Traveling the token world: A graph analysis of Ethereum ERC20 token ecosystem

Weili Chen
School of Data and Computer Science, Sun Yat-sen University
National Supercomputer Center in Guangzhou
Guangzhou, China
chenwli28@mail2.sysu.edu.cn

Tuo Zhang
School of Data and Computer Science, Sun Yat-sen University
National Engineering Research Center of Digital Life
Guangzhou, China
741554448@qq.com

Zhiguang Chen
School of Data and Computer Science, Sun Yat-sen University
National Supercomputer Center in Guangzhou
Guangzhou, China
zhiguang.chen@nscc-gz.cn

Zibin Zheng
School of Data and Computer Science, Sun Yat-sen University
National Engineering Research Center of Digital Life
Guangzhou, China
zhzibin@mail.sysu.edu.cn

Yutong Lu
School of Data and Computer Science, Sun Yat-sen University
National Supercomputer Center in Guangzhou
Guangzhou, China
yutong.lu@nscc-gz.cn

ABSTRACT

The birth of Bitcoin ushered in the era of cryptocurrency, which has now become a financial market attracted extensive attention worldwide. The phenomenon of startups launching Initial Coin Offerings (ICOs) to raise capital led to thousands of tokens being distributed on blockchains. Many studies have analyzed this phenomenon from an economic perspective. However, little is know about the characteristics of participants in the ecosystem. To fill this gap and considering over 80% of ICOs launched based on ERC20 token on Ethereum, in this paper, we conduct a systematic investigation on the whole Ethereum ERC20 token ecosystem to characterize the token creator, holder, and transfer activity. By downloading the whole blockchain and parsing the transaction records and event logs, we construct three graphs, namely token creator graph, token holder graph, and token transfer graph. We obtain many observations and findings by analyzing these graphs. Besides, we propose an algorithm to discover potential relationships between tokens and other accounts. The reported case shows that our algorithm can effectively reveal entities and the complex relationship between various accounts in the token ecosystem.

KEYWORDS

Blockchain, Ethereum, ERC20 token, Graph analysis

ACM Reference Format:

Weili Chen, Tuo Zhang, Zhiguang Chen, Zibin Zheng, and Yutong Lu. 2020. Traveling the token world: A graph analysis of Ethereum ERC20 token ecosystem. In *Proceedings of The Web Conference 2020 (WWW '20), April 20–24, 2020, Taipei, Taiwan.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3366423.3380215

WWW '20, April 20–24, 2020, Taipei, Taiwan

ACM ISBN 978-1-4503-7023-3/20/04.
https://doi.org/10.1145/3366423.3380215

1 INTRODUCTION

Since the creation of Bitcoin in 2009 [28], cryptocurrencies arouse great interest among researchers, developers, and investors. The technology underpinning cryptocurrencies is what we called blockchain technology [39]. Technically speaking, a blockchain is a distributed, append-only ledger database [38]. It usually maintains a native cryptocurrency that can be exchanged with fiat money through *cryptocurrency exchanges* (e.g., Binance[1]).

The interchangeability between cryptocurrency and fiat money makes it possible to raise funds based on cryptocurrency, which directly leads to the birth of the cryptocurrency economy. Thus, initial coin offerings (ICOs) become a new method of raising capital for start-ups in the cryptocurrency ecosystem. Different from traditional sources of start-up funding such as venture capital (VC) and angel finance, an ICO issuer raised cryptocurrency through selling blockchain-based cryptographically secured digital assets (usually called *tokens*) to any participants. According to one estimate, ICOs raised over $31 billion between January 2014 and August 2018, and the top token sell (EOS token) raised over $4 billion[2].

A token represents a programmable asset or access right provided by its issuer, managed by a smart contract and the underlying blockchain platform [37]. Thus, the choice of the blockchain platform is crucial to ICOs. Ethereum[3], an open-source platform for decentralized applications, is the first blockchain platform that simplifies the development of smart contracts. Based on Ethereum, one can create a token smart contract with just a few lines of code. Thus, Ethereum became the main platform for ICOs; and it accounts for

[1]https://www.binance.com/
[2]See https://www.coinschedule.com/stats/ALL?dates=Jan+01%2C+2014+to+Aug+10%2C+2019
[3]https://www.ethereum.org/

over 80% in the blockchain platform market share in ICO according to a statistic[4].

Just as stock investors might want to exchange their shares, token holders might also want to exchange their specific tokens. Thus, the Ethereum community launched the ERC20 token standard in November 2015 to facilitate token development and exchange (See Section 2.3 for detailed information). Although it's not a mandatory standard, it has become the de-facto standard that developers consciously adhere to since then. As a result, most of the tokens released through Ethereum are ERC20 compliant. As of September 1st, 2019, more than 160,000 ERC20 compatible tokens exist on Ethereum platform[5].

Undoubtedly, the token economy has become an economic phenomenon that cannot be ignored [37]. To study the new ecosystem has drawn great attention among researchers [1, 8, 10, 14, 16, 19, 27]. However, most of these studies focus on economic issues, there is little analysis of users and tokens in the token ecosystem. In fact, there may be complex relationships between tokens and users. Figure 1 shows an *entity* on Ethereum that might control a large number of accounts (i.e., addresses), including external owned accounts (EOA), smart contract owned accounts (COA) and token owned accounts (TOA). (Please refer to Section 2.2 for a detailed explanation of different accounts.) In the figure, *0x60* is a COA, which created 3,253 tokens; *0x19* is the address of the prediction market Augur[6]; And *0x51* is the address of a smart contract which is part of the augur implementation. (For writing convenience, in this paper, a four-character identifier beginning with *0x* represents an Ethereum address.) There are many other accounts. However, by using our proposed algorithm (Figure 7) and the defined relationships, we find that these seemingly unrelated accounts may belong to the same entity. The definition of relationships and more details will be discussed in section 4.4, however, this result indicates that the token ecosystem is more complicated than expected.

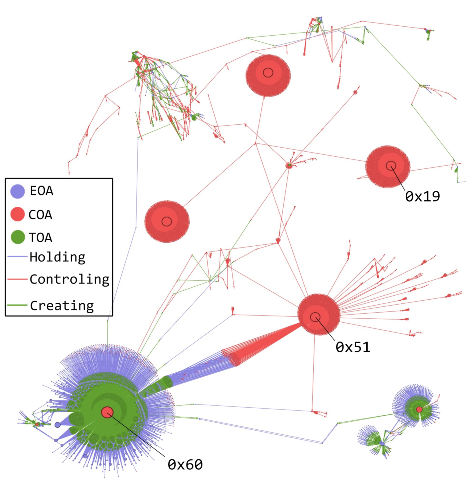


Figure 1: The relationship between accounts in the found Augur entity.

