BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2025
http://excel.fit.vutbr.cz

Automatická detekce nezákonných transakcí v Bitcoinu

Miroslav Šafář*

Abstrakt

Detekce ilegálních transakcí je jednou z hlavních priorit bezpečnostních složek chránících náš finanční systém. Tato práce se zaměřuje na využití hlubokého učení k forenzní analýze blockchainu Bitcoinu. Práce identifikuje problémy state-of-the-art metod, které brání efektivnímu nasazení těchto metod ve forenzní praxi, a analyzuje obsah nejpoužívanější anonymizované datové sady Elliptic. Za pomoci reverzního inženýrství se mi podařilo deanonymizovat 20 % vlastností, a díky tomu identifikovat 100 % transakcí obsažených v této datové sadě. Práce dále představuje otevřenou datovou sadu, která opravuje fundamentální chyby datové sady Elliptic, a poskytuje tak potřebný základ pro nasazení metod hlubokého učení ve forenzní praxi.

*xsafar23@stud.fit.vut.cz, *Faculty of Information Technology, Brno University of Technology*

1. Úvod

Pseudonymita a absence centrální kontrolní entity u kryptoměny Bitcoin přitahuje pozornost zlomyslných aktérů, kteří chtějí zneužít blockchainové technologie k praní špinavých peněz, k podvodům nebo k dalším ilegálním aktivitám. Detekce takovýchto transakcí je jednou z hlavních priorit bezpečnostních složek chránících finanční systém. Proto se tato práce zaměřuje na využití hlubokého učení pro automatickou detekci takovýchto transakcí.

Problém detekce transakcí spojených s ilegálními aktivitami lze reprezentovat jako problém klasifikace uzlů v grafu transakcí (viz Obrázek 1). Budeme-li chtít klasifikovat vektory vlastností těchto transakcí, lze využít klasických metod strojového učení, jako jsou Random Forest klasifikátor či XGBoost klasifikátor, nebo vícevrstvého perceptronu (MLP) jako zástupce hlubokého učení. Avšak ani jedna z těchto metod nevyužívá grafových informací a nevyužívá tak kontextu, ve kterém se transakce v rámci Bitcoinu nacházejí.

Jedním z možných řešení je využití grafových neuronových sítí, jak je využito v práci Weber a spol. [1], která představila doposud největší anotovanou datovou sadu transakčního grafu Elliptic, a dále v pracích [2, 3, 4, 5, 6]. Ačkoliv mají tyto práce velice slibné výsledky, všechny byly ověřovány na již dříve zmíněné datové sadě Elliptic, která je však anonymizována, a jejíž obsah je dosud neznámý, kromě prvních popsaných vlastností, které byly využity k částečné identifikaci obsažených transakcí (99,5 %).

Anonymita jednotlivých vlastností transakcí však přináší jeden kritický problém, který brání nasazení modelů natrénovaných na datové sadě Elliptic ve forenzní praxi. Tím je fakt, že není možné vytvořit tento vstupní vektor vlastností pro naši transakci zájmu.

Z tohoto důvodu se tato práce zaměřuje na vytvoření nové, otevřené datové sady, která vychází z datové sady Elliptic, jakožto z největší dostupné anotované datové sady transakcí, a jejíž cílem je umožnit integraci natrénovaných modelů grafových neuronových sítí do forenzních aplikací. Validita vytvořené datové sady je experimentálně ověřena a porovnána s datovou sadou Elliptic.

Mezi přínosy této práce lze zařadit:

- identifikaci hlavních výzev pro nasazení GNN ve forenzní praxi,
- publikaci otevřené datové sady Elliptic1.1,
- deanonymizaci 20 % vlastností datové sady Elliptic a identifikaci všech obsažených transakcí,
- sjednocené srovnání metod strojového učení v detekci ilegálních transakcí.

2. Datová sada Elliptic1.1

V rámci explorační analýzy datové sady Elliptic jsem zjistil, že všechny vlastnosti transakcí mají střed v bodě 0, a směrodatnou odchylku 1, z čehož vyplývá, že

jejich anonymizace spočívala v z-score normalizaci. Jedná se tedy o posun a škálování, které nijak nemění histogram jako takový, pouze jeho roztažení a umístění na ose x. Dále jsem si všiml, že jednotlivé vlastnosti transakcí jsou buď to zdola ohraničené, anebo ohraničené z obou stran, jak lze vidět na obrázku Obrázek 3. Z toho můžeme usoudit, že se v prvním případě jedná pravděpodobně o přirozená čísla (zdola ohraničena 0), která v rámci blockchainu reprezentují většinu atributů jednotlivých transakcí, anebo o nějaký poměr. V druhém případě budeme předpokládat, že se jedná o vlastnost s oborem hodnot $< 0, 1 >$.

K deanonymizaci normalizovaných přirozených čísel můžeme předpokládat, že nejmenší rozdíl mezi jednotlivými normalizovanými čísly bude odpovídat 1. Poté můžeme vlastnosti zrekonstruovat vydělením tímto rozdílem a posunem nejmenšího čísla na 0, předpokládáme-li, že alespoň jedna transakce měla nejmenší možnou hodnotu. V případě, kdy se jednalo o průměr nějakých celočíselných vlastností, nejmenší rozdíl jsem nahradil modusem.

Po deanonymizaci 20 % vlastností jsem se zaměřil na identifikaci transakcí obsažených v této datové sadě. Z již identifikovaných transakcí jsem identifikoval všechny bloky, ze kterých autoři čerpali, a následně jsem spočítal hodnoty deanonymizovaných vlastností pro všechny transakce v těchto blocích. Následně jsem spároval transakce s odpovídajícími vlastnostmi, čímž se mi podařilo dosáhnout identifikace všech obsažených transakcí.

