

# Tracking Counterfeit Cryptocurrency End-to-end

BINGYU GAO, Beijing University of Posts and Telecommunications, China  
 HAoyu WANG\*, Beijing University of Posts and Telecommunications, China  
 PENGCHENG XIA, Beijing University of Posts and Telecommunications, China  
 SIWEI WU, Zhejiang University, China  
 YAJIN ZHOU, Zhejiang University, China  
 XIAPU LUO, The Hong Kong Polytechnic University, China  
 GARETH TYSON, Queen Mary University of London, United Kingdom

The production of counterfeit money has a long history. It refers to the creation of imitation currency that is produced without the legal sanction of government. With the growth of the cryptocurrency ecosystem, there is expanding evidence that counterfeit cryptocurrency has also appeared. In this paper, we empirically explore the presence of counterfeit cryptocurrencies on Ethereum and measure their impact. By analyzing over 190K ERC-20 tokens (or cryptocurrencies) on Ethereum, we have identified 2, 117 counterfeit tokens that target 94 of the 100 most popular cryptocurrencies. We perform an end-to-end characterization of the counterfeit token ecosystem, including their popularity, creators and holders, fraudulent behaviors and advertising channels. Through this, we have identified two types of scams related to counterfeit tokens and devised techniques to identify such scams. We observe that over 7,104 victims were deceived in these scams, and the overall financial loss sums to a minimum of \$ 17 million (74,271.7 ETH). Our findings demonstrate the urgency to identify counterfeit cryptocurrencies and mitigate this threat.

CCS Concepts: • **Security and privacy** → **Intrusion/anomaly detection and malware mitigation**; **Web application security**; • **Information systems** → **Web mining**.

Additional Key Words and Phrases: counterfeit cryptocurrency; blockchain; ERC-20 token; scam

## ACM Reference Format:

Bingyu Gao, Haoyu Wang, Pengcheng Xia, Siwei Wu, Yajin Zhou, Xiapu Luo, and Gareth Tyson. 2020. Tracking Counterfeit Cryptocurrency End-to-end. *Proc. ACM Meas. Anal. Comput. Syst.* 4, 3, Article 50 (December 2020), 28 pages. <https://doi.org/10.1145/3428335>

## 1 INTRODUCTION

Since the first Bitcoin block was mined in 2009, cryptocurrencies have seen significant growth. This growth is mainly due to the rapid development of blockchain technologies and the digital economic system. Besides Bitcoin, thousands of cryptocurrencies have emerged. As of the end of 2019, the total market capitalization of cryptocurrencies is over \$180 billion [19].

\*Corresponding Author: Haoyu Wang ([haoyuwang@bupt.edu.cn](mailto:haoyuwang@bupt.edu.cn)).

Authors' addresses: Bingyu Gao, Beijing University of Posts and Telecommunications, Beijing, China; Haoyu Wang, Beijing University of Posts and Telecommunications, Beijing, China, [haoyuwang@bupt.edu.cn](mailto:haoyuwang@bupt.edu.cn); Pengcheng Xia, Beijing University of Posts and Telecommunications, Beijing, China; Siwei Wu, Zhejiang University, Hangzhou, China; Yajin Zhou, Zhejiang University, Hangzhou, China; Xiapu Luo, The Hong Kong Polytechnic University, HongKong, China; Gareth Tyson, Queen Mary University of London, London, United Kingdom.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2020 Association for Computing Machinery.

2476-1249/2020/12-ART50 \$15.00

<https://doi.org/10.1145/3428335>

*Where there is money, there are those who follow it.* Cryptocurrencies have attracted extensive attention from attackers. Attackers have exploited the vulnerabilities in smart contracts, cryptocurrency exchanges and wallets. According to endpoint security provider Carbon Black, \$1.1 billion in cryptocurrency was stolen in attacks during the first half of 2018 [1]. As reported in May 2019, attackers have stolen 7,000 bitcoins (worth \$41m) from Binance, one of the top leading exchanges [3]. Hundreds of popular Gambling Decentralized Applications (DApps), Defi Dapps, and other smart contracts were attacked recently, causing huge economic losses. Besides these known attacks, a number of newly emerging scams are taking advantage of cryptocurrencies to make a profit. For example, Marie Vasek and Tyler Moore [64] presented the first empirical analysis of Bitcoin-based scams in 2015, including high-yield investment programs, mining investment scams, scam wallet services and scam exchanges. After that, some other studies have characterized various scams including cryptocurrency Ponzi Schemes [25, 35, 65], blockchain honeypots [61], extortion emails [55], and cryptocurrency exchange phishing scams [71], etc.

However, an understudied attack is *counterfeit money* – the imitation of currency, which is produced without the legal sanction of the government [68]. Fraudsters strive to imitate the official currency so as to deceive its recipient. Following the same postulation, we ask *if such kinds of counterfeited money have appeared in the cryptocurrency ecosystem?* Evidence suggests that the answer is yes, with several news reports discussing cryptocurrency abuse [44, 52]. For example, since the Libra currency [10] was announced, scammers have designed fraudulent investment schemes involving the sale of fake “Libra tokens” that are unaffiliated with the actual Libra brand [52]. Similarly, recently, Chinese authorities have seized cryptocurrencies worth over \$10 million while bringing down a scam involving fake HuobiTokens [44]<sup>1</sup>.

Moreover, the ease of creating cryptocurrencies and launching Initial Coin Offerings (ICO), makes the cost of releasing counterfeit cryptocurrencies quite low. For example, Ethereum, as an open-source platform for decentralized applications (DApps), 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. By July 2020, there were over 200 thousand ERC-20 tokens created on Ethereum [18]. However, Ethereum does not enforce any restrictions on the names and symbols of the newly created tokens. Instead, the thing that identifies a token is its smart contract address. As shown in Figure 1, by searching Tether USD (USDT), a popular token that attempts to be tied to the US dollar, in the Etherscan (the most widely used Ethereum explorer) [18], there are over 170 tokens with the identical name “Tether USD” or symbol “USDT”. *This opens up a number of potential fraudulent avenues, with malicious parties potentially exploiting this fact to counterfeit cryptocurrencies.*

Despite this, to the best of our knowledge, the counterfeit cryptocurrency ecosystem has not been systematically investigated or measured. Thus, there is a general lack of an understanding of this attack, including: 1) *to what extent counterfeit cryptocurrencies exist*; 2) *what are the entities related to the counterfeit cryptocurrencies*, i.e., their targets, creators, distributors, and users; 3) *what are they used for*, i.e., whether counterfeit cryptocurrencies are involved in blockchain scams; and 4) *the advertising channels of counterfeit cryptocurrencies*, i.e., how do they reach and attract victims.

**This Work.** In this paper, we present the first systematic study of counterfeit cryptocurrencies on Ethereum. By analyzing over 170K ERC-20 tokens created on Ethereum (before March 2020), we have identified 2,117 counterfeit tokens that target 94 of the top-100 most popular cryptocurrencies (tokens). We then analyze the distribution and popularity of these counterfeit tokens, as well as the creators and holders of them (see **Section 4**). After identifying two types of fraudulent behaviors related to the counterfeit cryptocurrencies, we further measure the impacts of the counterfeit

<sup>1</sup>HuobiToken is an ERC-20 token on Ethereum, which is released by Huobi, one of the most popular exchanges.










Token Name	Symbol	Decimals	Official Site
 Tether USD	USDT	6	<a href="https://tether.to/">https://tether.to/</a> 
 Tether USD	USDT	6	-
 Tether USD	USDT	18	-
 Tether USD	USDT	18	-
 Tether USD	USDT	6	-
 Tether USD	USDT	6	-
 Tether USD	USDT	6	-
 Tether USD	USDT	6	-

Fig. 1. A number of ERC-20 tokens with the identical name and symbol of Tether USD (USDT).

cryptocurrency ecosystem, including the scale of the financial losses and the number of victims (see **Section 5**). Finally, to further understand how they are spread, we go a further step to investigate the advertising channels of these counterfeit cryptocurrencies (see **Section 6**). This paper reveals the ecosystem of counterfeit cryptocurrencies with some unexpected and interesting observations:

- **Counterfeit tokens are prevalent on Ethereum.** 94% (94/100) of the official tokens we studied in this paper have already been targeted by counterfeit tokens. Some counterfeit tokens are quite popular, with thousands of transactions and holders.
- **The scams related to counterfeit tokens cause huge financial losses.** We have characterized two types of scams related to counterfeit tokens, and quantified the direct financial impact. The overall volume is a minimum amount of \$17 million.
- **A number of reputable platforms are abused to help spread the fraudulent information of counterfeit tokens.** We have identified 935 pieces of advertising information related to counterfeit tokens from 103 well-known platforms. These include Telegram, Facebook, Bitcointalk (the official forum of Bitcoin), YouTube, etc. Various kinds of social engineering techniques are abused by attackers to attract victims.

## 2 BACKGROUND

### 2.1 Blockchain and Ethereum

Blockchain, which was invented in 2008 by Satoshi Nakamoto, is an open distributed ledger that stores transactions or related events among involved parties. It is maintained by a peer-to-peer network and secured by cryptographic design, thus it is resistant to data modification. By this design, each transaction in the block is verified by the confirmation of most participants in the system. Blockchain was originally served as a ledger for the Bitcoin, the first decentralized cryptocurrency. Bitcoin demonstrated the feasibility to construct a decentralized value-transfer system that can be shared across the world and virtually free to use. Bitcoin's blockchain design has inspired many other blockchain systems like Ethereum and EOSIO.

Ethereum is an open-source decentralized blockchain platform featuring smart contract functionality. It was proposed by Vitalik Buterin, and its development was funded by an online crowdsale. After that, it was initially released in 2015. *Ether* (ETH) is the cryptocurrency mined by Ethereum miners as a reward for computations and it is the second largest cryptocurrency based on volume.

## 2.2 Ethereum Account and Transactions

**Ethereum Account.** An account is the basic unit to identify an entity in Ethereum. An account is identified by a fixed-length hash-like address. Ethereum has two kinds of accounts: *external owned accounts* (EOAs) that are controlled by public-private key pairs (i.e. humans); and *contract owned accounts* (COAs) controlled by the code stored together with the account. An EOA is an ordinary account that can transfer tokens, invoke deployed smart contracts and store received tokens. Moreover, an EOA can deploy a smart contract into a COA account. All accounts are referred to by their addresses and denoted as an six-character identifier beginning with 0x in this paper.

**Transaction.** A transaction in Ethereum is a message sent from one account to another, which records the state changes of accounts. “Gas”, an internal transaction pricing mechanism, is used to protect the blockchain from spam and allocate resources on the network during transactions. A transaction can include binary data (called the “payload”) and Ether. There are two kinds of transactions depending on the message sender. The transactions sent from an EOA are called “external transactions”, which will be included in the blockchain and can be obtained by parsing the blocks. The other type, initiated by executing a smart contract, is called “internal transaction”. Internal transactions are usually triggered by external transactions and are not stored in the blockchain directly.

## 2.3 Smart Contract and ERC-20 Token

**Smart Contract.** Smart contracts are a kind of decentralized agreement built by computer programs, which are used to implement arbitrary rules as well as guaranteeing to produce the same result for decentralized parties. In Ethereum, a smart contract is a collection of code and data that reside in a Contract Account. A smart contract can be executed automatically, and it can control related events based on the terms built into its code. In the Ethereum platform, it is easy for people to build decentralized apps (DApps) and issue tokens for DApps or other purposes through smart contracts. Smart contracts are typically written in higher level languages (e.g., Solidity) then compiled to Ethereum Virtual Machine (EVM) bytecode. EVM is the runtime environment for smart contracts in Ethereum.

**Token.** In contrast to digital coins like Bitcoin and Ether, which are native to their own blockchain, tokens require existing blockchain platforms. Based on the function of tokens, they are usually classified into three types [2]: 1) *currency tokens*, which are entirely created as a method of payment; 2) *utility tokens*, which grant investors access to some kinds of products or services; and 3) *investment/asset tokens*, which are the assets that promise investors a return on their investment. The most famous investment token is the decentralized autonomous organization (DAO) token. It is an ERC-20 token which was a form of investor-directed venture capital fund, and whose vulnerabilities led to a hard fork of Ethereum. Note that every token that exists on the Ethereum is tied to a token contract, which defines a set of functions they use to perform tasks.

**ERC-20.** ERCs (Ethereum Request for Comments) are technical documents used by smart contract developers at Ethereum, which define a set of rules required to implement tokens for the Ethereum ecosystem. ERC-20 is by far the most recognizable token standard. It is proposed for developers to better handle different tokens on Ethereum. An ERC-20 token contract usually has properties including name, symbol, total supply and decimal, etc. An example of an ERC-20 token is shown in Figure 2, whose token name is “HuobiToken” and symbol is “HT”. Due to the ERC-20 standard, Ethereum has become one of the most popular token platforms — as of July 2020, there are over 200,000 ERC-20 tokens on Ethereum. Note that, in this paper, each token is represented in the form of `TokenName (SymbolName)`.

```
uint public totalSupply = 5*10**26;
uint8 constant public decimals = 18;
string constant public name = "HuobiToken";
string constant public symbol = "HT";
```

Fig. 2. The code snippet of an ERC-20 token smart contract.

