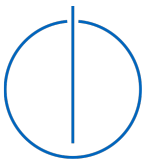
DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Information Systems

Towards an Online Service to Detect NFT Wash Trading Activities on the Ethereum Blockchain

Baran Kalkavan



DEPARTMENT OF INFORMATICS

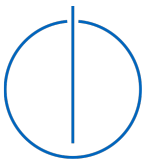
TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Information Systems

Towards an Online Service to Detect NFT Wash Trading Activities on the Ethereum Blockchain

Auf dem Weg zu einem Online-Dienst zur Erkennung von NFT Wash-Trading-Aktivitäten auf der Ethereum-Blockchain

Author:	Baran Kalkavan
Supervisor:	Prof. Dr. Florian Matthes
Advisor:	Burak Öz
Submission Date:	15.10.2022



I confirm that this bachelor's thesis in information systems is my own work and I have documented all sources and material used.

Munich, 15.10.2022 Baran Kalkavan

Acknowledgments

Foremost, I would like to give my warmest appreciation to my supervisor Burak Öz, who made this work possible. His support and guidance helped me throughout each stage of my thesis.

I also want to express my gratitude to Prof. Dr. Florian Matthes for allowing me to write my Bachelor's thesis in his chair and for his constructive feedback.

Abstract

NFTs have become very popular following the appreciation of Bitcoin and other cryptocurrencies. Being able to open an account on NFT marketplaces easily and the absence of laws regulating NFT trading also increased this popularity. With the increasing number of users, scams in the NFT ecosystem have also increased. One of the most popular scams is wash trading, in which the NFTs are traded between separate accounts belonging to the same owner or group to create artificial trading volume. This paper presents an algorithm that can detect wash trading and a methodology to create a web application that analyzes suspicious activities in the given NFT collection. According to our findings, most wash trading activities happen on the marketplace LooksRare, and 88% of the total trading volume of the biggest collection on this NFT marketplace is suspicious.

Contents

1 Introduction

1.1 Motivation

With the development of blockchain technology, cryptocurrencies and Non-fungible tokens (NFT) have started to gain popularity and are considered as a widespread trading instruments. NFTs enable the tokenization of assets, and they are, in contrast to cryptocurrencies, unique. Contrary to popular belief, NFTs are not just digital art. For example, NFT can also be created from an MP3 or a tweet and can still be very valuable. Exemplarily, the NFT of Twitter founder Jack Dorsey's first tweet was sold for $2.9 million in 2021 [Han22].

NFTs are mostly traded on the NFT marketplaces (NFTM). These marketplaces allow users to buy, sell and even create NFTs. The largest NFTM by total trading volume is OpenSea. It has more than 1 million active users in 2022, and more than 80 million NFTs are displayed on the marketplace [Wis22]. Other popular NFTMs are Rarible, SuperRare, Foundation, and LooksRare. Different NFTMs have different trading fees and divergent policies. For example, LooksRare rewards its users for every transaction they execute and for each NFT they list on the platform [Con22]. Such interventions disrupt the market dynamics and cause an increase in market manipulations.

Wash Trading is at the forefront of these market manipulations in the NFT ecosystem. A blockchain data platform, "Chainalysis," defines in their 2022 Crypto Crime Report wash trading as "*a transaction in which the seller is on both sides of the trade to paint a misleading picture of an asset's value and liquidity.*" This activity is illegal in the centralized finance markets. However, the lack of legal regulations and the diversity in NFTMs led to increased wash trading activities in the NFT ecosystem.

According to our research, existing literature on this topic focuses on analyzing the wash trading activities in the top collections by their trade volume of different NFTMs. There is a lack of tools where users can get insights about the wash trading activity of a collection that they desire to see. This thesis aims to fill that gap by implementing an efficient online service where users can gain insights into the wash trading volume of an NFT collection.

In this report, existing wash trading detection algorithms will be reviewed and the algorithm used in this work will be introduced. The methodology for creating the tool will be explained. The performance of the tool will be analyzed based on the wash

trading detection rate (compared to existing reports) and application speed.

1.2 Research Questions

This thesis addresses the following research questions:

1. **RQ1: Which research has been conducted so far on detecting wash trading on NFT marketplaces?**

   The first research question aims to understand the state of research on wash trading, specifically NFT wash trading, in order to understand what algorithms are used and what insights are gained. Additionally, results from the existing literature are used as a benchmark to assess the performance of the developed tool.

2. **RQ2: What insights can an online service provide regarding an NFT collection?**

   This question focuses on the potential findings that can be provided to the end user for a given collection, in an efficient manner. It aims to explore what are the constraints regarding the insights that can be provided (e.g., data).

3. **RQ3: What are the most common wash trading patterns?**

   The third research question helps to identify wash trading activities. Wash trading does not have a unique definition, and there are different ways to conduct wash trading, so different parameters should be considered when identifying wash trading activities.

4. **RQ4: Can wash trading be detected by an online service efficiently?**

   The fourth research question examines the efficiency of the web application. Fetching the transaction data from another service provider and analyzing this data can be time-consuming, and wash trading should be detected in a reasonable time.

5. **RQ5: Can wash trading activity be avoided/regulated by marketplaces?**

   The last research question will focus on the marketplaces. It is known that, while some marketplaces are trying to prevent wash trading activities, others' policies lead their users to wash trade, so it is important to have an overview of the NFTMs about their wash trading policies.

2 Background Information

To understand and analyze the problem of wash trading in the NFT ecosystem, it is necessary to know how blockchains and wash trading work.

2.1 Blockchain

A blockchain is a distributed database that can store any digital data using the concepts of decentralization and cryptographic hashing. The idea of a blockchain was first introduced in 1991 by the American researchers Stuart Haber and W. Scott Stornetta. A year later, the concept of Merkle trees is integrated into the system. Merkle trees are data structures which are used to store transactions and other state data (depending on the blockchain) in an efficient way. Later, in 2008, Satoshi Nakamoto released a white paper about the concept of a distributed blockchain, and the first Bitcoin purchase took place in 2010 [Jav].

Each block on a blockchain contains data, hash, and hash of the previous block. Hashes are unique identifiers of a block created by the hash functions. Hash functions are one-way functions, and that's why it is infeasible to find an input for a given output. Hashing ensures the chaining of the blocks is immutable. Thus, one cannot manipulate the data on the chain, once it is put there. Blockchains also use peer-to-peer networks and work as a distributed system. In 2022, more than 1000 blockchains exist [Law22].

2.2 Ethereum

Ethereum is an open-source, decentralized blockchain that supports smart contracts, founded by Vitalik Buterin and Gavin Wood. It is mainly known for its cryptocurrency, ether (ETH). Ethereum has introduced the concept of NFTs, and most of the NFTs are part of this blockchain. That is why this thesis is based on the Ethereum blockchain.

Ethereum Smart Contracts are programs written in the programming language "Solidity" that run on the blockchain and execute transactions if the predefined conditions are met. Interactions with smart contracts are irreversible. Smart contracts are executed by the Ethereum Virtual Machine(EVM), which holds Ethereum's accounts and balances.

Ethereum has numerous different use cases. Other than tokenizing real-world assets, it is widely used for Decentralized Finance (DeFi). With DeFi, one can create loans supported by smart contracts and create stable coins. DeFi also allows decentralized exchanges. Additionally, Ethereum can be used in health applications to store patient data securely, in digital identity, in voting systems, or to store data [22].

Table 2.1: Top 10 most expensive NFTs ever sold

	NFT	Price
1	Pak's 'The Merge'	$91.8m
2	Everydays: the First 5000 Days	$69.3m
3	Clock	$52.7m
4	Beeple's HUMAN ONE	$28.9m
5	CryptoPunk #5822	$23.7m
6	CryptoPunk #7523	$11.75m
7	CryptoPunk #4156	$10.26m
8	CryptoPunk #3100	$7.67m
9	CryptoPunk #7804	$7.6m
10	Beeple's Crossroad	$6.6m

2.3 NFT

Non-fungible tokens (NFT) are unique digital assets stored in a blockchain. NFTs have references to digital files, such as a jpeg of artwork, an mp3 of a song, or a web link. The process of turning real-world assets into NFTs is called "tokenizing."

ERC721 is the non-fungible token standard of Ethereum. It defines how a non-fungible token (NFT) can be issued. Each NFT has to have an id field that is unique in that collection. This standard has different capabilities, like transferring tokens between the accounts and getting the token balance of an account [Wac22].

In 2021 NFTs started to gain popularity, and dictionary publisher Collins chose "NFT" as the year's word. In the same year, the use of the term increased by 11,000% [Bha22]. There are several reasons behind the rise of NFTs. Firstly, Bitcoins and other crypto currencies' uptrend has played an important role. Secondly, Gen Z and X's tendency to easily invest online without the paperwork. The last reason was Corona Pandemic. People were stuck in their homes for a long time, and cultural experiences like concerts and exhibitions mostly took place online. As a part of this migration to online, art has transformed into digital art, and it has become reachable for many people with the

help of NFTs.

According to a report published by Verified Market Research (VMR) NFT market has a trading volume of $11.3 billion in 2021, and they predict this market value to pass $231 billion in the next ten years [Jen22]. Table 2.1 shows the most expensive NFTs sold in 2022.

