XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

ISBN

Detección de fraude en transacciones Blockchain usando procesos de Machine Learning, una Aproximación al Estado del Arte

Raúl Ocaña, Isaac Agudo, Javier López
Network, Information and Computer Security (NICS) Lab,
Universidad de Málaga, 29071
{roa, isaac, javier}@lcc.uma.es.

Las aplicaciones financieras basadas en tecnología blockchain son cada vez más comunes en el día a día de la economía regulada global. Con este precedente, y en una contextualización de la detección de fraude mediante el uso de técnicas de Machine Learning, el siguiente artículo de investigación en curso presenta una revisión sistemática de la literatura en la que se intenta dar respuesta a cuáles son las técnicas de detección más usadas actualmente y sus rendimientos, así como cuál es –si existe– la tendencia de uso de nuevas técnicas en este mismo contexto.

***Palabras Clave*—machine learning, blokchain, fraude, revisión, estado del arte**

I. INTRODUCCIÓN

La tecnología blockchain nació como la base para dar soporte al desarrollo de redes que permitiesen el intercambio de activos digitales. Esta era la idea original propuesta en el *white paper* de Bitcoin [1], y en un amplio sentido, se mantiene hasta la fecha.

Una de las características más innatas de las redes de cadena de bloques es su naturaleza transaccional. En ella se basan nuevas funcionalidades que han ido surgiendo en estos protocolos, como podrían ser la tokenización, la ejecución de un código remoto en formato *smart-contract*, o el registro de documentos digitales. Como resultado de esta evolución, las diferentes redes en activo están viendo crecer su volumen de operaciones y capitalización, alcanzando actualmente el trillón de dolares americanos [21].

Ante este tipo de situaciones, gobiernos e instituciones gubernamentales están tomando iniciativas de regulación para estas redes, las cuales están en proceso de tomar efecto o lo harán en periodos de tiempo cercano. Ejemplos de esto pueden ser la Ley MiCA [2], ratificada el 20 de abril de 2023 en el Parlamento Europeo; o el preaviso del Ministerio de Hacienda Español de la obligatoriedad de presentar registro de las operaciones realizadas en las plataformas de criptodivisas a partir de 2024, expuesta en el BOE Num. 81 [3], también de abril de 2023.

Con el objetivo de aportar valor al ecosistema investigador, y centrándonos en el estudio de la literatura hasta la fecha, se propone el desarrollo de una revisión sistemática basada en la metodología de Aproximación al Estado del Arte [4], mediante la cual se pretende dar respuesta a dos preguntas de investigación propuestas, las cuales indagan en las técnicas de Machine Learning más usadas, así como la evolución sufrida por las mismas en el contexto de la resolución del problema en la detección de fraude en transacciones blockchain.

II. METODOLOGÍA

La metodología de revisión seleccionada es la de Aproximación al Estado del Arte, que presentada por Gómez Vargas et al. [4], propone un proceso metodológico dividido en cinco fases: Indagación, Identificación, Selección, Clasificación, y Análisis. Como resultado, y esta es una de las principales diferencias con otras metodologías de revisión, se espera producir una recuperación y descripción del estado actual de la literatura, con la finalidad de transcender reflexivamente, generando una crítica en el caso de ser correspondiente.

La sección de Indagación será la encargada de dar contextualización a la temática a tratar, generando en el lector e investigador los conceptos necesarios para la correcta y objetiva evaluación de la literatura que será revisada posteriormente.

En segundo lugar, encontramos la fase de Identificación, en la que se definirán los términos PICOC, así como las preguntas de investigación y la cláusula de búsqueda que será utilizada en las bases de datos seleccionadas.

Además, continuaremos con parte de este proceso prerevisión en la tercera etapa, la Selección. En ella se propondrán los criterios de selección y filtrado, así como

se generará el cuestionario de evaluación de calidad y el de extracción de datos.

En cuarto lugar, procederemos con la etapa de Clasificación, en la que se realizará la importación de los resultados obtenidos, la evaluación de los mismos mediante el cuestionario de calidad, así como el proceso de extracción de datos.

Por último, finalizaremos con el Análisis, fase en la que se tratarán los datos obtenidos y se manipularán para su consumo. Así mismo, se discutirán los resultados obtenidos tanto cualitativa, como cuantitativamente. Como resultado, se presentará en este trabajo unas conclusiones formadas a partir de las reflexiones alcanzadas.

