electronics

MDPI

Communication

Detecting Phishing Accounts on Ethereum Based on Transaction Records and EGAT

Xuanchen Zhou [1,2], Wenzhong Yang [2,*] and Xiaodan Tian [2]

[1] School of Software, Xinjiang University, Urumqi 830046, China
[2] College of Information Science and Engineering, Xinjiang University, Urumqi 830046, China
* Correspondence: yangwenzhong@xju.edu.cn

Abstract: In recent years, the losses caused by scams on Ethereum have reached a level that cannot be ignored. As one of the most rampant crimes, phishing scams have caused a huge economic loss to blockchain platforms and users. Under these circumstances, to address the threat to the financial security of blockchain, an Edge Aggregated Graph Attention Network (EGAT) based on the static subgraph representation of the transaction network is proposed. This study intends to detect Ethereum phishing accounts through the classification of transaction network subgraphs with the following procedures. Firstly, the accounts are used as nodes and the flow of transaction funds is used as directed edges to construct the transaction network graph. Secondly, the transaction record data of phishing accounts in the publicly available Ethereum are analyzed and statistical features of Value, Gas, and Timestamp values are manually constructed as node and edge features of the graph. Finally, the features are extracted and classified using the EGAT network. According to the experimental results, the Recall of the proposed method from the article is 99.3% on the dataset of phishing accounts. As demonstrated, the EGAT is more efficient and accurate compared with Graph2Vec and DeepWalk, and the graph structure features can express semantics better than manual features and simple transaction networks, which effectively improves the performance of phishing account detection.

Keywords: ethereum; EGAT; phishing account; graph embedding

Citation: Zhou, X.; Yang, W.; Tian, X. Detecting Phishing Accounts on Ethereum Based on Transaction Records and EGAT. *Electronics* **2023**, *12*, 993. https://doi.org/10.3390/electronics12040993

Academic Editors: Andrea Prati, Luis Javier García Villalba and Vincent A. Cicirello

Received: 3 January 2023
Revised: 10 February 2023
Accepted: 14 February 2023
Published: 16 February 2023

1. Introduction

Blockchain, known as the infrastructure of web 3.0, integrates a P2P network, asymmetric encryption, consensus mechanism, and other technologies, and is characterized by distribution and anonymity [1]. In recent years, the rapid development of blockchain technology has attracted extensive attention from the industry and academia [2–4]. In particular, the combination of blockchain and smart contract technology has formed the Ethereum platform, enabling users to flexibly build decentralized applications. According to the statistics of Etherscan [5], the rapid development of the Ethereum platform has made Ethereum the second-largest cryptocurrency in the world.

The implementation of anonymity, on the one hand, protects the privacy of users, and on the other hand, it also lays a hidden danger for malicious deception. With the continuous rise of the comprehensive valuation of cryptocurrency, driven by huge interests, attackers can take advantage of the anonymity of the blockchain to carry out attacks, fraud, and other criminal activities even under market supervision, thus obtaining a large number of illegal funds [6]. For example, in the famous vulnerability attack The DAO event [7], the hacker used the address—0x F35e ... a77D to transfer 3.6 million ether, which is worth more than 60 million dollars. Malicious fraudsters use different technologies, such as malware or phishing websites, to induce users to send digital currencies to their pre-designed account addresses to complete fraud. In January 2019, according to a Chainalysis report [8], Ethereum has become the preferred encryption platform for fraudsters. Given all that, it can be seen that the anonymity of the blockchain provides an opportunity for cybercrime

Currently, Ethereum is the most prominent blockchain-based platform and the first that supports smart contracts. However, the number of phishing scam accounts are reportedly more than 50% of all cybercrimes in Ethereum [9]. In order to detect malicious phishing accounts hidden behind the address, a detailed analysis of the transaction behavior of the account is required [10].