[4]see https://icowatchlist.com/statistics/blockchain
[5]See https://etherscan.io/tokens
[6]https://www.augur.net/

Compared with the traditional investment ecosystem, the publicly available of the blockchain ledger provides us with unprecedented opportunities to study the new economic ecosystem. In an attempt to fill the gap in research and reveal the characteristics of the token ecosystem, this paper presents an approach for analyzing the ERC20 token ecosystem. Figure 2 shows the framework of our study. As can be seen, our approach consists of three phases, which are detailed in the following sections. The first phase, *data collection* (Section 3), collects and prepare all transaction records, event logs and some descriptive texts for the subsequent analysis. Then, we divided the accounts of Ethereum into three categories and discussed the token ecosystem characteristics by constructing the token creator graph (TCG) and token holder graph (THG) (Section 4). In the last phase, we focus our analysis on a special token by the token transfer graph (TTG) (Section 5).

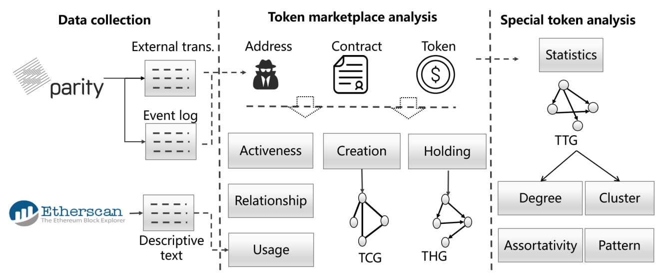


Figure 2: An overview of our framwork.

In summary, we make the following contributions.

(1) We conduct a systematic investigation on the whole Ethereum ERC20 token ecosystem (over 160,000 tokens) via graph analysis. We adopt a new method to collect all transaction records and event logs and then construct a token creator graph (TCG), toke holder graph (THG), and token transfer graph (TTG) to outline the characteristics of the token ecosystem. The data and code can be found on our home page xblock.pro[7].
(2) We obtain many new observations and findings of Ethereum token ecosystem by adopting graph analysis and other methods. They help us obtain a new understanding of the Ethereum token ecosystem. In particular, we find that in the decentralized exchange, the phenomenal of fake transaction volume also exists.
(3) We propose an algorithm to identify entities in the token ecosystem based on cross-graph analysis and the defined relationships. The reported case shows that our algorithm can effectively reveal the complex relationship between various accounts in the token ecosystem.

The rest of the paper is organized as follows. After providing some background on Ethereum and ERC20 standard in Section 2, we detail our data collection method in Section 3. Section 4 answer 5 different questions about the whole token marketplace by adopting different methods. Section 5 discusses the dynamic characteristics of tokens by analyzing the token transfer graph on the IDEX exchange.

[7]http://xblock.pro/research/fraud-detection-en/

After reviewing related work in Section 6, we conclude the paper and discuss future work in Section 7.

2 BACKGROUND

2.1 Blockchain and Ethereum

Simply speaking, a blockchain is a globally shared, distributed transactional database. Everyone can read and write transactions into the database by participating in the network. It uses a certain type of consensus mechanism to validate transactions and keep consistency. Records in the database are time-stamped and digitally signed, which make them immutable and traceable. In the blockchain system, there is no central authority, all system changes depending on the predetermined consensus mechanism, which makes it tamper-resistant.

Ethereum is the largest programmable blockchain platform. By introducing an Ethereum Virtual Machine (EVM), it can support smart contracts of any complexity. To motivate peers to participate in the system maintenance and prevent potential abuse of the system, a cryptocurrency called *ether* (ETH) is created. It is now the second-largest cryptocurrency by market capitalization.

2.2 Account and Transaction

Accounts are a core component of a blockchain system, and transactions are records that change the state of accounts. Ethereum has two kinds of accounts: 1) external owned accounts (EOAs) that are controlled by public-private key pairs (i.e., humans) and 2) contract accounts controlled by the code stored together with the account. Since tokens are controlled by contracts, in order to distinguish token contract accounts from other contract accounts, we call them token owned accounts (TOAs) and contract owned accounts (COAs) respectively. All accounts are referred by its *address* and denoted as a four-character identifier beginning with *0x* in this paper.

A *transaction* is a message sent from one account to another which can carry binary data (called "payload") and ether. The payload can be code to deploy, the function to invoke and the corresponding parameters, and the receiver of the message. There are two kinds of transactions depending on the message sender. The transactions sending from an external account are called "external transactions", which will be included in the blockchain and can be obtained by parsing the blocks. The other type, sending by a smart contract to another account, is called "internal transactions". Internal transactions are usually triggered by external transactions and are *not* stored in the blockchain.

2.3 Smart Contract, Token, ERC20 Standard

The idea of smart contracts is coined by Nick Szabo in 1994 [31]. Based on a blockchain platform, a smart contract can be seen as some lines of code. It will be auto-executed and can not be stopped when pre-set conditions are met. As mentioned, Ethereum platform is the biggest smart contract platform, it provides some high-level programming language, such as solidity, to implement a smart contract.

In the Ethereum platform, everyone can take advantage of the blockchain technology to build their projects or DAPPs (distributed applications) through smart contracts. Crowd-sale based on smart contract is usually called ICO (initial coin offering), in which one buys certain *tokens* of that DAPP with Ether. The token may act like a native currency or identify a sort of shareholder in the DAPP.

To make it easier for developers to handle different tokens, Ethereum community introduced the ERC20 token standard. The standard includes serveral unimplemented functions and events. Required by ERC20 standard, the event *Transfer* must be emitted every time when tokens being transferred. Thus, if we parse the event logs, we can know how the tokens are transferred, where they go, and by whom they are held.

3 DATA COLLECTION

We launch Parity[8], an Ethereum client, on our server to download the ledger of Ethereum. We first download all the blocks before July 6th, 2019 (from the very first block to block 8,099,999). Then, by using Parity's APIs, we reorganize the data extracted from the blocks into 4 parts: external transactions, internal transactions, contract information, and contract callings.

From the contract information data set, we can get bytecodes and creators of all contracts. By scanning these bytecodes, we can determine whether a contract implements a token. Through this method, we find over 160,000 ERC20 tokens and their creators on Ethereum.

Each external transaction contains all the events emitted by the transaction. That means if a transaction calls a contract function, and the contract emits an event, the event will be logged with the transaction. Each event contains a few *topics*, which facilitate users to subscribe specific events. Figure 3 shows an example log of a standard ERC20 token transfer event. Topic 0 is always the hash of the event type. As shown in the figure, the hash of the event *Transfer* is *0xdd..ef*. Topic 1 and topic 2 are the address of the sender and the receiver. In the picture, the sender address is *0x8d..31* and the receiver address is *0x08..2e*. The amount of transferred tokens is stored in the data field, which is *0x1c..00* in the example, and it equals 33.4108612 tokens (or $33.4108612 \times 10^{18}$ in its smallest token unit). To find out all the ERC20 transfer events, we go through the external transaction data and parse all the event logs whose topic 0 matches *0xdd..ef*. Once we get those logs, we know how the tokens are transferred.

```
> Transfer
(index_topic_1 address from, index_topic_2 address to, uint256 value)
[topic0] 0xddf252ad1be2c89b69c2b068fc378daa952ba7f163c4a11628f55a4df523b3ef
[topic1] 0x0000000000000000000000008d738b3ddf8425e155f52caaaedc65ab28234f31
[topic2] 0x0000000000000000000000008313a839d119c5c06c0a3aa527beb7dcb56ad2e
Hex → 000000000000000000000000000000000000000000000001cfab393e39e1a000
```

Figure 3: A log of ERC20 token transfer event.