## 2.4 Counterfeit Cryptocurrency

Ethereum does not enforce any restrictions on the names and symbols of newly created tokens, even if the names have been used by existing tokens. *This, however, could be abused by attackers to create counterfeit cryptocurrencies.* Just like producing counterfeit money by imitating fiat money (e.g., US dollars), attackers may use same identifiers (e.g., token name and symbol) or confusingly similar identifier names to deceive inexperienced investors. Thus, we consider the following two types of counterfeit tokens in this paper:

- **Type-1** The counterfeit token has an identical identifier name to an imitated official cryptocurrencies, but is released by different creators.
- **Type-2** The attacker adopts combo-squatting techniques [47]<sup>2</sup> to create a counterfeit token. This involves the combination of a recognizable token name (e.g., USDT) with other characters or keywords (e.g., USDT-2, USDT New). Such counterfeit tokens often have confusingly similar names to official cryptocurrencies, leading people to believe they are the new version of the official tokens or at least released by the same team.

## 3 STUDY DESIGN

We present the details of our characterization study on counterfeit cryptocurrencies in this section. We first describe our research questions, and then present the dataset used for our study. Last, we discuss our rationale for selecting the tokens that are most likely to be abused.

### 3.1 Research Questions

Our study aims to investigate the overall ecosystem of counterfeit cryptocurrencies, from their creation, transaction and circulation, to the scams related to them and their advertising channels. To this end, our study is driven by the following research questions (RQs):

- RQ1** *Are counterfeit cryptocurrencies prevalent in the cryptocurrency ecosystem?* Although a few reports in the media mentioned the existence of counterfeit cryptocurrencies, their scale remains unknown. Furthermore, it is necessary to investigate: RQ1.1) *Which tokens are their targets?* Considering there are thousands of cryptocurrencies, we seek to explore whether adversaries predominantly target tokens with greater popularity (volume). RQ1.2) *who create counterfeit cryptocurrencies?* It is interesting to study whether some malicious campaigns habitually create a number of counterfeit cryptocurrencies to deceive unsuspecting users or investors. RQ1.3) *who hold these counterfeit cryptocurrencies?* Analyzing the holders of these counterfeit tokens can help us understand the overall scale of the ecosystem, i.e., how many accounts have been involved in counterfeit tokens. RQ1 will be studied in Section 4.
- RQ2** *What are the fraudulent behaviors related to counterfeit cryptocurrencies?* We are still unaware of the usage of the counterfeit tokens. Thus, it is interesting to investigate how

<sup>2</sup>Combo-squatting is a specific type of domain squatting, in which attackers register domains that combine a recognizable brand name (e.g., Paypal) with other keywords (e.g., login). Combo-squatting attacks are prevalent in domain [47] and even mobile apps [45].

users can be scammed by counterfeit cryptocurrencies, whether there are other collusion addresses and even malicious campaigns involved in the scams, and how many users were scammed by them? RQ2 will be studied in Section 5.

**RQ3 *What are the advertisement channels of counterfeit cryptocurrencies?*** It is interesting to analyze how counterfeit currencies reach users, especially the advertising channels, tricks and social engineering techniques adopted. This can help us better identify and trace scams and malicious campaigns behind. RQ3 will be studied in Section 6.

### 3.2 Datasets

Since our goal is to measure counterfeit cryptocurrencies in Ethereum, we require both 1) a complete list of the *ERC-20 tokens*, which is used for detecting counterfeit tokens; and 2) the whole *Ethereum transaction dataset*, which is used to analyze transactions related to counterfeit cryptocurrencies.

Thus, we take advantage of Geth<sup>3</sup>, a widely-used Ethereum client to synchronize the ledger of Ethereum. We have synchronized all the blocks until March 18th, 2020, with over 9.6 million blocks in total. The data extracted from the blocks contains *external transactions*, *internal transactions*, *contract information*, and *contract calling information*. We then get the bytecode and creator information for all the smart contracts. Then, we analyze the bytecode to determine whether a contract implements an ERC-20 token. Based on this method, we have identified over 176K ERC-20 tokens alongside their creator information. We further obtain the metadata of these tokens (e.g., website, total supply, holder, etc.) from either the code or Etherscan. To facilitate the analysis, we use ElasticSearch to store these structured data, which provides a query interface for our characterization study.

### 3.3 Target Cryptocurrency Selection

While all the tokens could be the subject of counterfeit cryptocurrencies, it is arguably not in the best interest of an attacker to use a less known token for abuse (e.g., an official token with less user and volume). Furthermore, as there are thousands of official tokens on Ethereum (mixed with counterfeit tokens), it is hard for us to compile a complete list of all the official tokens. We do this to reduce the number of potential counterfeit currencies that must be measured. Consequently, we compile a list of the top-100 tokens based on market capitalization from Etherscan. Table 1 shows the information of these official tokens. Column 1 (#CAP) shows the ranking of the tokens based on their market capitalization by the time of our study.

## 4 MEASUREMENT OF COUNTERFEIT CRYPTOCURRENCIES

### 4.1 Detection Method

According to the previous definition of the counterfeit token (see Section 2.4), we adopt a two-step and semi-automated approach to perform accurate detection of counterfeit tokens. *The first step is to identify all the possible counterfeit token candidates by keyword matching.* For the selected 100 official tokens, we search their token names and symbol names in the ERC-20 token dataset, to flag tokens that contain such keywords. However, we find that the keyword matching based method may introduce a number of false positives. Thus, *for the second step, we propose to remove false positives* according to the following rules:

**Rule1 Migrated tokens.** As smart contracts in Ethereum cannot be modified once deployed, some tokens migrate their addresses due to security concerns or new features being introduced. For example, the token HEDG has migrated from 0xb6B6Bd<sup>4</sup> to 0x3363D5<sup>5</sup>, due to new features

<sup>3</sup><https://geth.ethereum.org/>

<sup>4</sup>0xb6B6Bd3c75c4237089b5ED518A1809C297CC2e6B

<sup>5</sup>0x3363D570f6DF3c74d486BB8785d3EbFB9E2347D3

Table 1. The target official tokens and their counterfeit tokens identified (CTokens). Column 5 (#Trans) shows the aggregated number of all transactions related to counterfeit tokens for the corresponding official token.

# CAP	Token Name	Symbol	# CTokens	# Trans	# CAP	Token Name	Symbol	# CTokens	# Trans
1	Tether USD	USDT	171	12,188	51	Ampleforth	AMPL	5	202
2	BNB	BNB	90	781	52	Pundi X Token	NPXS	5	31
3	ChainLink Token	LINK	15	24	53	Trace	TRAC	6	1,262
4	HuobiToken	HT	545	12,883	54	Tellor Tributes	TRB	21	89
5	Bitfinex LEO Token	LEO	20	306	55	MXCToken	MXC	0	0
6	Crypto.com Coin	CRO	1	0	56	Uquid Coin	UQC	3	10
7	-	HEDG	5	69	57	BandToken	BAND	1	6
8	Maker	MKR	10	102	58	Synth sUSD	sUSD	3	61
9	USD Coin	USDC	70	3,303	59	Insolar	INS	1	2
10	OKB	OKB	60	1,451	60	OriginToken	OGN	10	202
11	Ino Coin	INO	1	41	61	Melon Token	MLN	8	316
12	VeChain	VEN	8	20	62	Utrust Token	UTK	9	68
13	BAT	BAT	69	277	63	Wrapped BTC	WBTC	8	353
14	Paxos Standard	PAX	21	417	64	Rocket Pool	RPL	5	14
15	ZRX	ZRX	25	369	65	Pinakion	PNK	0	19
16	Insight Chain	INB	3	3	66	Ankr Network	ANKR	14	177
17	ICON	ICX	9	19	67	QuarkChain Token	QKC	39	1,297
18	OMG Network	OMG	54	552	68	DATAcoin	DATA	21	554
19	Republic	REN	7	78	69	Chimpion	BNANA	1	1
20	Baer Chain	BRC	19	6,456	70	Numeraire	NMR	7	511
21	ZBToken	ZB	13	73	71	Reserve Rights	RSR	5	9
22	Synthetix Network Token	SNX	17	38	72	SingularityNET	AGI	12	273
23	TrueUSD	TUSD	52	573	73	BTU Protocol	BTU	1	1
24	HoloToken	HOT	16	298	74	SwissBorg	CHSB	2	1
25	Dai Stablecoin	DAI	69	991	75	Paxos Gold	PAXG	4	0
26	Mixin	XIN	3	3	76	Polymath	POLY	45	1,973
27	Theta Token	THETA	10	21	77	Fantom Token	FTM	20	260
28	Nexo	NEXO	12	157	78	Ocean Token	OCEAN	3	713
29	Cryptonex	CNX	0	0	79	Gnosis	GNO	13	66
30	Kucoin Shares	KCS	6	12	80	Bancor	BNT	18	32
31	Sai Stablecoin v1.0	SAI	11	510	81	Aragon	ANT	5	136
32	Bytom	BTM	33	446	82	HarmonyOne	ONE	40	124
33	EnjinCoin	ENJ	8	417	83	Storj	STORJ	7	17
34	MCO	MCO	3	29	84	Io TeX Network	IOTX	5	14
35	DGD	DGD	4	132	85	Fetch	FET	11	104
36	IOSToken	IOST	14	198	86	STASIS EURS Token	EURS	0	0
37	Centrality Token	CENNZ	1	0	87	UniBright	UBT	4	58
38	Zilliqa	ZIL	16	378	88	Enigma	ENG	19	157
39	KyberNetwork	KNC	9	120	89	Celsius	CEL	3	11
40	WAX Token	WAX	3	260	90	LoopringCoin V2	LRC	7	15
41	StatusNetwork	SNT	8	42	91	WaykiCoin	WIC	2	5
42	Golem	GNT	12	125	92	PowerLedger	POWR	32	994
43	SeeleToken	Seele	20	30	93	EthLend	LEND	6	23
44	NOAHCOIN	NOAH	17	76	94	AION	AION	4	30
45	Reputation	REP	16	7	95	Matic Token	MATIC	19	1,643
46	RLC	RLC	1	0	96	Decentraland	MANA	49	168
47	Cryptoindex 100	CIX 100	0	0	97	chiliZ	CHZ	2	202
48	Banker Token	BNK	5	4	98	ELF	ELF	7	39
49	Binance USD	BUSD	18	225	99	DxChain Token	DX	0	0
50	EXMER FDN.	EXMR	1	7	100	Swipe	SWP	4	10
-	-	-	-	-	-	<b>Sum</b>		<b>2,117</b>	<b>56,764</b>

being introduced. In this case, we manually analyze the 100 selected tokens to verify whether they have migrated addresses and further remove such false positives.

**Rule2 Official tokens created by trustworthy creators.** It is a common practice that, before releasing a new token officially, some creators release some test tokens to check whether their tokens will perform as expected. These attempts will lead to the creation of a few tokens with the same/similar identifier names. Thus, we manually check the creators of the filtered tokens to remove such false positives.

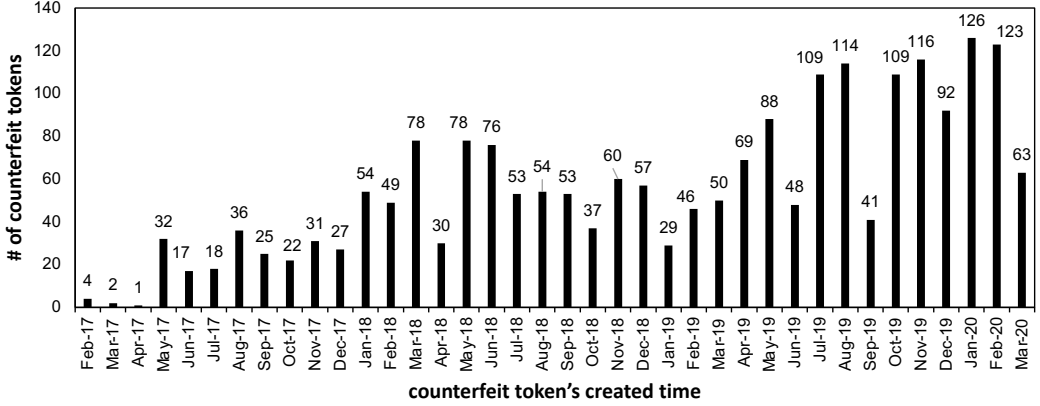


Fig. 3. The creation time of counterfeit tokens (till March 18th, 2020).

