USO DE UNA RED NEURONAL EN LA DETECCIÓN DE TRANSACCIONES FRAUDULENTAS REALIZADAS SOBRE UNA PLATAFORMA EN LÍNEA

USE OF A NEURAL NETWORK IN THE DETECTION OF FRAUDULENT TRANSACTIONS CARRIED OUT ON AN ONLINE PLATFORM

Guillermo De Ita Luna
Diego Saldaña Ulloa*

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

https://orcid.org/0009-0004-1428-6625
https://orcid.org/0000-0001-7948-8253

Año 10 No. 28
Recibido: 31/mayo/2023
Aprobado: 30/noviembre/2023
Publicado: 30/enero/2024

Facultad de Ciencias de la Computación, Edif. CCO1 – 14 Sur y
Av. Sn. Claudio, C.U.
Doctorado en Ingeniería del lenguaje y del conocimiento
Benemérita Universidad Autónoma de Puebla
guillermo.deita@correo.buap.mx
diego.ulloa13@hotmail.com

Resumen

El problema de los fraudes cibernéticos ha ido en aumento y es ya una problemática económica para las empresas que usan pagos electrónicos. Se han propuesto modelos y algoritmos dentro del área del aprendizaje automático con la finalidad de detectar patrones en las transacciones digitales que pudiesen exhibir las transacciones fraudulentas. Explicamos aquí una propuesta del uso de redes neuronales que usan estructuras de grafos para modelar y realizar la clasificación de usuarios fraudulentos.

Palabras clave: Aprendizaje automático, red neuronal de grafos, detección de fraude, transacciones en línea.

ABSTRACT

The problem of cyber fraud has been increasing and is already an economic problem for companies that use electronic payments. Models and algorithms have been proposed within machine learning to detect patterns in digital transactions that could exhibit fraudulent transactions. Here, we suggest using neural networks that use graph structures to model and classify fraudulent users

Keywords: Machine learning, Graph neural network, Fraud detection, Transactions on line.

Introducción

En un proceso antagónico, los defraudadores continúan buscando diferentes formas de allegarse de recursos mal habidos, mientras que el sistema legal intenta reconocer y proteger al público de las transacciones fraudulentas.

Con el advenimiento de las compras por internet, surgió todo un sistema de defraudación digital donde se aplican diferentes mecanismos. Desde el engaño vía páginas web falsas, suplantaciones de identidad donde se espera que el usuario digital sea quien realiza las primeras transacciones vía el engaño. Hasta los fraudes donde el usuario ya no es el que realiza las primeras transacciones que llevan al fraude, dado que, más bien, sufrió del robo de su información personal sin que participara de forma directa en el proceso de robo de su información.

En este último caso, existe una red oscura donde se comercializan datos para aquellos cibercriminales que buscan acceder a datos confidenciales del usuario digital. O bien, grupos de cibercriminales, o hackers solitarios, que se quieren allegar de aquella información crítica de usuarios para realizar fraudes cibernéticos.

En este artículo trataremos solo de uno de este tipo de sistemas fraudulentos, y comentaremos algunas de las estrategias digitales que se están usando en busca de reconocer y detener los movimientos digitales fraudulentos. Habremos de comentar, que la propuesta presentada es parte de los trabajos en la aplicación de algoritmos de aprendizaje que se están realizando recientemente en la Facultad de Ciencias de la Computación (FCC) de la Benemérita Universidad Autónoma de Puebla. Trataremos en este artículo de las llamadas transacciones monetarias digitales fraudulentas. Las transacciones digitales que se realizan a través de la banca en línea, las compras por internet y las transferencias monetarias vía sistemas de pago digitales son las principales fuentes del intento de cibercriminales por realizar transacciones falsas que les reditué en ganancia monetaria.

Con el acceso a las apps proporcionadas por las principales instituciones bancarias, las apps de la nueva industria llamada Fintech (empresas relativamente nuevas que utilizan la tecnología digital para brindar servicios financieros), o incluso cualquier plataforma digital en línea, se ha incrementado la industria de los fraudes en las transacciones digitales. El fraude transaccional en las plataformas digitales puede resultar del uso no autorizado de tarjetas bancarias y del acceso a cuentas de usuarios para realizar transacciones no autorizadas. Estos procesos generalmente se originan fuera de las plataformas digitales, mediante la venta de información bancaria (producto de hackeos a estas entidades) o en foros ilegales como en la red oscura.