Early research mainly conducted behavior clustering based on transaction behavior combined with heuristic conditions to analyze transaction rules [11]. This method is highly subjective and has a certain false positive rate. As the research progresses, researchers utilize the determined account type to analyze the transaction mode and employ machine learning algorithms to detect the phishing accounts. However, with the continuous growth of Ethereum transactions, traditional detection methods fail to identify malicious accounts neither quickly nor effectively.

In recent years, various graph-based models have been created to detect phishing scams on Ethereum. To construct a network from Ethereum transaction records, an address is regarded as a node while a transaction is regarded as an edge. This approach improves the accuracy and efficiency, but it also has major drawbacks. Most graph-based studies treated the phishing accounts detection problem as a node classification task [12,13]. In this case, it is usually necessary to build large graph structures, which are extremely resource-intensive. To overcome this problem, we treated it as a task of graph classification, extracting a subgraph of accounts to represent it.

By contrast, EGAT [14] can quickly aggregate the characteristics of neighboring nodes to obtain the characteristic details of the account transaction subgraph. Therefore, this paper proposes a phishing account detection method based on EGAT after a manual analysis of public datasets.

The main contributions are in the following two aspects.

1. According to the anonymity of the Ethereum trading network, a star subgraph structure is designed, and the node features as well as edge features are constructed manually;
2. A double-layer EGAT network with feature weight is constructed, which efficaciously aggregates edge features for graph classification tasks.

2. Related Work

The identity information of the blockchain account is strictly encrypted, but due to the special data structure and transaction mode of the blockchain, the transaction information is open and transparent. Moreover, the tamper proofing and traceability of the blockchain also laid the foundation for detecting the types of cryptocurrency accounts. Therefore, researchers can identify different accounts by analyzing transaction activities and extracting features.

Zhao [15] analyzed Bitcoin accounts through the heuristic clustering method, which can speculate user identity to a certain extent. Monaco [16] includes 12 transaction characteristics, such as transaction time interval and timestamp, to detect Bitcoin users. Experiments show that the transaction behavior is nonrandom with a power law distribution. Androulaki [17] and others found that 40% of student identities can be successfully matched with Bitcoin users even if users adopt the privacy protection methods recommended by Bitcoin. Reid et al. [18] cluster the transaction network and user network topology derived from the bitcoin transaction data with the bitcoin users, which can effectively identify and track some user activities. The above research methods are based on historical transaction data and heuristic clustering technology is introduced to study the behavior differences among different accounts. Although much valuable information can be obtained, such methods require rather rich prior knowledge and are vulnerable to the interference of outliers.

In recent years, large data analysis companies and the Ethereum block browser have released some transaction account types. Researchers can train machine learning detection models to classify and predict cryptocurrency accounts according to the identified account extraction features. Yin et al. [19] selected 98 features, such as the number of

transactions, average value, and standard deviation, and used the supervised machine learning algorithm to classify 957 Bitcoin accounts. The experiment showed that the F1 value was 79.64%. In order to improve the classification performance, Lin [20] and others further added time features on the basis of transaction frequency, average value, and other features, and Micro-F1 reached 87%.

Thanks to the continuous development of blockchain, cryptocurrency has been used in many fields, such as banking and medical care. Driven by substantial profits, cryptocurrencies such as Bitcoin have become tools for cybercrime. Some outlaws take advantage of the anonymity of blockchain to conduct illegal transactions without scruple, such as online extortion, online fraud, currency theft, and other malicious acts [21]. To tackle the problem, many researchers have been continuously making efforts in blockchain security. For example, Toyaoda et al. [22] extracted the frequency of the pattern as the key feature. The experimental results show that around 83% of the High Yield Investment Project (HYIP) accounts using the XGBoost algorithm can be correctly classified. Kanemura et al. [23] proposed a voting-based identification method for dark network transactions, selected 73 account features such as sending and receiving transaction values, and applied a random forest classifier to detect dark network transactions. Chen et al. [24] established a classification model based on the characteristics of transaction behavior and contract bytecode, which can effectively detect Ponzi fraud in Ethereum.