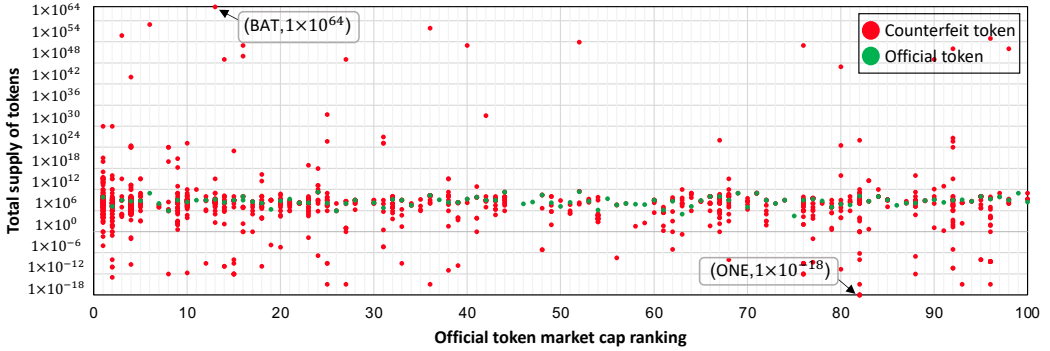


Fig. 4. The total supply of counterfeit tokens. Each vertical line represents the statistics of a type of token (100 lines in total). The dots shown on each vertical line indicate the total supply of an official token (in green) and its counterfeit ones (in red)

**Rule3 Official tokens that have similar names with our target tokens.** As the symbol name is usually short (e.g., 3 to 5 characters), it is quite possible that some official tokens have the same or similar symbol names. For example, there are three ERC-20 tokens, HEDG (with only a symbol but no token name)<sup>6</sup>, Hedgie (HDG)<sup>7</sup> and Hedge (HDG)<sup>8</sup>, that have similar names and symbols. Despite this, they are all independent and official tokens. An official token usually has its own official site, and we can collect detailed information about the token from Google. Thus, if a token has active transactions and we can collect its legal information online, we will regard it as a false positive and thus remove it from our dataset.

## 4.2 Overall Results

Using the above method, we have identified 2,117 counterfeit tokens, targeting 94 of the 100 popular tokens (94%), as shown in Table 1. Unsurprisingly, HuobiToken (HT), Tether USD (USDT), and

<sup>6</sup>0xF1290473E210b2108A85237fbCd7b6eb42Cc654F

<sup>7</sup>0x452B2bc7c94515720b36d304CE33909a8323F3e3

<sup>8</sup>0xfFe8196bc259E8dEDc544d935786Aa4709eC3E64

BNB (BNB) are the most popular targets. For example, there are 545 counterfeit tokens targeting HuobiToken (HT), with 12,883 transactions in total. HuobiToken (HT) is released by a famous cryptocurrency exchange with the same name Huobi, which is easier for unsuspecting users to believe its authenticity. We observe that, *in general, the counterfeit tokens are more likely to target popular tokens with high market capitalization rank*. However, not all high ranking tokens are their prior targets. For example, we only observe one counterfeit token that targets Crypto.com Coin (CRO), which ranks 6 by the time of our study.

Figure 3 shows the creation time of counterfeit tokens on a monthly basis. The first counterfeit token was created at February 12nd, 2017. After that, counterfeit tokens have become increasingly prevalent, especially after July 2019. For example, in Jan 2020, 126 counterfeit tokens were created. *This result suggests that, the counterfeit tokens are prevalent in the cryptocurrency ecosystem*. There are 56,762 transactions related to these counterfeit tokens in total, involved 56,057 unique Ethereum addresses. Figure 4 shows the comparison of *total supply* between official tokens and their counterfeit tokens. Total supply refers to the number of coins or tokens in existence right now and are either in circulation or locked somehow [20], which is defined by the token creator in the corresponding token contract. In Figure 4, the dots on each line represent the total supply of a corresponding token (in green) and its counterfeit ones (in red). For each official token, its total supply is usually between  $1e + 6$  and  $1e + 11$ , but obviously, the total supply of some counterfeit tokens are very high. For example, the total supply of 100 counterfeit tokens exceed  $1e + 12$ . The highest one is BAT (BAT) <sup>9</sup> which targets the official BAT (BAT), with a total supply of  $1e + 64$ , while the total supply of official BAT (BAT) is only  $1.5e + 9$ .

### 4.3 Lexical Characteristics

We next analyze the lexical characteristics of these 2,117 identified counterfeit tokens. As mentioned in Section 2.4, we consider two types of counterfeit tokens based on their naming strategies.

**Type-1.** As shown in Table 2, almost 80% of counterfeit tokens (1,674) have identical token names or symbols with official tokens. Among them, 23.6% of counterfeit tokens (499) have exactly the same token names and symbols with official tokens, while 10.1% of counterfeit tokens (214) have only the same token names with official tokens (but different symbols). Further, 45.4% of counterfeit tokens (961) have the same symbols with official tokens (but different token names).

**Type-2.** Approximately 77% of counterfeit tokens have adopted combo-squatting strategies in either their token names or their symbols. We characterize them into two categories. First, a number of counterfeit tokens combine the official identifier names with special characters like spaces, parentheses, underscores, or insert some abbreviations. For example, for the symbols of HuobiToken (HT), we have identified a number of confusingly similar symbols, including HT Coin, HT\_huobi, and Token HT, etc. Second, the counterfeit tokens take advantage of different string encoding methods to mislead users, such as UTF-8, ASCII, and GBK, etc.

Figure 5 shows an example of the word cloud extracted from the counterfeit tokens of HuobiToken (HT). Note that the token names and symbols are shown separately. Besides the identical names with the official token, they always adopt a number of variants as aforementioned.

### 4.4 Popularity Analysis

We next inspect the scale of these counterfeit currencies, looking at both *the number of transactions* and *the active period* of the tokens.

**4.4.1 The number of transactions.** The number of transactions can reflect the popularity of counterfeit tokens to some extent. Figure 6 (a) presents the number of transactions per token. We see

<sup>9</sup>0x031f053a6b4e49A4A64450B2ca8A0b0ef6335134

Table 2. The lexical characteristics of counterfeit tokens.

	# Combo-squatting (%)	# Identical (%)
Symbol	214 (10.1%)	961 (45.4%)
Token name	961 (45.4%)	214 (10.1%)
Both	443 (20.9%)	499 (23.6%)
Sum	1,618 (76.4%)	1,674 (79.1%)
All	2,117	

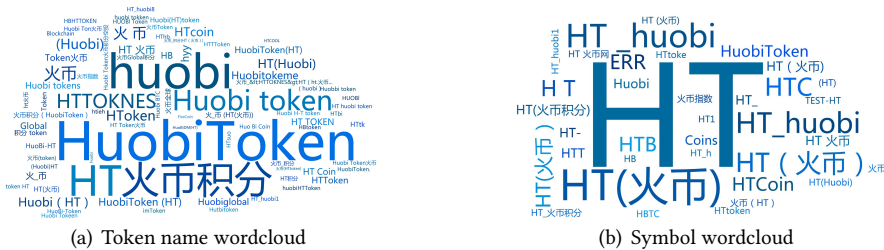


Fig. 5. An example of the word cloud extracted from the 545 counterfeit tokens of HuobiToken (HT).

that it varies greatly across the counterfeit tokens. A large portion of the counterfeit tokens are less popular, i.e., 579 (27.35%) counterfeit tokens have never been transferred, and over 90% of counterfeit tokens have been transferred no more than 45 times. In contrast, some counterfeit tokens are very active with thousands of transactions. To be specific, 7 counterfeit tokens have over 1,000 transactions. Table 3 shows the top-5 counterfeit tokens with the most transactions. The most heavily used counterfeit token, brc (brc), is an imitation token of Baer Chain (BRC), has over 5,000 transactions in total. The counterfeit token Tether USD (USDT), also has over 4,500 transactions. It is interesting to further investigate the activities related to these “popular” counterfeit tokens, which we will explore in Section 5.

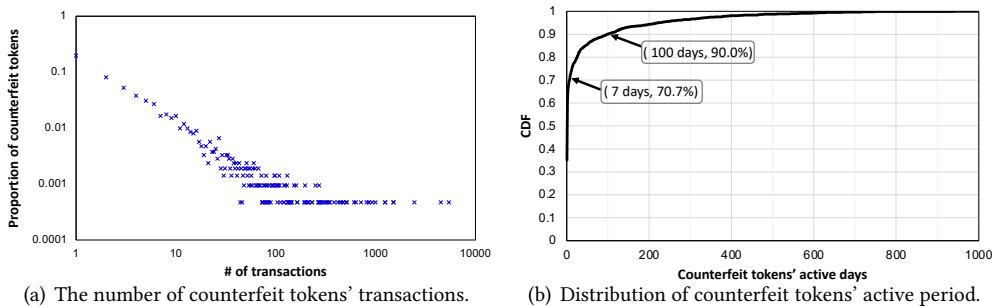


Fig. 6. Distribution of counterfeit tokens' transactions and active period.

**4.4.2 Active Period.** We further analyze the active period of the counterfeit tokens. As the cost of creating a counterfeit token is quite low in Ethereum, we wonder *whether the counterfeit tokens*

Table 3. The top-5 counterfeit tokens with the most number of transactions.

Counterfeit Token	Target Token	Address	Transactions
brc (brc)	Baer Chain (BRC)	0xB64555C4fEb7Dbe8584cb3b10D8993d1B3572f7e	5,429
Tether USD (USDT)	Tether USD (USDT)	0x68771C9d7F6A498743Aa167967627b198B97e9E2	4,511
USDT (Tether USD)	Tether USD (USDT)	0x3936eE7369e9c278A78a44dE7b272dEB97bc6253	2,446
USDT (USDT)	Tether USD (USDT)	0xfdf6b11baA0A17d39a378f0c97bC93dE8303f338	1,512
POLYMATH (POLYM)	Polymath (POLY)	0xe518098BB49354bc4372d48D3474d8C1F2eddF8	1,503

Table 4. The Top 5 Counterfeit Token Creators Ranked by the Number of Created Counterfeit Tokens.

Address	# CTokens	# CToken Types
0x2468293D8059bc578CF312F09Ed78D6CE1005dCb	25	23
0xB2BDBb9Cc6583D10C3c043DE3AC6bE07A074dd16	25	1
0xB2f0dbb7FF8f1A451fF9486756EA53Ff2F654633	24	1
0x1f7A74F06359e08Cc82a5c60cFa21D277f5f4181	23	21
0xb9B6885D0Af9914d432871DcBeB20DAa8282A763	23	5

are only utilized for a short period. Figure 6 (b) presents the distribution of the active time for all the counterfeit tokens. Over 70% of the counterfeit tokens (1,497) are active for less than 7 days. However, to our surprise, some tokens remain active for a long time. Particularly, 9.97% of the counterfeit tokens (211) are active over 100 days. For example, the counterfeit token 0x9900E9<sup>10</sup>, remains active for over 940 days, which is a BeraCoin (BRC) token that targets Baer Chain (BRC).

#### 4.5 Counterfeit Token Creator Analysis

We next analyze the creators of the counterfeit tokens. Overall, 1,210 creators created 2,117 tokens. Over 95% (2,016) of the tokens are created by external addresses (EOA, i.e., by humans, see Section 2), while the remaining tokens (101) are created automatically (COA, i.e., by smart contracts).

**Token Creator Graph.** To further investigate the token creator relationship, we introduce the token creator graph (TCG), as shown in Figure 7 (a). In the TCG, each node denotes an address on Ethereum, i.e., the orange one represents the EOA creator address, the green one represents the COA creator address and the purple one is the counterfeit token address. The TCG is a directed graph, with each edge represents the creation relationship, i.e., from EOA to counterfeit token, or from COA to counterfeit token. Note that, the COA need to be called by an EOA account to create a token, thus each COA has a connection with an EOA. The size of the creators' node denotes the number of counterfeit tokens they created.

We observe that almost 30% of creators (362) have released more than one counterfeit token. This shows that *certain actors may specialize in the creation of such tokens*. Table 4 presents the top-5 addresses that create the most counterfeit tokens. The most aggressive addresses are 0xB2BDBb<sup>11</sup> and 0x246829<sup>12</sup>, which have both created 25 counterfeit tokens. By further analyzing the creation of counterfeit token, we find that *most creators (1,107) exclusively counterfeit the same currency*. However, 103 creators counterfeit multiple currencies. For example, the most aggressive one, 0x246829<sup>13</sup>, has counterfeited 23 official tokens (with 25 counterfeit tokens released in total).