4 MARKETPLACE CHARACTERISTCIS

In this section, we provide an overview of the token ecosystem before discussing special tokens. By parsing the blocks and the event logs, we find out that there are 165,955 ERC20 tokens, 30,008,087 users participated in the ecosystem and 227,698,645 token transfer transactions. It should be pointed out that there are over 200,000 tokens according to Etherscan.io. However, we found that some

[8] https://www.parity.io/ethereum/

of them are not standard ERC20 tokens. Besides, there are almost 170,000 tokens listed on eth.btc.com[9], which indicates that our method finds all the ERC20 tokens. In the following, we focus our study on the found ERC20 tokens and try to answer five questions. Based on these analyses and results, we obtain the following findings.

- **Finding 1.** Although there are a lot of tokens, which makes it seems that the token economy is very prosperous, but in fact, only a few tokens are active, and most of them do not have much value.
- **Finding 2.** Tokens can be created either by a person or automatically with a smart contract. Some addresses create a large number of tokens, possibly by creating token contracts to attack the Ethereum network.
- **Finding 3.** A small number of accounts hold a large number of tokens, while a large number of accounts only hold a small number of tokens. Similarly, a small number of tokens have a large number of holders, while a large number of tokens have a small number of holders.
- **Finding 4.** There are very complex relationships in the token ecosystem. By analyzing these relationships, we can have a deep understanding of the tokens' behavior and even reveal the entity.

4.1 How Active Are These Tokens?

The degree of activeness of a token is an indication for the degree of health of the economy using the token. The number of transfer transactions of a token can be seen as the activeness of the token to some extent. To this end, we present the distribution of the activeness of tokens in Figure 4.

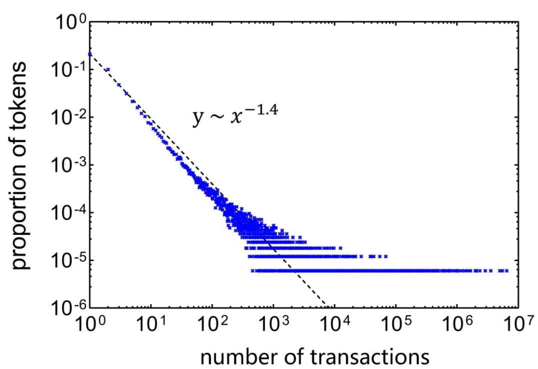
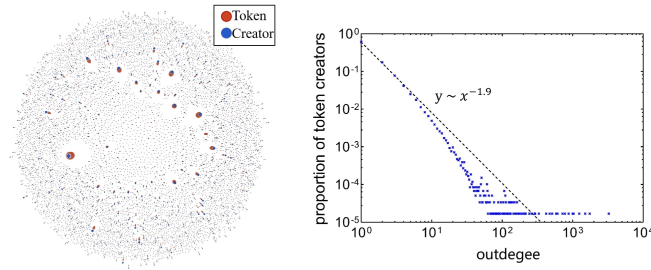


Figure 4: Distribution of activeness of tokens.

As can be seen, many tokens have never been transferred (35.67%) and over 90% tokens transferred no more than 65 times. This result indicates that the majority of the token is not succeeded in the aspect of user activity. Furthermore, there is an obvious power-law distribution of user activeness. That is to say, there are a few very active tokens and many inactive tokens. We plot the fitted line $y \sim x^{-\alpha}$ for the distribution. The larger the α, the less variable of the tokens' level of activeness. When focusing on the most active tokens, we found that the most active token is *MGC* token (6,448,768 transactions). According to its website[10], MGC token, officially launched on April 26, 2019, is "a decentralized payment application, creating a multi-functional digital business system that can be applied across international fields." Compared with other tokens, it is undoubtedly successful in terms of user activity. The second most active token is LIVEPEER TOKEN (LPT). Livepeer, launched two years ago, is a blockchain-based service platform for developers who want to add live or on-demand videos to their project according to its website[11]. An in-depth analysis of user behavior in these tokens may be an interesting topic, but it is out the range of this paper and will be discussed in our future work.

[9]https://eth.btc.com/



(a) TCG (b) Outdegree distribution of TCG

Figure 5: Visualization of Token Creator Graph (TCG) and its outdegree distribution.

4.2 Who Created These Tokens?

Due to the anonymous nature, it is difficult to reveal the identity of a token creator, as an address is enough to create a token on the Ethereum platform. To investigate the token creator relationship, we introduce the token creator graph (TCG) and try to infer the identity of some creators.

TCG Definition and Construction. $TCG = (V, E)$, where V is a set of accounts (i.e., EOAs, COAs, and TOAs) and E is a set of edges. $E = \{(v_i, v_j)|v_i, v_j \in V\}$ is a set of ordered pairs of nodes (i.e., accounts or addresses); The order of the edge indicates the creation relationship, i.e., an address V_i created a token V_j (we treat a token and its address as the same). Thus, TCG is a directed graph. By going through the contract information data set, the TCG is easily constructed.

To get an overall impression of TCG, we randomly select 10,000 edges and show the result in Figure 5(a). In the graph, token nodes are in red and creators are in blue, the size of the nodes represents the number of tokens they created. As can be seen, some nodes created a large number of tokens. It is surprising for a user to create so many tokens, as a token usually represents some rights or used as a currency in a DAPP.

To further analyze TCG, we plot the outdegree distribution of it. The outdegree of a node in TCG indicates the number of tokens created by it. Figure 5(b) shows the outdegree distribution of TCG. As can be seen, it follows the power-law distribution, which

[10]http://www.mgctoken.io/
[11]https://livepeer.org/

means that there are a few large-outdegree nodes and many small-outdegree nodes. The fitted line $y \sim x^{-\alpha}$ for the distribution is plotted. Generally, the larger the α, the less variable of the nodes' outdegree. By carefully analyzing the relevant data, we found that many creators (60.3%) created only one token, 92.3% of the creators created no more than 5 tokens, and 99.04% of the creators created less than 20 tokens. However, as can be seen, some accounts created more than 1000 tokens.