2.4 NFT Marketplace

NFT Marketplaces(NFTM) are digital platforms where NFT collections are listed. In an NFT collection, a group of a limited number of unique NFTs is contained. These marketplaces give detailed information about the NFT collections and their transaction history. They are built user-friendly so it is possible to sort collections by price or seller. In order to use these marketplaces, one has to have an account on these marketplaces and a digital wallet. As there exist different types of NFTs, there are also different types of NFTM. Some marketplaces are for digital collectibles and game characters, while others are for trading real estate and music tokens.

Table 2.2: Top 10 NFTM with the highest number of users

	NFTM	Number of Users	Trading Volume
1	Axie Marketplace	2139,8k	$4,2b
2	OpenSea	2091k	$39,91b
3	AtomicMarket	1071k	$427,35m
4	Magic Eden	889,2k	$1,6b
5	NBA Top Shot	566k	$964,35m
6	Solanart	233,1k	$656,69m
7	BloctoBay	135,1k	$400,14m
8	Rarible	106,9k	$298,68m
9	Lookrare	99k	$1,59b
10	X2Y2	98k	$730,96m

Table 2.2 shows the 10 NFTMs with the most users. Axie marketplace is the NFTM with the most traders, but OpenSea has the highest trading volume. The trading volume is a better metric to compare NFTMs as one can open multiple accounts in the same marketplace. This should be considered when looking at the number of traders.

2.5 NFT Trading

1. **The process of creating an NFT is called *NFT Minting*. When an NFT is minted, it can be listed on an NFTM.**
2. **To buy an NFT, one has to have an account on the NFTM and a digital wallet to pay and store the NFT.**
3. **When the NFT is purchased, sale and transfer events are triggered in most of the NFTMs. The sale event shows the money transfer between the buyer and seller accounts. The transfer event is responsible for the ownership rights of the NFT, and the ownership is transferred from the seller (from_address) to the buyer (to_address).**
4. **When one has ownership of the NFT, he can again sell it to someone else. This process is called *Flipping an NFT*.**

2.6 Frauds

The lack of legal regulations and the diversity in the marketplaces lead to increasing scams in the NFT ecosystem. Different scams exist, like rug-pull scams, phishing, plagiarized NFTs, and wash trading.

With the *rug-pull method*, scammers try to convince traders to buy an NFT. When the transaction occurs, and scammers receive their money, scammers leave the market with their money.

Phishing is another common scam in the NFT Ecosystem. Scammers are using fake advertisements, pop-ups, messages, or emails to get passwords and user details of the victim. The obtained information allows scammers to access the victim's account and transfer coins and NFTs to their wallets.

Plagiarized NFTs are the biggest problem of the artists in the NFT market. One can easily steal and copy the image of an NFT and create a new NFT using the same image. The biggest NFTM has reported that 80% of the NFTs on the platform are fake [Coh22].

2.7 Wash Trading

A wash trade is a market manipulation in which the same person or connected traders regularly buy and sell a financial instrument to create an artificial trading volume. Wash trading has been illegal in the United States of America since 1936, but this ban does not apply to NFTMs. That is why it is widespread in the NFT ecosystem.

There is not a single wash trading pattern, but an example wash trading scenario in the NFT ecosystem works as follows: a scammer mints an NFT and sells it to his second account for 5eth. Then he buys it back for 20eth. Now it looks like the NFT is worth 20eth, but it is only an artificial trading volume. If a third-party user is deceived by the apparent value of the NFT and buys it, the user becomes a victim of the scam, and the scammer gets more money than its NFT is worth.

3 Literature Research

Existing literature on wash trading activities in the NFT ecosystem is mostly focused on the most popular NFTMs or the most popular NFT collections. Additionally, there is various research about the wash trading activities on decentralized cryptocurrency exchanges, which uses relevant methodology and has relevant results for our work.

Friedhelm et al. analyze in their paper "Detecting and Quantifying Wash Trading on Decentralized Cryptocurrency Exchanges" wash trading instances in two decentralized exchanges IDEX and Etherdelta to a time frame between 02.09.2017 and 05.04.2020 [VW21].

In the paper, they model the trades of a specific type of token as G(V,E), which is a directed multigraph. V is the set of trading accounts, and E is the set of trades.

Their wash trade detection algorithm consists of two parts. The first part is called "Account Candidate Set Generation," which counts iteratively the strongly connected components in each token trade graph to identify wash traded accounts. The second part of the algorithm is called "Trade Volume Matching." This part of the algorithm helps to identify trade subsets where there is no position change for the traders.

In their paper, researchers conclude that wash trading is more frequent at the beginning and at the end of a token's life. Their results highlight that 31,54% of the tokens on IDEX and 42,12% of the tokens on EtherDelta were part of at least one wash trading activity.

In the paper "NFT Wash Trading- Quantifying suspicious behavior in NFT markets," Wachter et al. inspect wash trading activities in the 52 largest NFT collections on the Ethereum blockchain by volume. They obtain the transaction data for the time period January 2019 to mid-November 2021 using OpenSea API [Wac+21].

Researchers build each NFT as a directed multigraph G_{NFT}=(N,E) . Identical to the graphs in Friedhelm et al.'s paper, N represents the set of accounts, and E is the set of transactions. The direction of the edges is from the seller to the buyer.

Their paper utilizes Dept-First-Search-Algorithm to identify closed cycles in the multigraphs. These closed cycles help to identify wash trading accounts. This algorithm has a time complexity $\mathcal{O}((n+e)(c+1))$. n represents the addresses, e represents the transactions, and c stands for cycles. After utilizing this algorithm on their data set, they also search for path-like transaction patterns, which are frequently repeated trades in the same order that does not form a closed cycle. In contrast to cryptocurrencies,

the search for wash trading activities in the NFT ecosystem doesn't require volume matching since NFTs are unique and can be easily identified with their id.

According to their findings, 75% of the closed cycles take place within 30 days. Their results also show that 3,93% of the total accounts involved in a wash trading activity and 2,04% of the total transactions are considered as wash traded transactions.

In the paper "Understanding Security Issues in the NFT Ecosystem," Sas et al. analyze different security problems in NFTMs, and one of the main problems mentioned in the paper is wash trading. They analyze the top seven NFTMs on the Ethereum Blockchain by their trading volume to detect possible wash trading activities [Das+22].

In their paper, they build different graphs to analyze wash trading. A sales graph G_{st} (sale), a payment graph G_{pt} (paid), and an asset transfer graph G_t (transfer).

Sas et al. are using strongly connected components(SCC) and weakly connected components(WCC) to detect wash trading activities. An SCC of a directed graph is a maximal connected sub-graph, which means that every vertex is reachable from every other vertex in this sub-graph, considering the direction of the edges. WCC of a directed graph is a sub-graph where all nodes are reachable from every other node irrespective of the direction of the edges. In the paper, they identify an address as a wash trade conducted address if: SCC(u1,u2,G_{st}) V WCC(u1,u2,G_t) V WCC(u1,u2,G_{pt}) with the limitation that both users u1 and u2 should be in the same SCC more than 10 times.

Using this methodology, Sas et al. find 9393 wash trading instances in 5297 NFT collections, which have 17821 users. These wash trading activities have generated $96,858,093.

Chainalysis also has a section about wash trading activities in NFTs on their "2022 Crypto Crime Report". They do not publish their methodology and the algorithm that they use to identify wash trading activities [TEA22]. They analyze the NFT sellers by the number of sales to self-financed addresses. Contrary to the previous articles mentioned, Chainanlysis analyzes if the wash trading activities are profitable or not. By analyzing the historical transaction data, they identify 262 users that have sold NFTs to their self-controlled addresses more than 25 times. 110 addresses are profitable, and they have a profit of $8,875,315, but 152 addresses are not profitable and have a loss of $8,458,331.

3.1 Wash Trading Policies of NFTMs

Most of the NFTMs don't have anti-wash trading policies, and they disregard the wash trading activities. There is a lack of literature about the wash trading prevention mechanisms of the NFTMs, but some websites that track NFTs have some small reports

on that topic.

In 2020 Nonfungible.com finds out that many NFT sales on the platform Rarible, which have a value of more than $1000, are part of a wash trading activity because Rarible rewarded the users with the token RARI for those transactions. After this finding, Rarible responded with a Tweet that wash traders are being banned from the marketplace upon detection. But the explicit number of users and Rarible's definition of wash trading are uncertain [Hoo20].

LooksRare is the NFTM where the most wash trading activities occur. According to its documentation, this NFTM rewards its users with Looks coins when they trade an NFT on the platform. This system leads users to trade between the self-controlled accounts because both the buyer and seller of an NFT earns Looks coins. As a result of this, the platform witnesses an increased number of wash trading activities, and NFT collections on the platform have fake trading volumes.

CryptoSlam's article "Wash Trading: Who, What, Why, and What Should We Do About It?" is one of the first reports on this topic. According to their finding, most of the wash trading activities occur on the so-called royalty-free collections, where the secondary sale fee is zero. Normally, traders have to pay additionally to the platform fee (2%) royalty fee of between 5%-10%. CrptoSlam identifies $8,3 billion worth of wash trading on the platform. Meebits, Terraforms, Loot, and CrptoPunks are the collections where most of the suspicious activities are identified [Cof22]. In Figure 3.1, the transaction history of the token 7385 of the collection Terraforms is shown. It can be seen that the latest transactions only occur between the same pair of addresses, which are probably controlled by the same person/group.