Cabe destacar, además, que se utilizará la herramienta Parsifal [5] para el desarrollo de esta revisión.

III. APROXIMACIÓN AL ESTADO DEL ARTE

A. Indagación

Blockchain es principalmente y en su definición más objetiva, una base de datos distribuida en la que se almacenan registros de transacciones, y la cual es mantenida y validada por una red de ordenadores repartidos por todo el mundo [6]. Esta red es la que fundamenta los principios de confianza y escalabilidad, a través de protocolos como el de consenso, o el cifrado de clave público-privada.

Fundamentadas en esta capacidad transaccional nacen las criptomonedas, que son en esencia, una unidad monetaria digital cuyo valor, a diferencia de cualquier otro activo financiero, no depende de ningún activo físico que lo respalde. Fomentar la participación, así como colaborar en el mantenimiento de la propia red, fueron las principales razones para la creación de estos activos, que sin embargo, han excedido tales responsabilidades, alcanzando en impacto la economía global y sus procesos.

Movimientos político-económicos como las regulaciones comunitarias [2] o nacionales [3] están ocurriendo, demostrando un claro interés por mantener bajo control una economía de escala que podría convertirse en parte de una transición digital.

B. Identificación

Durante el desarrollo de esta fase se ha realizado la definición de los términos PICOC, de los cuales nos hemos ayudado posteriormente para la formulación de nuestras preguntas de investigación. Una vez obtenidas las mismas, se han expandido con la inclusión de algunas palabras claves, además de sinónimos de las mismas, con el objetivo de generar una cláusula de búsqueda más completa. Por último, se han seleccionado las bases de investigación que se utilizarán durante este estudio.

1) Términos PICOC: Provenientes de población, intervención, comparación, salida y contexto, nos ayudan a entender de forma aislada cuales son los 5 factores principales que afectan a nuestro estudio, y así, acotar de una forma más semántica el proceso de generación de las preguntas de investigación. En nuestro caso, la definición realizada ha sido la siguiente:

- Población: *Software Engineers*, *Researchers*
- Intervención: *Blockchain*
- Comparación: *Machine Learning techniques*
- Salida: *Fraud detection*
- Contexto: *Cripto networks*, *Decentrilized Finance*

2) Preguntas de investigación: A través de los términos PICOC, así como teniendo en cuenta el interés en la temática de estudio a tratar, se han formulado las siguientes preguntas de investigación:

- RQ1: *Which are the main Machine Learning techniques applied to blockchain technologies for fraud detection in cripto networks?*
- RQ2: *How have fraud detection Machine Learning techniques evolved on decentralized finance blockchains applications over the years?*

3) Extensión de palabras clave: Tras la extensión de nuestras palabras clave usando sinónimos de los términos PICOC, así como las redes de criptomonedas más prolíferas actualmente, se ha generado como resultando en el siguiente listado:

- *Software Engineers, Data Engineers, Machine Learning Engineers, Researchers*
- *Blockchain, Ledger, Bitcoin, Ethereum, Tether, Dogecoin, Cardano, Polygon, Solana*
- *Machine Learning techniques, Artificial Intelligence techniques*
- *Fraud detection, Fraud prevention, Anti-fraud*

4) Cláusula de búsqueda: Usando la herramienta Parsifal, y basándonos en los términos previamente extendidos, se ha generado la siguiente cláusula de búsqueda:

(“Researchers” OR “Software Engineers” OR “Data Engineer” OR “Machine Learning Engineers”) AND (“Bitcoin” OR “Cardano” OR “Dogecoin” OR “Ethereum” OR “Polygon” OR “Solana” OR “Tether” OR “Blockchain” OR “Ledger”) AND (“Machine Learning techniques” OR “Artificial Intelligence techniques”) AND (“Fraud detection” OR “Anti-fraud” OR “Fraud prevention”)

5) Bases de datos de investigación: Para este estudio, se ha decidido utilizar dos bases de datos de investigación a las que se tiene acceso gracias a la red nacional de investigación de la que participa la Universidad de Málaga, siendo las seleccionadas:

- Scopus
- Web of Science

C. Selección

En esta sección de selección se definirán los criterios de filtrado que determinaran la amplitud de los resultados obtenidos en nuestras búsquedas. Además, estos serán los encargados de permitirnos replicar de la forma más rigurosa posible las diferentes búsquedas en las bases de datos de investigación seleccionadas.