Most of the previous studies have been concentrated on Bitcoin [25], but have in recent years shifted the focus onto Ethereum account detection. For instance, Tingke et al. [26] proposed a model based on Hybrid Deep Neural Network to detect phishing scam accounts. Kabla et al. [27] proposed a detection mechanism called Ethereum Phishing Scam Detection that attempts to detect phishing scam-related transactions with a novel machine learning-based approach.

3. Methodology

Aiming at the phenomenon that malicious attackers take advantage of the anonymity of blockchain to carry out illegal and fraudulent activities on the Ethereum platform, this chapter describes the specific methods of fishing account detection based on EGAT, mainly including four stages: transaction record acquisition, graphic construction, feature embedding, and model extraction and classification. The overall framework is shown in Figure 1.

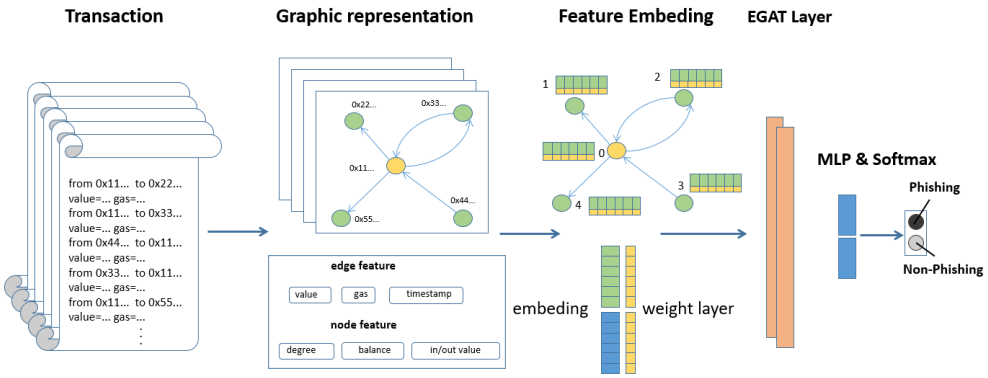

Figure 1. The framework of phishing account detection method based on EGAT.

3.1. Graph Structure

In this research, the graph is constructed as a star-shaped subgraph of the overall transaction network based on the transaction records of a single tagged account, as presented in Figure 2. The graph has a star structure, with the center point being the label account and the other points being the transaction account. In the graph structure, the sending address, and receiving address in the transaction record are the vertices, and the capital flow direction is the edge. Because of the particularity of the transaction network, there are

multiple edges between two nodes. In this study, the characteristics of edges are used as the basis for distinguishing the same node and direction edges at different times.

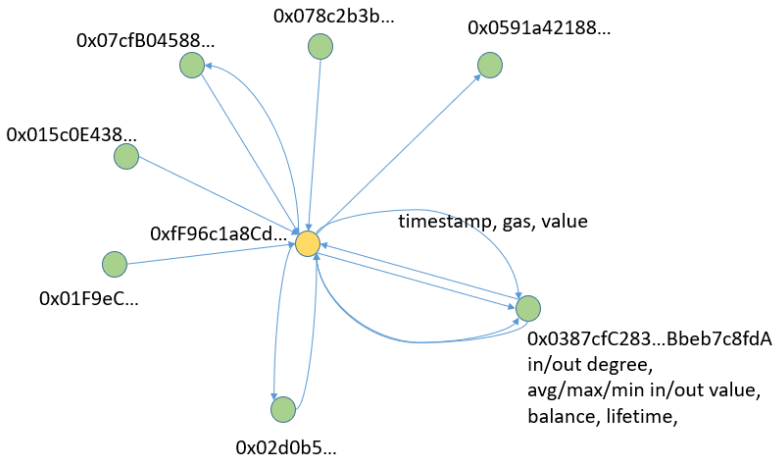

Figure 2. The graph structure based on the transaction network.