From		To	TokenID
0x4b33887d2647dbe37d...	→	0x86cdba7e3b63a9fb72...	7385
0x86cdba7e3b63a9fb72...	→	0x4b33887d2647dbe37d...	7385
0x4b33887d2647dbe37d...	→	0x86cdba7e3b63a9fb72...	7385
0x86cdba7e3b63a9fb72...	→	0x4b33887d2647dbe37d...	7385
0x4b33887d2647dbe37d...	→	0x86cdba7e3b63a9fb72...	7385
0x86cdba7e3b63a9fb72...	→	0x4b33887d2647dbe37d...	7385
0x4b33887d2647dbe37d...	→	0x86cdba7e3b63a9fb72...	7385
0x86cdba7e3b63a9fb72...	→	0x4b33887d2647dbe37d...	7385
0x4b33887d2647dbe37d...	→	0x86cdba7e3b63a9fb72...	7385
0x86cdba7e3b63a9fb72...	→	0x4b33887d2647dbe37d...	7385
0x4b33887d2647dbe37d...	→	0x86cdba7e3b63a9fb72...	7385
0x86cdba7e3b63a9fb72...	→	0x4b33887d2647dbe37d...	7385

Figure 3.1: Transaction history of the token 7385 on Terraforms

On the platform Dune, another important report was published in February 2022 [hil22]. According to this report, transactions that take place on LooksRare have some spikes at certain hours. This can be a result of automated software that wash traders are

using to earn rewards. This report was published according to the data in February, but in September 2022 the results are still similar. These results are visualized in Figure 3.2. In the report, a methodology to filter back and forth trades (where the same NFT is traded several times between the same pair of wallets) is presented and some collections have 50 times lower trade volume when this filter is applied. The trade volume of the top 4 collections on LooksRare is in Table 3.1.

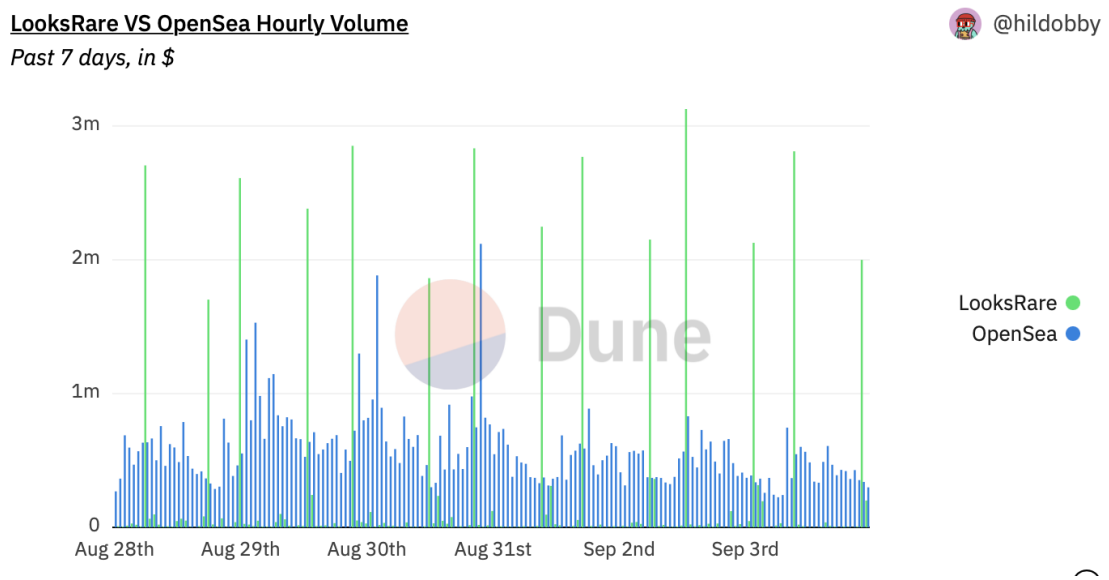


Figure 3.2: LooksRare VS OpenSea Hourly Volume

Table 3.1: Top 4 collections on LooksRare with and without Dune's filter

	Collection	Total Trade Volume(TTV)	TTV with "back and forth trades" filtered
1	Meebits	$6.279.054.549	$114.688.013
2	Terraforms	$4.982.091.516	$50.212.078
3	dotdotdot	$1.195.801.609	$2.668.913
4	Loot	$664.159.682	$107.937.106

4 Wash Trade Detection

To detect wash trading activities, the legal definitions of wash trading are identified, and the existing algorithms and methodologies are compared. After our analysis, different methodologies and algorithms are combined. The reason behind this decision is to detect wash trading activities as accurately as possible. If some suspicious transactions cannot be identified in the first part of the algorithm, they can be identified in the second one.

4.1 Trade Graph

To analyze the wash trading activities of an NFT collection, the trade graphs of all of the tokens on the given NFT collection are constructed. G_{NFT}=(V,E) is a directed multigraph, where V is the user address, and E is the set of trades. The direction of the edges is from the sender address to the receiver address. The weight of the edges is represented by the value of the transactions in eth. Figure 4.1 shows an example trade graph, which is a directed multigraph. In this graph, 4 different addresses take part in the trade of the token. While there are single trades between nodes 1,2, and 2,3, there are 3 trades between 3,4. Node 3 sells the NFT twice to node 4, and node 4 sells it one time to node 4.

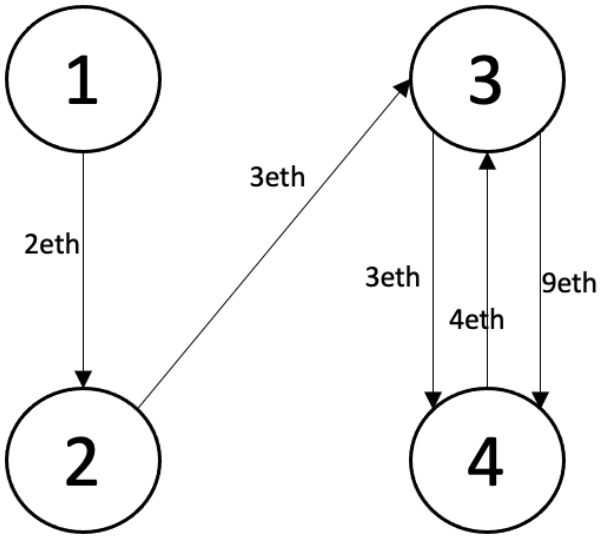


Figure 4.1: An example trade graph of an NFT

4.2 Detection Algorithm

In our algorithm, we use the idea of iteratively counting the strongly connected components in each NFT trade graph from the paper "Detecting and Quantifying Wash Trading on Decentralized Cryptocurrency Exchanges" [VW21] and improved it by additionally counting the weakly connected components(WCC) in the transfer graph.

The first step of the algorithm is detecting the strongly connected components of the trade graph of each token on the given NFT collection. Identifying SCCs helps us to find the path of transactions that form a cycle. Addresses, which are part of frequently repeated closed cycles, are mostly controlled by one person or connected people. Therefore finding frequent SCCs is important to identify suspicious trades and possible wash trading activities. The first step to identifying SCCs is to *simplify* the directed multigraphs G_{NFT}, which means that these multigraphs are transformed into directed graphs. The nodes still show the addresses, but the edges indicate the number of transactions between the nodes for the given direction.

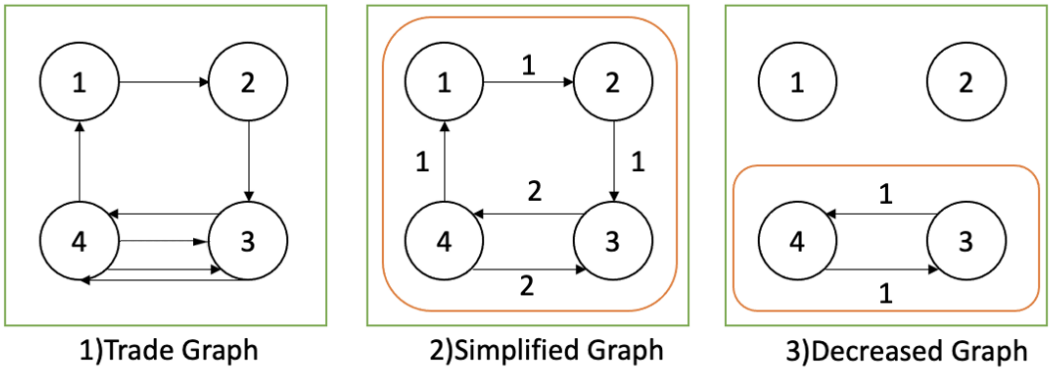


Figure 4.2: SCC detection process

After detecting the SCCs of the simplified graph, the weight of the edges decreases by 1, so that the subcycles of it can also be identified. If a weight reaches the value 0, that edge is removed from the graph. The iteration of decreasing the weights goes on until all the edges have a value of 0. In each iteration, SCCs of the current graph are identified and stored on a list.