1) Criterios de selección: Además de la cláusula de búsqueda, las bases de datos de investigación nos permiten determinar parámetros de filtrado. En este estudio, se han decidido aplicar los siguientes criterios de selección:

- Área de estudio: Ciencias de la Computación o Ingeniería del Software

- Tipo de documento: Artículo
- Keyword(s) indexada(s): "Blockchain", "Machine Learning" o "Machine-Learning"
- Idioma: Inglés
- Política de disponibilidad: Acceso público

2) Cuestionario de evaluación: Una vez realizadas las búsquedas, es necesario evaluar los resultados obtenidos, con el objetivo de asegurarnos que estos aportan valor a la temática de estudio seleccionada. Para esto utilizaremos el formulario de evaluación de calidad, formado de las preguntas de evaluación, así como tres respuestas tipo puntuadas de forma ponderada. Las preguntas definidas son las siguientes:

- Q1: *Does the document expose any kind of fraudulent action, behaviour or pattern detection?*
- Q2: *Does the document contain relation with Blockchain, Ledger, or any other mix of related technologies?*
- Q3: *Does the document expose the usage of one or more Machine Learning techniques for fraud detection?*
- Q4: *Does the document contains direct relation with at least one criptocurrency or tokenize asset?*

En cuanto a las respuestas tipo, estas tienen una puntuación asociada entre 0 y 1 puntos, pudiéndose obtener un máximo de 4 puntos y un mínimo de 0 puntos. Se ha situado el umbral de corte en 2.0 puntos, de forma que solo los artículos que superen dicha puntuación pasarán a la fase de extracción de datos. Las respuestas definidas son las siguientes:

- Yes (1.0 puntos)
- Partially (0.35 puntos)
- No (0.0 puntos)

3) Formulario de extracción de datos: En el se detallan las diferentes preguntas que se realizarán durante una segunda revisión del artículo, de la que se esperará extraer algunos de los detalles clave necesarios para el correcto análisis del mismo. Las preguntas seleccionadas han sido:

- Q1: *What is the study objective regarding fraud detection?*
- Q2: *What is the study definition of fraud for the presented use case?*
- Q3: *Is the study focused on proactive/preemptive measures to avoid fraud, or counteracting measures to analyse intents?*
- Q4: *What are the Machine learning techniques used in the study?*
- Q5: *Which was the Machine Learning technique with the best results?*
- Q6: *Which was the accuracy (%) of the most performer Machine Learning technique used?*
- Q7: *Is the most performer Machine Learning technique a supervised or unsupervised one?*

D. Clasificación

Durante el proceso de clasificación se trabajará con los resultados obtenidos de la búsqueda en las bases de datos de investigación, los cuales serán depurados y canalizados a través del cuestionario de evaluación, así como el formulario de extracción de datos.

1) Importación y eliminación de duplicados: Tras la importación de los ficheros *BibText* con las referencias de los resultados de búsqueda en la herramienta Parsifal, se realizó una primera limpieza de duplicados, que decrementó nuestro dataset de trabajo en un 25,73%: de 171 artículos inicialmente, a 127.

2) Realización de la evaluación de calidad: Tras una primera revisión, y la resolución del formulario de calidad, encontramos que solo 34 de los artículos (un 26.77% con respecto a la fase anterior) pasan nuestro umbral de corte, continuando hacia la siguiente fase.

3) Extracción de datos: Durante el proceso de extracción de datos, se completaron los registros correspondientes a las respuestas recuperadas de los diferentes artículos revisados [22]. Estos fueron transformados en tres tablas a utilizar durante el proceso de análisis. En esta fase no se produce merma en el número de artículos.

E. Análisis

El proceso de análisis se ha dividido en dos grandes bloques, atendiendo el primero de ellos al análisis de las características cualitativas y los objetivos de estudio de los artículos revisados (cuestiones Q1-Q2 del formulario de extracción de datos), mientras que el segundo se centra en el análisis más cuantitativo de las técnicas utilizadas, sus rendimientos y las tendencias temporales observadas en las mismas (cuestiones Q3-Q7 del mismo).

