Propuesta para la Implementación de una Herramienta de Análisis y Visualización de Redes

Diego Sarceño

diego.sarceno@chn.com.gt
Compliance Quality Control
Gerencia de Cumplimiento

27 de enero de 2025

Índice

1. Introducción

En el contexto actual, las instituciones financieras enfrentan desafíos cada vez más complejos para monitorear actividades sospechosas que puedan estar relacionadas con fraude bancario, lavado de dinero o financiamiento al terrorismo. Estos riesgos no solo amenazan la seguridad financiera, sino que también comprometen la reputación del banco y el cumplimiento normativo.

El manejo de transacciones financieras genera un vasto volumen de datos interconectados que contienen patrones complejos. Entender y analizar estas conexiones es esencial para identificar actividades irregulares. La teoría de grafos, el aprendizaje automático y el deep learning en grafos ofrecen herramientas avanzadas para mapear relaciones, detectar anomalías y prever posibles fraudes.

Esta propuesta presenta un enfoque formal y detallado para desarrollar una herramienta que integre estas técnicas, permitiendo visualizar las interacciones transaccionales de clientes y detectar actividades sospechosas de manera efectiva.

2. Objetivos

2.1. General

Diseñar e implementar una herramienta inteligente que permita:

1. **Mapear** la información transaccional de los clientes del banco en un grafo, destacando conexiones relevantes entre cuentas, productos y personas asociadas.
2. **Detectar y clasificar** actividades sospechosas relacionadas con fraude bancario, lavado de dinero y financiamiento del terrorismo.
3. **Proveer visualizaciones interactivas** y alertas automáticas para facilitar el análisis y la toma de decisiones por parte de los analistas de riesgos.

2.2. Específicos

1. Modela transacciones financieras y relaciones legales entre clientes como un grafo dinámico.
2. Identificar comunidades y patrones de comportamiento utilizando algoritmos de detección de comunidades y clustering.
3. Detectar transacciones y relaciones anómalas mediante técnicas de clasificación y modelos de anomalías.
4. Aplicar redes neuronales en grafos para mejorar la precisión en la detección de actividades sospechosas.
5. Utilizar herramientas de visualización (o desarrollar una herramienta) para explorar relaciones y emitir alertas de riesgo.

3. Metodología

3.1. Modelado de los Datos

La base de la herramienta será un grafo que represente las relaciones entre los elementos clave.

- **Nodos:**
  - Clientes (individuales y jurídicos)
  - Cuentas y productos.
  - Representantes legales, notarios, accionistas, beneficiarios, etc.
- **Aristas:**
  - Transacciones
  - Relación de propiedad o representacion entre personas y entidades jurídicas.
  - Conexiones entre productos financieros asociados a un cliente.
- **Propiedades de los nodos y aristas:**
  - Monto y frecuencia de transacciones
  - Categoría del producto financiero
  - Relación temporal (historial de transacciones).

3.2. Análisis de Grafos y Redes Sociales

Para identificar patrones y relaciones críticas en el grafo, se aplicarán métodos de teoría de grafos y análisis de redes sociales.

1. **Detección de Comunidades:**
   - Louvain, Leiden, Label propagation
   - Identificar grupos de cuentas con conexiones densas, que podrían corresponder a esquemas sospechosos (por ejemplo, estructuras de lavado de dinero).
2. **Métricas de Red:**
   - Centralidad (degree, closeness, betweenness)
   - Conectividad y densidad de nodos.
   - Identificar nodos clave en posibles cadenas de transacciones sopechosas.
3. **Detección de subgráfos anómalos:**
   - Anomalous Subgraph Detection (ASD), OddBall.
   - Descubrir patrones transaccionales inusuales, como cuentas que actúan como intermediarios en flujos atípicos.