This process is visualized in Figure 4.2. The trade graph consists of nodes from 1 to 4, and it is a directed multigraph. In the second graph, the trade graph is simplified, and its edges have the number of transactions in the given direction as their weight. In the first iteration, the SCC that consists of nodes 1,2,3,4 is identified. The sub-cycle between nodes 3 and 4 can't be identified in the first iteration because an SCC is a maximal strongly connected sub-graph. To solve this problem, the edges of the simplified graph will decrease by 1, and only the edges between nodes 3 to 4 will remain because all the other edges had a weight of 1. As the remaining edges reach weight 0, they are deleted from the graph. SCC of the new graph consists of nodes 4 and 3. There can't be any other sub-cycles because there will be no edges left in the next iteration of decreasing the weight. At the end, SCCs of {3,4}, and {1,2,3,4} are stored once on the list. The pseudocode for applying this process is given in algorithm 1.

After applying the mentioned SCC detection algorithm to every trade graph of NFTs in a collection, the next step is identifying the suspicious accounts. An SCC that occurs only one or two times can happen by chance. According to the definition of wash trading, accounts should frequently buy and sell NFTs. To reduce the probability of false positive labeling, we filter out the SCCs that occur less than 5 times. For all the other SCCs, we label every distinct address as suspicious. Any trade that occurs between two addresses that exist in the same SCC is then labeled as a wash trade.

Some transactions in our data set have no value attached to them. This means there

had been an NFT transfer without any ether transfer. This is called unconditional asset transfer as defined in the paper "Understanding Security Issues in the NFT Ecosystem" [Das+22]. Unconditional transfers are suspicious and need to be analyzed in detail. Weakly connected components of the transfer graphs of each token on the given NFT collection have to be identified to detect possible wash trading activities amongst the transfer events.

This process is identical to detecting the SCCs on the trade graphs, with the only difference being that the chosen empirical threshold here is 2. This is because in a transfer from_address sends NFTs to the to_address without getting paid. This shows that there has to be a relationship between the pair of users if a transfer between them occurs. As a result of this, we find WCCs of transfers more suspicious than SCCs and mark them as suspicious if they are repeated more than 2 times.

After receiving the suspicious SCCs and WCCs, all the addresses on these suspicious trades are obtained. Transactions, which's to- and from_addresses are part of the same suspicious SCC/WCC are marked as possible wash trading activity.

Data: NFT Collection C
Result: List of suspicious SCCs
$L \leftarrow$ *Create empty list*;
$SCC \leftarrow$ *Create empty set*;
foreach $NFT \in C$ **do**
  $G \leftarrow simplify(G_{NFT}, weight = number\ of\ edges)$;
  while $|E| > 0$ **do**
    L.add(getSCCs(G));
    foreach $e \in E$ **do**
      $weight(e) \leftarrow weight(\mathrm{e}) - 1$;
      if $weight(e) = 0$ **then**
        $E <- E/e$;
      end
    end
  end
end
foreach $l \in L$ **do**
  if $count(l) >= 5$ **then**
    $SCC.insert(l)$;
  end
end
return SCC;

Algorithm 1: SCC detection algorithm

5 Methodology

Creating a web application, which can apply the wash trading detection algorithm mentioned in the previous chapter, requires several steps. The first step is obtaining the transaction history of the given NFT collection.

5.1 Input Data

In this paper, transaction data is achieved through Moralis API. Web scraping and blockchain parsing require overhead, such as implementing a scraper script or running a full node. Since the thesis is time-constrained, we didn't use these methods.

Moralis is a web3 developing platform where users can create decentralized apps using the APIs, SDKs, and data provided by itself. It supports different Blockchains like Ethereum, Binance Smart Chain, Polygon, Avalanche, Cronos, Solana, and Fantom.

The NFT API of Moralis is used in this paper the most. This API supports both ERC721 and ERC1155. ERC-721 is a token standard for non-fungible tokens, and ERC-1155 is a multi-token standard that can be used to create non-fungible and semi-fungible tokens. Endpoints provide much information, including the metadata, ownership data, transfer data, and price data of the given NFT. NFT API has several endpoints to provide this information. Only the ones that are used in the web application will be mentioned in the following sections.

5.2 Retrieving Token List

"/nft/address" endpoint gets the NFTs of a collection. We are using this endpoint to get all token_id's of the collection to create their transaction graph and search for wash trading instances. It has the input fields; chain, format, limit, range, cursor, and address. Only the address field is compulsory. As an output it returns the list of tokens with their following information: token_address, token_id, owner_of, token_hash, block_number, block_number_minted, contract_type, token_uri, metadata, last_token_uri_sync, last_metadata_sync, amount, name, and symbol. Figure 5.1 shows an example result of this endpoint.

```
{
  "total": 2000,
  "page": 2,
  "page_size": 100,
  "cursor": "string",
  "result": [
    {
      "token_address": "0xb47e3cd837dDF8e4c57F05d70Ab865de6e193BBB",
      "token_id": "15",
      "owner_of": "0x9c83ff0f1c8924da96cb2fcb7e093f78eb2e316b",
      "token_hash": "502cee781b0fb40ea02508b21d319ced",
      "block_number": "88256",
      "block_number_minted": "88256",
      "contract_type": "ERC721",
      "token_uri": "string",
      "metadata": "string",
      "last_token_uri_sync": "string",
      "last_metadata_sync": "string",
      "amount": "1",
      "name": "CryptoKitties",
      "symbol": "RARI"
    }
  ]
}
```


Figure 5.1: An example result of the endpoint "/nft/address"

5.3 Retrieving Transfers

The second endpoint that is used is "nft/address/token_id/transfers/," which gets the transfers of an NFT given a contract address and token ID. When a sale takes place on an NFTM, it triggers not only the transfer event but also the sale event. While this paper was written, Moralis only offered the endpoint to get transactions that triggered a transfer event. Thus, if there is a money transfer independently from the token transfer (i.e., in a different transaction), our current data set does not distinguish it. The transactions with a value of 0 are potential candidates for this. Another possibility is that, since some NFTMs allow payment with ERC20 tokens, the value of the transfer transaction may still be 0, while there are actually ERC20 tokens transferred to the seller. In this case, our algorithm still includes this transaction in the WCC graph, while this is actually not an unconditional asset transfer. These cases are potential false positive labelings of our approach due to insufficient data.

This endpoint has the input fields; chain, format, limit, order, cursor, address, and token_id. Only the contract address and token_id are obligatory fields. Besides these obligatory fields, the cursor field is used as each request is limited to 100 transfers, and

if a token has more than 100 transfers, a new request has to be sent using the cursor of the previous request. As output, many different attributes are returned, but for this thesis, only *from_address*, *to_address*, *value*, *transaction_hash*, and *block_timestamp* are used. The value field is shown in "Wei"(One ether = 1,000,000,000,000,000,000 wei), and block_timestamp has format "YYYY-MM-DD-HH-MM-SS". Using this endpoint, each token's transfer history is obtained separately.

5.4 Transaction Data Storage

Transaction data storage is not necessary but to improve the web application's performance, the NFT collections' transaction history is stored at MongoDB.

MongoDB is an open-source, cross-platform, NoSQL database. Data stored on MongoDB has JSON format. This data is named as a document. Each document can contain several fields, but it has a restriction that a document's size should be less than 16MB. This means that if the transaction history of an NFT collection is larger than 16MB, it can't be analyzed with our web app.


```
{
_id":"0x4becbdf97747413a18c5a2a53321d09198d3a100",
token_list":[
"token": "15"[
{"from_address": "0x057Ec652A4F150f7FF94f089A38008f49a0DF88e",
 "to_address": "0x057Ec652A4F150f7FF94f089A38008f49a0DF88e",
 "value": "1000000000000000",
 "transaction_hash": "0x057Ec652A4F150f7FF94f089A38008f49a0DF88e",
 "block_timestamp": "2021-06-04T16:00:15"
},
{"from_address": "0x057Ec652A4F150f7FF94f089A388688f49a0DF88e",
 "to_address": "0x057Ec652A4F150f7FF94f089A38008f49a0DFt86868",
 "value": "196696000000000000",
 "transaction_hash": "0x057Ec652A4F150f7FF949A38008f49a0DF88e",
 "block_timestamp": "2022-07-08T17:00:30"
}
```


Figure 5.2: An example of a collection stored on MongoDB

The transaction data fetched from Moralis is stored on MongoDB because to obtain the transaction data from Moralis, "the number of total tokens/100" requests have to be sent to get all of the tokens. After getting the list of tokens, one has to send "the

total number of transactions of the token/100" requests for each token to obtain their transaction data. That is why it is time-consuming to get the data from Moralis.

On MongoDB, first, the collection_id is stored. Then, the transaction array is stored under each token's name. This field is named "Token." The array of transactions contains the from_address, to_address, transaction_hash, value, and time. It is not necessary to store all the attributes fetched from Moralis. The mentioned attributes are enough to analyze wash trading activities. In Figure 5.2 an example of transaction data of a collection stored as a document on MongoDB is shown.

On the other hand, storing the data for better performance also has some disadvantages. Since the app won't fetch new data every time, it could be the case that new trades of a specific collection won't get analyzed. However, this shouldn't comprise a significant problem as the app can easily be extended to fetch the missing transactions (based on the last available transaction on the DB).