Otro sector que también participa en este tipo de transacciones es el comercio online, es decir, empresas dedicadas a prestar un servicio no financiero que utiliza métodos de pago digitales. Este fenómeno afecta no solo a comercios o entidades financieras, sino también a usuarios de plataformas digitales debido a la vulnerabilidad de la información o situaciones en las que el dinero robado tiene que ser reintegrado por el mismo usuario. Las pérdidas económicas globales acumuladas por esta situación fueron de 38 mil millones de dólares americanos para el cierre de 2023. Además, las tendencias indican que para 2028, la cantidad aumentará a

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

362 mil millones de dólares americanos (Malone, 2023).

En el caso de fraude debido al acceso no autorizado a cuentas de usuarios, las causas pueden ser diversas. Desde vulnerabilidades informáticas en el lado comercial que se aprovechan para extraer información de acceso a cuentas, robo de dispositivos de usuario (que contienen acceso a cuentas digitales), hasta programas maliciosos que infectan los dispositivos de los usuarios para extraer información personal. Una vez que el acceso a una cuenta se ha visto comprometido, existe el riesgo de que se realicen transacciones fraudulentas.

Actualmente, las empresas absorben la mayor parte de las pérdidas monetarias a medida que los bancos les transfieren esta deuda. Si no se aborda el problema, la empresa puede adquirir una gran cantidad de deuda debido a este tipo de transacciones. Además, la reputación empresarial se ve afectada en diferentes niveles, desde la perspectiva social hasta los mecanismos implementados por los bancos que perjudican la tasa de aceptación de las transacciones comerciales.
El fenómeno de las transacciones fraudulentas ha provocado que las entidades que utilizan pagos online establezcan medidas para combatir este comportamiento. Entre las herramientas digitales que han permitido tener resultados sobresalientes en la detección y prevención de transacciones fraudulentas, ha sido el modelado matemático de las transacciones digitales, por ejemplo, usando grafos, y el uso de algoritmos de aprendizaje automático para reconocer patrones en los modelos gráficos que caracterizan transacciones fraudulentas.

Grafos

El origen de la palabra grafo es griego y su significado etimológico es 'trazar'. Un grafo se refiere a un conjunto de entidades (aristas y vértices) que puede ser utilizado para modelar relaciones entre ellos. Los grafos tienen utilidad al analizar problemas de diversa índole (ciencias sociales, química, física, biología, etc.), en donde aristas y vértices adquieren diferentes significados dependiendo del área de estudio.
Un grafo puede considerarse como un objeto geométrico, aunque en realidad es un objeto combinatorio, es decir, se conforma por un conjunto de puntos (vértices) y un conjunto de líneas (aristas) que conectan a esos puntos.

Formalmente, un grafo se define como un par ordenado formado de vértices y aristas, en donde ambos son un conjunto de elementos numerables. El tamaño del grafo se determina de acuerdo al número de vértices. Las aristas expresan relaciones entre vértices.

Un grafo puede ser dirigido si las aristas poseen una dirección (las aristas conectan un vértice fuente con un vértice destino). Un grafo no dirigido corresponde al caso donde no existe distinción en el orden de conexión. Por otro lado, se pueden obtener particiones de los vértices y aristas que forman un grafo de tal forma que estos elementos forman una representación más pequeña del grafo original. Esto es llamado un subgrafo y es un concepto útil, ya que en ocasiones se necesita operar solo sobre subgrafos de un grafo principal.

Los grafos tienen dos tipos de representación tomando en cuenta la información para construir el grafo, esto es, mediante una matriz de adyacencia o una lista de adyacencia. La Figura 1 muestra

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

un ejemplo de ambos tipos de representación. La matriz de adyacencia es una tabla en donde las filas y columnas dan como referencia la conexión entre vértices de un grafo. Por otro lado, las listas de adyacencia, como su nombre lo indica, son listas que almacenan los vértices adyacentes unos de otros. Ambos tipos de representación pueden ser útiles al momento de implementar algoritmos basados en grafos.