3.1.1. Edge Feature

The nature of the transaction network makes it possible for two nodes to have multiple duplicate records of the same direction transactions. Therefore, in this study, we chose to use the three pieces of information of Timestamp, Gas, and Value as the characteristics of edges and also as the differentiation of isotropic edges.

Value (V): Value at the time of transaction.

Timestamp (T): Timestamp of the transaction.

Gas (G): Gas fees at the time of transaction.

3.1.2. Node Feature

In terms of node features, we chose to use the Value at the time of the transaction as the primary information after analyzing the transaction records of the phishing smart contract and constructing it.

The manually collected statistical features collected are as follow:

Balance (B): The amount remaining in the account of the current node of the Ethernet when the transaction takes place.

In-degree (ID): In-degree of each node in the graph structure.

Out-degree (OD): Out-degree of each node in the graph structure.

Lifetime (LT): The difference value between the timestamp of the last transaction record and the first transaction record.

$$\mathrm{LT}_i = T_{i\_\mathrm{last}} - T_0 \tag{1}$$

Average Transaction Gap (ATG): The average of the time interval between each transaction.

$$\mathrm{ATG} = \frac{\sum_0^n (T_{i+1} - T_i)}{n} \tag{2}$$

Sending Value (SV): The value sent during the transaction.

Receiving Value (RV): The value received during the transaction.

Average Send Value (ASV): The average of the value sent by the node over all transaction records.

$$\mathrm{ASV} = \frac{\sum_0^n SV_i}{n} \tag{3}$$

Average Receiving Value (ARV): The average of value received by the node's overall transaction records.

$$\text{ARV} = \frac{\sum_0^n RV_i}{n} \tag{4}$$

Min Received Value (Min_RV): The minimum of the value received by the node over all transaction records.

Max Received Value (Max_RV): The maximum of the value received by the node over all transaction records.

Min Sent Value (Min_SV): The minimum of the value sent by the node over all transaction records.

Max Sent Value (Max_SV): The maximum of the value sent by the node over all transaction records.

3.2. EGAT

In this research, the use of a two-layer EGAT for information aggregation is proposed for the constructed multi-loop star graph structure. Since it is star-shaped, the two-layer structure can aggregate information from neighboring nodes to the current node, while edge features that point to other nodes can also be aggregated back in the second layer.

3.2.1. Feature Weight

We transform each feature into a tensor of 1 bit and end up with 13 bits for node features and 3 bits for edge features. We attach a layer of Feature self-Attention to each feature value. The feature set of edges is $e = \{\vec{e}_1, \vec{e}_2, \vec{e}_3, \ldots, \vec{e}_N\}, e_i \in R^F$, where N is the number of nodes, and $n = \{\vec{n}_1, \vec{n}_2, \vec{n}_3, \ldots, \vec{n}_K\}, n_i \in R^f$, K is the number of edges. $w = \{\vec{w}_1, \vec{w}_2, \vec{w}_3, \ldots, \vec{w}_M\}$, M is the number of weights, $M = N + K$. Finally, we can get the weighted edge features $e' = \vec{e}_i \cdot w_i$, and the weighted node features $n' = \overrightarrow{n_i} \cdot w_i$. We will use the obtained features with learnable weights for message passing in EGAT.

3.2.2. EGAT Layer

In GATconv, the input to our layer is a set of node features $h = \vec{h}_1, \vec{h}_2, \vec{h}_3, \ldots, \vec{h}_N, \vec{h}_i \in R^F$, where N is the number of nodes, and F is the number of features in each node. The layer produces a new set of node features (of potentially different cardinality F'), $h' = \overrightarrow{h'}_1, \vec{h}_2, \overrightarrow{h'}_3, \ldots, \vec{h}'_N, h' \in R^F$, as its output. In order to obtain sufficient expressive power to transform the input features into higher-level features, at least one learnable linear transformation is required. To that end, as an initial step, a shared linear transformation, parametrized by a weight matrix, $W \in R^{FxF'}$ is applied to every node. W is a weight matrix to be trained, equivalent to a fully concatenated layer, for the purpose of transforming the features, from low-order features to high-order features. We then performed self-attention on the nodes—*a* shared attentional mechanism illustrated as:

a: $R^F \times R^{F'} \rightarrow R$ computes attention coefficients.