To reveal the characteristics of the creators, we counted the number of creators of each account category. Table 1 shows the creator category and the number of tokens created by it. The same as we guessed, a lot of tokens are created by external accounts (i.e., by humans), some tokens are created automatically (i.e., by smart contracts). The most surprising result is that some tokens are created by token contracts. Of course, a token contract is also a smart contract, thus it can automatically create tokens when necessary. But by doing so it might make the token contract unreadable. Due to there are a large number of creators, we then focused our study on the accounts created the most tokens. The account with the most number of token creation is *0x60*[12]. It created 3,253 tokens. Because this is weird, revealing the identity of the creator is of great interest to us. To this end, we try to find all the relevant information related to *0x60*. First of all, by searching the address with Google we cannot find any useful information. Then, we try to look for possible clues in the contract source code and comment lines in Etherscan. We find a keyword *augur* in the contract code, which suggests to us that this contract may be a part of the prediction market Augur[13]. Augur was created on July 2018, which allows users to bet tokens to predict the outcome of certain events.

To confirm the relationship between *0x60* and the Augur contract, we propose an algorithm (i.e., Figure 7) to find all the addresses associated with an account (i.e., the entity [20]). The algorithm will be explained after analyzing various graphs that we defined. By using the *finding entity* algorithm, we find that *0x60* associated with Augur contract[14]. Furthermore, by analyzing the transaction records and event logs, we found that every transaction which creates a new token by *0x60* triggered an Augur token transfer transaction. These facts make us confident that *0x60* belongs to Augur.

The account with the second most number of token creation is *0x67*[15]. It is an external owned account (i.e., a human) who created 1,740 tokens. It is difficult to determine the identity of the creator, however, by analyzing the token contracts created, we found that none of the contracts had source code and the bytecode was very similar, so the tokens created were likely to have very similar functions. Strangely, someone would create so many tokens that not only do not provide the source code but are likely to have very similar functions. One possible reason is that someone attacked the Ethereum network by creating tokens.

4.3 Who Hold These Tokens?

If a token can be regarded as a small economy, its holders should be as widely as possible. This section discusses the token holder characteristics. To this end, we first define and construct a token holder graph (THG) as follows.

Table 1: Token creator category and the number of created tokens.

Category	Number
EOA	55656
COA	2298
TOA	1269

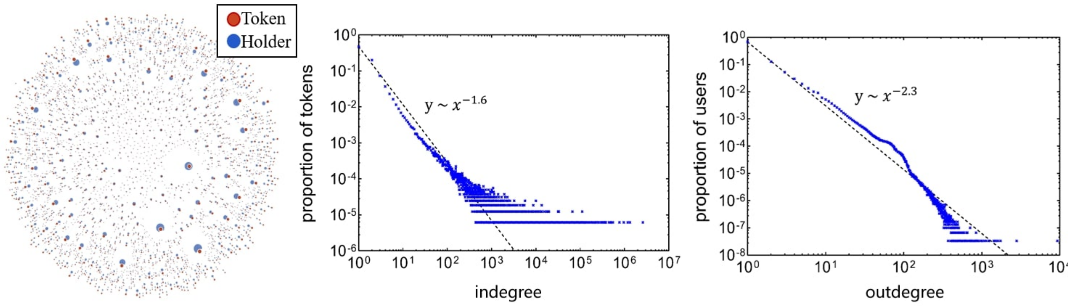
THG Definition and Construction. $THG = (V, E, w)$, where V is a set of accounts and E is a set of edges, the same as in TCG definition. $E = \{(v_i, v_j)|v_i, v_j \in V\}$ is a set of ordered pairs of nodes; The ordered of the edge indicates the holding relationship (i.e., a node V_i holds a token V_j). $w : E \rightarrow \mathcal{R}^+$ associates each edge with a weight, which indicates the node holds w shares of the token. Hence, THG is a weighted directed graph. To construct the graph, we parsed all the token transfer logs. For each address and each token in the logs, we add up all their received tokens and subtract all their sent tokens. This allows us to calculate the most recent balance of a certain token an address holds. If the balance is larger than 0, the address is considered a holder of the token. By doing so, we can get both all the holders of a token and how many tokens and its shares are held by an account. Based on this information, it is easy to construct THG.

Figure 6(a) shows a sampled THG. The blue points denote accounts (please note that an account may hold many kinds of tokens) and the red points denote tokens. The size of the points indicates the number of holders (for token nodes) or the number of held tokens (for an account). As can be seen, some tokens have a lot of holders. The number of tokens held by an account may be much less than the number of holders of a token, thus the size of the blue points are very small.

Figure 6(b) and 6(c) show the indegree and outdegree distribution of THG. In THG, the indegree of a token represents the number of holders of the token, the outdegree of an account is the number of tokens it holds (not the share of a special token). As shown in Figure 6(b), the distribution also follows a power law. Almost half of the tokens (47%) have only one holder or no holder. 90% of the tokens have no more than 33 holders. This shows from another angle that most tokens are not active enough, and the token economy is still in its initial stage. The token with the most holders is *0x58*[16], the second most active token mentioned in Section 4.1. It has 2,601,321 holders. Consider that it is a distributed live or on-demand video service platform, such holder and transaction size suggests that it has constructed a new ecology as compared with other tokens.

Similarly, as shown in Figure 6(c), the distribution of outdegree also conforms to the power law. An in-depth analysis shows that 68.2% of the accounts hold only one token. More than 90% of accounts hold no more than 5 tokens. The largest holder is a distributed exchange, called EtherDelta, which hold 9,312 tokens.

[12]0x60a977354a6ba44310b2ee061bcf19632450e51d
[13]https://www.augur.net/
[14]0x1985365e9f78359a9b6ad760e32412f4a445e862
[15]0x67c838cd6e0ad4487a279f8286ee8673968bd615
[16]0x58b6a8a3302369daec383334672404ee733ab239



(a) THG (b) Indegree distribution of THG (c) Outdegree distribution of THG

Figure 6: Visualization and the indegree/outdegree distribution of THG.

EtherDelata was very popular amongst traders and was the go-to platform for ERC20 traders in 2017[17]. However, the exchange broke down after the SEC charges its founder with operating an unregistered exchange[18].

4.4 What is the Relationship Between Tokens?

In the previous study, we found that a user may create multiple tokens, and a token may also have multiple holders. Then, is there any relationship between different tokens, such as whether they are created or controlled by the same person? In this subsection, we try to answer this question. To this end, we provide an algorithm named *finding entity* as shown in Figure 7.