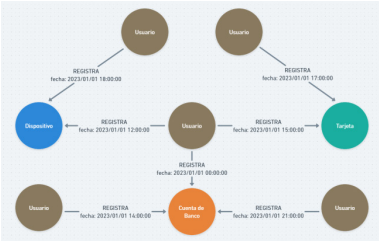


Figura 1. Ejemplo esquemático de un grafo temporal (TG) con la información que consideramos. Para este trabajo, siete diferentes tipos de grafos se construyeron. Aquellos con solo un tipo de evento de interacción (por ejemplo registro de tarjeta) contienen como tipos de vértices a Usuarios y Tarjetas. En esta figura el grafo corresponde a un registro de tarjeta-dispositivo-cuenta de banco.

Los grafos permiten representar objetos matemáticos y sus relaciones, por ejemplo, representar relaciones binarias, la topología de redes de carreteras o de enlaces ferroviarios, redes aéreas, o la red eléctrica de una localidad. Los grafos han permitido modelar, visualizar y analizar problemas de diferentes índoles. En nuestro caso, hemos utilizado la estructura de grafos para representar transacciones digitales que se realizan a través de una plataforma online. Por ejemplo, en nuestro modelo de grafos, los vértices representan; usuarios, dispositivos, direcciones de IP, tarjetas y cuentas bancarias.

Normalmente se asocian etiquetas para identificar cada uno de los vértices y aristas en un grafo. Pero además, dependiendo de la aplicación, es posible asociar a cada uno de los vértices del grafo un vector de características que representa el estado actual de ese vértice. Similarmente, para cada una de las aristas del grafo se le puede asociar también un vector de características. Por ejemplo, asociar a cada arista información estructural sobre la transacción representada por la arista. Así, los grafos son estructuras de datos no lineales que tienen una naturaleza generalmente dinámica.

En la Figura 2 se ilustra una vista parcial del grafo que modela las transacciones digitales que pueden realizarse en una plataforma de pago en línea.

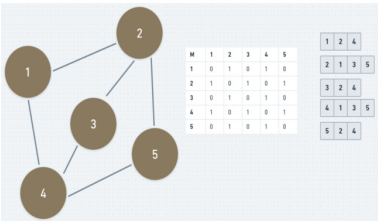


Figura 2. Matriz de adyacencia (izquierda en blanco) y lista de adyacencia (derecha en gris) de un grafo.

En nuestra aplicación, se desea estudiar la topología de cada una de las transacciones fraudulentas. Es por esto que necesitamos usar el concepto de vecindad de un vértice v que se define, para grafos dirigidos, como todos los vértices adyacentes al vértice v. Diremos que un grafo es bipartito si puede dividirse en dos subconjuntos de forma que cada arista tiene a sus vértices extremos en un subconjunto distinto.

Además de lo anterior, los grafos pueden dividirse dependiendo de las características que los definen. Por ejemplo, si el grafo posee un conjunto de diferentes tipos de vértices y diferentes tipos de aristas, es llamado un grafo heterogéneo. En contraposición,

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

grafos con un único tipo de vértice y arista se definen como grafos homogéneos. Dentro de los grafos heterogéneos podemos encontrar a los grafos bipartitos mencionados anteriormente. En general los grafos heterogéneos tienen un gran número de aplicaciones en la vida cotidiana, desde interacciones sociales para modelar fenómenos económicos, sistemas de recomendación de compras en línea, hasta el mismo proceso de detección de fraude tratado en este trabajo.

De manera adicional a la clasificación anterior, los grafos también pueden ser del tipo cuyos elementos están asociados a marcas temporales, es decir, existe una secuencia temporal sobre la forma en que vértices y aristas se conectan entre sí. Este tipo de grafos son llamados grafos dinámicos. Corresponden a un tipo más general de grafos, a su vez, los grafos dinámicos pueden ser homogéneos o heterogéneos.

Redes Neuronales

Las redes neuronales son un conjunto de algoritmos que tienen como finalidad el aproximar alguna función, tomando en consideración una serie de datos de entrada, de tal modo que con base a esa información se identifique que función podría genéralos. Esta es una forma sencilla de entender a una red neuronal, sin embargo, de manera histórica el término se concibió como una analogía al funcionamiento de las neuronas, formando redes y conexiones entre sí para el intercambio de información.