<sup>10</sup>0x9900E95AE292e264B517f1979eB30C6c4D6458ab

<sup>11</sup>0xB2BDBb9Cc6583D10C3c043DE3AC6bE07A074dd16

<sup>12</sup>0x2468293D8059bc578CF312F09Ed78D6CE1005dCb

<sup>13</sup>0x2468293D8059bc578CF312F09Ed78D6CE1005dCb

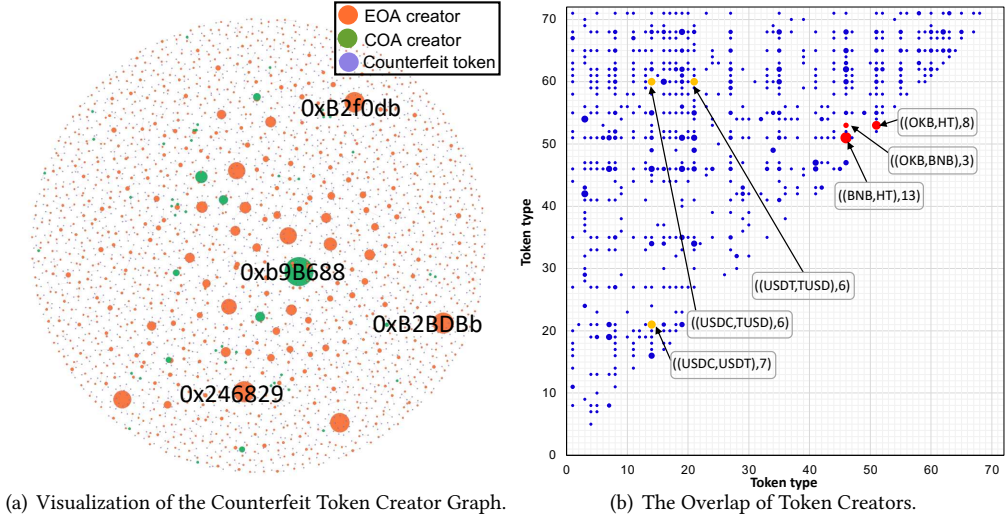


Fig. 7. Visualization of Counterfeit Token Creators.

Thus, we analyze the co-occurrence of identical token creators across different types of counterfeit tokens. As shown in Figure 7 (b), there are 71 types of targeted tokens<sup>14</sup>. All of the 71 official tokens have counterfeit tokens whose creators have generated other kinds of counterfeit tokens. The size of the circle represents the number of creators counterfeited both of the corresponding tokens. We mark two sets of tokens with the most identical creators. They are HuobiToken (HT), BNB (BNB) and OKB (OKB) in red, and Tether USD (USDT), True USD (TUSD) and USD Coin (USDC) in orange. Among them, HuobiToken (HT) and BNB (BNB) share the most counterfeit token creators (13). The following is HuobiToken (HT) and OKB (OKB), with 8 counterfeit token creators in common. We speculate that scammers tend to create counterfeit tokens with tokens with the same characteristics. First, they are all top-50 tokens with high popularity. For HuobiToken (HT), BNB (BNB) and OKB (OKB), they are all released by well known cryptocurrency exchanges that share identical names: HuobiToken (HT) is released by Huobi, BNB (BNB) is released by Binance (with same Chinese pronunciation and spelling), and OKB (OKB) is released by OKB. For Tether USD (USDT), True USD (TUSD) and USD Coin (USDC), they are all stablecoins (stable-value cryptocurrency) that mirror the price of the U.S. dollar. *This result suggests that malicious actors tend to counterfeit more than one type of official token, likely due to the low cost of creating tokens on Ethereum.*

#### 4.6 Counterfeit Token Holder Analysis

We next analyze the characteristics of token holders, the cornerstone of the ecosystem. Overall, there are 28,861 unique holders in total. As the number of holders is much higher than the number of creators, we use a sampled (20% of all the holders) counterfeit token holder graph (THG) for clear illustration, as shown in Figure 8 (a). Note that our sampling method is based on the proportion of

<sup>14</sup>For 23 official tokens, their counterfeit tokens do not share creators with other tokens, thus we eliminate them in the figure to save space. The orders of tokens in Figure 7 (b) are: BRC, OECAN, QKC, SNT, DGD, ZB, BTM, LRC, REN, HEDG, INB, TRAC, AION, USDC, HOT, PAX, NPXS, LINK, OMG, THETA, USDT, AGI, GNO, LEO, LEND, XIN, ELF, OGN, NEXO, MATIC, KNC, TRB, VEN, DAI, IOST, CEL, ANKR, IOTX, ONE, STORJ, ZIL, SEELE, UBT, CRO, ENJ, BNB, ICX, FET, SAI, UTK, HT, BNT, OKB, FTM, MKR, DATA, KCS, POWR, ENG, TUSD, GNT, BAT, MANA, NOAH, REP, SXP, POLY, ZRX, WBTC, CHZ, WAX.

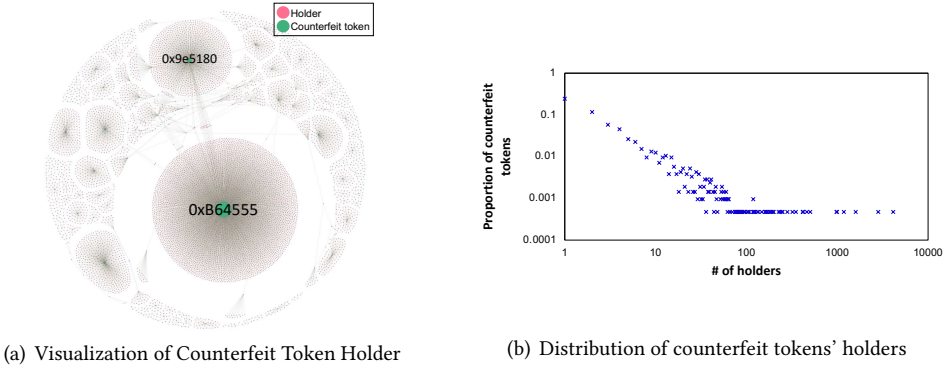


Fig. 8. Visualization of Counterfeit Token's Holders.

holders for each counterfeit token. Figure 8 (b) presents the distribution of holders of the counterfeit tokens. Over half of the tokens (64.58%) have no more than two holders. Over 90% of the tokens have no more than 20 holders. It further suggests that, most tokens are only used for a short time, thus have limited holders and transactions to prevent victims from reporting or causing suspicion from regulators. The token with the most holders (4,147) is 0xB64555<sup>15</sup>. We further observe that, although over 96% of holders only possess a single kind of counterfeit token, roughly 1,189 holders have multiple kinds of counterfeit tokens. Among them, 506 holders have multiple kinds of counterfeit tokens that target more than one official currency. Table 5 shows the top-5 holders ranked by the number of different counterfeit tokens. The largest holder address, 0x8d12A1<sup>16</sup>, holds a remarkable 85 different counterfeit tokens that target 47 types of official currencies.

Table 5. The top-5 token holders ranked by the number of different counterfeit tokens.

Address of counterfeit token holder	# CTokens	# CToken Types
0x8d12A197cB00D4747a1fe03395095ce2A5CC6819	85	47
0xB2BDBb9Cc6583D10C3c043DE3AC6bE07A074dd16	24	1
0x09f91Ce790dbb36ddA4EC507A4754b7a4e2b0e55	21	21
0xa4fd7ACa0A39e1c70d464D6380e293761d64FA63	20	19
0xb9A8436700cbaBbf855d03a00513502515C49B83	18	4

We further investigate the type of users that are generally affected by counterfeit cryptocurrencies. Note that it is non-trivial to determine the experience of counterfeit token holders, yet we can gain insights from the number of transactions per holder. In general, the more transactions related to the holder address, the more experienced the holder is. Figure 9 shows the distribution of the number of the transactions and the balance of counterfeit token holders. Note that, some holders are verified as victims (see Section 5), which are labelled in red. Almost 50% of holders (13,812) have fewer than 5 transactions, and nearly 95% of holders (27,134) have a balance under 0.2 ETH. Thus, the observation is inline with our expectation that most of the holders (victims) are novices without much experiences.

<sup>15</sup>0xB64555C4fEb7Dbe8584cb3b10D8993d1B3572f7e<sup>16</sup>0x8d12A197cB00D4747a1fe03395095ce2A5CC6819

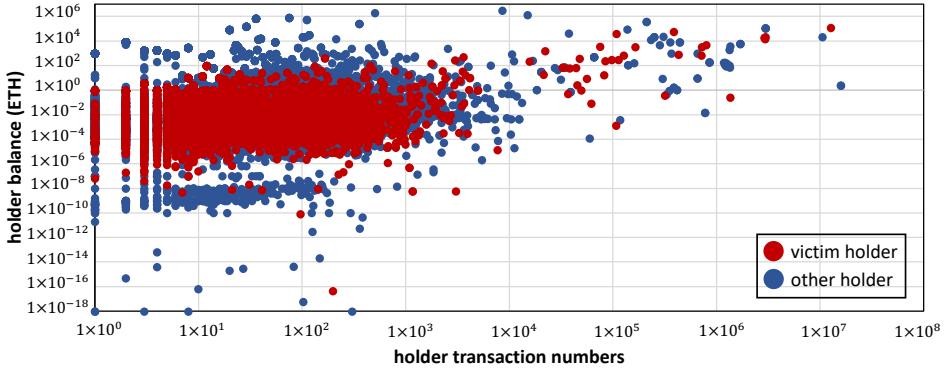


Fig. 9. The distribution of the balance and the number of transactions of counterfeit token holders.

**Answer to RQ1:** Counterfeit tokens are indeed prevalent in the cryptocurrency ecosystem. We have identified 2,117 counterfeit tokens targeting 94 out of the 100 tokens we studied. Although most of the tokens have very few transactions, some of them are quite popular, with thousands of transactions and holders. Malicious actors tend to target more than one type of official token, mainly due to the cost of creating a counterfeit token is quite low in Ethereum.

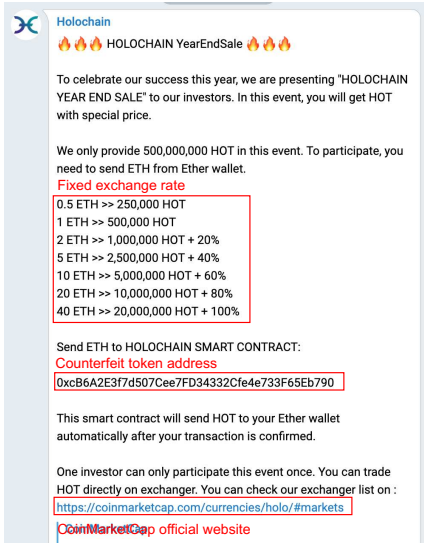
## 5 FRAUDULENT BEHAVIORS OF COUNTERFEIT CRYPTOCURRENCIES

Our prior exploration suggests the prevalence of counterfeit cryptocurrencies in Ethereum. However, we are still unaware of the usage of counterfeit tokens. Although the token names and symbols can be counterfeit in Ethereum, the token addresses cannot. The token address is the unique identifier to represent a token. As long as users input the official address during the transaction, the counterfeit cryptocurrency scams will not succeed. Thus, in this section, we seek to explore the scams and social engineering attacks related to counterfeit tokens.

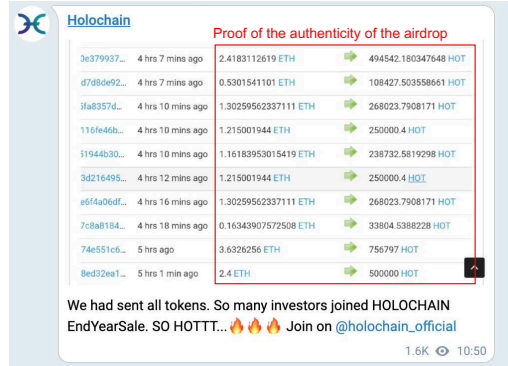
### 5.1 Types of Fraudulent Behaviors

To understand the characteristics of scams related to counterfeit tokens, we first resort to existing scam reports to identify the types of scam activities related to counterfeit tokens. To be specific, we first resort to the following scam repositories, and use keyword searching (e.g., “fake”, “token” and “counterfeit”) to identify related scams for manual verification.

- 1) **CoinHunter.** CoinHunter [7] is a crowd-sourcing platform that collects cryptocurrency related scams reported from users. We have implemented a crawler to get all the scam reports from the CoinHunter. After keyword filtering and manual inspection, 52 reports related to counterfeit tokens on Ethereum are identified.
- 2) **Blockchain Forums.** The imTokenFans [9] is a forum operated by the official team of imToken, a popular cryptocurrency wallet. BitcoinTalk [5] is an online forum devoted to the discussion of Bitcoin and other cryptocurrencies. Both of them host the “Scam Accusations” board for users to report scams. Thus, we have crawled all the related posts on the “Scam Accusation” board based on aforementioned keywords and further perform a manual verification. At last, we have identified 18 reports on counterfeit token scams.
- 3) **Other scam reports from search engines.** We also resort to search engines to identify scams related to the identified counterfeit tokens. We have implemented an automated crawler, to feed the 2,117 counterfeit token addresses to Google and get all the search results.



(a) The information the fraudster posted



(b) The screenshot the fraudster provided

Fig. 10. An example of a fake HOLOTOKEN (HOT) airdrop scam.

After eliminating unrelated ones (e.g., blockchain address searching services), we manually verify whether there are user complaints, and media reports on these counterfeit token addresses. This helps us identify 10 more scam reports.