Při deanonymizaci jednotlivých vlastností se mi podařilo zjistit, že globální charakteristiky adres, jako je například jejich životnost či počet transakcí, nereflektují stav v době vzniku dané transakce, ale stav odpovídající době vytvoření datové sady. To znamená, že tyto vlastnosti odrážejí pozdější období, než které pokrývá samotná testovací sada. V důsledku toho dochází k průsaku informací z testovacích dat do trénovacích dat prostřednictvím vlastností adres. Shoda těchto vlastností mezi trénovací a testovací sadou dále způso-buje, že trénované neuronové sítě se nemusí učit obecné vzorce ilegálních transakcí, ale místo toho si mohou pouze „zapamatovat“ specifické charakteristiky adres spojených s ilegálními transakcemi a na základě jejich výskytu v testovací sadě označit transakci jako ilegální. Tento jev zásadně zpochybňuje hypotézu, že datová sada Elliptic je vhodná pro detekci ilegálních transakcí, jejichž adresy se v anotovaných datech dosud nevyskytly.

Datová sada Elliptic1.1 byla sestavena následovně (viz Obrázek 2). Nejprve byla původní množina transakcí z datové sady Elliptic iterativně rozšířena o své okolní transakce. Následně byl spuštěn nástroj *Blockchain Parser*, jehož úkolem bylo shromáždit statistické údaje o adresách obsažených v těchto transakcích. Tento krok byl nezbytný pro zajištění toho, aby transakce obsahovaly statistiky adres odpovídající době své realizace, nikoliv době tvorby vlastností jako u datové sady Elliptic. Po vytvoření databáze se statistikami adres následovala fáze konstrukce vektorů vlastností jednotlivých transakcí. Celý proces byl zakončen post-processingem, konkrétně aplikací normalizace pomocí z-score standardizace.

3. Experimenty

V rámci této práce jsem provedl experimenty s modely Random Forest (RF), XGBoost, vícevrstvého perceptronu a dále s modely grafových neuronových sítí – konvoluční neuronovou sítí (GCN), grafovou attention sítí (GAT) a modelem DGA-GNN [5]. Experimenty byly prováděny na dvou datových sadách, na datové sadě Elliptic a na mnou vytvořené datové sadě Elliptic1.1. Pro každý z modelů bylo provedeno několik experimentů s různým nastavením hyperparametrů dle literatury a byl vybrán model s nejvyšší dosaženou hodnotou validačního F1 skóre. Vzhledem k výrazné nevyváženosti datové sady byly neuronové sítě a klasifikátor XGBoost trénovány s váženou ztrátovou funkcí křížové entropie. U algoritmu Random Forest bylo štěpení uzlů stromů prováděno na základě Shannonovy entropie, přičemž váhy tříd byly nastaveny na 0,5 a 1.

Výsledky modelů RF, XGBoost, MLP a 1-layer GAT s residuálním spojením a využitím kódování založeném na rozdělení do košů [5] lze vidět v tabulce Tabulka 1. Z těchto výsledků lze konstatovat, že vytvořená validační sada poskytuje srovnatelné výsledky s datovou sadou Elliptic a může se tak stát její náhradou.

4. Závěr

Za úspěch považuji identifikaci kritických nedostatků dosavadního poznání v oblasti využití neuronových sítí k forenzní analýze blockchainu Bitcoinu. Reverzním inženýrstvím datové sady Elliptic se mi podařilo deanonymizovat 20 % původních vlastností transakcí, díky čemuž se mi podařilo identifikovat 100 % transakcí obsažených v této datové sadě. Dále se také podařilo sestavit otevřenou datovou sadu Elliptic1.1, která umožní trénování grafových neuronových sítí a jejich integraci do forenzních nástrojů určených k detekci ilegálních transakcí v Bitcoinu, což dosud prakticky nebylo možné.

Literatura

[1] Mark Weber, Giacomo Domeniconi, Jie Chen, Daniel Karl I. Weidele, Claudio Bellei, Tom Robinson, and Charles E. Leiserson. Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics, July 2019.

[2] Aldo Pareja, Giacomo Domeniconi, Jie Chen, Tengfei Ma, Toyotaro Suzumura, Hiroki Kanezashi, Tim Kaler, Tao Schardl, and Charles Leiserson. Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 5363–5370, 2020.

[3] Ismail Alarab and Simant Prakoonwit. Graph-based lstm for anti-money laundering: Experimenting temporal graph convolutional network with bitcoin data. *Neural Processing Letters*, 55:1–19, 06 2022.

[4] Ziwei Chai, Siqi You, Yang Yang, Shiliang Pu, Jiarong Xu, Haoyang Cai, and Weihao Jiang. Can Abnormality be Detected by Graph Neural Networks? In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pages 1945–1951, Vienna, Austria, July 2022. International Joint Conferences on Artificial Intelligence Organization.

[5] Mingjiang Duan, Tongya Zheng, Yang Gao, Gang Wang, Zunlei Feng, and Xinyu Wang. DGA-GNN: Dynamic Grouping Aggregation GNN for Fraud Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 11820–11828, March 2024.

[6] Qianyu Wang, Wei-Tek Tsai, and Bowen Du. RM-GANets: reinforcement learning-enhanced multi-relational attention graph-aware network for anti-money laundering detection. *Complex & Intelligent Systems*, 11(1):5, November 2024.