De manera típica, una red neuronal está formada por el encadenamiento de múltiples funciones, de ahí el por qué recibe el nombre de red. Cada una de estas funciones es llamada capa de la red y son comúnmente conocidas como redes neuronales multicapas. El término "neuronal", en analogía a las neuronas del cerebro, se debe a que cada uno de los datos de entrada se asocia al concepto de neurona. De esta forma, las neuronas intercambian esta información mediante las diferentes capas (funciones) de esta red.

Para poder obtener la aproximación de esta función, las redes neuronales necesitan de un conjunto de datos de los cuales aprender o extraer información. El método mediante el cual una red neuronal lleva a cabo este proceso se denomina entrenamiento. Durante el entrenamiento, los parámetros de las funciones que forman parte de la red se inicializan y los datos de entrada se procesan por cada una de estas funciones hasta obtener un resultado. Este resultado es comparado con el resultado real, de tal modo que se puede obtener un error o diferencia entre ambos. El algoritmo continúa mediante el método llamado propagación hacia atrás, en donde los parámetros de cada función se modifican tomando en consideración la retroalimentación debida a los resultados predichos y los resultados reales. Esta retroalimentación hace que durante la siguiente iteración del algoritmo, los resultados predichos se vayan acercando cada vez más a los resultados reales.

La forma específica en que están organizadas la secuencia de operaciones en una red neuronal se denomina arquitectura. Existen diferentes tipos de arquitecturas de redes neuronales, sin embargo, existen algunas que están plenamente identificadas y que se ha probado que funcionan para procesar la información de diferentes fuentes (por ejemplo texto, imágenes o datos estructurados como tablas).

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

Se pueden clasificar las redes neuronales de acuerdo a su arquitectura. Por ejemplo, las redes neuronales convolucionales (LeCun, Bengio, Hinton, 2015) generalmente se utilizan en tareas que implican imágenes. Esto es debido a que las convoluciones (operaciones de transformación) ayudan a hacer más eficiente el procesamiento y a su vez generalizan conceptos como la dependencia de los datos sobre su vecindario. En el caso de las imágenes los datos de entrada corresponden a los valores de los píxeles, entonces existe una dependencia de unos píxeles respecto de la información de sus píxeles vecinos). También se tienen las redes neuronales formadas por más de una capa, estas son llamadas redes neuronales profundas. Este tipo de redes son las de interés para el área de aprendizaje profundo.

Otro tipo de arquitectura de red neuronal son las llamadas redes neuronales recurrentes (Jordan, 1986; Rumelhart, Hinton, Ronald, 1985; Hochreiter, Schmidhuber, 1997) que están enfocadas en el procesamiento de datos secuenciales, por ejemplo, texto o información estructurada de manera histórica (clima, información económica, etc.). Las redes neuronales recurrentes son llamadas de esta manera porque aplican operaciones de manera recursiva sobre sí misma, de esta manera también son útiles para tener un tipo de contexto sobre los datos de entrada. Por ejemplo, en tareas de texto donde el contexto de una frase u oración son de suma importancia, de igual manera en tareas de pronósticos temporales. Los transformadores son otro tipo de arquitectura basada en mecanismos de atención (Vaswani, et al, 2017). Generalmente este tipo de arquitecturas trabajan sobre datos de texto. Los mecanismos de atención ayudan a mejorar el proceso del contexto de la información ya que otorgan mayor o menor importancia a las secuencias de información.

De esta forma, las redes neuronales se pueden aplicar para diferentes propósitos que van desde la clasificación de imágenes, detección de rostros, sistemas de recomendación, generación o clasificación de texto, entre otras aplicaciones. En nuestro caso, nos interesa aplicar las redes neuronales para realizar un proceso de clasificación automática de transacciones fraudulentas, esto mediante el uso de información relacionada al comportamiento de un usuario.

Por ejemplo, un enfoque sencillo consiste en utilizar algoritmos de machine learning convencionales (como árboles de decisión o redes neuronales), obtener características extraídas de un grafo y combinarlas con otro tipo de características relacionadas a transacciones. Sin embargo, a pesar de utilizar información extraída directamente del grafo, este enfoque falla en considerar la información estructural del contexto del vecindario de vértices y aristas, es decir, a qué tipos de vértices o aristas están conectados unos con otros.

