Detección de anomalías basadas en grafos para detección de usuarios fraudulentos en comercio por Internet

Lorena Luraghi
Maestría en Ciencia de Datos
Facultad de Ingeniería, Universidad de la República
Montevideo, Uruguay
loreluraghi@gmail.com

I. Introducción

El desarrollo del sistema financiero junto con el avance de la tecnología ha convertido al comercio por internet como el sistema de compras más popular, consiste en la posibilidad de realizar compras desde un ordenador, sin interacción persona a persona y en la mayoría de los casos, a través del uso de tarjetas de crédito, sin la necesidad de manipular dinero en efectivo. Si bien la seguridad es una de las características que describe este tipo de pagos y la probabilidad de fraude es pequeña, el mismo implica un costo significativo para las empresas por el monto del fraude y también por la pérdida de reputación hacia los clientes, por lo tanto, son suficientes los motivos para intentar impedirlos.

Una de las principales casuísticas de fraude en este tipo de comercio, es el uso de numeración de tarjetas de crédito obtenida de manera ilícita para la compra de productos.

En este trabajo nos centraremos en las transacciones de un comercio de venta de ropa, el hecho de que se venda un producto físico, el cual se envía a una dirección detallada por el usuario es no menor, ya que luego de la compra existe una ventana de tiempo antes de que el producto sea enviado. Actualmente un equipo de analistas analiza las transacciones recibidas de forma manual encontrando usuarios vinculados a patrones fraudulentos y enviándolos a una lista negra dónde se cancela la compra si está pendiente de envío o de lo contrario, si ya se concretó, se bloquea para futuras compras.

Se dispone de una base de datos relacional con el registro de 156456 transacciones en la tienda del primer semestre de 2021. El objetivo del trabajo es, partiendo de la tabla disponible, construir un grafo que disponga la información personal de los usuarios que compran en esta tienda, de manera de localizar aquellos que estén interconectados y facilitar la búsqueda al equipo de analistas.

La elección de utilizar bases de datos de grafos para llevar a cabo el proyecto se debe a su flexibilidad para representar relaciones binarias entre pares de objetos, la cual se vuelve casi imposible de visualizar utilizando el formato tabla. También es un puntapié inicial a futuros proyectos que requieran una clasificación inmediata de la transacción (fraude, no fraude), en este posible caso el uso de bases de datos de grafos podría ser beneficioso dadas las características que las identifican: mayor rendimiento en comparación con las bases de datos relacionales y otras NoSQL y flexibilidad en la estructura incluso con grandes volúmenes de información. [Washington A. Velásquez Vargas(2020)]

El proyecto se divide en dos etapas, por un lado, se definirá para nuestra información la clasificación de la información planteada por [Tahereh Pourhabibi, Kok-Leong Ong, Booi H. Kam, Yee Ling Boo(2020)] que nos ayudará a la construcción del grafo que necesitamos para extrer la información valiosa para nuestro objetivo. Por otro lado, utilizaremos el paquete de Python Networkx[1] para localizar aquellas comunidades de usuarios que comparten información.

II. Trabajos relacionados

Este trabajo toma como guia el artículo "Fraud detection: A systematic literatura review of graph-based anomaly detection approachesrealizado por [Tahereh Pourhabibi, Kok-Leong Ong, Booi H. Kam, Yee Ling Boo(2020)] . El artículo seleccionado realiza una investigación de los enfoques de detección de anomalía basada en grafos (GBAD) que consideren interdependencias entre diferentes objetos de datos en un grafo para detectar actividades fraudulentas. El objetivo es comprender la evolución de estos métodos en la última generación (entre 2007 y 2018), los cuales han contribuido en los últimos años a analizar relaciones y patrones de conectividad de redes para identificar patrones inusuales vinculados al fraude.

Para llevar a cabo el objetivo han desarrollado un marco de clasificación de los artículos seleccionados para la revisión, se presentan los puntos a tener en cuenta a continuación:

- Disponibilidad de etiquetas: Nos determinan si estamos frente a un problema supervisado, no supervisado o semi-supervisado.
- Naturaleza de la red de entrada: Refiere a cómo se propaga la información dentro de la red (por ejemplo la dirección de los enlaces y el tiempo en el que se

[1] https://networkx.org/

establecen los mismos), las características de los nodos y si existen estructuras dentro de los enlaces y atributos.
- Tipos de anomalías: Los enfoques analizados se han diversificado para detectar distintos tipos de anomalías dentro de las redes, puede ser de interés detectar un nodo anómalo, bordes o subgrafos.
- El método que se utiliza para la detección de la anomalía: la elección del método depende de los puntos anteriormente mencionados. Por ejemplo si se quiere detectar un subgrafo, se desea encontrar subgrupos de nodos fuertemente conectados entre sí analizando sus interconexiones.
- Representación de la estructura: Es necesario definir las medidas que mejor pueden mapear una red.

