скорость
Обнаружение паттернов	Статистические аномалии	Структурные паттерны
Пример запроса	5-7 JOIN для связей 3-го уровня	MATCH path=(a)-[*3]-(b)

Алгоритмы обхода графа для обнаружения аномалий

Графовые базы данных представляют собой мощный инструмент для анализа структуры сети, особенно при использовании специализированных алгоритмов.

PageRank: выявление ключевых узлов

PageRank — алгоритм, изначально созданный для ранжирования веб-страниц, может быть эффективно применён для определения наиболее влиятельных или значимых узлов в сети финансовых транзакций. В данном контексте речь идёт не о популярности, а о центральности потока денежных средств.

Принцип работы алгоритма заключается в рассмотрении транзакций как своеобразных «голосов». Счёт, получающий большое количество транзакций от других счетов, особенно от тех, которые сами являются источниками множества транзакций, получает высокий рейтинг PageRank.

Использование в борьбе с мошенничеством.

Выявление счетов-концентраторов: мошенники часто создают промежуточные счета, через которые проходят средства от различных источников перед выводом. Такие счета будут иметь аномально высокий рейтинг PageRank.

Определение организаторов схемы: счета, координирующие деятельность в рамках схем, подобных «пирамидам», будут иметь высокий уровень

значимости. Пример запроса с использованием Neo4j GDS (Graph Data Science Library):

```
cypher
// Создание проекции графа для алгоритма
CALL gds.graph.project(
  'transactionGraph',
  ['Client', 'Account'],
  {
    SENT: {orientation: 'NATURAL'},
    RECEIVED: {orientation: 'REVERSE'}
  }
)
// Запуск алгоритма PageRank
CALL gds.pageRank.stream('transactionGraph')
YIELD nodeId, score
RETURN gds.util.asNode(nodeId).name AS name, score
ORDER BY score DESC LIMIT 10
```

Данный запрос позволит нам определить 10 счетов/клиентов с наиболее высоким показателем PageRank. Эти клиенты представляют интерес для более детального анализа.

Центральность по посредничеству (Betweenness Centrality)

Данный алгоритм позволяет оценить, насколько часто определённый узел встречается на кратчайших путях между другими узлами в графе. Узел с высокой посреднической центральностью выполняет функцию «моста» между различными частями сети.

Алгоритм работает следующим образом: он рассчитывает количество кратчайших путей между всеми парами узлов, проходящих через данный узел.

Применение алгоритма к мошенническим схемам

Данный алгоритм может быть полезен в следующих ситуациях, связанных с мошенническими схемами.

Обнаружение связующих узлов: В сложных мошеннических сетях, состоящих из отдельных кластеров, связующие узлы соединяют их между собой. Блокировка таких узлов может привести к разрушению всей сети. Выявление «слабых звеньев». Для правоохранительных органов такие узлы

являются ключевыми фигурами, через которые можно получить доступ к информации о всей сети.

Обнаружение сообществ (Community Detection). Алгоритмы обнаружения сообществ, такие как алгоритм Лувена, предназначены для группировки узлов (объектов) в кластеры или сообщества. Узлы в одном сообществе имеют более тесные и частые связи друг с другом, чем с узлами в других сообществах. Алгоритм работает в несколько этапов, объединяя узлы в сообщества с целью максимизации модулярности — метрики, которая оценивает качество разделения на группы. В контексте противодействия мошенничеству алгоритмы обнаружения сообществ играют ключевую роль в выявлении организованных групп и схем кругового перемещения средств.

В частности, алгоритмы могут автоматически выявлять группы счетов и клиентов, которые взаимодействуют друг с другом более интенсивно, чем с остальными клиентами. Это может свидетельствовать о наличии мошеннической схемы или картеля.

Кроме того, алгоритмы способны обнаруживать замкнутые циклы транзакций, в рамках которых денежные средства перемещаются между группой счетов, создавая видимость активности или позволяя обойти ограничения.

Рассмотрим практический пример: выявление схемы обналичивания денежных средств через несколько счетов.

В ходе анализа финансовой системы были выделены следующие ключевые элементы:

1. Счета-сборщики — множество счетов, принимающих небольшие переводы от клиентов.

2. Счета-агрегаторы — счета, которые собирают средства от счетов-сборщиков.

3. Основные счета-концентраторы — один или два основных счёта, получающие крупные суммы от счетов-агрегаторов.