Para la detección de transacciones fraudulentas, entre las técnicas del área de aprendizaje automático que han obtenido mejores resultados, se encuentran las redes neuronales que usan estructuras de grafos (Zhao, Fu, Wu, Li, & Li, 2019) para la detección de transacciones fraudulentas dentro de un sistema de grafos que modele transacciones digitales.

El aplicar algoritmos de aprendizaje tiene la intención de detectar particularidades, que en el área de aprendizaje automático se le llama: 'reconocer patrones'. La idea es identificar qué pa-

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

trones pueden exhibir las transacciones fraudulentas. Así, por ejemplo, a través del análisis de todo tipo de transacciones, se quiere características como; tipos de cuentas, tipos de compras, tipos de usuarios, etc. Que son los que pueden llegar a tener relación como parte de un proceso fraudulento.
El análisis incluye comparar transacciones válidas versus fraudulentas e identificar los patrones más comunes que tienen estas últimas. Este análisis no es sólo de tipo estadístico, sino precisamente los algoritmos de aprendizaje permiten correlacionar diferentes atributos asociados a las transacciones con la clase de fraudulentas, en búsqueda de los patrones subyacentes en este tipo de transacciones.

Al considerar un enfoque de aprendizaje automático con estructuras de grafos para la detección de la clase fraude, se deben tomar en cuenta consideraciones específicas relacionadas con la heterogeneidad del grafo y su evolución en el tiempo. El enmascaramiento del comportamiento fraudulento como comportamiento a reconocer, el método de entrenamiento (ya que en aplicaciones reales los grafos son masivos) y el considerar el problema de la baja disponibilidad de observaciones fraudulentas en comparación con las observaciones normales (clases no balanceadas)

Redes Neuronales de Grafos

Los tipos de vértices que se pueden utilizar para un proceso de detección de fraude, dependen de la plataforma, pero generalmente corresponden a usuarios, tarjetas, dispositivos o direcciones de IP. Los tipos de aristas pueden ser aquellas que representen una interacción con alguna tarjeta, dispositivo, dirección IP o inclusive transacciones monetarias. De manera adicional, cada vértice y arista puede ser representado por vectores de características relacionados con el comportamiento transaccional de las entidades involucradas (como los usuarios) dentro de la plataforma.

La intención de combinar grafos con redes neuronales es encontrar los patrones que los estafadores pudiesen exhibir (entre estos patrones, las semejanzas que se pudiesen dar de acuerdo a la topología del grafo que modela las transacciones), con el propósito de hacer identificable las transacciones fraudulentas. A pesar de esto, el comportamiento fraudulento cambia con el tiempo y hay situaciones en las que los estafadores logran mimetizarse como buenos usuarios (camuflaje) (Zexuan, Guodong, Yang, Wei, Bailing, 2022).

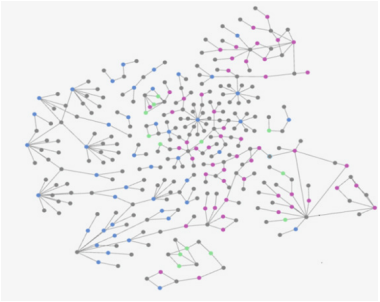


Figura 3. Subgrafo que contiene Tarjetas-Dispositivos-Cuentas de Banco. Los vértices grises representan a Usuarios, los vértices azules representan Dispositivos, los vértices verdes representan Tarjetas y los vértices púrpuras corresponden a Cuentas de Banco.

Para identificar transacciones fraudulentas se puede utilizar las denominadas Redes Neuronales de Grafos (Graph Neural Networks - GNN), utilizando la información de la topología del grafo y combinándola con redes neuronales para obtener una herramienta de detección de fraude (transacciones y usuarios). Esta tarea se puede realizar teniendo en cuenta algunas consideraciones importantes como la estructura dinámica del grafo, el problema del

camuflaje del fraude, la baja disponibilidad de eventos fraudulentos (en comparación con los eventos regulares), la perspectiva heterogénea del grafo y la estrategia de muestreo a seguir durante el entrenamiento del algoritmo (debido al uso de datos reales, es decir, un grafo masivo), así como a la dependencia temporal existente.