5.5 Creating Token Trade Graphs

To build the G_{NFT}=(V,E), the Python library "networkX," which enables the creation and study of graphs and networks, is being used. This library includes data structures for graphs and some basic graph algorithms. NetworkX provides the following types of graphs as Python classes: Graph, DiGraph, MultiGraph, and MultiDiGraph. The prefix "di-" stands for the directed graphs, and multi means that several edges between the same pair of nodes are allowed. The nodes of a graph should be hashable objects. Edges can contain different information like weight and labels. This library also contains several algorithms, like shortest path, breadth-first search, depth-first search, and strongly connected components.

5.6 Applying the Algorithm

Firstly, our algorithm checks if the given NFT collections (contract address) transaction history is already stored in our database or not. If it does not exist on the database, then the list of tokens is fetched from Moralis using the "/nft/address" endpoint. After obtaining the token list, the transaction history of each token is fetched using the "nft/address/token_id/transfers/" endpoint and stored on MongoDB.

Secondly, the trade graphs of every token on the given collection are built to identify the SCCs. "strongly_connected_components(G)" method of the library "networkX" is applied to the graphs iterativly, and the returned list of nodes for every SCC is stored in a single list. The SCCs, which are repeated more than 5 times, are marked as suspicious, and stored separately on another list.

Thirdly, transfer graphs have to be built to identify the WCCs. Transfer graphs contain the transactions with the value 0. That's why only the transactions of a collection with value 0 are fetched from the stored transaction data on MongoDB. With the fetched transactions transfer graphs are built. After building the transaction graphs, WCCs are identified iterativly with the "weakly_connected_components(G)" method of "networkX". This method returns sets of nodes for every WCC. All the results are stored together on a list, and the ones that are repeated more than 3 times are stored separately because they are suspicious.

After obtaining the suspicious SCCs/WCCs, the list of addresses that were part of the suspicious activities is stored on a set. Every address is stored only once on the set. The length of the set represents the number of suspicious addresses that were part of at least one suspicious activity. This length divided by the number of total addresses represents the ratio of suspicious accounts.

The last step is identifying suspicious transactions. A transaction is marked as suspicious if both to- and from-addresses of it are located in the same suspicious SCC or WCC. The value of every suspicious transaction is summed up to obtain the total value of the suspicious transactions and divided by the total trade volume to have the ratio.

Analysis results and token list of every NFT collection, analyzed with our web application, are stored on the web APP's database.

5.7 The Web Application

The web application is created with the Python library "Django" and the source code is stored on Github.[1] Django is an open-source web framework that eases the creation of websites with the help of the reusability and pluggability of components. Figure 5.3 displays the home page of the web application. On the left side is a search bar where the user can enter the collection_address of the collection he wants to analyze. On the right side is a dropdown to get the list of collections already stored in our database.

[1] https://github.com/baranimo7/Wash-Trading-Detection-Tool

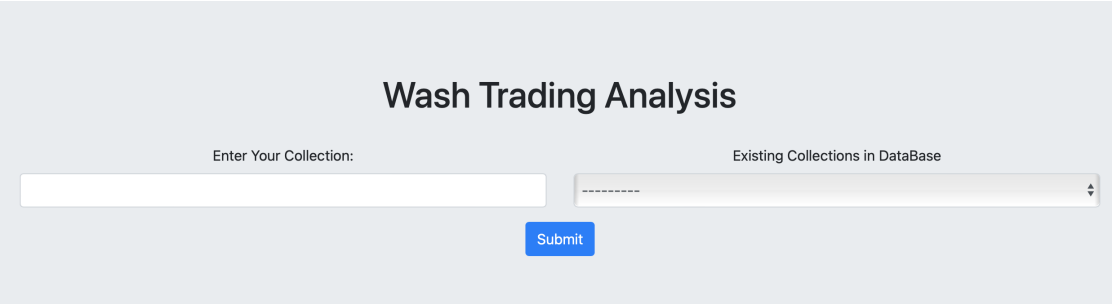


Figure 5.3: Homepage of the web APP

When the wanted collection_address is entered, the user has to wait until the transaction data is fetched. After the data is obtained, the web application displays the following information about an NFT collection: *name*, *collection address*, *number of total addresses*, *date of the last wash trading activity*, *number of addresses that were part of a wash trading activity*, *address ratio*, *total volume in ETH and USD*, *wash trading volume in ETH*, and *volume ratio*. On the right side, there exists a dropdown to get the list of tokens of the collection. An example results page can be seen in Figure 5.4.

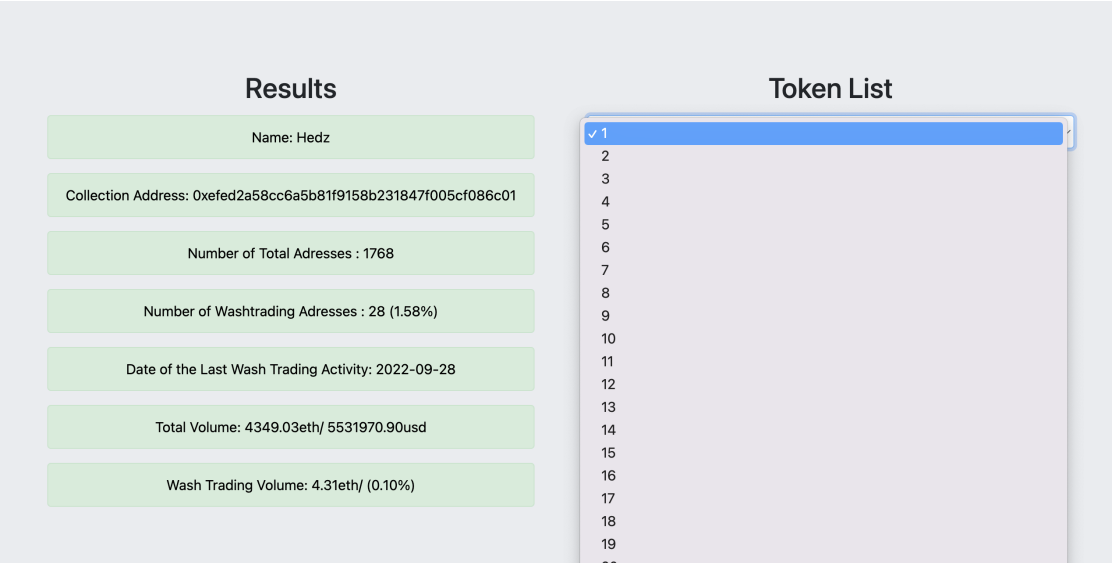


Figure 5.4: Results page of the web APP

After the collection is analyzed, a user can select a token from the dropdown to get its trade- and transaction history graphs. Transaction history of the selected token is obtained from MongoDB and its trade graph is created, and displayed on the screen.

To have a better view of the graph, it is created as an undirected graph, where the nodes represent the addresses and the labels of the edges show the total volume(in eth) and the number of transactions between the pair of nodes. We have chosen not to use a multigraph because there can be hundreds of edges between a pair of nodes, and the visualization can be very complicated. For a better view, a node is identified with the first 4 characters of the respective address. The edges that were identified as suspicious in the algorithm are marked red. The Python library "networkX" is used to visualize the graph. It has a method called draw(), which visualizes the graph. An example of a trade graph is shown in Figure 5.5.