4. Счета-обналичиватели — счета, оперативно выводящие денежные средства с основных счетов-концентраторов.

Анализ финансовых данных с использованием графов открывает новые перспективы в борьбе с мошенничеством. Он позволяет выявить ключевые элементы схемы — «концентраторы» и «агрегаторы» — как узлы с высоким уровнем влияния в сети.

С помощью алгоритмов обнаружения сообществ мы можем объединить этих лиц в единое сообщество, изолированное от законопослушных клиентов. Анализ путей в сети мгновенно визуализирует путь денежных средств от любого из «сборщиков» до «обналичивателей».

Применение графовых баз данных для борьбы с мошенничеством выходит за рамки внедрения новых технологий. Это скорее новый подход к анализу финансовых данных, который переосмысливает традиционный метод изучения отдельных элементов и связей между ними.

Данный подход позволяет выявлять не только подозрительные транзакции, но и целые схемы мошенничества. Алгоритмы, такие как PageRank, центральность и обнаружение сообществ, предоставляют мощный математический инструментарий для количественной оценки аномалий в структуре финансовых сетей.

Внедрение этих методов в системы мониторинга позволяет финансовым организациям перейти от реактивного подхода к борьбе с мошенничеством, который действует только после инцидента, к проактивному. Такой подход помогает прогнозировать и предотвращать преступления до того, как они нанесут значительный ущерб.

В современном мире, где связи имеют такое же значение, как и сами данные, графовые базы данных становятся стратегически важным инструментом для обеспечения финансовой безопасности.

Перспективы развития графовых технологий в области финансовой безопасности связаны с интеграцией методов искусственного интеллекта и машинного обучения для прогнозирования emerging-угроз и автоматизации принятия решений. Особый потенциал демонстрируют гибридные подходы, сочетающие графовый анализ с алгоритмами обработки естественного языка для выявления сложных паттернов в неструктурированных данных. Дальнейшие исследования должны быть направлены на разработку стандартизированных метрик оценки эффективности графовых методов и создание адаптивных систем, способных эволюционировать вместе с изменяющимися тактиками мошенников. Важным направлением является также обеспечение баланса между эффективностью обнаружения и соблюдением требований к защите персональных данных, что особенно актуально в условиях ужесточающегося регуляторного контроля. Внедрение графовых технологий должно сопровождаться разработкой комплексных методологий, включающих не только технические решения, но и

организационные меры по подготовке специалистов и оптимизации бизнес-процессов.

Список литературы

1. Робинсон Я., Веббер Д., Айфрем Э. Графовые базы данных. Новые технологии для работы с связанными данными. - 2-е изд. - Москва: Диалектика, 2023. - 382 с.

3. Внуков А.А. Защита информации в банковских системах: учеб. пособие для бакалавриата и магистратуры. - 2-е изд., перераб. и доп. - Москва: Юрайт, 2017. - 246 с.

4. Баланов А.Н. Комплексная информационная безопасность. Полный справочник специалиста: Практическое пособие. - Москва: Инфра-Инженерия, 2024. - 156 с.

5. Миллер Дж. Дж. Введение в анализ графов с использованием Neo4j: обнаружение мошенничества и анализ финансовых схем. - Москва: ДМК Пресс, 2021. - 254 с.

6. Внуков А.А. Защита информации в банковских системах: учеб. пособие для бакалавриата и магистратуры. - 2-е изд., перераб. и доп. - Москва: Юрайт, 2017. - 246 с.

References

1. Robinson Ya., Webber D., Efrem E. Graph databases. New technologies for working with related data. - 2nd ed. - Moscow: Dialectics, 2023. - 382 p.

3. Vnukov A.A. Information protection in banking systems: textbook. handbook for bachelor's and master's degrees. - 2nd ed., revised and add. - Moscow: Yurait, 2017. - 246 p.

4. Balanov A.N. Complex information security. The complete Specialist's Guide: A Practical Guide. - Moscow: Infra-Engineering, 2024. - 156 p.

5. Miller J. J. Introduction to graph analysis using Neo4j: Fraud Detection and Financial Scheme Analysis. - Moscow: DMK Press, 2021. - 254 p.

6. Vnukov A.A. Information protection in banking systems: textbook. handbook for bachelor's and master's degrees. - 2nd ed., revised and add. - Moscow: Yurait, 2017. - 246 p.