$$e_{ij} = a\left(Wh_i, Wh_j\right) \tag{5}$$

In this research, the graph attention layer handles edge features from the Rossmann–Toolbox. The difference lies in how the unnormalized attention score e_{ij} is obtained:

$$e_{ij} = \overrightarrow{F}\left(f'_{ij}\right) \tag{6}$$

$$f'_{ij} = \text{LeakyReLU}\left(A\left[h_i \| f_{ij} \| h_j\right]\right) \tag{7}$$

To illustrate, f'_{ij} are the edge features. A is the weight matrix and F is the weight vector. After that, the generated node features h' are updated in the same way as in regular GAT. The message passing of EGAT is shown in Figure 3.

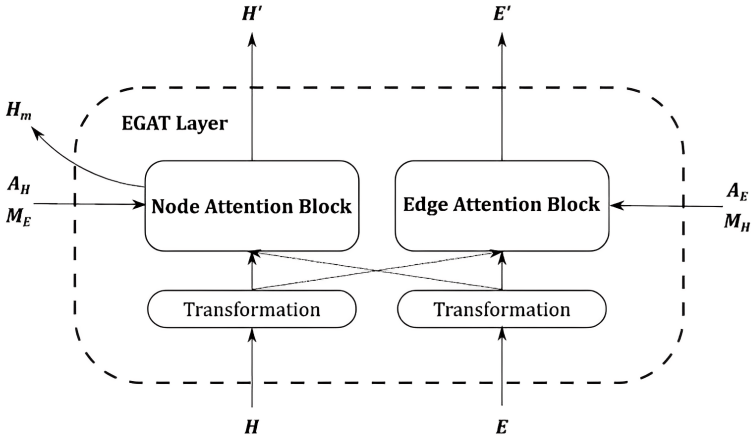

Figure 3. Message passing of EGAT.

4. Experimental Results

4.1. Dataset

The phishing account data in this article comes from Etherscan.io [5], which obtained the phish-tagged accounts, filtered, and deleted invalid accounts with a transaction value of zero, and finally, obtained 1659 phishing accounts. In order to figure out the difference between the transaction modes of malicious accounts and benign accounts, 5805 accounts including contract and wallet users were collected, and 1700 accounts were sampled down to form benign account data. The composition of the experimental dataset is shown in Table 1.

Table 1. Phishing smart contracts dataset.

Type	**Train**		**Test**	
label	Phishing	Normal	Phishing	Normal
$\|Y\|$	1327	1360	332	340
$\|V\|$	55,777	34,846	13,925	7408
$\|E\|$	187,119	82,860	46,344	16,347

Basic statistics of the Ethereum datasets. $|V|$ and $|E|$ are the numbers of nodes and edges, respectively, and $|Y|$ is the number of labeled phishing nodes. The dataset in this experiment is a binary classification dataset, and is divided into two parts: the training set and the test set. In the training set, for example, category 1 is a collection of 1327 star-shaped subgraphs centered on phishing accounts, and the total number of nodes in all phishing account subgraphs is 55,777 and the number of edges is 187,119. There are also 1360 normal account-centered subgraphs in the training set, the total number of nodes in all subgraphs is 34,846, and the total number of edges is 82,860.

4.2. Experimental Settings

To find the most suitable parameters for the performance of the model, adjustments were also made while referring to the best parameters of other GCN models. To facilitate the comparison of the experimental results, the graph neural network model involved in the experiment was restricted to a single variable. In addition, to verify the performance of the improved model in graph embedding, several different standard datasets were trained, tested, and verified as well. In the experiment, the initial network learning rate was 0.01. Adaptive Adam Optimizer was used for model optimization. The Dropout rate was uniformly set to 0.5. After adopting the dynamic learning rate adjustment method, the learning rate decreased by 0.1 for every 20 batches. The model training iterated 100 times. After the test, with the increase in the number of iterations, the results will appear smooth. In the meantime, the multi-head attention mechanism was used to improve the stability of the model training.