At last, by analyzing the collected scam reports, we have collected 204 counterfeit token addresses, all of which have been included in our 2,117 counterfeit token dataset. The identified scams can be classified into two types: *airdrop scam* (22) and *arbitrage scam* (182). *Note that the identified scams are only used as the ground truth for our following study. We will further propose approaches to investigate how many counterfeit tokens have such fraudulent behaviors in the Section 5.2 and Section 5.3, as the scams are typically under-reported by users in the wild.*

**5.1.1 Airdrop Scam.** An airdrop is the distribution of a cryptocurrency token, *usually for free*, to numerous user wallet addresses [67]. Airdrops are primarily implemented as a way of gaining attention and new followers. As airdrops are quite popular for well-known tokens, counterfeit tokens are also taking advantage of this opportunity to perform airdrop scams. In general, the attackers promise that, after sending a certain amount of ETH to the (counterfeit) token address, the victim will get (imitated) official tokens according to a *fixed exchange rate* (far more than the actual value). After victims send the ETH, they likely only receive counterfeit tokens that have no value at all. Figure 10 shows an airdrop scam of the counterfeit token HOLOTOKEN (HOT), with the same icon and account name of the official token. To deceive unsuspecting users, they usually embed the official link of the imitated token in the airdrop information (see Figure 10 (a)). Furthermore, the scammer will also post screenshots of the relevant tokens returned to the user to enhance the credibility (but the tokens they returned are fake) (see Figure 10 (b)).

**5.1.2 Arbitrage Scam.** Arbitrage is an investment method that capitalizes on imbalances in prices between markets, i.e., buy at a low price and sell at a slightly higher price [29]. From the perspective of amateur investors, the arbitrage does not require too much professional knowledge, which is relatively safer compared with other investment methods. Thus, cryptocurrency arbitrage is

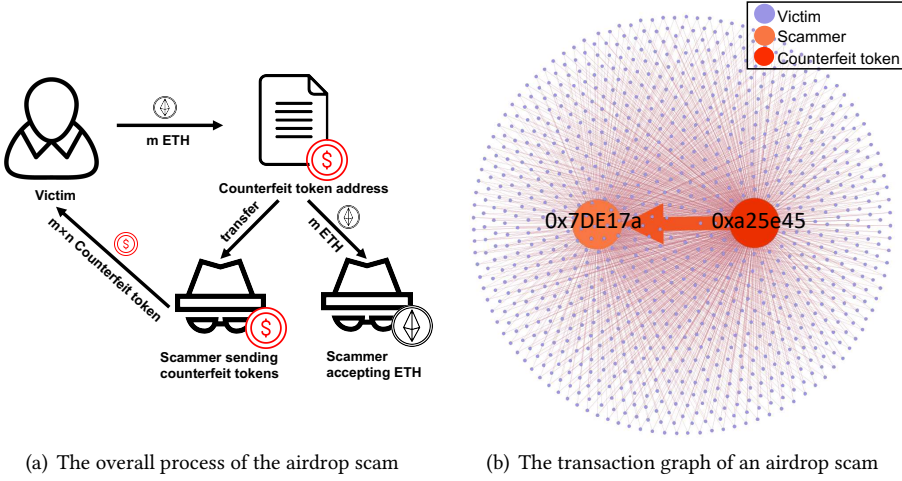


Fig. 11. Visualization of the Airdrop Scam.

popular among many investors. However, our exploration suggests that the arbitrage can be abused by attackers, i.e., *the arbitrage can be combined with counterfeit tokens to carry out well-designed scams*. In our collected scams, the scammers usually use fake Telegram groups (that imitate the official token) as advertising channels (see Section 6), providing scam addresses for victims to send ETH. After victims send ETH to the specified scam address, as promised, they should get official tokens (far more than the actual value) in a few minutes. However, victims usually received counterfeit tokens of no worth.

Next, we will analyze in detail the workflow of these two scams, devise approaches to detect such scams, and measure the prevalence of the scams in our 2,117 counterfeit token dataset.

## 5.2 Analyzing and Detecting Airdrop Scam

**5.2.1 The Workflow of Airdrop Scams.** Figure 11 (a) shows the workflow of airdrop scams, which consists of four main roles: *the victim*, *the counterfeit token contract*, *the scam address that accepts ETH for money laundering*, and *the scam address that sends counterfeit tokens*. Note that in the airdrop scam, all these steps are fulfilled in one transaction. When *the victim* sends  $m$  ETH to *the counterfeit token contract*, the contract transfers all the received ETH to the scam address that accepts ETH, and then calls the transfer function to return  $m \times n$  counterfeit tokens to the victim, distributed by the scam address that sends counterfeit tokens. The exchange rate of ETH to counterfeit token is fixed (the  $n$ ) in Figure 11 (a). Note that the scam address that accepts ETH and the scam address that sends counterfeit tokens can be either same or different addresses.

**5.2.2 Detecting Airdrop Scams.** Based on the characteristics of airdrop scams, we seek to investigate whether the 2,117 counterfeit tokens we identified have such behaviors. To be specific, for a transaction involving airdrop scams, ① *the victim address must send some ETH to the counterfeit token contract first*; and ② *the counterfeit token distributor must transfer some fixed percentage of counterfeit tokens to the victim*. Furthermore, ③ *the counterfeit token contract should transfer the received ETH to a scam address*. Thus, we have implemented a script to analyze all the transactions related to counterfeit tokens. A counterfeit token whose transactions meet the aforementioned behaviors will be considered as part of an airdrop scam.

Table 6. The top-10 counterfeit tokens involved in airdrop scams ranked by the number of ETH received.

Token	Contract	Token/ETH	Victims	Scam Address	ETH Received	Est. Value (\$)
QuarkChain Token (QKC)	0xd7bb68B0cE5893983e5a2511b87E083609eB6fF9	32,000	16	0x7F83284C0cE0906F58eFF0FC433967e50c4d607E	180.46	42,163.14
Matic Token (MATIC)	0x4256117a02aC880335f8bFbEed63F92eC0001A5a	53,992	10	0xCC9CD9d0CA055616093c52741F19a01Fb8fD709c	107.76	25,176.51
HOLOToken (HOT)	0xcB6A2E3f7d507Cee7FD34332Cfe4e733F65EB790	1,125,000	24	0xe4F349f54e8490b7aE9EFB090Be7EAA41b08D965	75.0	17,523.00
Enjin Coin (ENJ)	0x7e534b4192daea6559c08c7147364b00a7ce697	5,000	109	0xe937910a2A748296eE884fC2beb1A041789DA50F	70.2	16,401.53
Matic Token (MATIC)	0xd7E6460a5ff2e51BD1583808d1C19d521CF43DeF	91,255	20	0x3A3870d9066fB88F1127E78103B038939B58b63e	62.62	14,630.84
Matic Token (MATIC)	0x41993b3F7979dcd8A15e1448CACCbC6873803	91,255	13	0xBb733A9c379cA22829cb689C9eafC04c0ea8A51f	58.19	13,596.12
Matic Token (MATIC)	0x4b67BE266D81BF5CEE60FC12b3bded87fCDaC783	47,015	49	0xB663b4BeA1197e58D69F27B04f8ee2B7531e9668	57.16	13,355.20
Holotoken (HOT)	0x5566E98b8Bce420E943C7366ab0F15ba3D181F10	1,000,000	27	0x23705f9b708eE6363Dcc93Cc3d8056098a3c5787	36.2	8,457.77
Polymath (POLY)	0xa25e457ae666A67C9C906d1052fAE39710Af63CC	10,000	133	0x7DE17a49f857dA070D4a6c6d272Def7B6e603015	34.74	8,117.57
OmiseGo (OMG)	0xFdDe0a2bFfE2d95695d3240A0A8F5e530dC67	330	33	0xd0531B689856db9B745f804Bd0506f443244E4981	33.9	7,920.40
Kyber Network (KNC)	0x29ad674d180D33f0391b7b034d4880918b66b72e	1,215	36	0x97FECFC329f5213DAba173519F8062c33A3eD	30.0	7,009.20

**5.2.3 Results and Observations.** This method can achieve 100% accuracy in our collected ground truth dataset, i.e., all the 22 collected airdrop scams are correctly detected. As airdrop scams have explicit patterns, i.e., the exchange rate of ETH to counterfeit token is fixed, our detection has no false positives. We further apply this method to all 2,117 counterfeit tokens we identified, i.e., analyzing their related transactions to detect airdrop scams. As a result, 87 counterfeit tokens have shown the behaviors of airdrop scams, targeting 44 official tokens. Overall, 2,037 victims were deceived in airdrop scams, and the attackers have received 970.8 ETH in total (\$226, 817.71)<sup>17</sup>. Table 6 shows the top-10 counterfeit tokens with airdrop scams ranked by the number of ETH they received. Note that, for each of them, the scam address used to receive ETH and the scam address used to distribute counterfeit tokens are identical (see Column “scam address” in Table 6). The counterfeit token 0xd7bb68<sup>18</sup> gains the most profit in the airdrop scam, reaching 180.46 ETH (\$42, 163.14). The counterfeit token Polymath (POLY), whose address is 0xa25e457<sup>19</sup>, has the most number of victims (133). Figure 11 (b) shows its transaction graph. The red point denotes the Polymath (POLY) counterfeit token address, the orange point denotes the scam address used to receive ETH and distribute counterfeit tokens, and the purple points denote the victims. The size of points represent the amount of ETH involved, and the thickness of the edge represents the amount of ETH or counterfeit tokens transferred. Besides, it is obvious that the exchange rate of ETH to counterfeit tokens is stable (ranging from 330 to 1,125,000, see the Column “Token/ETH” in Table 6).

### 5.3 Analyzing and Detecting Arbitrage Scams

**5.3.1 The Workflow of Arbitrage Scams.** Figure 12 shows the workflow of arbitrage scams, which consists of four main roles: *the victim*, *the scam address that accepts ETH*, *the scam address that sends counterfeit tokens*, and *the scammer posing as the official customer service* (e.g., a fake Telegram account). Based on our collected scams, we have summarized two types of arbitrage scams: 1) scammers return counterfeit tokens directly after receiving the ETH; 2) scammers return official tokens at first, but then return counterfeit tokens in the following transactions.

**Type 1:** The victim sends ETH to the address that the scammer provides, and receives counterfeit tokens in a few minutes or does not receive tokens at all. Since counterfeit tokens are manual sent to the victim by the scammer, the scammer can easily conduct a *secondary scam*. If the victim does not receive tokens at all, the fake imToken (cryptocurrency wallet) customer service will explain that the smart contract address cannot read the victim address successfully. The victim needs to send the same amount of ETH to the same scam address for rollback to get the ETH back or provides the private key for help (see secondary scam 1 and 3 in Figure 12). If the victim receives counterfeit tokens and identifies that they could not return counterfeit tokens to trade on

<sup>17</sup>As the price of Ether fluctuates everyday, we estimate the profit using the closing price of Ether on July 16th.

<sup>18</sup>0xd7bb68B0cE5893983e5a2511b87E083609eB6fF9

<sup>19</sup>0xa25e457ae666A67C9C906d1052fAE39710Af63CC7

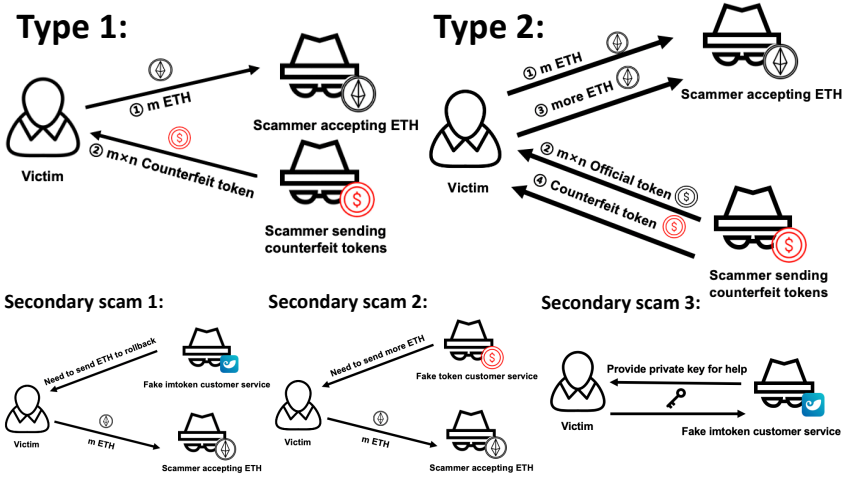


Fig. 12. The workflow of arbitrage scams.

the exchange, the counterfeit token customer service will explain that the amount of tokens is too small to use the exchange for withdrawal, and the victim needs to send more ETH to get tokens.

**Type 2:** The victim sends ETH to the address that the scammer provides, and receives official tokens in a few minutes for the first time. The victim would then believe the arbitrage's authenticity and send more ETH to the scam address, but will only receive counterfeit tokens the second time.

**5.3.2 Detecting Arbitrage Scams.** Unlike airdrop scam, we cannot identify the arbitrage scam directly through the counterfeit token contract, as they do not accept ETH directly. Thus, we devise an approach that consists of the following steps. First, we regard every transfer of the counterfeit token as a *candidate transfer from the scammer to the victim*. Thus, the receiver address in the transfer could be a potential victim address. Note that one major feature of the arbitrage scam is that, before victims receive counterfeit tokens, they should have already sent the ETH to the scam address a while ago (we found that almost all the victims received counterfeit tokens from the confirmed scam addresses within 2 hours after they sent ETH in our collected scams). Thus, we further identify the corresponding ETH transfer transaction of the potential victims. The ETH transfer transaction should be the most recent one just before the counterfeit token transfer. As a result, the receiver address in the ETH transfer should be the scam address published in Telegram or other platforms (see Section 6). In this way, we can accurately identify the arbitrage scams and label both the victim addresses and scam addresses.