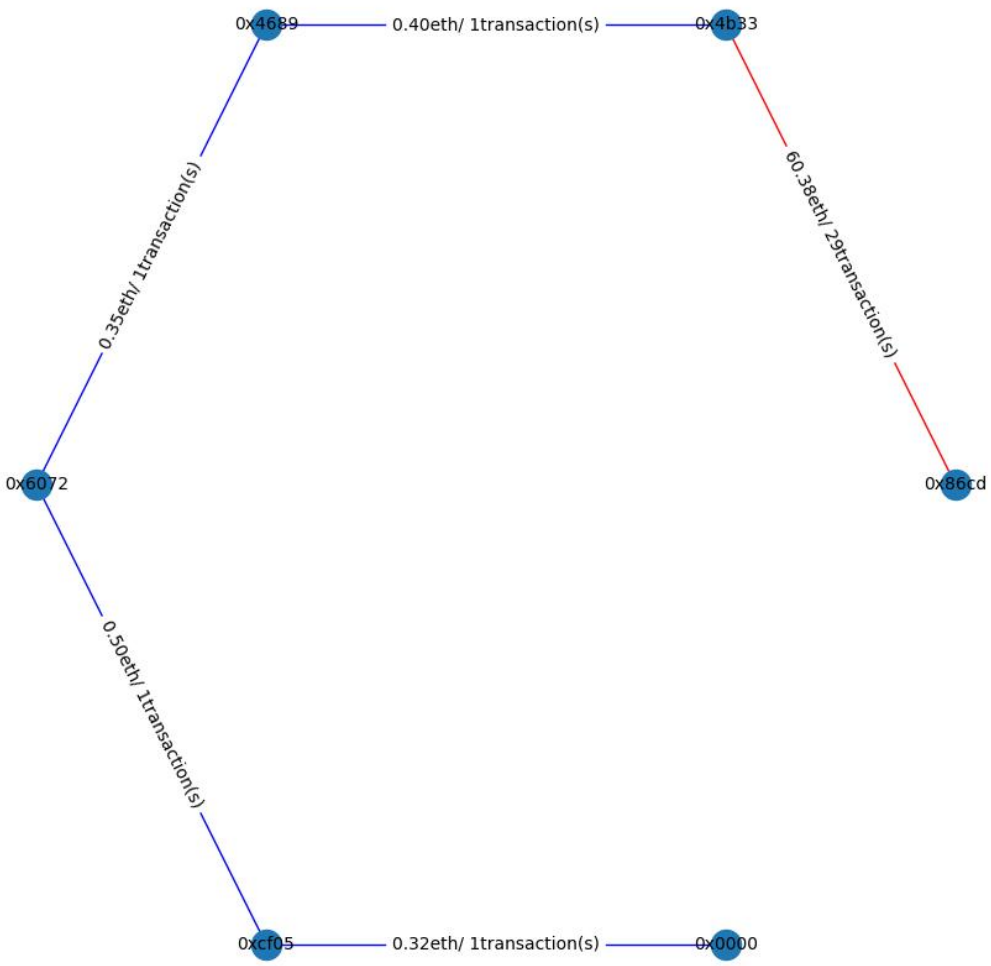


Figure 5.5: An example of a trade graph

The mentioned display of the trade graph gives valuable information about the addresses and transactions that were part of a possible wash trade. However, it is not enough to analyze the position of wash trading activities in the life span of the token. For this purpose, transaction history graph (a scatter plot) is also created. On the x-axis, the time of the transaction is represented in the format YYYY-MM. On the y-axis, the value of the transaction is shown in eth. Suspicious transactions are shown with blue stars, and non-suspicious transactions are represented with green dots. This graph

is created with the Python library "Matplotlib" and saved as a png. An example of a transaction history graph is shown in Figure 5.6

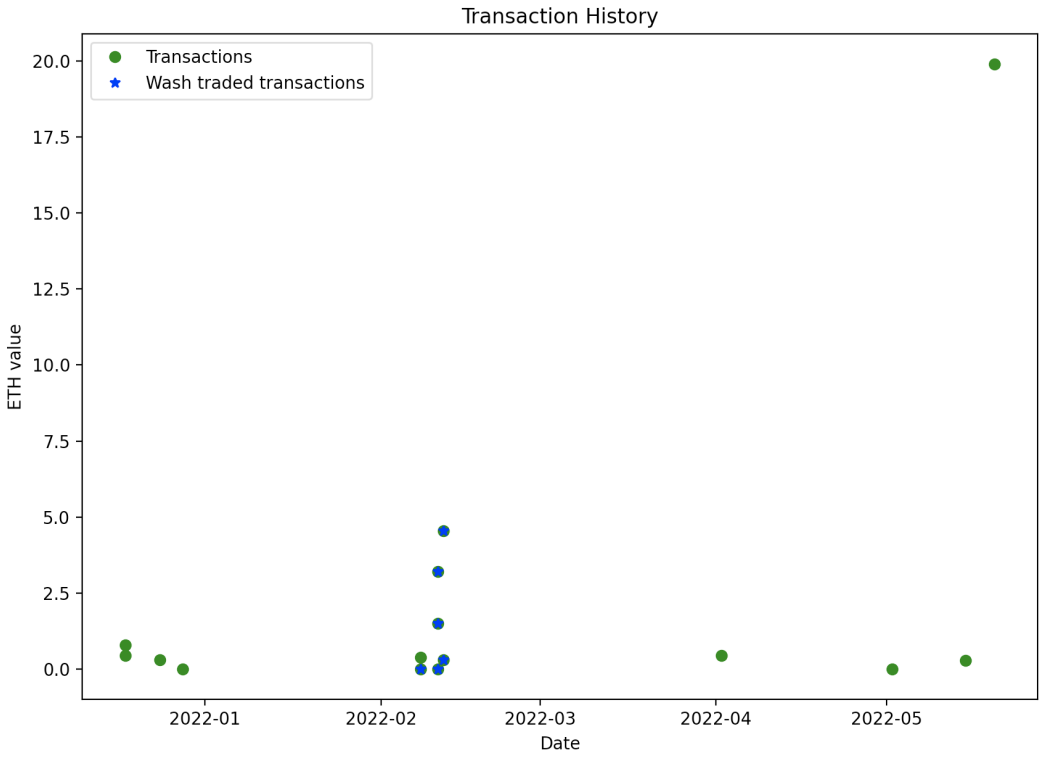


Figure 5.6: An example of a transaction history graph

6 Results

With the web application mentioned in the paper, we wanted to analyze some NFT collections in detail to have a better overview of some NFTM, compare our results with the current literature findings, and see if the wash trading activities are still a big problem.

6.1 Selection Criteria

The top 5 NFT collections by their whole trading volume on OpenSea, and LooksRare, and the top 5 NFTs by the trading volume since the last 30 days on Rarible are chosen to analyze with the selection criteria that the collection should be on the Ethereum blockchain and have a maximum of 20k items. The mentioned NFTMs are selected because of several reasons. Firstly, OpenSea is the most popular NFTM, and it has the most active users. Secondly, existing literature highlights that there were many wash trading activities on the NFTMs Rarible and LooksRare. With our analysis, it can be identified if the problem goes on or if the marketplaces have prevented these activities. Thirdly, the last 30 days are considered for Rarible because the whole time largest NFT collections on the platform are the same as the ones on Opensea. Our aim was to analyze different collections, so this criterion is applied. The reason behind the limitation of 20k items is that only collections smaller than 16MB can be stored on MongoDB.

The selected top 5 collections with their total volume, number of items, and number of owners from Looksrare are shown in Table 6.1 and the ones from OpenSea in Table 6.2. Trading volume since the last 30 days, total trading volume, and the number of items/owners of the selected collections from Rarible are shown in Table 6.3.

6.2 Volume and Address Ratios

When analyzing the wash trading activities in an NFT collection, it is crucial not only to detect the number of suspicious addresses and the amount generated from the wash trading activities, but it is also important to analyze the ratio of these mentioned calculations, because NFT collections have divergent sizes.

Table 6.1: Top 5 LooksRare collections

	NFT Collection	Total Volume	Items	Owners
1	Terraforms	4282k eth	9,9k	2k
2	Meebits	3114k eth	20k	6,5k
3	dotdotdot	961k eth	4,8k	2,6k
4	Loot	230k eth	7,8k	2,5k
5	Bored Ape Yacht Club	95k eth	10k	6,5k

Table 6.2: Top 5 OpenSea collections

	NFT Collection	Total Volume	Items	Owners
1	CryptoPunks	983k eth	10k	3,6k
2	Mutant Ape Yacht Club	447,7k eth	19k	13k
3	Azuki	260,1k eth	10k	5,1k
4	CLONE X - X TAKASHI MURAKAMI	228,8k eth	19,4k	9,5k
5	Moonbirds	169,7 eth	10k	6,6k

Table 6.3: Top 5 Rarible collections

	NFT Collection	Last 30 Day Volume	Total Volume	Items	Owners
1	RareApepeYachtClub	6,7k eth	9,3k eth	10k	3k
2	DigiDaigaku Genesis	6,5k eth	8,3k eth	2k	760
3	8liens	5,4k eth	7,4k eth	10K	2.7K
4	PudgyPenguins	5,2k eth	64,1k eth	8.9K	4.5K
5	Women Ape Yacht Club	5,0k eth	5,3k eth	8,8k	2,1k

Table 6.4: Top 15 collections

	NFTM	NFT Collection	Suspicious Trade Ratio	Suspicious Address Ratio
1	LooksRare	Terraforms	88%	9.5%
2	LooksRare	dotdotdot	78.4%	3.24%
3	LooksRare	Meebits	76.3%	6.33%
4	LooksRare	Bored Ape Yacht Club	28.74%	2.65%
5	LooksRare	Loot	22.71%	8.95%
6	Rarible	8liens	21.6%	1.65%
7	Rarible	RareApepeYachtClub	15.49%	6%
8	OpenSea	Azuki	12.57%	6.25%
9	Opensea	CryptoPunks	7.57%	5.5%
10	OpenSea	CLONE X	6.32%	3.64%
11	Rarible	PudgyPenguins	6.2%	3.91%
12	OpenSea	Mutant Ape Yacht Club	5.05%	1.72%
13	OpenSea	Moonbirds	0.86%	3.61%
14	Rarible	Women Ape Yacht Club	0.65%	5.38%
15	Rarible	DigiDaigaku Genesis	0.56%	2.2%

In Table 6.4 the suspicious trade ratios and the ratio of the suspicious addresses are shown. These ratios give the first insights into the wash trading activities. According to our findings, we observe most of the wash trading activities in the collections from LooksRare. Terraforms has not only the highest suspicious trade ratio with 88%, but it also has the highest suspicious address ratio with 9.5%. DigiDaigaku Genesis has the lowest suspicious trade ratio with 0.56%. On the other hand, 8liens has the least suspicious address ratio, with 1.65%.