4.3. Evaluation

To facilitate a comparison with the experiments in other papers, the experimental evaluation metrics in this paper also apply Accuracy, Precision, Recall, and *F*1-score, which are commonly used in classification models [28,29]. The specific formulas are as follows.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

$$F1 - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

Among them, *TP* (True Positive) indicates a benign account that is correctly predicted; *TN* (True Negative) indicates a phishing account that is predicted to be benign; *FP* (False Positive) indicates a phishing account that is predicted to be benign, and *FN* (False Negative) indicates a benign account that is predicted to be phishing.

4.4. Evaluation on Phishing Detection

For the graph structure we built, we repeated a variety of common graph embedding methods and carried out comparative experiments. The results are shown in Table 2.

Table 2. The phishing detection results of this study.

Model	Accuracy	Precision	Recall	F1-Score
Graph2vec	0.731	0.743	0.742	0.732
Node2vec	0.605	0.614	0.614	0.609
DeepWalk	0.724	0.599	0.792	0.682
Sub2vec	0.510	0.509	0.520	0.514
GraphSage	0.818	0.986	0.832	0.832
GCN	0.819	0.868	0.694	0.853
GAT	0.924	0.935	0.888	0.911
EGAT	0.981	0.966	0.993	0.979

As Table 2 shows, this EGAT model achieves the best performance on Accuracy, Precision, Recall, and *F*1-score among all models. The embedding methods of Sub2vec, DeepWalk, and Node2vec are similar in that they all extend from the basic method of the random walk so that the feature representation of the whole graph is weaker in the subgraph classification task. Graph2vec gets better performance than node2vec and sub2vec since it can better preserve structural information from both local and global similarities among graphs. However, it still lacks utilizing the rich attribute information within transaction graphs. GCN and GAT can make effective use of node features, but edge features are only used as auxiliary information to participate in the message passing and updating of the graph. Our EGAT method overcomes this deficiency by synchronizing the update of node and edge features, thus capturing more information about transaction attributes while preserving structural information. As is shown in Table 3, we implemented some simple detection algorithms based on existing data and compared them with the identification results in other papers. Apparently, our experimental results showed superiority compared to the results of the related works.

Table 3. A comparison with the phishing detection results in other papers.

Method	Accuracy	Precision	Recall	F1-Score
RF	0.821	0.823	0.821	0.821
LightGBM	0.852	0.832	0.852	0.853
DElightGBM [30]		0.819	0.805	0.812
LightGBM [31]		0.943	0.956	0.949
[12]		0.813	0.827	0.819
MP-GCN [13]		0.932	0.932	0.932
EGAT	0.981	0.966	0.993	0.979

5. Conclusions

In this study, a simple star subgraph was designed for the detection of phishing account nodes in Ethereum, and the improved two-layer EGAT model was used for classification, with desirable results. The manually made features using basic transaction information effectively manifested the properties of subgraphs. EGAT with feature attention can automatically learn the feature weight of the graph and avoid the loss of edge features. The final experimental results show that the improved scheme in this study excels in Accuracy, Precision, Recall, and *F*1-score, respectively, especially compared with other baseline algorithms. Therefore, it is feasible to employ the EGAT phishing detection in Ethereum, which can successfully identify the phishing accounts so as to improve the security and stability of Ethereum.