The input of the algorithm is a token creator (x), the transaction records (TR), and the two constructed graphs (TCG and THG). By using the proposed algorithm, we hope that we can find an address set which "controlled or owned" by the same entity. Three kinds of relationships between addresses are considered: *creating*, *holding* and *controlling*. The creating relationship between addresses is the same as in the token creator graph (TCG). That is to say, if token *A* is created by another address *B*, there is a creating relationship between address *A* and *B*. Similarly, the holding relationship comes from the token holder graph (THG), with a slight modification. Specifically, if there is a holding relationship between an address and a token, it means that the address holds more than 50% of the share of the token-in-circulation. Please note that we only consider the token-in-circulation, because for some tokens, the total supply of the token is hard to obtain and it can be changed after the token deployed. If an address *A* can trigger another address *B* to transfer its token to the third address *C* through a transaction, then we assume that address *A* controls address *B*, or there are *controlling* relationship between addresses *A* and *B*. Similarly, if address *A* can create tokens by invoking one or more related contracts (with the same transaction hash), we assume that there is a controlling relationship between address *A* and those contracts. This is not always true, as there may be a contract that can be used for all users to create tokens. However, in general, to create a useful token, a user must know how to control it. This requires the user to have an in-depth understanding of the contracts that create tokens, indicating that the user may control these contracts (or create these contracts).

[17]https://thebitcoinnews.com/etherdelta-exchange-review-2019-guide/
[18]https://www.sec.gov/news/press-release/2018-258

Besides, for some contracts without publishing its source code, only its creator (and thus the controller) knows the full details of its contract. If any of the three relationships exist between any two addresses, we consider they may belong to or may be controlled by the same entity. Although it is impossible to verify that an address set belongs to an entity due to the anonymous nature, this algorithm gives us a better understanding of the potential relationship between tokens and other accounts.

Next, let's briefly explain what each part of the algorithm means. To find all the tokens and other accounts that may be owned or controlled by the same entity, we start the algorithm with the address of a creator, denoted as *x*. The *associated_set* records all the addresses associated with *x*. Next, in line 3, we construct the *invocation_graph* by find all addresses (i.e., nodes) and corresponding invocation transactions (i.e., edges) associated with token creation. The whole while loop (line 4 ∼ 14), based on *invocation_graph* , iterates continuously to find all addresses that have a controlling relationship with *x* (direct or indirect). The code after line 15 mainly divides the association set (i.e., *associated_set*) into different subsets. The creators set (i.e., *creator_set*) is constructed by checking whether the address is a token creator (line 15 ∼ 16); The token set (i.e., *token_set*) consists of all tokens created by a creator in *creator_set*; The holder set (i.e., *holder_set*) composed of addresses that have a holding relationship with at least one token in the token set.

Figure 1 shows the relationship between the accounts of the Augur entity found by our algorithm with start address *0x60*. As can be seen, in addition to *0x60*, the entity contains several contracts creating tokens. This result vividly shows that our algorithm can discover various potential relationships between accounts, which is beneficial to reveal the internal relationship between tokens. To understand the distribution of other entities, we applied our algorithm to addresses which created more than 1000 tokens. By doing this, we found quite several large entities. Some of them are token-issuing smart contracts, and others we cannot infer the identity. Since we do not intend to conduct an in-depth analysis of the entities, the relevant results will not be shown to save space.

4.5 What Are These Tokens Used For?

As there are over 160,000 tokens and many tokens do not have relevant introductory information, it is impossible to verify the

```
Inputs: x, a given token creator address to start with
    TR, transaction records
    TCG, token creation graph
    TKG, token holding graph
Outputs: Creators Set, a set of token creators associated with x
    Associated Set, a set of accounts or contracts associated with x
    Tokens Set, a set of tokens created by Creators Set
    Holders Set, a set of token holders that have control on one of the tokens in Tokens Set,
which means each holder holds more than 50% of total supply of a token
associated_set = [x];
n1, n2 = 0, 1;
invocation_graph = [nodes and eages for each edge in TR if edge.transactionHash in TCG]
while n1 < n2
    n1 = size(associated_set);
    for each address a1 in associated_set and each address a2 in invocation_graph.nodes
        if hasedge(a1,a2, invocation_graph) and a2 not in associated_set
            if a1.type == contract and a2.type == contract
                //a1 invocates a2 or a2 invocates a1
                associated_set += a2;
            else if a1.creator == a2 or a2.creator == a1
                //a1 is creator of a2 or a2 is creator of a1
                associated_set += a2;
    n2 = size(associated_set);
for each address a in associated_set if a is token creator
        creators_set += a, associated_set -= a;
tokens_set = getTokens(creators_set, TCG);
for each address t in tokens_set
    h = getTopHolder(t, TKG);
    if h.holdAmount(t) > t.totalSupply * 0.5
        holders_set += h;
return creators_set, associated_set, tokens_set, holders_set;
```

Figure 7: Pseudo-code of finding entity algorithm.

functionality of each token. Thus, we focus our classification on the ERC20 top Tokens listed on the blockchain explorer etherscan.io. To understand the purpose of the tokens, we try to classify the tokens by reading the descriptive texts of the token. To this end, we first crawl on September 1st, 2019 the descriptive texts of all the tokens. There are 968 out of 165, 955 tokens having descriptive text. For a first impression, we draw a word cloud graph of the text as shown in Figure 8 (The most frequent word "blockchain" is deleted). As can been seen, the tokens may have different usage scenarios. However, some typical scenarios may be popular in the token ecosystem. For example, platforms, digit currency, exchanges, and markets.

A taxonomy of Tokens. Next, we read the descriptions of all the tokens and their web site when unsure of their category. However, it is impossible to present an accurate classification, as many tokens have multiple functions. Therefore, we give a rough classification of tokens based on the description texts. Roughly speaking, we divide the tokens into five categories in the following:

- *Digital money:* A digital money aims to create a digital currency based on blockchain for the digital world. Among these tokens, the stable coin is an important type. A stable coin is an attempt to solve the problem of extreme fluctuations in the price of cryptographic digital currency and create a digital currency with stable value [30].
- *Exchange:* Exchanges allow users to buy, sell and trade with tokens, cryptocurrencies, fiat money and so on. It could be centralized, running by a company or organization. It also could be decentralized, with the whole transaction process controlled by some smart contracts. A token of exchange may represent shares of the exchange or the medium within the exchange.

Figure 8: Word Cloud for descriptive texts of the Tokens.

- *Service platform:* Service platform contains the most extensive content. In essence, any supporting platform that utilizes blockchain technology to solve industry problems can be regarded as a service platform. Broadly speaking, the exchange mentioned above can also be viewed as a service platform. Because the keyword exchange appears explicitly in the description text s, we put it into a more accurate classification. A service platform usually involves the development of a chain; the token is a fundraising tool and may represent the holder's future interests.
- *Wallet:* The wallet's function is relatively clear. A wallet is a digital bank of the blockchain era, which allows users to manage their cryptographic digital assets (e.g., Bitcoin, tokens). The token of the wallet is more a financing tool of the project side.
- *Marketplace:* A marketplace token represents a project that realizes commodity trading based on blockchain technology. The token is not only the medium of commodity exchange but also an incentive for all parties to participate in it.