III. Desarrollo

En esta sección se mostrará el desarrollo del proyecto. En primer lugar se presentará el conjunto de datos y los objetivos de acuerdo a la realidad planteada. Luego partiendo de la información disponible se construirá el grafo y por último, el experimento para la detección de anomalías.

III-A. Problemática y aspectos a considerar

Se trabajará en la detección de fraude de un sitio de internet de venta de ropa. Se cuenta con una base de datos relacional con información de 156456 compras del primer semestre del 2021. Dado que el objetivo es localizar redes de usuarios que estén utilizando información compartida no legítima, nos centraremos en los campos que contengan información personal del usuario dejando de lado las características propias de las compras (como el monto de la transacción). Los campos a tener en cuenta para la construcción del grafo son:

- User_mail: El correo electrónico con el cual el usuario se registra para realizar la compra.
- Hash: Numero de tarjeta de crédito tokenizado. Nos interesa determinar aquellas tarjetas de crédito que no pertenecen a quien está realizando la transacción.
- User_address: La dirección de envió es un dato clave para la detección del fraude, ya que es el aquel que nos conecta con el fraudulento, aunque se haya utilizado una tarjeta robada aquella persona que reciba el producto deberá tener relación con el fraude.

Algunos aspectos a considerar y problemas a la hora de construir el grafo:

- Por tratarse de un comercio con envíos, los cuales deben ser planificados luego de la compra, no es necesario construir un sistema de detección de fraude en tiempo real. En este caso se quiere localizar transacciones sospechosas para guiar y facilitar el trabajo que realiza un equipo de analistas para bloquear usuarios con comportamientos sospechosos e impedir las transacciones actuales o futuras compras.
- La dirección de envío es un campo *string* escrito por el usuario, por lo tanto para una mejor solución debería aplicarse algún procesamiento de texto o geolocalización para poder detectar cuando dos direcciones escritas de modo distinto identifican el mismo lugar.
- Los campos tarjeta y dirección de correo electrónico son identificadores 100 % confiables del usuario. Sin embargo la dirección debe ser analizada de manera particular ya que puede tratarse de direcciones populares como puede ser un lugar público (oficinas, edificios grandes, entre otras). Por lo tanto previo al análisis se realizó un análisis exploratorio de las direcciones, excluyendo 17 direcciones del grafo.

III-B. Construcción del grafo

Se utilizará el enfoque planteado por el artículo que guía este proyecto para la construcción del grafo en función de la información disponible y la naturaleza del problema. Repasaremos los items mencionados en la sección **??**

- Disponibilidad de etiquetas: En este caso no se tiene disponible una etiqueta que indique si una transacción del pasado fue fraudulenta o no. Por lo tanto, estamos frente a un problema de clasificación no supervisada.
- Naturaleza de la red de entrada: Nos encontramos frente a un grafo no dirigido, la dirección entre los atributos no es de interés. Dado que se utilizará un periodo corto de tiempo y que la información disponible es información personal que no suele ser modificada con alta frecuencia, el tiempo tampoco es un factor que importe.
- Tipos de anomalías: En este caso las anomalías que queremos detectar son comunidades que compartan información personal más allá de lo que puede ser un núcleo familiar.
- Método para la detección de la anomalía: Detectaremos aquellos subgrafos con mayor cantidad de interconexiones.

Partiendo de estos criterios sugeridos por el artículo [Tahereh Pourhabibi, Kok-Leong Ong, Booi H. Kam, Yee Ling Boo(2020)] se opta por construir un grafo no dirigido donde los nodos son la información personal y las aristas indican si esa información está relacionada o no. Dado que el objetivo es obtener una lista de usuarios interconectados entre sí, al momento de construir el grafo se le asignan atributos que pueden ser: 'hash', 'mail' o 'address' que posibilitan clasificar los nodos de la red que se encuentre.

IV. Experimentación

En la siguiente sección se muestra cómo se llevó a cabo la construcción del grafo principal y la detección de los subgrafos de usuarios con mayores interconexiones entre sí. Para la evaluación del método se analizó los primeros 10 subgrafos de forma manual para determinar si existe evidencia de fraude o no.