1) Análisis de los objetivos de estudio: En este primer bloque, se detectan tres grandes categorías en cuanto a objetivos de estudio, quedando dentro de ellas el 88.23% de los estudios revisados. Estas son: Detección de comportamientos y transacciones anómalos (38.23%), detección de estafas (29.41%), detección de vulnerabilidades (20.59%).

En el caso de la detección de anomalías, se encuentra que algunas de las definiciones encontradas hace referencia de forma explícita al comportamiento (Chuyi Yan et al. [8], Ruchi Mittal et al. [9], Xiao Fan Liu et al. [10]) del usuario en la red, sin embargo, algunos estudios desligan esta componente del patrón transaccional, atendiendo únicamente a los fines u objetivos de dichos intercambios. En este sentido, el 60% de los estudios centrados en la detección de transacciones anómalas tenían el fin de detectar patrones de lavado de dinero (Wai Weng Lo et al. [11], Johrha Alotibi et al. [12], Steven Farrugia et al. [13]).

En referencia a la detección de estafas, se encuentra un especial interés (70% revisiones) en los tipos de estafa piramidal o en planes de inversión con altos rendimientos. En estos, los estafadores aprovechan el pseudoanonimato de esta tecnología para implementar fraudes financieros [14], que sustentados en el desarrollo de contratos inteligentes, mantienen el proceso de estafa autónomamente.

El 57.14% de los estudios en esta categoría encuentra más rendimiento en la resolución de este problema a través de la clasificación mediante el uso de Redes Neuronales (Emad Badawi et al. [9], Lingyu Bian et al. [14], Yizhou

Chen et al. [15], Shuhui Zhang et al. [16]); sin embargo, en el caso del estudio de Xiajiong Shen et al. [17] se decidió transformar un claro problema de clasificación en uno de detección de anomalías.

En términos de detección de vulnerabilidades, se encuentra de especial interés en las contribuciones realizadas por Kabla et al. [18], en su implementación de un Sistema de Detección de Intrusiones con el fin de reducir las amenazas de la red Ethereum; o la de Khan et al. [19] en su investigación en la detección de ataques DDoS que amenacen con congestionar el ancho de banda transaccional de una red blockchain.

2) Análisis de las técnicas seleccionadas y sus rendimientos: Durante el estudio analítico de los resultados de los artículos revisados se detecta que estos quedan contenidos entre los años 2018 y 2023, siendo el año 2022 el que más contribuciones nos aporta: un 52.94% del total. Esto nos muestra la creciente tendencia de la temática entre la comunidad investigadora.

Es también interesante destacar que el 65.62% de los estudios aportan una solución preventiva, siendo además un 86.2% del total soluciones basadas en técnicas de aprendizaje supervisado.

Tras analizar las técnicas reportadas como mejores candidatas en cada uno de los artículos revisados, podemos ver como Random Forest (RF) se coloca en primera posición, seguida de Redes Neuronales Convolucionales (CNN), y con XGBoost (XGB) en tercera posición. El rendimiento promedio de RF es del 92%, mientras que el de XGBoost se acerca al 98%. También vemos como las CNN, que tienen un rendimiento algo superior al 96%, empiezan a jugar un papel fundamental apareciendo principalmente en artículos del año 2021 (Yizhou Chen et al. [15], Lingyu Bian et al. [14]), 2022 (Shuhui Zhang et al. [16]) y 2023 (Zijian Zhang et al. [20]), es decir, en la última mitad del periodo de estudio.

IV. CONCLUSIONES

Tras la recopilación analítica realizada, se observó que el objetivo principal del uso de las técnicas utilizadas era resolver un problema de clasificación. Dando respuesta a la primera de las preguntas de investigación (RQ1), se encuentra que Random Forest es la técnica de Machine Learning más usada; así como XGBoost, la cual presenta resultados de media un 5% mejores que Random Forest, y llegando a precisiones cercanas al 98% de acierto; o las Redes Neuronales Convolucionales, las cuales con precisiones medias cercanas al 96%, han demostrado ser muy efectivas en la extracción automática de características de clasificación.