6.3 Number of Addresses in an SCC/WCC

A wash trade can always occur between the same pair of addresses, but this can be easily identified when looking at the transaction history of the token. Many NFTMs show the trade history of the tokens, so scammers can use more complex and bigger trade patterns to conduct wash trading. One person can open several accounts on an NFTM and control these accounts to trade between them, so having larger SCCs/WCCs is not a problem for the scammers. In Figure 6.1, the average number of addresses in an SCC/WCC of each collection is shown. On the X-axis, the NFT collections are displayed in the same order as in Table 6.4. On the Y-axis, the average numbers are

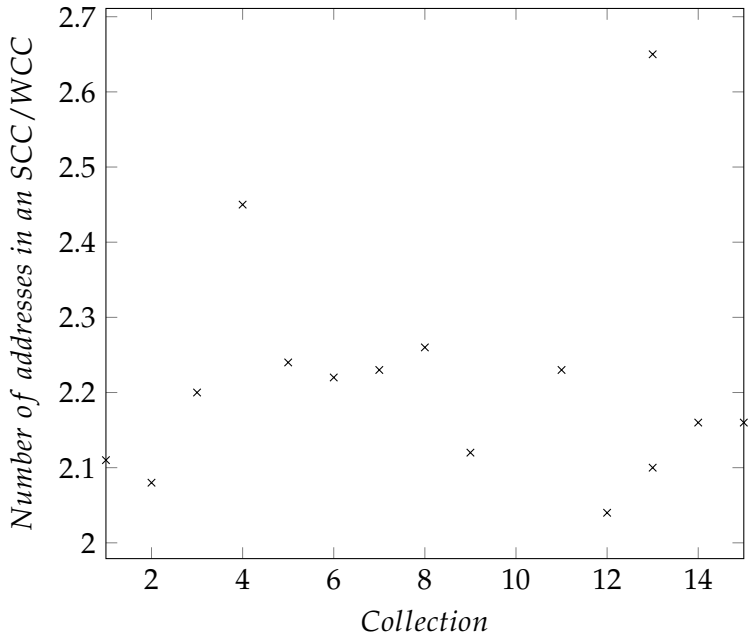


Figure 6.1: Number of addresses in an SCC/WCC of the top collections

shown. Looking at the results, it can be seen that the average number of addresses in an SCC/WCC is between 2.04 and 2.65. This shows that wash trading mainly occurs between the same pair of addresses. Most scammers don't try to open several accounts and trade between them, but they only control the minimum amount of addresses, which saves time and effort.

6.4 Frequency of Suspicious SCCs/WCCs

SCCs and WCCs, which are repeated frequently, make up the most critical part of the wash trading detection algorithm used in this paper and in the web app. The more an SCC/WCC is repeated, the more it is possible that these suspicious transactions are part of an organized wash trading activity. Therefore, analyzing the frequency of SCCs/WCCs of collections is vital when searching for wash trading.

In Figure 6.2 average frequency of the suspicious SCCs/WCCs of each NFT collection is displayed. On the X-axis, the NFT collections are displayed in the same order as in Table 6.4. On the Y-axis, the average frequencies are shown. Again the NFT collection Terraforms has the highest result with an intermediate frequency of 21.53 because the higher the trading volume is, the more suspicious trades occur. Our results range

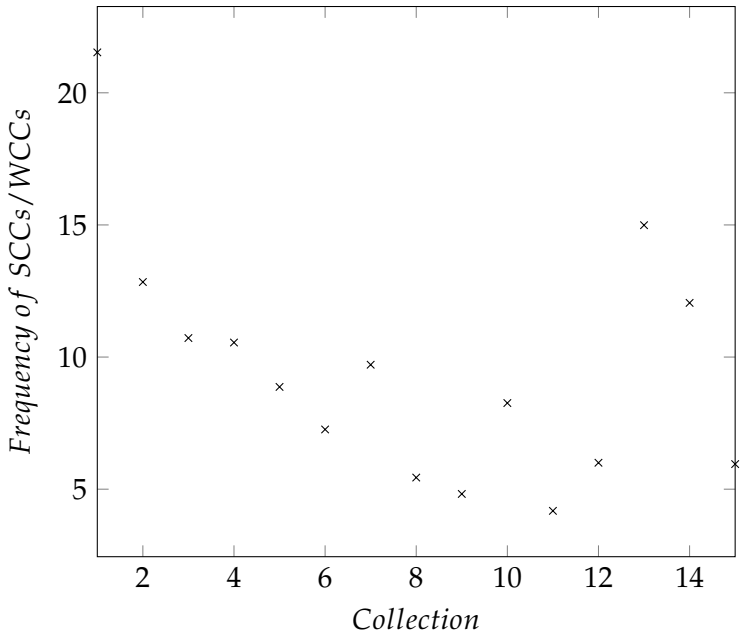


Figure 6.2: Frequency of suspicious SCCs/WCCs of the top collections

between 4.18 and 21.53. This range is extensive compared to the results of the average number of addresses in SCCs/WCCs.

6.5 Frequency of Addresses

The detected SCCs/WCCs that are repeated frequently are helping to identify the suspected addresses. An address can be a part of one frequent cycle or different ones.

Figure 6.3 shows the average of how many different SCCs/WCCs a single address is located. On the X-axis, the NFT collections are displayed in the same order as in Table 6.4. On the Y-axis, the average frequencies of addresses are shown. The results are between 1.08 and 1.47, which highlights that a single address is mostly part of a single SCC/WCC. Some addresses can be part of subcycles, which also increases the average frequency. Our findings show no positive relationship between the ratio of suspicious trades and the frequency of addresses.

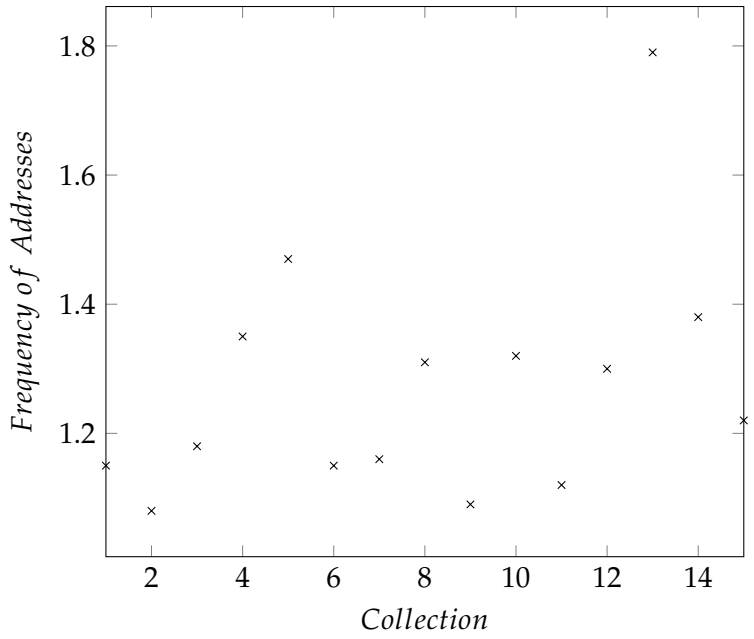


Figure 6.3: Frequency of addresses of the top collections

6.6 Suspicious Token Ratio

For traders, it is not only essential to know if the interested token is part of wash trading activities, but the whole collection, in general, plays an important role because the repetition of the collection influences the value of the tokens in that collection. Secondly, if a collection contains some wash trading activities, it doesn't mean that each token on that collection contains such trades. One scammer can buy a token from the collection and trade it between his own accounts to create an illusion of demand, while all other tokens are traded without wash trading activities. In this situation, the wash trading algorithm will detect wash trading activities in the collection without highlighting if a single token contains suspicious activities or most of the tokens, so analyzing the ratio of tokens that were part of a wash trade is needed.

In Figure 6.4, the ratio of tokens that contain at least one suspicious trade is shown. On the X-axis, the NFT collections are displayed in the same order as in Table 6.4. On the Y-axis, the token ratios are shown. The transaction history of many tokens only consists of the mining transaction. That's why the ratio of suspicious tokens is not high. For example, 35.60% of the tokens on the most suspicious NFT collection Terraforms are marked as suspicious. This shows that not every token in the collection was part of the wash trading activities, whereas other tokens in the collection contain frequent

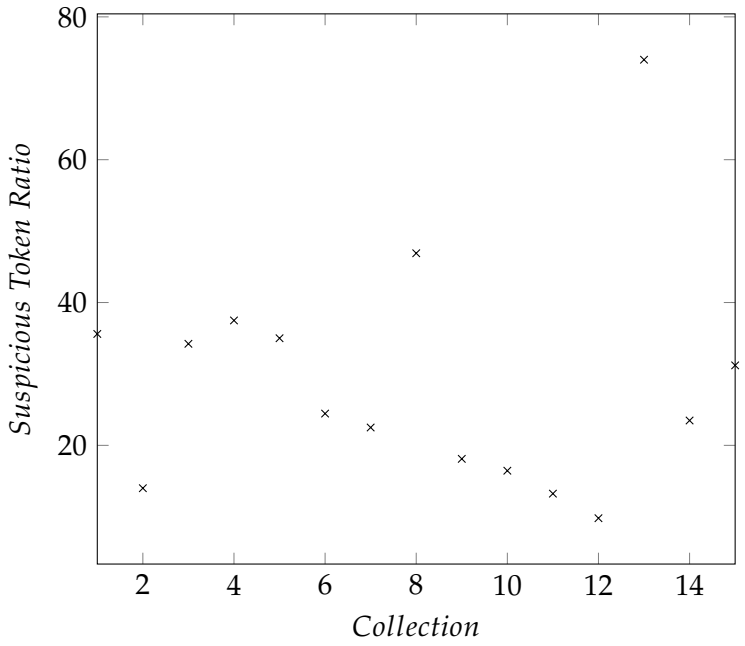


Figure 6.4: Suspicious token ratios of the top collections

suspicious activities.