Although the classification may not be accurate and comprehensive, it is good for us to understand the general status of token ecology. Based on the above classification, we counted the number of tokens, the average market capitalization, the average number of holders, and the average of 24-hour trading volume in each category. The original data was crawled from ethescan.io on September 7, 2019. Table 2 shows the statistics. As can be seen, in terms of the number of tokens, various service platforms based on blockchain account for the overwhelming majority (73%). But in terms of average market capitalization, exchanges are the highest. This indicates that although many blockchain projects have been proposed to solve the industry problems, the exchange is probably the most favored by the capital market. As for the average number of holders, the marketplace is undoubtedly the largest, because, for many marketplaces, users must hold the corresponding token to participate in the market. However, in terms of 24-hour trading volume, digital money token is significantly higher than other types. This is due to the very high demand for stability coins among cryptographic digital currency participants, as it provides a "safe harbor" for participants to store and exchange their digital assets. Although the capitalization share of stable coins is very low in the cryptocurrency market, the stable currency trading volume is huge. For example, tether[19], a stable coin which converts cash into digital currency by anchoring the value to the price of national currencies like the US dollar, has become the most heavily traded cryptocurrency according to coinmarketcap.com.

5 SPECIAL TOKEN ANALYSIS

In the cryptocurrency economy, exchanges are important to activate the market. However, many illegal activities such as price manipulation exist in centralized exchange (see [11, 17]), thus decentralized exchanges are considered to be fair trading venues, as all their trades are automatically completed by smart contracts. To see if decentralized exchanges are really what cryptocurrency investors are hoping for, we did a preliminary analysis of IDEX[20], a famous decentralized ERC20 token exchange according to stateofthedapps.com. we obtain the following findings by analyzing the IDEX exchange.

- **Fingding 5.** Although various tokens aspire to be a digital currency, *ether* is the absolute payment currency in the exchange.
- **Finding 6.** Decentralised exchanges may be fairer, but there are also fake transactions. The reason, of course, may be to raise the profile of the exchange. This shows from another angle, decentralized exchange is not the first choice of ordinary traders.

5.1 Data parsing

Unlike tokens, there is no standard for exchange smart contracts. The structures of event logs and contract functions differ between exchanges, which makes it almost impossible to establish a unified exchange data extraction method. To trade on IDEX, users need to deposit their tokens, or ether into their balance in IDEX's smart contract. Then, they can initiate an order to sell their tokens or buy other tokens; They can also satisfy existing orders by calling functions. Finally, users can withdraw whatever kinds of tokens in their balance into their Ethereum accounts. Besides, the contract only emit events when users deposit or withdraw their tokens. The function *trade*, which transfers tokens between sellers and buyers, does not emit any event. To find the data of token transferring, we need to lookup data of contract calling instead of event logs. By scanning the contract calling records, we extract the following useful information: *taker* and *maker* (or buyer and seller), *buyToken* and *sellToken* (the token to be traded), *buyAmount* and *sellAmount* (trading volume).

By parsing the function calling records, we obtain 3,642,588 trading records. Each record can be represented as a 6-tuple (*taker*, *maker*, *takerToken*, *makerToken*, *takerAmount*, *makerAmount*), which means the *taker* gives *takerAmount* of *takerToken* to the *maker*, and in return, he/she gets *makerAmount* of *makerToken* from the *maker*. Users of IDEX can make orders of buying or selling tokens at any price of ETH. They can also fill the orders made by other users. If a user makes an order, the user will be seen as an order maker, and if he fills an order, he will be seen as an order taker.

[19]https://tether.to/
[20]https://idex.market/

5.2 Token Transfer Analysis

5.2.1 What is the medium in the exchange? Ether, the token provided by Ethereum itself, is the most traded in IDEX. In fact, Ether is traded 3,642,346 times, nearly the same number of trading records. That is to say, almost all of the trading is either selling ether or buying with ether. Thus, ether is de facto "money" in the exchange; Direct exchange between tokens without ether is difficult to achieve.

5.2.2 What token is being bought and sold? 1,130 different tokens (except ether) are traded on IDEX and each token is traded 3,223 times on average. The most traded token is QNT (whose address is 0x4a..75), and it is traded 101,264 times. To compare the difference in the intention of buying or selling a token, we divide the records into bought records or sold records, depending on whether the order was originally initiated by the buyer or the seller. The difference between sold times and bought times of most tokens is small. However, the token LOT (whose address is 03d..4d) is sold 61,894 times and bought 22,947 times, which makes it the most sold token and the 14th bought token. It is difficult to understand the reason for the difference in buying and selling intentions, but it may be related to the trend of the token price, which we will further explore in future work.

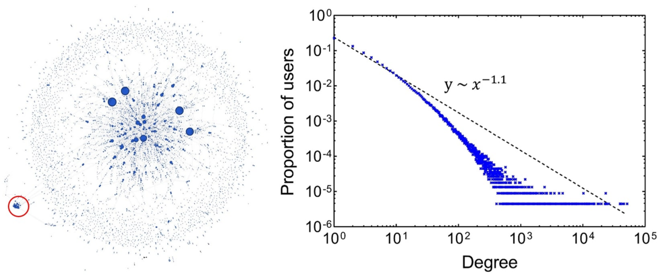
5.2.3 Who is buying and selling? Next, we introduce the graph analysis to study the behavior characteristics of users participating in token buying and selling.

TTG Definition and Construction. $TTG = (V, E, w)$, where V is a set of nodes (all the users in the exchange) and $E = \{(v_i, v_j)|v_i, v_j \in V\}$ is a set of ordered edges. Each edge indicates that node V_i sold a certain amount of some token to node V_j. $w : E \rightarrow \mathcal{R}^+$ associates each edge with a weight, which represents the total number of records. Hence, THG is a weighted directed graph. Please note that we ignore the type and number of token transferred and only count the number of transfer records in the construction of TTG since the tokens are not comparable. Figure 9 shows a sample of 10,000 edges and the degree distribution of TTG. We also analyze the indegree/outdegree distribution of TTG, but it is similar to the degree distribution. Thus, we do not show them to save space. Similar to centralized exchange [11], some accounts are very active. It's worth noting that there is a small group of nodes on the left-down side of Figure 9(a)(denoted as a red circle). These nodes connect to each other strongly while their connection to the whole graph is weak. This is a strange phenomenon, and we will study it further in the future.

Table 3 shows some statistics (the number of nodes and edges) and metrics of TTG. The clustering coefficient describes whether nodes cluster together. The clustering coefficient of TTG is 0.093, nearly 0. It means that if user A has traded with user B and user C, then user B and C is slightly likely to trade with each other. The assortativity coefficient shows whether a node with a large/small

Table 2: Token classification and its statistics.