Con respecto a la segunda de nuestras preguntas de investigación (RQ2), se encuentra la posible existencia de una tendencia en el uso de técnicas basadas en Redes Neuronales, la cual se obtiene de la distribución temporal encontrada en los artículos que utilizan esta solución de implementación. Dichos artículos suponen el 57% de la categoría de detección de vulnerabilidades, y todos ellos se encuentran aglutinados en la segunda mitad del periodo de estudio (2021-2023); aún así, parece necesaria una revisión futura para confirmar esta tendencia, asumiendo un eje temporal más largo.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por la Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía a través del proyecto BIG-PrivDATA (UMA20-FEDERJA-082) del Programa Operativo FEDER Andalucía 2014-2020.

REFERENCIAS

[1] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Decentralized business review, 21260.
[2] Legislative Observatory - European Parliament, Procedure File: 2021/0241(COD), URL: https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?reference=2021/0241(COD)&l=en (visitado 29-04-2023).
[3] Agencia Estatal Boletín Oficial del Estado, BOE Num. 81, URL: https://www.boe.es/boe/dias/2023/04/05/ (visitado 29-04-2023).
[4] Gómez Vargas, M., Galeano Higuita, C., & Jaramillo Muñoz, D. A. (2015). El estado del arte: una metodología de investigación.
[5] Parsifal: About Parsifal. 2023. URL: https://parsif.al/about/ (visitado 29-04-2023)
[6] Sarmah, S. S. (2018). Understanding blockchain technology. Computer Science and Engineering, 8(2), 23-29.
[7] CoinMarketCap: Global Cryptocurrency Charts. 2023. URL: https://coinmarketcap.com/charts/ (visitado 29-04-2023)
[8] Chuyi Yan et al. Blockchain abnormal behavior awareness methods: a survey. En: Cybersecurity 5.1 (2022), pág. 5
[9] Ruchi Mittal y Mahinder Pal Singh Bhatia. Detection of Suspicious or Un-Trusted Users in Crypto-Currency Financial Trading Applications. En: International Journal of Digital Crime and Forensics (IJDCF) 13.1 (2021), págs. 79-93.
[10] Xiao Fan Liu et al. Characterizing key agents in the cryptocurrency economy through blockchain transaction analysis. En: EPJ Data Science 10.1 (2021), pág. 21.
[11] Wai Weng Lo et al. Inspection-L: self-supervised GNN node embeddings for money laundering detection in bitcoin. En: Applied Intelligence (2023), págs. 1-12.
[12] Johrha Alotibi et al. Money Laundering Detection using Machine Learning and Deep Learning. En: International Journal of Advanced Computer Science and Applications 13.10 (2022)
[13] Steven Farrugia, Joshua Ellul y George Azzopardi. Detection of illicit accounts over the Ethereum blockchain. En: Expert Systems with Applications 150 (2020), pág. 113318
[14] Lingyu Bian et al. Image-based scam detection method using an attention capsule network. En: IEEE Access 9 (2021), págs. 33654-33665.
[15] Yizhou Chen et al. Improving Ponzi scheme contract detection using multi-channel TextCNN and transformer. En: Sensors 21.19 (2021), pág. 6417.
[16] Shuhui Zhang et al. Ethereum Ponzi Scheme Detection Based on PD-SECR. En: Security and Communication Networks 2022 (2022).
[17] Xiajiong Shen, Shuaimin Jiang y Lei Zhang. Mining bytecode features of smart contracts to detect Ponzi scheme on blockchain. En: Computer Modeling in Engineering & Sciences 127.3 (2021), págs. 1069-1085.
[18] Arkan Hammoodi Hasan Kabla et al. Applicability of intrusion detection system on Ethereum attacks: a comprehensive review. En: IEEE Access (2022).
[19] Zulfiqar Ali Khan y Akbar Siami Namin. A Survey of DDOS Attack Detection Techniques for IoT Systems Using BlockChain Technology. En: Electronics 11.23 (2022), pág. 3892.
[20] Zijian Zhang et al. A Multi-Dimensional Covert Transaction Recognition Scheme for Blockchain. En: Mathematics 11.4 (2023), pág. 1015.
[21] CoinMarketCap - Capitalización total del mercado de criptomonedas. URL:https://coinmarketcap.com/es/charts/
[22] Referencias de artículos revisados - Tablas de clasificación. URL:https://github.com/raulocana/crypto-fraud-detection-using-ml-slr/blob/58c6feb4e4f6b397838df35809c2400f9f3a4caf/apendix_and_references.pdf