**5.3.3 Results and Observations.** The above method can achieve a recall of 97.8% in our collected ground truth dataset, with 4 missed cases. The reason is that we track the scam address through the transfer of counterfeit tokens, thus we cannot confirm the scam address that did not send the counterfeit token to the victim, i.e., the victim receives nothing at all in this scenario. We further apply this method to all the 2,117 counterfeit tokens. We have identified 7,617 transfer transactions related to 486 counterfeit tokens (targeting 10 official tokens) that are considered to be arbitrage scams. We further identify 1,879 scam addresses that are used to accept ETH. To measure the potential false positives of our arbitrage scam detection, we manually sampled 100 identified scam addresses to investigate whether some of them are scams. On the one hand, we check the identified scam addresses on Google to find whether they were involved in any scam posts (e.g., telegram groups).

Table 7. Counterfeit tokens that involved in arbitrage scams.

Targeted token	# CTokens	# transactions	# victims	ETH received	USD received
HuobiToken (HT)	373	4,291	2,610	60,824.3	14,210,989.50
USD Coin (USDC)	17	809	578	4,559.1	1,065,188.12
OKB (OKB)	20	295	245	2,948.7	688,934.27
BNB (BNB)	12	258	156	2,265.7	529,358.15
TrueUSD (TUSD)	4	162	95	1,428.7	333,801.47
Paxos Standard (PAX)	5	130	82	967.0	225,929.88
Tether USD (USDT)	36	1,377	1,074	226.2	52,849.37
Bytom (BTM)	13	265	254	68.6	16,027.70
Zilliqa (ZIL)	3	18	15	12.0	2,803.68
IOSToken (IOST)	3	12	10	0.6	140.18
Sum	486	7,617	5,087	73,300.9	17,126,022.30

On the other hand, we manually analyzed their transactions to see whether they follow some specific patterns. For example, the real scam addresses usually have transactions with multiple victims. If most of their transactions follow the patterns we summarized (note that normal addresses would not have such behaviors), we believe they are definitely scams. Our manual verification does not flag any false positives. The total financial losses is 73,300.9 ETH (\$17, 126, 022.30). Table 7 shows the statistics of the counterfeit tokens that are involved in arbitrage scams. Obviously, HuobiToken (HT) has the largest scale of counterfeit token arbitrage scams, whose campaigns have received 60,824 ETH (\$14, 210, 919.40).

**Secondary scam.** We observe that over 29% of victims (1473) have been deceived more than once. Note that the real proportion can be higher as the attacker sometimes asks the victim to transfer ETH to other addresses that we cannot track. The victim that was scammed the most is 0xE5D1Ef<sup>20</sup>, who has transferred ETH to the scam address 19 times, with a total amount of 2,799.1 ETH (\$635, 981.72).

**Type-1 and Type-2 scams.** We are also interested in the victims who have ever received official tokens from confirmed scam addresses. 4.6% of victims have received official tokens from the scam address (*Type-2 scam*). Then, 81% of them send ETH to the scam addresses for the second time. This result suggests that *sending official tokens to the victim can greatly increase the probability for victims to send ETH to the scam address again*. When sending ETH for the second time, over 90% of victims send more ETH compared with the amount they sent at first time.

For example, the victim 0xf8a6aa<sup>21</sup> sent 1 ETH (\$223.64) to 0xa6C678<sup>22</sup> (the scam address) for the first time, and it received 55 official HuobiToken (HT). After that, the victim sent 115.18 ETH (\$26, 910.66) to the scam address one hour later, but only received counterfeit tokens in the second time.

## 5.4 Summary of the Scams

**5.4.1 The Impact of the Scams.** We summarize the overall impact of the scams, including the *involved addresses*, the *financial losses*, and the *number of victims*. As shown in Table 8, we see 565 counterfeit tokens are involved in an airdrop (87) or arbitrage scam (486). As we mentioned in Section 4.4.1, over 27% of counterfeit tokens (579) have never been transferred, and almost 75% of counterfeit tokens (1,584) were transferred fewer than 9 times. We have manually analyzed many

<sup>20</sup>0xE5D1Ef3297896FAA8B118031edDDC7a372655932

<sup>21</sup>0xf8a6aa3fcEE296DE9c388492c496aAa85EA66eee

<sup>22</sup>0xa6C678Ed8b54521Bf0DC46933aFb48005f543411

Table 8. Summary of the scams we identified.

Type	# Transactions	Scam Address		# Victims	# ETH	# USD
Airdrop Scam	2,872	Counterfeit Token Address	87	2,037	970.8	226,817.71
		Counterfeit Token Creator	71			
		ETH Received Address	70			
		Counterfeit Token Distributed Address	56			
		Sum	166			
Arbitrage Scam	7,617	Counterfeit Token Address	486	5,087	73,300.9	17,126,022.30
		Counterfeit Token Creator	293			
		ETH Received Address	1,879			
		Counterfeit Token Distributed Address	869			
		Sum	2,904			
Sum	10,489	3,053 (565 counterfeit tokens involved)		7,104	74,271.7	17,352,840.00

less popular counterfeit tokens and found it hard to justify their usage based on no or very few transactions. These less popular tokens may be used for testing or used in some scams, but no users were successfully cheated. Thus, we cannot infer their original intentions based on no or few transactions. We classify scam addresses into four categories based on their roles: *counterfeit token contract*, *counterfeit token creator*, *the money laundering address that receives ETH*, and *the counterfeit token distribution address*. From all aspects, the scale of arbitrage scam is much larger than that of airdrop. In general, the airdrop scam addresses are more likely to play more than one role. On average, each scam address has 1.71 roles. While for the arbitrage scam addresses, their roles are more specific (i.e., each scam address has 1.21 roles on average). Overall, 7,104 victims have been successfully cheated, with an overall financial loss of 74,271.7 ETH (\$17,352,840.00), which is a lower bound estimate of the criminal profits. Interestingly, 20 victims were deceived by both of the scams. We further tagged all victims who held counterfeit tokens in Figure 9. It is interesting to see that only 67.2% of victims (4,775) still hold counterfeit tokens after they have been cheated. We infer that after victims receive counterfeit tokens, they usually try to transfer these tokens to other addresses to obtain ETH. We observe that most of the victims are novices to Ethereum, i.e., have limited transactions and balance. Among these victims, over 5% of victims (2,518) have less than 20 transactions, and 92.8% of victims (4,431) have balance less than 0.2 ETH.

**5.4.2 Money flow of scam addresses.** Next, we track the money flow of scam addresses. As an example, Figure 13 shows the money flow of 180 randomly selected scam addresses. Note that, we have studied all the scam addresses indeed. There are three types of addresses in the graph: 1) *the identified scam address* (shown in orange); 2) *the fund transfer address* (shown in purple), which is served as the money laundering channel, i.e., receive money from scam addresses and help the attackers transfer the money they have scammed; and 3) *the exchange address* that belongs to certain known cryptocurrency exchanges. Note that we have purchased a premium service from an anonymous leading blockchain company to label whether an address belongs to an exchange or not. Each edge represents the direct or indirect relationship between the addresses. There are 3,351 fund transfer addresses and 44 exchange addresses related to these 180 scam addresses. Obviously, the scam addresses have transferred the money via a number of fund transfer addresses to exchanges finally. This observation is inline with all the scam addresses. For example, 0x6cC5F6<sup>23</sup>, the exchange address of OKEx, has received ETH from 27 scam addresses we identified. The exchange address of Tokenlon, 0xdc6c91<sup>24</sup>, has received ETH from 17 scam addresses.

<sup>23</sup>0x6cC5F688a315f3dC28A7781717a9A798a59fDA7b

<sup>24</sup>0xdc6c91b569C98F9F6f74d90F9BEFF99FDAf4248b

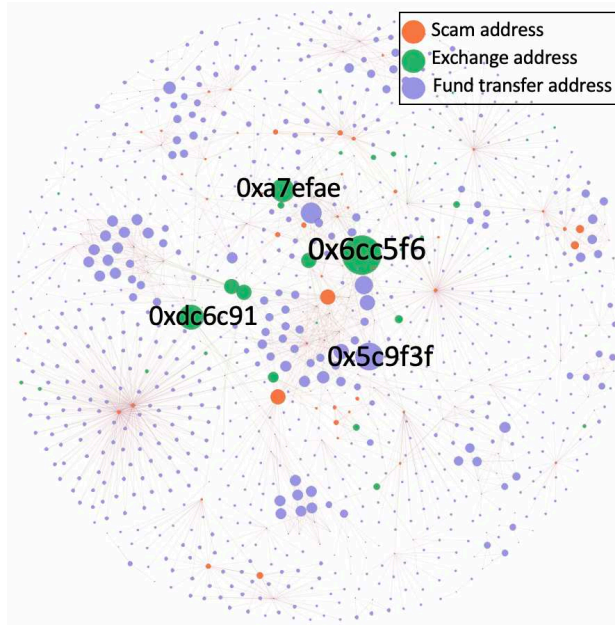


Fig. 13. The money flow of scam addresses (sampled 180 scam addresses).

**Answer to RQ2:** We have identified two types of scams related to counterfeit tokens, i.e., airdrop scam and the arbitrage scam. In total, 565 active counterfeit tokens have been found involved with scams. Over 7100 victims were deceived in these scams, and the overall financial loss sums to a minimum amount of \$ 17 million (74,271.7 ETH).

## 6 ADVERTISING CHANNELS

In this section, we seek to explore how these counterfeit tokens are advertised to trick users.

### 6.1 Approach

To reach the potential victims, an attacker usually provides their scam address or the counterfeit token address when they advertise the scam information. Thus, we have implemented an automated crawler, to harvest the advertising information of these counterfeit tokens by searching the address directly. To be specific, we first feed the 2,117 counterfeit token addresses and their corresponding scam addresses (see Table 8) to Google<sup>25</sup>, and collect all the related information. Note that Google also indexes information from social network platforms including Telegram, Twitter and Facebook, etc. Thus, we did not crawl information from these social networking platforms separately. For the crawled results, we further eliminate information irrelevant to scams. By manually browsing the collected information, we find that there are many blockchain explorers and services (e.g., Etherscan and bloxy.info) that index all the Ethereum addresses (including the scam ones). Thus, we remove the search results from these services. For the remaining results, we further manually verify whether they are advertising information published by attackers.

<sup>25</sup>It is different with the step in the scam report harvesting (see Section 5.1), as here we focus on the advertising information, rather than the scam reports. Besides, we further search the identified scam address to identify the advertising information.

Table 9. A Summary of the advertising channels of counterfeit tokens.

Channel	Arbitrage	Airdrop	Sum
Telegraph Page	589	4	593
Telegram Group	52	76	128
Twitter Page	0	25	25
Facebook Page	0	22	22
YouTube	0	5	5
Others	20	142	162
Sum of Pages	661	274	935
Channel Type	11	96	103

## 6.2 The Advertising Platforms

**6.2.1 Overall Results.** We have identified 935 pieces of advertising information from 103 advertising platforms, most of which are reputable ones, including Telegraph [15], Telegram [14], Facebook [8], V2EX [22], Bctalk [4], Bytechats [6], Bitcointalk [5], Telemetr [16], Sina [12], Twipu [21], Youtube [23], Steemkr [13], Zhihu [24], Tgchannels [17], Medium [11], etc. Indeed, social networking platforms and blockchain forums are the most widely used advertising channels. Table 9 summarizes the results. For example, we have identified 593 Telegraph pages and 128 Telegram Groups involved in the counterfeit token scams. Obviously, airdrop and arbitrage scams have shown different advertising strategies. Although we have identified a large number (661) of arbitrage scam advertising information, they are only active on 11 platforms. However, we have identified 96 advertising channels exploited by airdrop scams. Next, we use some case studies (see Figure 14) to illustrate how the attackers use social engineering techniques to trick users.

**6.2.2 Telegram.** Due to the anonymity of Telegram, attackers most often use Telegram groups as advertising channels. Here, we take the HuobiToken (HT) arbitrage scam as an example (see Figure 14 (a)) to summarize how the Telegram groups are exploited. **1) Group name and icon.** The arbitrage group always pretends to be an official group, so its name is usually like “Huobi Official Arbitrage Tutorial Group”, and it uses the HuobiToken (HT) official icon. **2) Advertising proposal.** The *pinned message* in the Telegram group is used to publish the (fraudulent) arbitrage tutorial. The tutorial explains in detail the workflow of arbitrage and the popularity of arbitrage tokens. Besides, scammers will insert official and reputable website links in the tutorial, which increases its credibility. **3) Customer Service.** Every Telegram group has a fake official customer service. The customer service also imitates the real HuobiToken (HT) telegram official group. As aforementioned, their main purpose is to lead the victim to the secondary scam. Besides, once the victim found out that they have been deceived, the fake customer service will kick out those who expose the scam from the Telegram group. **4) Bot Accounts.** For the scam Telegram groups we identified that every group has thousands of members, but most of them are robots. The robots in the group falsify the conversation, making victims mistakenly believe that the arbitrage is real and profitable.