Author Contributions: Data curation, X.Z.; funding acquisition, W.Y.; investigation, X.Z. and X.T.; methodology, X.Z.; project administration, X.Z.; software, X.Z.; supervision, X.Z., X.T.; validation, X.Z. and X.T.; writing—original draft, X.Z.; writing—review and editing, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research work is supported by the National Natural Science Foundation of China, grant Number: No.U1603115, Science and Technology Project of Autonomous Region, grant number: No. 2020A02001-1, and Research on short-term and impending precipitation prediction model and accuracy evaluation in Northern Xinjiang Based on deep learning, grant number: 2021D01C080.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shao, Q.F.; Jin, C.Q.; Zhang, Z.; Qian, W.N.; Zhou, A.Y. Blockchain: Architecture and Research Progress. *Chin. J. Comput.* **2018**, *41*, 969–988.
2. Guo, D.; Dong, J.; Wang, K. Blockchain: Graph Structure and Statistical Properties of Ethereum Transaction Relationships. *Inf. Sci.* **2019**, *492*, 58–71. [CrossRef]
3. Puthal, D.; Malik, N.; Mohanty, S.P.; Kougianos, E.; Das, G. Everything You Wanted to Know About the Blockchain: Its Promise, Components, Processes, and Problems. *IEEE Consum. Electron. Mag.* **2018**, *7*, 6–14. [CrossRef]
4. Zhu, L.H.; Gao, F.; Shen, M.; Li, Y.D.; Zheng, B.K.; Mao, H.; Wu, Z. Survey on Privacy Preserving Techniques for Blockchain Technology. *J. Comput. Res. Dev.* **2017**, *54*, 2170–2186.
5. Etherscan.io. Available online: https://etherscan.io/ (accessed on 10 December 2022).
6. Yuan, Y.; Wang, F.Y. Blockchain: The State of the Art and Future Trends. *Acta Autom. Sin.* **2016**, *42*, 481–494.
7. Han, L.; Yang, Z.; Jiang, Y.; Zhao, W.; Sun, J. Enabling Clone Detection for Ethereum via Smart Contract Birthmarks. In Proceedings of the 2019 IEEE/ACM 27th International Conference on Program Comprehension (ICPC), Montreal, QC, Canada, 25–26 May 2019; pp. 105–115.
8. The Chainalysis Crypto Crime Report is Here. Download to Learn Why 2019 Was the Year of the Ponzi Scheme. Available online: https://blog.chainalysis.com/reports/cryptocurrency-crime-2020-report/ (accessed on 10 February 2023).
9. Conti, M.; Kumar, E.S.; Lal, C.; Ruj, S. A Survey on Security and Privacy Issues of Bitcoin. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 3416–3452. [CrossRef]
10. Sun, G.Z.; Wang, J.T.; Gu, Y. Security Threat Analysis of Blockchain Technology. *J. Nanjing Univ. Posts Telecommun.* **2019**, *39*, 48–62.
11. Huang, B.; Liu, Z.; Chen, J.; Liu, A.; Liu, Q.; He, Q. Behavior Pattern Clustering in Blockchain Networks. *Multimed. Tools Appl.* **2017**, *76*, 20099–20110. [CrossRef]