Category	Number	$Marketcap(Mean)	Holder(Mean)	$Volume(Mean)
Digital money	116	61595396	16851.41	153237854
Exchange	67	96393303	18939.87	8312161
Service Platform	710	8361742	19484.63	1093462
Wallet	28	4595837	12341.82	317768
Marketplace	43	3657813	38570.58	244501



(a) TTG (b) Degree distribution of TTG

Figure 9: Visualization of TTG and its degree distribution.

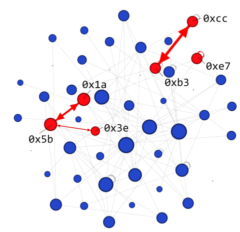


Figure 10: A small trading group and some abnormal trading patterns.

Table 3: Statistics and metrics of TTG.

Nodes	Edges	Clustering	Assortativity	Pearson
223827	3642588	0.093	-0.095	0.96

degree is likely to connect to another node with a large/small degree. The assortativity coefficient of our graph, which is negative and close to 0 (i.e., -0.095), reveals that large-degree-nodes slightly tend to connect to nodes with a small degree. Both the clustering coefficient and the assortativity coefficient indicate that users of IDEX tend to choose their counterparty randomly. Pearson coefficient is calculated by the indegrees and out degrees of nodes in TTG. The high Pearson coefficient (i.e., 0.96), meaning that there is a strong connection between the indegrees and outdegrees of the nodes. In other words, if a user buys a lot, he/she is very likely to sell a lot.

To see whether there are special trading patterns between users, we apply the Pagerank algorithm to the nodes and select the top 2000 nodes in TTG. There are 162,871 edges between these nodes, and the sum of the weights is 603,243, which indicates that these top traders have traded with each over 600 thousand times (accounting for 16.6% of the total trading records). We also calculate the clustering coefficient of these nodes, which is 0.192, far higher than that of TTG, meaning that these top traders tend to group.

By further analyzing the transfer records between them, we find some interesting patterns. To show the patterns, we randomly select 50 accounts from the top 2, 000 nodes and display the graph in Figure 10. There are some abnormal patterns. For example, the users *0x3e* and *0x1a* trade with *0x5b* 15,557 times; User *0xcc* and user *0xb3* trade with each other 16,554 times; User *0xe7* trades with himself 3, 012 times, including various tokens. We consider these records *abnormal*, because in traditional financial markets, it is impossible for investors to trade with themselves, nor for two different investors to trade with each other so many times. These results suggest that decentralized exchanges may not be as good as they seem and that while price manipulation cannot be proved, for the time being, these trading volume is likely to be fake.

6 RELATED WORK

Since the birth of Bitcoin, a lot of literature based on blockchain data have emerged. Three kinds of research are closely related to our study. The first type focused on mining the blockchain to reveal the characteristics of users in the system, from discussing user privacy issues [3, 20, 21, 29], to identifying various user behaviors [23, 24, 26], to reveal illegal activities [5, 6, 12, 33–35]; The second type mainly discusses smart contract, the core element of blockchain 2.0. Many topics about the blockchain-based smart contract are discussed. For example, the security of smart contract [4, 18, 32], code analysis [2, 22], and applications [7, 9, 15, 25]. The third, which is also the most relevant to our research, is an economic analysis of the cryptocurrency market. Paper [13] performs a comprehensive measurement analysis of Silk Road which uses Bitcoin as its exchange currency. The paper [10] explores how entrepreneurs can use initial coin offerings to fund venture start-up costs. The authors of the paper [14, 19] discussed the characteristics of a successful token. It is worth mentioning that paper [36] analyzes the top 1,000 ERC20 tokens. Different from it, this research is aimed at the whole token ecosystem, more than 160,000 tokens have been analyzed, thus we can have a more comprehensive understanding of the token ecosystem.

7 CONCLUSION AND FUTURE WORK

We conduct a systematic study to characterize the ERC20 token ecosystem. By using the Parity client, we collect all transactions and event logs of the Etehreun blockchain and then construct three graphs. By using the parsed records and the constructed graphs, we study the activity of tokens, the characteristics of tokens and its creators and holds, and the usage of tokens. Through these studies, we obtain many new observations and findings, which help people have a deep understanding of the ERC20 token ecosystem. This study raised, but did not answer, many interesting and important questions that we commend to future researchers (including ourselves):

- Why do some users create so many tokens? What are the differences and connections between these tokens? Is this a form of attack on Ethereum? Who are the developers behind it?
- Why do some users hold many kinds of tokens? What effect do holders have on the token price fluctuations? Do the holders of different tokens have different intrinsic characteristics?
- Whether different tokens have unique transfer characteristics? Do token buying and selling affect the price of tokens? What's the impact? Is there market manipulation in decentralized exchanges? Whether there is money laundering in the transfer of tokens?

This is only part of the problems. Tokens are numerous and the transaction volume is huge; Based on these data, we can deeply analyze a new economical world and then promote the development of blockchain technology.

ACKNOWLEDGMENTS

The work described in this paper was supported by the National Key R&D Program of China (2018YFB0204303), the National Natural Science Foundation of China (61722214, U1611261, 61872392), the Natural Science Foundation of Guangdong (2018B030312002), project funded by China Postdoctoral Science Foundation (Grant NO. 2019TQ0372, 2019M660223), the Program for Guangdong Introducing Innovative and Entrepreneurial Teams under Grant NO. 2016ZT06D211 and Pearl River S & T Nova Program of Guangzhou under Grant NO. 201906010008. Zhiguang Chen and Zibin Zheng are the corresponding authors.

REFERENCES

[1] Saman Adhami, Giancarlo Giudici, and Stefano Martinazzi. 2018. Why do businesses go crypto? An empirical analysis of initial coin offerings. *Journal of Economics and Business* 100 (2018), 64–75.

[2] Sidney Amani, Myriam Bégel, Maksym Bortin, and Mark Staples. 2018. Towards verifying ethereum smart contract bytecode in Isabelle/HOL. In *Proceedings of the ACM SIGPLAN International Conference on Certified Programs and Proofs*. ACM, 66–77.

[3] Elli Androulaki, Ghassan O Karame, Marc Roeschlin, Tobias Scherer, and Srdjan Capkun. 2013. Evaluating user privacy in bitcoin. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. Springer, 34–51.

[4] Nicola Atzei, Massimo Bartoletti, and Tiziana Cimoli. 2017. A survey of attacks on ethereum smart contracts (sok). In *Proceedings of the International Conference on Principles of Security and Trust*. Springer, 164–186.

[5] Massimo Bartoletti, Salvatore Carta, Tiziana Cimoli, and Roberto Saia. 2020. Dissecting Ponzi schemes on Ethereum: Identification, analysis, and impact. *Future Generation Computer Systems* 102 (2020), 259–277.

[6] Massimo Bartoletti, Barbara Pes, and Sergio Serusi. 2018. Data mining for detecting Bitcoin Ponzi schemes. In *Proceedings of the Crypto Valley Conference on Blockchain Technology*. IEEE, 75–84.