IV-A. Implementación del grafo

Se utilizó la librería Networkx [2] para llevar a cabo el grafo.

Cómo se menciono anteriormente, primero se definen los nodos del grafo y como se muestra en la tabla 1. Se utilizó

[2]https://networkx.org/

la función *add_nodes_from* para asignar los atributos 'hash', 'mail' y 'address' a cada uno de los nodos.

Listado 1 Creación de los nodos del grafo

```
import networkx as nx

G = nx.Graph()

#Construimos los nodos con su correspondiente atributo

nodos_hash = df_sin_na[['Hash']].drop_duplicates()
.apply(tuple, axis=1).values

nodos_address = df_sin_na[['User_address']]
.drop_duplicates().apply(tuple, axis=1).values

nodos_mail = df_sin_na[['User_mail']]
.drop_duplicates().apply(tuple, axis=1).values

G.add_nodes_from(nodos_hash, name = "hash")
G.add_nodes_from(nodos_address, name = "address")
G.add_nodes_from(nodos_mail, name = "mail")
```

Luego, mediante el uso de la función *add_edges_from* se crearon dos enlaces por transacción: el enlace que relaciona la tarjeta con la dirección de correo electrónico y el enlace que relaciona la tarjeta con la dirección de envío, tal como se muestra en la tabla 2

Listado 2 Creación de los enlaces del grafo

```

#La columna hash_address guarda las
#relaciones entre hash y address

df_sin_na['hash_addres']=df_sin_na[['Hash','User_address']]
.apply(tuple, axis=1)


#Idem para hash_mail

df_sin_na['hash_mail'] = df_sin_na[['Hash', 'User_mail']]
.apply(tuple, axis=1)


G.add_edges_from(df_sin_na['hash_address'].values)

G.add_edges_from(df_sin_na['hash_mail'].values)
```

Se obtiene entonces un grafo conformado por 222218 enlaces y 597892 nodos.

IV-B. Detección de subgrafos

Luego de que se tiene el grafo construido, se procede a la detección de las componentes con mayor cantidad de interconexiones, es decir, usuarios que comparten mayor cantidad de información personal como se muestra en la siguiente tabla 3.

Listado 3 Extracción de subgrafos más grandes

```

S=sorted(nx.connected_components(G),key=len,reverse=True)
```

Como se menciono anteriormente, es de esperar que cada usuario forme un subgrafo muy pequeño entre sus datos personales y (posiblemente) su núcleo familiar por medio de la dirección. Por lo tanto, nos enfocaremos en aquellos subgrafos anómalos en cantidad de interconexiones.

El cuadro I muestra el tamaño de los primeros 10 subgrafos de mayores en cuanto a cantidad de enlaces. Se evidencia la presencia de comunidades.

Número de subgrafo	Cantidad de enlaces
1	417
2	198
3	97
4	58
5	33
6	23
7	22
8	22
9	21
10	20

Cuadro I: Top 10 subgrafos

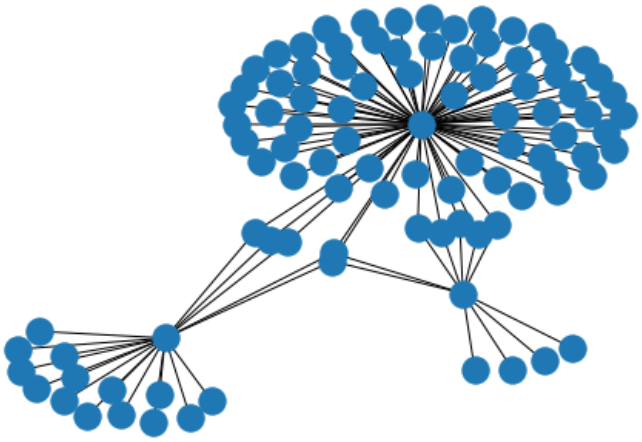

Figura 1: Grafo de tarjetas compartidas por gran cantidad de usuarios

IV-C. Resultados

Se presenta a continuación los resultados del análisis manual de la evidencia de fraude dentro de los subgrafos más grandes localizados en la red de usuarios. Se muestra la morfología de 3 de los 10 subgrafos analizados, dado que estamos trabajando con información personal y sensible, nos limitamos a no mostrar las etiquetas de cada uno de los grafos.