Por otro lado, un grafo puede modelarse mediante un enfoque estático o dinámico. Si los datos contienen marcas de tiempo que representan el momento en que una arista conecta dos vértices, y esto se encuentra asociado a un evento de interacción (por ejemplo, en una plataforma en línea; un evento puede ser el momento en que un usuario se hace amigo de otro usuario o un usuario realiza un pago a otro usuario) entonces la información puede modelarse mediante un Grafo Temporal Basado en Eventos (ETG por sus siglas en inglés). Este tipo de representación tiene la ventaja de incorporar el parámetro temporal como un parámetro adicional que puede resultar útil en un proceso de detección, como es el fraude.

El uso de una red neuronal de grafos temporales (Temporal Graph Neural Network - TGNN) con datos de un ETG enfocado a prevención de fraude, puede realizarse de manera directa mediante la clasificación de cuentas de usuarios fraudulentos. En este caso, el algoritmo se alimenta de todos los tipos de eventos de interacción disponibles. Las marcas de tiempo de múltiples eventos, incluidos los mencionados anteriormente son almacenadas por las plataformas en línea en diferentes bases de datos. Una forma típica al trabajar con los algoritmos de aprendizaje, en este caso con una ETG, es dividir el conjunto de transacciones en subconjuntos para realizar el entrenamiento, la validación y el conjunto de pruebas. Por ejemplo, se puede dividir el conjunto de transacciones mediante una partición cronológica; 70%, 15%, 15%, esto es, el conjunto de entrenamiento corresponde al 70% inicial de los datos en orden cronológico, el siguiente 15% corresponde al conjunto de validación de igual forma en orden cronológico y el 15% restante al conjunto de prueba. Durante cada iteración del proceso de entrenamiento, el algoritmo toma una muestra de los datos para realizar el proceso de manera más óptima, tal y como se propone en (Hamilton, Ying, Leskovec, 2017).

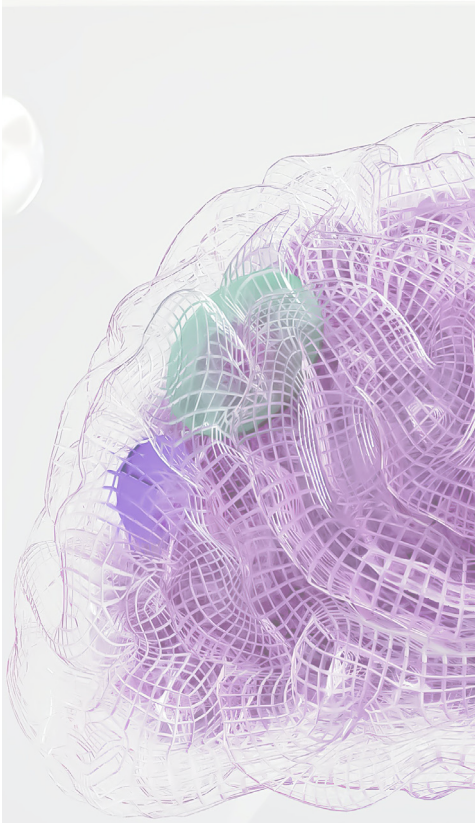
La forma más común para construir un ETG es mediante listas de adyacencia ya que mediante este tipo de estructura es más óptimo realizar el proceso de entrenamiento de una TGNN. De manera adicional, la estructura debe incluir las marcas de tiempo de cada evento de interacción en las duplas que representan los vértices adyacentes. El proceso para la construcción del algoritmo TGNN puede realizarse en cualquier lenguaje de programación, sin embargo, en la actualidad el área de redes neuronales en su mayoría se aborda con el lenguaje Python e implementaciones que permiten el trabajo sobre operaciones matriciales, tal como torch o tensor flow.

En general, los algoritmos de GNN hacen uso de la información de la red local mediante el 'paso de mensaje', que es un tipo de implementación que permite compartir información hacia vértices y aristas vecinas respecto de un vértice objetivo. En una TGNN este procesamiento se combina con una serie de módulos (memoria, agregación y actualización) que permiten obtener información del contexto temporal de un vértice. La memoria es una lista en donde se guarda la información del contexto temporal local de un vértice.

RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

Durante cada iteración del algoritmo, se agrega información del vecindario local de un vértice tomando en consideración la memoria. Esta información es agregada y sintetizada de manera conjunta y posteriormente actualizada de manera recursiva en la memoria. De esta forma el entrenamiento de una TGNN toma en consideración siempre la información temporal del grafo.