**6.2.3 YouTube.** The content sharing platforms can also be exploited to distribute scam information. We use scams on YouTube to show the advertising process (see Figure 14 (b)). They usually pretend to be the official tutorial of airdrop, which illustrate the complete airdrop process for the potential victims. To increase their credibility, 1) they show some fake information on the video which states

**Huobi BM Detailed tutorial** Reply

火币网HT搬砖套利图文教程: <https://telegra.ph/Huobi-07-19>

你所未了解的币圈“闷声发大财”的搬砖军团! 这期间的疯狂已有人赚得百万

暴涨暴跌不知道如何操作? 再不来就out了! 跟上搬砖大军, 成为币圈资深交易员!

2020年比特币即将减半, 牛市初期引起各平台价差, 以太坊区块链上基于0X协议的去中心化交易数字货币-火币网平台 HT, 已在全球最大火币网交易所上形成了套利空间, 赶快来撸! 欢迎进入火币智能合约中文社区电报群

【简略版】↓不懂请咨询管理员→@HuobiBM

- 1、注册火币交易所官网: <https://www.huobi.me>(免翻墙) 或 <https://www.huobi.io> 或 <https://www.huobi.com>或者自己去手机下载火币网APP <https://www.huobi.com/zh-cn/download/#exchange>
- 2、下载ETH钱包, 没有的请去官网 <https://edge.app/> (不会下载请私聊管理员 发安装包给你)
- 3、添加HT代币: (不会的请看链接中图文教程)
- 4、领取HT币: 打开edge钱包, 发送ETH (至少1个ETH) 至矿池地址:

**Scam address**

**0xCAC83a7dFd3d4439aE1592010FaF72A60cfd5**

领取比例为1:100, 即1个ETH可以领取100个HT, 也就是目前利润最在15%左右单次! (HT会在发送ETH后30分钟内到达钱包, HT到账后会有声音提示, 手机要记得开VPN翻墙哦(#!^#))

**Exchange rate (ETH:HT)**

4、将领取到的HT币, 转到自己的火币交易所, 选择我的资产HT, 充值进交易所, 到账后卖掉换成ETH, 多出来的就是这一次的收益了

PS: 合约一天只能搬三次, 切记。牛市初期, 搬砖为主, 套利空间没几天时间, 抓紧时间获利! 教程如果打不开, 手机可以看及时预览, 如有不懂的地方请点击私聊管理员我, 无法私聊的请在群里@管理员→@HuobiBM)

(a) Arbitrage advertisement on Telegram

**MyEtherWallet**

Send Ether & Tokens

To Address: 0x2a778af8a1a1778e0736b3388322a3a333f

Amount to Send: 0.001 ETH

Generate Transaction

Get 2000 THUG Coin Airdrop Alert 2018 ICO

Send 0.001 ETH Minimum for 2000 THUG

If you send less then 0.001 ETH you lose your eth you didn't get any amount and its limited time offer (End in 15-April-2018)

**Scam address**

**0x2a778af8a1a1778e0736b3388322a3a333f**

(b) Airdrop advertisement on YouTube

**ProgrammarSought**

Search

**Scam address**

0xCAC83a7dFd3d4439aE1592010FaF72A60cfd5

Will get the airdrop: 500 BNT (Okeas, the price of the coins is around 0.2 yuan)

Use the imToken wallet to transfer 0.2 ETH to this address, and the miner fee is adjusted to the lowest

Will receive 1888 BNT airdrop (Okeas, the price of the coins is around 0.2 yuan)

Use the imToken wallet to transfer 0.2 ETH to this address, and the miner fee is adjusted to the lowest

Will receive 500 BNT (Okeas, 8 yuan)

Use the imToken wallet to transfer 0.2 ETH to this address, and the miner fee is adjusted to the lowest

Will receive 400 B N T empty coins. (The price of the currency is about 10 yuan.)

Use the imToken wallet to transfer 0.2 ETH to this address, and the miner fee is adjusted to the lowest

(c) Airdrop advertisement on ProgrammarSought

Fig. 14. Examples of Counterfeit Token Scam Advertising Channels

that many users have received tokens; 2) they always imitate famous people (e.g., Bill Gates, and Vitalik Buterin, the founder of Ethereum); and 3) official websites are embedded in the video.

**6.2.4 Blockchain Forums.** Online forums are also the main targets for attackers to reach potential victims, mostly for airdrop scams, based on our observations. We take the airdrop scam we found from ProgrammarSought as an example (see Figure 14 (c)). The article published by the attacker is also a tutorial of the airdrop. To increase its credibility, 1) it provides a list of airdrop tokens, which might be mixed with real official airdrops. Some of them are free airdrops, thus victims can get tokens for free, which makes victims believe the credibility of the listed other (scam) airdrops. 2) it also shows the screenshots of tokens received after the airdrop is completed.

**Answer to RQ3:** A number of reputable platforms have been abused by attackers to help spread fraudulent information on counterfeit tokens. Social network platforms like Telegram, Twitter and Facebook are the main targets of attackers. Various social engineering techniques have been adopted to trick users.

## 7 DISCUSSION

### 7.1 Implication

Our observations are of key importance to stakeholders in the community.

**The governance of the cryptocurrency.** Considering the large number of counterfeit tokens and scams we discover, the governance of cryptocurrency needs to be improved. There is a need to design policies to regulate cryptocurrency naming schemes. However, like domain squatting issues that remain in the wild for years, it is not easy to fully address counterfeit token issues. Even if Ethereum disallows creators to release ERC-20 tokens with the same identifier names, a number of other kinds of attacks in the domain-squatting field, including typosquatting (e.g., “yuotube.com”), bit-squatting (accidental random bit flip, e.g., “yo5tube.com”), homograph-based squatting (e.g., “y0utube.com”), and sound squatting (e.g., “yewtube.com”) would be easily applied to the token names/symbols to mislead users.

**Cryptocurrency wallets, exchanges and blockchain browsers.** Cryptocurrency wallets, exchanges and blockchain browsers have the responsibility to detect counterfeit tokens and protect users from being scammed. Our approach can be integrated within major exchanges and wallets to stop such scams. For example, once a new counterfeit token is found, our approach can help flag all the suspicious scam addresses related to it and warn users before they interact with these addresses. We observe that major blockchain browsers (e.g., Etherscan and Bloxy) have started to flag scam addresses using their own approaches, thus our work can also be implemented in these browsers to help flag counterfeit token related scams and remind users timely.

**Cryptocurrency creators** The official cryptocurrency creators should be aware of the counterfeit token abuse. They should take the responsibility to search and identify counterfeit tokens and even fake social networking accounts (e.g., Telegram and Twitter). In such cases, cryptocurrency creators could then take actions to mitigate possible abuses (e.g., by reporting them to regulators and investors). Furthermore, they should regularly post public announcement to remind users.

**Investors** Awareness should also be raised among investors. For educational purposes, we commit to post regular tutorials and reports to provide a means for regulators, cryptocurrency creators and investors to learn more about counterfeit tokens. More importantly, investors should keep in mind that there is no such thing as a free lunch in the cryptocurrency world.

**Advertising Channels.** Finally, as we have identified a number of advertising channels exploited by attackers, including reputable ones, it is also urgent to regulate the contents published on these platforms, which can help reduce the propagation of scams.

### 7.2 Limitation

Our work carries several limitations. First, we only study counterfeit tokens related to the top-100 official cryptocurrencies on Ethereum. Although our observation suggests that the attackers are more likely to target popular tokens with a high market capitalization rank, it is quite possible that there are some counterfeit tokens targeting other cryptocurrencies beyond our study. Second, counterfeit cryptocurrencies can target official tokens on any cryptocurrency platforms, while we only focus on the counterfeit ERC-20 tokens on Ethereum, as ERC-20 is the most popular token standard, accounting for over 90% of alternative tokens in the blockchain ecosystem. Nevertheless, we agree that the counterfeit token might exist in other blockchain platforms like EOSIO and Tron. We leave them for future work. Third, we have characterized two typical scams related to the counterfeit tokens by resorting to existing scam reports. However, it is possible that there are other scams related to counterfeit tokens we did not cover. Finally, although we tried our best to understand the overall workflow of counterfeit cryptocurrency scams, we lack a deep understanding of the attackers behind the scams. The social engineering techniques and the advertising posts we

identified may be just the tip of the iceberg. Nevertheless, this paper presents the lower bound impact analysis of the counterfeit cryptocurrency scams.

## 8 RELATED WORK

### 8.1 Blockchain Scams

Since the birth of blockchain, various kinds of scams have emerged. A number of studies have characterized blockchain scams. Vasek and Moore [64] surveyed the presence of Bitcoin scams, including Ponzi schemes, mining scams, scam wallets, and fraudulent exchanges. After that, some other studies have characterized various scams including Ponzi schemes [25, 26, 33, 35, 62, 63, 65], scam Initial Coin Offerings (ICOs) [37, 49, 53, 72], market manipulation of cryptocurrencies [32, 33, 41–43], blockchain honeypots [61], and phishing scams [57, 69, 70].

ICO scams are most relevant to this paper. For example, Alexander et al. [39] built a predictive model by applying natural language processing (NLP) and machine learning techniques to detect ICO scams. Shuqing et al. [27] created ICORating, a learning-based cryptocurrency rating system. They have analyzed 2,251 cryptocurrencies from a number of perspectives, including whitepapers, founding teams, Github repositories, websites, etc. *Counterfeit cryptocurrency, as a new emerging threat, has not been systematically studied yet*. In this work, we take the first step to characterize counterfeit tokens and study their relevant scams.

### 8.2 ERC-20 Tokens

A few studies have characterized the ERC-20 token ecosystem [30, 34, 38, 54, 66]. For example, Chen et al. [34] investigated the Ethereum ERC20 token ecosystem to characterize the token creator, holder, and transfer activity. Friedhelm Victor et al. [66] provided an overview of more than 64,000 ERC20 token networks and analyzed the top 1,000 from a graph perspective. Besides, there are some studies dedicated to optimizing the ERC-20 token standard [51, 58] or using ERC-20 token contract to address practical issues [36, 48]. For example, Mayer et al. [51] proposed a proxy scaling solution for ERC-20 tokens named BatPay, which is suitable for micropayments in one-to-many and few-to-many scenarios and can reduce gas cost of transactions. Christodoulou et al. [36] designed a smart contract that can interact with any ERC-20 token to help decentralised organizations run public voting campaigns and engage token holders in voting decision.

### 8.3 Blockchain Transaction Analysis

A number of studies have investigated blockchain systems by performing transaction-based analyses. Several studies are focused on Bitcoin [28, 40, 50, 56, 59, 60, 73], including de-anonymization and money laundering detection, by using graph-based approaches. Researchers have also investigated Ethereum and EOSIO by using transaction-based analyses. For example, Chen et al. [31] performed a graph-based analysis of Ethereum to characterize activities including money transfer, smart contract creation and smart contract invocation. Huang et al. [46] analyzed the transactions on EOSIO blockchain, and developed techniques to automatically detect bots and fraudulent activities.

## 9 CONCLUSION

This paper has presented the first in-depth measurement study of counterfeit tokens on Ethereum. Our study has revealed that counterfeit tokens are prevalent in the cryptocurrency ecosystem, thereby motivating the need for more efforts to identify and prevent cryptocurrency abuses. We have characterized two kinds of scams related to counterfeit tokens, and designed methods to identify airdrop scams and arbitrage scams. At least 7,104 victims have been scammed and the

overall profit is over \$17,352,840.00. By studying the advertising channels for the counterfeit tokens and scams, we find 103 platforms have been exploited to spread fraudulent information.

## ACKNOWLEDGMENT

We sincerely thank our shepherd Prof. Stefan Schmid (University of Vienna) and all the anonymous reviewers for their valuable suggestions and comments to improve this paper. This work was supported by the National Natural Science Foundation of China (grants No.61702045 and No.62072046), Hong Kong RGC Project (No. 152193/19E), the Fundamental Research Funds for the Central Universities, Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (2018R01005). Haoyu Wang (haoyuwang@bupt.edu.cn) is the corresponding author.