6.7 Comparison With Existing Research

Our results indicate that the NFTM "Rarible" has prevented most of the wash trading activities because the top 5 largest collections on the marketplace by their trading volume since the past 30 days have relatively low suspicious trade volume.

On the other hand, our findings show that the "community first" marketplace LooksRare, created as a decentralized marketplace to compete with OpenSea, failed to prevent wash trading activities on the platform. Different reports point out that the collections, especially the “no-royalty” collections on this marketplace, contain increased wash trading activities. Meebits and Terraforms are one of those royalty-free collections [Hu22]. CryptoSlam has analyzed the transaction history of Meebits for the week of 1/11/22 - 1/19/22 and found out that almost all of the volume was created by wash trading activities. Our results, generated with our detection algorithm, also support these findings about the no-royalty collections. Meebits and Terraforms have the most suspicious trade volume ratios among the 15 NFT collections that we analyze. Similar to CryptoSlam’s results, our findings show that wash trading activities form more than

75% of Meebits' total trading volume [Cof22].

In conclusion, the results generated with our detection algorithm are very similar to the results from existing literature and reports. This fact shows that our wash trading detection algorithm can detect suspicious activities correctly in an efficient way and provide the first insights about the fake volume.

6.8 Performance

The performance of the web application is tested with a MacBook Pro, which has a 16Gb ram and uses Intel Core i5. For testing purposes, 3 collections with different sizes are used.

The first collection is a small-sized collection containing 1000 tokens and has 3187 transactions in total. Fetching the data from Moralis takes 477s, and analyzing the transaction data takes 5s. Trade- and transaction history graphs of a token in this collection can be created on average in 2.5s.

The second collection is medium-sized, which has 4890 tokens and 19431 transfers. Getting the transaction history takes 2339s, and analyzing the data set takes 15s. Creating a token's trade- and transaction history graphs in this collection takes an average of 6.5s.

The last collection is large-sized and contains 9000 tokens, and has 30633 transfers. It takes 4100s to get the transfer data from Moralis and 28s to analyze the data. Creating the trade and transaction history graphs of a token in this collection takes, on average, 11s.

In conclusion, our web app can fetch the transaction history of the given small-sized NFT collection and analyze it in less than 10 minutes, but if the collection contains around 10k items, the process takes more than an hour, which is not a reasonable time. Fetching the data from Moralis takes up to 99% of this time. To reduce waiting time and increase performance, our web app stores the transaction history of the given collection on MongoDB and the analysis results in its database. If a user wants to get the analysis of a collection, which is already analyzed and stored, it takes less than two seconds to get the analysis regardless of the collection size. To improve the web app's performance, one can also run a full node and index the data instead of using an API like Moralis. Using this method and then querying this data would be much faster.

7 Conclusion

Taking everything into consideration, as wash trading activities in the NFT ecosystem become more and more problematic for the traders, it is crucial to identify these activities before investing in a token. This paper presents a wash trading detection algorithm and steps to create a web app that can detect wash trading activities. At the end, it analyzes the top NFT collections. After the analysis and process of creating the web app, the five research questions have been answered.

1. **RQ1: Which research has been conducted so far on detecting wash trading on NFT marketplaces?**

   There are only a few papers about the wash trading activities in the NFT ecosystem. They mostly focus on using the same idea of detecting closed cycles or strongly connected components. Additionally, some websites that analyze the NFTs also published reports on specific issues like crypto crime and the wash trading problem on LooksRare.

2. **RQ2: What insights can an online service provide regarding an NFT collection?**

   The web application, developed by us, can detect suspicious addresses and trades, which can be a part of a wash trade, and calculate the amount and ratio of these trades. Similar calculations can be done for each of the tokens in the given collection. Additionally, the trade graphs and the timeline of the transactions of both regular and suspicious trades can be displayed.

3. **RQ3: What are the most common wash trading patterns?**

   Most of the suspicious trades occur between the same addresses, and an address is mostly part of a single SCC/WCC. If the suspicious trade ratio of a collection is high, then the SCCs/WCCs occur more frequently in that collection.

4. **RQ4: Can wash trading be detected by an online service efficiently?**

   Our web application can detect wash trading activities for the given NFT collection in less than 10 minutes, if the given collection is a small-sized collection (contains no more than 1000 tokens) or if the given collection is already stored in our

database. For the large-sized collections with more than 10k tokens, it takes more than an hour to obtain the transaction data from Moralis and this is not a user-friendly waiting time.

5. **RQ5: Can wash trading activity be avoided/regulated by marketplaces?**

   Some NFTMs are threatening to ban the users that conduct wash trading from their marketplace but are non-transparent about the number of users who are banned from the marketplace. Some marketplaces are rewarding their users that buy and sell NFTs on their platform with their own tokens, and this encourages users to conduct wash trading.

As a part of future work, the algorithm can be extended by also detecting the path like transactions, which happen frequently but not forming a cycle. Additionally, the relation of the suspicious pair of trade partners can be analyzed in different collections to confirm they are organized wash traders. Furthermore, a data source can be developed to distinguish between sales and transfer events.

List of Figures

List of Tables

List of Algorithms

Bibliography

[22] *What Can I Do with Ethereum? The Use Cases*. 2022. URL: https://www.finsmes.com/2022/01/what-can-i-do-with-ethereum-the-use-cases.html.

[Bha22] A. Bhattacharya. *Why is 'NFT' the Collins Dictionary's word of the year?* 2022. URL: https://www.weforum.org/agenda/2021/12/non-fungible-token-collins-dictionary-word-of-the-year-2021/.

[Cof22] C. Coffman. *Wash Trading: Who, What, Why, and What Should We Do About It?* 2022. URL: https://blog.cryptoslam.io/wash-trading-who-what-why-and-what-should-we-do-about-it/.

[Coh22] A. Cohen. *OpenSea Self-Reports That More Than 80% of Its NFTs Minted for Free Were Unoriginal or Fake*. 2022. URL: https://www.sporttechie.com/opensea-self-reports-that-80-of-its-nfts-were-unoriginal-or-illegtimate.

[Con22] Contributor. *What is LooksRare: The Upstart Giving OpenSea A Run for Its Money*. 2022. URL: https://phemex.com/academy/what-is-looksrare-looks-coins.

[Das+22] D. Das, P. Bose, N. Ruaro, C. Kruegel, and G. Vigna. "Understanding Security Issues in the NFT Ecosystem." In: (2022).

[Han22] S. Handagama. *'Jack Dorsey's First Tweet' NFT Went on Sale for $48M. It Ended With a Top Bid of Just $280*. 2022. URL: https://www.coindesk.com/business/2022/04/13/jack-dorseys-first-tweet-nft-went-on-sale-for-48m-it-ended-with-a-top-bid-of-just-280/.

[hil22] hildobby. *LooksRare VS OpenSe*. 2022. URL: https://dune.com/hildobby/LooksRare-VS-OpenSea.

[Hoo20] R. Hoogendoorn. *Rarible Hampered by Wash Trading*. 2020. URL: https://www.playtoearn.online/2020/07/28/rarible-hampered-by-wash-trading/.

[Hu22] E. Hu. *Clever NFT traders exploit crypto's unregulated landscape by wash trading on LooksRare*. 2022. URL: https://cointelegraph.com/news/clever-nft-traders-exploit-crypto-s-unregulated-landscape-by-wash-trading-on-looksrare.

[Jav] Javatpoint. *History of Blockchain*. URL: https://www.javatpoint.com/history-of-blockchain.

[Jen22] G. Jenkinson. *NFT market worth $231B by 2030? Report projects big growth for sector*. 2022. URL: https://cointelegraph.com/news/nft-market-worth-231b-by-2030-report-projects-big-growth-for-sector.

[Law22] G. Lawton. *Top 9 blockchain platforms to consider in 2022*. 2022. URL: https://www.techtarget.com/searchcio/feature/Top-9-blockchain-platforms-to-consider.

[TEA22] C. TEAM. "Crime and NFTs: Chainalysis Detects Significant Wash Trading and Some NFT Money Laundering In this Emerging Asset Class." In: (2022).

[VW21] F. Victor and A. M. Weintraud. "Detecting and Quantifying Wash Trading on Decentralized Cryptocurrency Exchanges." In: (2021).

[Wac+21] V. von Wachter, J. R. Jensen, F. Regner, and O. Ross. "Quantifying suspicious behaviour in NFT markets." In: (2021).

[Wac22] P. Wackerow. *ERC-721 NON-FUNGIBLE TOKEN STANDARD*. 2022. URL: https://ethereum.org/en/developers/docs/standards/tokens/erc-721/.

[Wis22] J. Wise. *OPENSEA STATISTICS 2022: USERS, REVENUE MARKET SIZE*. 2022. URL: https://earthweb.com/opensea-statistics/.