Con estas consideraciones se tomaron los datos provenientes de una plataforma de pagos en línea y se construyeron diferentes ETG para eventos como el registro de tarjetas, dispositivos y cuentas de banco. Adicionalmente, se formaron diferentes combinaciones de los eventos anteriores con la finalidad de procesar un algoritmo TGNN en cada uno de estos grafos. La idea se centró en averiguar si la incorporación de diferentes eventos ayudaba al proceso de clasificación de usuarios fraudulentos. Los resultados corroboraron esta información debido a que la incorporación de más eventos ayuda a tener más información estructural que ayuda a diferenciar entre usuarios fraudulentos y usuarios normales.



RD-ICUAP Año 10, No. 28, 2024, pp. 192-203 ISSN 2448-5829

Conclusiones

En este artículo se aborda cómo modelar las transacciones digitales de una plataforma de pagos en línea mediante una Red Neuronal de Grafos Temporal (Temporal Graph Network - TGN). La TGN considera un conjunto de eventos de interacción que representan el registro de tarjetas, dispositivos y cuentas de banco por parte de los usuarios, es decir, tres tipos de aristas fueron tomadas en cuenta.

Mediante el uso de la TGN se utilizó un algoritmo de TGNN, tomando como base el algoritmo propuesto en (Rossi, et al, 2020), para realizar un proceso de clasificación de usuarios fraudulentos en una plataforma de pagos en línea. Con los eventos mencionados anteriormente, se pudo formar diferentes combinaciones de TGN que incorporan eventos y sus tiempos de acción, esto con la finalidad de determinar si la cantidad de eventos era determinante para los resultados de la clasificación.

Debido a que las clases no se encontraban balanceadas (usuarios fraudulentos y usuarios normales) se incorporó una ponderación sobre cada una de las clases con el objetivo de otorgar mayor importancia sobre la clase con menor número de observaciones. Los resultados arrojaron que la información estructural y temporal de la combinación de diferentes tipos de eventos logra mejorar el proceso de clasificación.

Declaración de privacidad

Los datos personales facilitados por los autores a RD-ICUAP se usarán exclusivamente para los fines declarados por la misma, no estando disponibles para ningún otro propósito ni proporcionados a terceros.

Declaración de no Conflicto de intereses

Los autores declaran que no existe conflicto de interés alguno

Agradecimientos

Agradecemos a Moneypool por los datos proporcionados para la realización de este artículo.

REFERENCIAS

Hamilton WL, Ying R, Leskovec J (2017) Inductive representation learning on large graphs. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates Inc., Red Hook, NY, USA, NIPS 17, p 1025–1035

Hochreiter, S., & Schmidhuber, J. (1997) Long Short-Term Memory. Neural Comput. 9, 8, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jordan, Michael I. (May 1986). Serial order: a parallel distributed processing approach. Tech. rep. ICS 8604. San Diego, California: Institute for Cognitive Science, University of California.

LeCun, Y., Bengio, Y. & Hinton, G. (2015) Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Malone, C. (2023). Online payment fraud: Market forecasts, emerging threats and segment analysis 2023-2028. Accessed: 2023-06-08 12:00 CST. Retrieved from https://www.juniperresearch.com/researchstore/fintechpayments/online%-5C%5C-payment-fraudresearch-report

Rossi E, Chamberlain B, Frasca F, et al (2020) Temporal graph networks for deep learning on dynamic graphs. CoRR abs/2006.10637. URL https://arxiv.org/abs/2006.10637, 2006.10637

Rumelhart, David E; Hinton, Geoffrey E, & Williams, Ronald J (Sept. 1985). Learning internal representations by error propagation. Tech. rep. ICS 8504. San Diego, California: Institute for Cognitive Science, University of California.

Zexuan, D., Guodong, X., Yang, L., Wei, W., & Bailing, W. (2022). Contrastive graph neural network-based camouflaged fraud detector. Information Sciences, 618, 39–52. doi:10.1016/j.ins.2022.10.072
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017) Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 6000–6010.

Zhao, P., Fu, X., Wu, W., Li, D., & Li, J. (2019). Network-based feature extraction method for fraud detection via label propagation. 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), 1–6.