3.3. Algoritmo de Clustering y Clasificación

1. **Clustering no Supervisado:**
   - K-Means, DBSCAN, HDBSCAN
   - Agrupar transacciones según características como monto, frecuencia y destino.
2. **Combinacion de KNN y DNN:**
   - Combinar K-Nearest Neightbors con redes neuronales profundas para mejorar la detección.

3.4. Redes Neuronales en Grafos (GNNs)

1. **Modelos:**
    - GraphSAGE: generar embeddings de nodos basados en sus vecinos.
    - GCN (Graph Convolutional Networks): capturar relaciones complejas en el grafo.
    - GAT (Graph Attention Networks): priorizar relaciones más reelevantes.

3.5. Visualización y Sistema de Alertas

Para la visualización de los datos transformados en un grafo, existen varias opciones. Una de ellas es utilizar herramientas ya existentes como Power BI o Gephi, que ofrecen funcionalidades robustas para la visualización de grafos. Power BI, aunque generalmente orientado a la visualización de datos tabulares, tiene la capacidad de integrarse con bases de datos de grafos a través de conectores y permite crear dashboards interactivos. Por otro lado, Gephi es una herramienta específicamente diseñada para la visualización de redes, permitiendo crear representaciones interactivas de grafos de manera fácil, con una amplia gama de algoritmos de análisis de redes.

Sin embargo, también es posible desarrollar una herramienta personalizada simple que permita hacer visualizaciones interactivas mediante frameworks como Dash o Streamlit, los cuales son más flexibles y personalizables para crear aplicaciones web interactivas. Estas herramientas permiten integrar visualizaciones de grafos generadas con librerías como NetworkX o PyVis, y brindan una experiencia dinámica para explorar las relaciones transaccionales y detectar patrones sospechosos.

Además, al desarrollar una herramienta propia, se obtiene una mayor escalabilidad, ya que se puede adaptar y expandir conforme las necesidades del banco evolucionen, integrándose de manera más fluida con otras herramientas y sistemas existentes en el banco, ya sea dentro de la misma gerencia o en otras áreas relacionadas, como el monitoreo de riesgos o la gestión de productos financieros. Esta integración optimiza los flujos de trabajo y facilita el acceso a la información relevante, mejorando la eficiencia en la toma de decisiones y en los procesos de auditoría interna.

4. Utilizada y Aplicaciones

1. **Para analistas:**
    - Visualizar relaciones complejas y patrones anómalos de forma intuitiva.
    - Identificar de manera eficiente las cuentas y transacciones de mayor riesgo.
2. **Para el banco y autoridades regulatorias:**
    - Reducir riesgos.
    - Mejorar la reputación al demostrar un enfoque proactivo.
    - Generar reportes que cumplan con los estándares.

Bibliografía

[1] *Como usar Neo4j en la Prevención del Blanqueo de Capitales.*

[2] *Las relaciones, la "Materia Oscura" de sus Datos.*

[3] *ML y Grafos.*

[4] Alotibi, J., Almutanni, B., Alsubait, T., Alhakami, H., and Baz, A. (2022). Money laundering detection using machine learning and deep learning. *International Journal of Advanced Computer Science and Applications*, 13(10).

[5] Assumpção, H. S., Souza, F., Campos, L. L., de Castro Pires, V. T., de Almeida, P. M. L., and Murai, F. (2022). Delator: Money laundering detection via multi-task learning on large transaction graphs. In *2022 IEEE International Conference on Big Data (Big Data)*, pages 709–714. IEEE.

[6] Branco, B., Abreu, P., Gomes, A. S., Almeida, M. S., Ascensão, J. T., and Bizarro, P. (2020). Interleaved sequence rnns for fraud detection. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 3101–3109.

[Rzayeva and Malekzadeh] Rzayeva, D. and Malekzadeh, S. A combination of k-nearest neighbor and deep neural networks for credit card fraud detection.

[8] Tian, Y., Liu, G., Wang, J., and Zhou, M. (2023). Transaction fraud detection via an adaptive graph neural network. *arXiv preprint arXiv:2307.05633*.