[7] Massimo Bartoletti and Livio Pompianu. 2017. An empirical analysis of smart contracts: platforms, applications, and design patterns. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. Springer, 494–509.

[8] Cristiano Bellavitis, Douglas Cumming, and Tom R Vanacker. 2019. The cross-country diffusion of new entrepreneurial practices: The case of initial coin offerings. In *Proceedings of the Academy of Management*, Vol. 2019. Academy of Management Briarcliff Manor, NY 10510, 13098.

[9] Andreas Bogner, Mathieu Chanson, and Arne Meeuw. 2016. A decentralised sharing app running a smart contract on the ethereum blockchain. In *Proceedings of the International Conference on the Internet of Things*. ACM, 177–178.

[10] Christian Catalini and Joshua S Gans. 2018. *Initial coin offerings and the value of crypto tokens*. Technical Report. National Bureau of Economic Research.

[11] Weili Chen, Jun Wu, Zibin Zheng, Chuan Chen, and Yuren Zhou. 2019. Market Manipulation of Bitcoin: Evidence from Mining the Mt. Gox Transaction Network. In *Proceedings of the IEEE Conference on Computer Communications*. IEEE, 964–972.

[12] Weili Chen, Zibin Zheng, Jiahui Cui, Edith Ngai, Peilin Zheng, and Yuren Zhou. 2018. Detecting ponzi schemes on ethereum: Towards healthier blockchain technology. In *Proceedings of the 2018 World Wide Web Conference*. International World Wide Web Conferences Steering Committee, 1409–1418.

[13] Nicolas Christin. 2013. Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the International Conference on World Wide Web*. ACM, 213–224.

[14] John P Conley et al. 2017. Blockchain and the economics of crypto-tokens and initial coin offerings. *Vanderbilt University Department of economics working papers* 17-00008 (2017).

[15] Gianni Fenu, Lodovica Marchesi, Michele Marchesi, and Roberto Tonelli. 2018. The ICO phenomenon and its relationships with ethereum smart contract environment. In *Proceedings of the International Workshop on Blockchain Oriented Software Engineering*. IEEE, 26–32.

[16] Christian Fisch. 2019. Initial coin offerings (ICOs) to finance new ventures. *Journal of Business Venturing* 34, 1 (2019), 1–22.

[17] Neil Gandal, JT Hamrick, Tyler Moore, and Tali Oberman. 2018. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics* 95 (2018), 86–96.

[18] Ilya Grishchenko, Matteo Maffei, and Clara Schneidewind. 2018. A semantic framework for the security analysis of Ethereum smart contracts. In *Proceedings of the International Conference on Principles of Security and Trust*. Springer, 243–269.

[19] Sabrina T Howell, Marina Niessner, and David Yermack. 2018. *Initial coin offerings: Financing growth with cryptocurrency token sales*. Technical Report. National Bureau of Economic Research.

[20] Marc Jourdan, Sebastien Blandin, Laura Wynter, and Pralhad Deshpande. 2018. Characterizing entities in the bitcoin blockchain. In *Proceedings of the International Conference on Data Mining Workshops*. IEEE, 55–62.

[21] Merve Can Kus Khalilov and Albert Levi. 2018. A survey on anonymity and privacy in bitcoin-like digital cash systems. *IEEE Communications Surveys & Tutorials* 20, 3 (2018), 2543–2585.

[22] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, and Aquinas Hobor. 2016. Making smart contracts smarter. In *Proceedings of the ACM SIGSAC conference on computer and communications security*. ACM, 254–269.

[23] Damiano Di Francesco Maesa, Andrea Marino, and Laura Ricci. 2016. An analysis of the bitcoin users graph: Inferring unusual behaviours. In *Proceedings of the International Workshop on Complex Networks and their Applications*. Springer, 749–760.

[24] Damiano Di Francesco Maesa, Andrea Marino, and Laura Ricci. 2017. Detecting artificial behaviours in the Bitcoin users graph. *Online Social Networks and Media* 3 (2017), 63–74.

[25] Patrick McCorry, Siamak F Shahandashti, and Feng Hao. 2017. A smart contract for boardroom voting with maximum voter privacy. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. Springer, 357–375.

[26] Dan McGinn, David Birch, David Akroyd, Miguel Molina-Solana, Yike Guo, and William J Knottenbelt. 2016. Visualizing dynamic bitcoin transaction patterns. *Big data* 4, 2 (2016), 109–119.

[27] Paul Momtaz, Wolfgang Drobetz, and Henning Schroeder. 2018. Investor sentiment and initial coin offerings. *The Journal of Alternative Investments* 21, 4 (2018).

[28] Satoshi Nakamoto et al. 2008. Bitcoin: A peer-to-peer electronic cash system. (2008). Retrieved September 1, 2019 from https://bitcoin.org/bitcoin.pdf

[29] Fergal Reid and Martin Harrigan. 2013. An analysis of anonymity in the bitcoin system. In *Security and privacy in social networks*. Springer, 197–223.

[30] EL Sidorenko. 2019. Stablecoin as a New Financial Instrument. In *Proceedings of the International Scientific Conference "Digital Transformation of the Economy: Challenges, Trends, New Opportunities"*. Springer, 630–638.

[31] Nick Szabo. 1996. Smart contracts: Building blocks for digital markets. (1996). Retrieved September 1, 2019 from http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_2.html

[32] Christof Ferreira Torres, Julian Schütte, et al. 2018. Osiris: Hunting for integer bugs in ethereum smart contracts. In *Proceedings of the Annual Computer Security Applications Conference*. ACM, 664–676.

[33] Kentaroh Toyoda, Tomoaki Ohtsuki, and P Takis Mathiopoulos. 2017. Identification of high yielding investment programs in Bitcoin via transactions pattern analysis. In *Proceedings of the IEEE Global Communications Conference*. IEEE, 1–6.

[34] Marie Vasek and Tyler Moore. 2015. There's no free lunch, even using Bitcoin: Tracking the popularity and profits of virtual currency scams. In *Proceedings of the International conference on financial cryptography and data security*. Springer, 44–61.

[35] Marie Vasek and Tyler Moore. 2018. Analyzing the Bitcoin Ponzi scheme ecosystem. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. Springer, 101–112.

[36] Friedhelm Victor and Bianca Katharina Lüders. 2019. Measuring Ethereum-based ERC20 token networks. In *Proceedings of the International Conference on Financial Cryptography and Data Security*.

[37] Shermin Voshmgir. 2019. *Token economy: How blockchains and smart contracts revolutionize the economy*. BlockchainHub Berlin.

[38] Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang. 2017. An overview of blockchain technology: Architecture, consensus, and future trends. In *Proceedings of the IEEE International Congress on Big Data*. IEEE, 557–564.

[39] Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang. 2018. Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services* 14 (2018), 352–375.