12. Xia, Y.; Liu, J.; Wu, J. Phishing Detection on Ethereum via Attributed Ego-Graph Embedding. *IEEE Trans. Circuits Syst. II Express Briefs* **2022**, *69*, 2538–2542. [CrossRef]
13. Yu, T.; Chen, X.; Xu, Z.; Xu, J. MP-GCN: A Phishing Nodes Detection Approach via Graph Convolution Network for Ethereum. *Appl. Sci.* **2022**, *12*, 7294. [CrossRef]
14. Wang, Z.; Chen, J.; Chen, H. EGAT: Edge-Featured Graph Attention Network. In *Artificial Neural Networks and Machine Learning (ICANN)*; Springer: Cham, Switzerland, 2021; pp. 253–264.
15. Featuretools. Available online: https://docs.featuretools.com/ (accessed on 27 February 2021).
16. Graph-Based Forensic Investigation of Bitcoin Transactions. Available online: https://dr.lib.iastate.edu/handle/20.500.12876/28432 (accessed on 25 February 2021).
17. Monaco, J.V. Identifying Bitcoin users by transaction behavior. In Proceedings of the Biometric and Surveillance Technology for Human and Activity Identification XII, Baltimore, MD, USA, 14 May 2015; p. 945704.
18. Androulaki, E.; Karame, G.O.; Roeschlin, M.; Scherer, T.; Capkun, S. Evaluating User Privacy in Bitcoin. In *Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, 1–5 April 2013, Revised Selected Papers*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 34–51.
19. Reid, F.; Harrigan, M. *An Analysis of Anonymity in the Bitcoin System*; Springer: New York, NY, USA, 2012; pp. 179–223.
20. Sun Yin, H.H.; Langenheldt, K.; Harlev, M.; Mukkamala, R.R.; Vatrapu, R. Regulating Cryptocurrencies: A Supervised Machine Learning Approach to De-anonymizing the Bitcoin Blockchain. *J. Manag. Inf. Syst.* **2019**, *36*, 37–73. [CrossRef]
21. Lin, Y.J.; Wu, P.W.; Hsu, C.H.; Tu, I.P.; Liao, S.W. An Evaluation of Bitcoin Address Classification Based on Transaction History Summarization. In Proceedings of the 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Seoul, Republic of Korea, 14–17 May 2019; pp. 302–310.
22. Toyoda, K.; Ohtsuki, T.; Mathiopoulos, P.T. MultiClass Bitcoin-Enabled Service Identification Based on Transaction History Summarization. In Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canada, 30 July–3 August 2018; pp. 1153–1160.
23. Toyoda, K.; Ohtsuki, T.; Mathiopoulos, P.T. Identification of High Yielding Investment Programs in Bitcoin via Transactions Pattern Analysis. In Proceedings of the GLOBECOM 2017–2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–6.
24. Kanemura, K.; Ohtsuki, T.; Toyoda, K. An Evaluation of Bitcoin Address Classification Based on Transaction History Summarization. In Proceedings of the International Conference on Blockchain and Cryptocurrency (ICBC), Seoul, Republic of Korea, 14–17 May 2019; pp. 302–310.
25. Chen, W.; Zheng, Z.; Ngai, E.C.H.; Zheng, P.; Zhou, Y. Exploiting Blockchain Data to Detect Smart Ponzi Schemes on Ethereum. *IEEE Access* **2019**, *7*, 37575–37586. [CrossRef]
26. Wen, T.; Xiao, Y.; Wang, A.; Wang, H. A novel hybrid feature fusion model for detecting phishing scam on Ethereum using deep neural network. *Expert Syst. Appl.* **2023**, *211*, 118463. [CrossRef]
27. Kabla, A.H.H.; Anbar, M.; Manickam, S.; Karupayah, S. Eth-PSD: A Machine Learning-Based Phishing Scam Detection Approach in Ethereum. *IEEE Access* **2022**, *10*, 118043–118057. [CrossRef]
28. Ju, Y.; Sun, G.; Chen, Q.; Zhang, M.; Zhu, H.; Rehman, M.U. A Model Combining Convolutional Neural Network and LightGBM Algorithm for Ultra-ShortTerm Wind Power Forecasting. *IEEE Access* **2019**, *7*, 28309–28318. [CrossRef]
29. Jourdan, M.; Blandin, S.; Wynter, L.; Deshpande, P. Characterizing Entities in the Bitcoin Blockchain. In Proceedings of the 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018; pp. 55–62.
30. Chen, W.; Guo, X.; Chen, Z.; Zheng, Z.; Lu, Y. Phishing Scam Detection on Ethereum: Towards Financial Security for Blockchain Ecosystem. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, Yokohama, Japan, 7–14 January 2021; pp. 4506–4512.
31. Bian, L.; Zhang, L.; Zhao, K.; Shi, F. Detection Method of Ethereum Malicious Account Based on LightGBM. *Inf. Netw. Secur.* **2020**, *20*, 73–80.