## REFERENCES

- [1] \$1.1 billion in cryptocurrency has been stolen this year, and it was apparently easy to do. <https://www.cnbc.com/2018/06/07/1-point-1b-in-cryptocurrency-was-stolen-this-year-and-it-was-easy-to-do.html>, 2018.
- [2] Three types of cryptocurrency tokens explained as quickly as possible. <https://thenextweb.com/hardfork/2018/11/19/cryptocurrency-tokens-explained/>, 2018.
- [3] Binance exchange hackers steal bitcoins worth \$41m. <https://www.bbc.com/news/technology-48199375>, 2019.
- [4] Bctalk. <https://bctalk.info/>, 2020.
- [5] Bitcointalk, 2020. <https://bitcointalk.org/>.
- [6] Bytechats. <https://bytechats.info/>, 2020.
- [7] Coin hunter. <https://coinhunter.io>, 2020.
- [8] Facebook. <https://www.facebook.com/>, 2020.
- [9] Imtoken wallet security. <https://imtoken.fans/c/18-category/33-category>, 2020.
- [10] Libra (digital currency), 2020. [https://en.wikipedia.org/wiki/Libra\\_\(digital\\_currency\)](https://en.wikipedia.org/wiki/Libra_(digital_currency)).
- [11] Medium. <https://medium.com/>, 2020.
- [12] Sina. <http://blog.sina.com.cn/>, 2020.
- [13] Steemkr. <https://steemkr.com/>, 2020.
- [14] Telegram. <https://telegram.org/>, 2020.
- [15] Telegraph. <https://telegra.ph/>, 2020.
- [16] Telemetr. <https://telemetr.me/>, 2020.
- [17] Tgchannels. <https://ru.tgchannels.org/>, 2020.
- [18] Token tracker. <https://cn.etherscan.com/tokens>, 2020.
- [19] Total crypto market capitalization and volume. <https://www.tradingview.com/markets/cryptocurrencies/global-charts/>, 2020.
- [20] Total supply, 2020. <https://academy.binance.com/en/glossary/total-supply>.
- [21] Twipu. <https://www.twipu.com/>, 2020.
- [22] V2ex. <https://www.v2ex.com/>, 2020.
- [23] Youtube. <https://www.youtube.com/>, 2020.
- [24] Zhihu. <https://www.zhihu.com/>, 2020.
- [25] Massimo Bartoletti, Salvatore Carta, Tiziana Cimoli, and Roberto Saia. Dissecting ponzi schemes on ethereum: identification, analysis, and impact. *Future Generation Computer Systems*, 102:259–277, 2020.
- [26] Massimo Bartoletti, Barbara Pes, and Sergio Serusi. Data mining for detecting bitcoin ponzi schemes. In *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pages 75–84. IEEE, 2018.
- [27] Shuqing Bian, Zhenpeng Deng, Fei Li, Will Monroe, Peng Shi, Zijun Sun, Wei Wu, Sikuang Wang, William Yang Wang, Arianna Yuan, Tianwei Zhang, and Jiwei Li. Icorating: A deep-learning system for scam ico identification. *arXiv preprint arXiv:1803.03670*, 2018.
- [28] Alex Biryukov, Dmitry Khovratovich, and Ivan Pustogarov. Deanonymisation of clients in bitcoin p2p network. In *Proceedings of CCS 2014*. ACM, 2014.
- [29] Connor Blenkinsop. Arbitrage trading in crypto, explained. <https://cointelegraph.com/explained/arbitrage-trading-in-crypto-explained>, 2019.
- [30] Ting Chen, Yufei Zhang, Zihao Li, Xiapu Luo, Ting Wang, Rong Cao, Xiuzhuo Xiao, and Xiaosong Zhang. Tokenscope: Automatically detecting inconsistent behaviors of cryptocurrency tokens in ethereum. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS '19*, page 1503–1520, 2019.

- [31] Ting Chen, Yuxiao Zhu, Zihao Li, Jiachi Chen, Xiaoqi Li, Xiapu Luo, Xiaodong Lin, and Xiaosong Zhang. Understanding ethereum via graph analysis. In *IEEE International Conference on Computer Communications (INFOCOM)*, 2018.
- [32] Weili Chen, Jun Wu, Zibin Zheng, Chuan Chen, and Yuren Zhou. Market manipulation of bitcoin: evidence from mining the mt. gox transaction network. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 964–972. IEEE, 2019.
- [33] Weili Chen, Yuejin Xu, Zibin Zheng, Yuren Zhou, Jianxun Eileen Yang, and Jing Bian. Detecting "pump & dump schemes" on cryptocurrency market using an improved apriori algorithm. In *2019 IEEE International Conference on Service-Oriented System Engineering (SOSE)*, pages 293–2935. IEEE, 2019.
- [34] Weili Chen, Tuo Zhang, Zhiguang Chen, Zibin Zheng, and Yutong Lu. Traveling the token world: A graph analysis of ethereum erc20 token ecosystem. In *Proceedings of The Web Conference 2020*, pages 1411–1421, 2020.
- [35] Weili Chen, Zibin Zheng, Jiahui Cui, Edith Ngai, Peilin Zheng, and Yuren Zhou. Detecting ponzi schemes on ethereum: Towards healthier blockchain technology. In *Proceedings of the 2018 World Wide Web Conference*, pages 1409–1418, 2018.
- [36] Panayiotis Christodoulou and Klitos Christodoulou. A decentralized voting mechanism: Engaging erc-20 token holders in decision-making. In *2020 Seventh International Conference on Software Defined Systems (SDS)*, pages 160–164. IEEE, 2020.
- [37] Patrick Schueffel Daniel Liebau. Cryptocurrencies&initial coin offerings: Are they scams? - an empirical study. volume 2, pages 47–55, 2019.
- [38] Simon F Dyson, William J Buchanan, and Liam Bell. Scenario-based creation and digital investigation of ethereum erc20 tokens. *Forensic Science International: Digital Investigation*, 32:200894, 2020.
- [39] Dürr, Alexander, Griebel, Matthias, Welsch, Giacomo, Thiesse, and Frédéric. Predicting fraudulent initial coin offerings using information extracted from whitepapers. *the 28th European Conference on Information Systems(ECIS),An Online AIS Conference, June 15-17, 2020*.
- [40] Michael Fleder, Michael S. Kester, and Sudeep Pillai. Bitcoin transaction graph analysis, 02 2015.
- [41] Neil Gandal, JT Hamrick, Tyler Moore, and Tali Oberman. Price manipulation in the bitcoin ecosystem. *Journal of Monetary Economics*, 95:86–96, 2018.
- [42] JT Hamrick, Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek. The economics of cryptocurrency pump and dump schemes. 2018.
- [43] JT Hamrick, Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek. An examination of the cryptocurrency pump and dump ecosystem. *Available at SSRN 3303365*, 2018.
- [44] Kevin Helms. Chinese authorities confiscate \$15 million in cryptocurrencies, arrest 10 scammers. <https://news.bitcoin.com/chinese-authorities-confiscate-15-million-cryptocurrencies/>, 2020.
- [45] Yangyu Hu, Haoyu Wang, Ren He, Li Li, Gareth Tyson, Ignacio Castro, Yao Guo, Lei Wu, and Guoai Xu. Mobile app squatting. In *Proceedings of The Web Conference 2020*, pages 1727–1738, 2020.
- [46] Yuheng Huang, Haoyu Wang, Lei Wu, Gareth Tyson, Xiapu Luo, Run Zhang, Xuanzhe Liu, Gang Huang, and Xuxian Jiang. Understanding (mis)behavior on the eosio blockchain. *Proc. ACM Meas. Anal. Comput. Syst.*, 4(2), June 2020.
- [47] Panagiotis Kintis, Najmeh Miramirkhani, Charles Lever, Yizheng Chen, Rosa Romero-Gómez, Nikolaos Pitropakis, Nick Nikiforakis, and Manos Antonakakis. Hiding in plain sight: A longitudinal study of combosquatting abuse. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS '17*, page 569–586, 2017.
- [48] Mirko Koscina, Mariusz Lombard-Platet, and Pierre Cluchet. Plasticcoin: an erc20 implementation on hyperledger fabric for circular economy and plastic reuse. In *IEEE/WIC/ACM International Conference on Web Intelligence-Companion Volume*, pages 223–230, 2019.
- [49] Dan Liebau and Patrick Schueffel. Crypto-currencies and icos: Are they scams? an empirical study. *An Empirical Study (January 23, 2019)*, 2019.
- [50] Damiano Di Francesco Maesa, Andrea Marino, and Laura Ricci. An analysis of the bitcoin users graph: inferring unusual behaviours. In *International Workshop on Complex Networks and their Applications*, 2016.
- [51] Hartwig Mayer, Ismael Bejarano, Daniel Fernandez, Gustavo Ajzenman, Nicolas Ayala, Nahuel Santoalla, Carlos Sarraute, and Ariel Futoransky. Batpay: a gas efficient protocol for the recurrent micropayment of erc20 tokens. *arXiv preprint arXiv:2002.02316*, 2020.
- [52] Rachel McIntosh. Nearly a year after libra's launch, fake libra token scams persist. <https://www.financemagnates.com/cryptocurrency/news/nearly-a-year-after-libras-launch-fake-libra-token-scams-persist/>, 2020.
- [53] Adrian Gepp& Kuldeep Kumar Milind Tiwari. The future of raising finance-a new opportunity to commit fraud: a review of initial coin offering(icos) scams. In *Crime Law Soc Change* 73, pages 417–441. Springer, 2020.
- [54] Alfredo J Morales, Shahar Somin, Yaniv Altshuler, and Alex'Sandy' Pentland. User behavior and token adoption on erc20. *arXiv preprint arXiv:2005.12218*, 2020.

- [55] Masarah Paquet-Clouston, Matteo Romiti, Bernhard Haslhofer, and Thomas Charvat. Spams meet cryptocurrencies: Sextortion in the bitcoin ecosystem. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies (AFT '19)*, page 76–88, 2019.
- [56] Silivanxay Phetsouvanh, Frédérique Oggier, and Anwitaman Datta. Egret: Extortion graph exploration techniques in the bitcoin network. In *IEEE International Conference on Data Mining Workshops (ICDMW)*, 2018.
- [57] Ross Phillips and Heidi Wilder. Tracing cryptocurrency scams: Clustering replicated advance-fee and phishing websites. *arXiv preprint arXiv:2005.14440*, 2020.
- [58] Reza Rahimian, Shayan Eskandari, and Jeremy Clark. Resolving the multiple withdrawal attack on erc20 tokens. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 320–329. IEEE, 2019.
- [59] Fergal Reid and Martin Harrigan. An analysis of anonymity in the bitcoin system. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. IEEE, 2011.
- [60] Dorit Ron and Adi Shamir. Quantitative analysis of the full bitcoin transaction graph. In *Financial Cryptography (FC)*, 2013.
- [61] Christof Ferreira Torres, Mathis Steichen, et al. The art of the scam: Demystifying honeypots in ethereum smart contracts. In *28th {USENIX} Security Symposium ({USENIX} Security 19)*, pages 1591–1607, 2019.
- [62] Kentaro Toyoda, P Takis Mathiopoulos, and Tomoaki Ohtsuki. A novel methodology for hyip operators' bitcoin addresses identification. *IEEE Access*, 7:74835–74848, 2019.
- [63] Kentaro Toyoda, Tomoaki Ohtsuki, and P Takis Mathiopoulos. Identification of high yielding investment programs in bitcoin via transactions pattern analysis. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.
- [64] Marie Vasek and Tyler Moore. There's no free lunch, even using bitcoin: Tracking the popularity and profits of virtual currency scams. In *International conference on financial cryptography and data security*, pages 44–61. Springer, 2015.
- [65] Marie Vasek and Tyler Moore. Analyzing the bitcoin ponzi scheme ecosystem. In *International Conference on Financial Cryptography and Data Security*, pages 101–112. Springer, 2018.
- [66] Friedhelm Victor and Bianca Katharina Luders. Measuring ethereum-based erc20 token networks. In *International conference on financial cryptography and data security*. Springer, 2019.
- [67] Wikipedia. Airdrop (cryptocurrency). [https://en.wikipedia.org/wiki/Airdrop\\_\(cryptocurrency\)](https://en.wikipedia.org/wiki/Airdrop_(cryptocurrency)), 2020.
- [68] Wikipedia. Counterfeit money. [https://en.wikipedia.org/wiki/Counterfeit\\_money](https://en.wikipedia.org/wiki/Counterfeit_money), 2020.
- [69] Jiajing Wu, Qi Yuan, Dan Lin, Wei You, Weili Chen, Chuan Chen, and Zibin Zheng. Who are the phishers? phishing scam detection on ethereum via network embedding. *arXiv preprint arXiv:1911.09259*, 2019.
- [70] Pengcheng Xia, Haoyu Wang, Xiapu Luo, Lei Wu, Yajin Zhou, Guangdong Bai, Guoai Xu, Gang Huang, and Xuanzhe Liu. Don't fish in troubled waters! characterizing coronavirus-themed cryptocurrency scams, 2020.
- [71] Pengcheng Xia, Bowen Zhang, Ru Ji, Bingyu Gao, Lei Wu, Xiapu Luo, Haoyu Wang, and Guoai Xu. Characterizing cryptocurrency exchange scams. *arXiv preprint arXiv:2003.07314*, 2020.
- [72] Dirk A Zetsche, Ross P Buckley, Douglas W Arner, and Linus Föhr. The ico gold rush: It's a scam, it's a bubble, it's a super challenge for regulators. *University of Luxembourg Law Working Paper*, (11):17–83, 2017.
- [73] Chen Zhao and Yong Guan. A graph-based investigation of bitcoin transactions. In *11th IFIP International Conference on Digital Forensics (DF)*, 2015.

Received August 2020; revised September 2020; accepted October 2020