1. Se observó la presencia de comunidades que comparten información de uso personal dentro de las transacciones analizadas
2. Luego de analizar los grafos más grandes, se encontraron diversos patrones de fraude que, se detallan a continuación:

   Tarjetas compartidas por una cantidad muy grande de usuarios y direcciones, como se muestra en la figura 1 dónde los elementos centrales del grafo son atributos de tarjeta de crédito.

   Usuarios utilizando tarjetas compartidas y realizando modificaciones en el string de la dirección para

tratar de ser identificados como usuarios distintos. Por ejemplo, en el subgrafo que se muestra en la figura 2 las direcciones presentes son del siguiente estilo: *Calle1 102, 1er sector cale 1 102, nuevo amanecer cale1 102, casa calle 1 102, casa casa calle1 102*

En el subgrafo 3 se puede observar una nueva casuística de fraude, la presencia de patrones de fraude en los nombres de los correos electrónicos, por ejemplo: *camaronesenchiladitos, cabezahumada,churrosmieleros o elposhasdoblas*

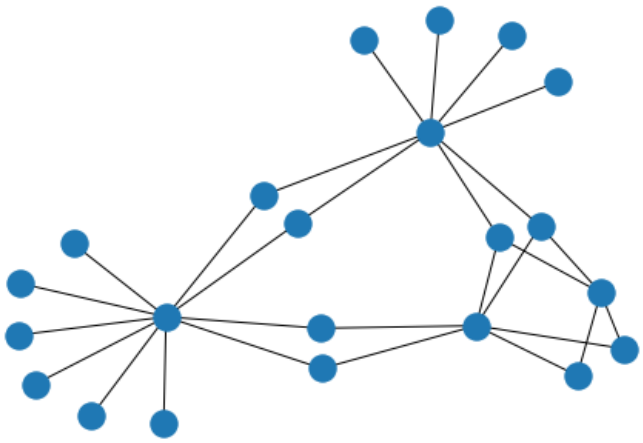

Figura 2: Grafo con tarjetas compartidas y direcciones falsas

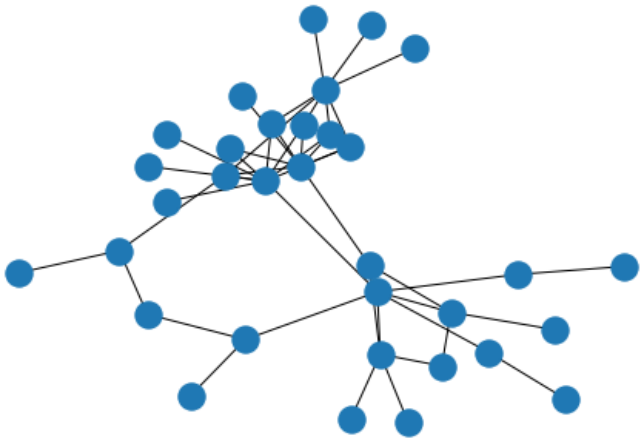

Figura 3: Grafo con patrones de mail

IV-D. Conclusiones y trabajo futuro

Se enumeran a continuación las siguientes conclusiones:

1. Las comunidades de usuarios que comparten información personal están relacionadas a casuísticas de fraude
2. El uso de bases de datos basadas en grafo nos permite identificar interconexiones independientemente del grado de las mismas de forma sencilla, lo cual con bases de datos relacionales es engorroso de ver y requiere de utilizar funciones como la función *join*, la cual puede ser costosa computacionalmente.
3. La detección de anomalías en forma de subconjuntos de gran tamaño de información compartida es útil para obtener una lista de usuarios negativos y tarjetas de origen dudoso.

Trabajos a futuro:

1. Etiquetar el fraude y utilizar la detección de anomalías en bases de datos de grafos para clasificar las nuevas transacciones.
2. Incluir el factor tiempo y realizar una clasificación del histórico transaccional que se pueda ir actualizando con las transacciones del día a día, sin tener que construir el grafo de todas las transacciones cada vez que se quiera obtener nueva información.

REFERENCIAS

[Tahereh Pourhabibi, Kok-Leong Ong, Booi H. Kam, Yee Ling Boo(2020)] Tahereh Pourhabibi, Kok-Leong Ong, Booi H. Kam, Yee Ling Boo. Fraud detection: A systematic literature review of graph-based anomaly detection approaches, 2020.

[Washington A. Velásquez Vargas(2020)] Washington A. Velásquez Vargas. Bases de datos orientadas a grafos y su enfoque en el mundo real, 2020.