Miracle or Mirage? A Measurement Study of NFT Rug Pulls

JINTAO HUANG, Huazhong University of Science and Technology, China
NINGYU HE, Peking University, China
KAI MA, Huazhong University of Science and Technology, China
JIANG XIAO, Huazhong University of Science and Technology, China
HAOYU WANG*, Huazhong University of Science and Technology, China

NFT rug pull is one of the most prominent type of NFT scam, whose definition is that the developers of an NFT project abandon it and run away with investors' funds. Although they have drawn attention from our community, to the best of our knowledge, the NFT rug pulls have not been systematically measured. To fill the void, this paper presents the first in-depth measurement study of NFT rug pulls. Specifically, we first compile a list of 253 known NFT rug pulls as our initial confirmed rug pulls (i.e., ground truth), based on which we perform a pilot study, highlighting the key symptoms of NFT rug pulls. Then, we design an effective rule-based detector to measure the prevalence of NFT rug pulls in the ecosystem. We have labelled 7,487 happened NFT rug pull projects which were not revealed by our community. To eliminate the potential damage brought by rug pull scams, we take a step further towards designing a real-time prediction model to proactively identify the potential rug pull projects in an early stage ahead of the scam happens. We have implemented a prototype system, and deployed it in the real-world setting for over 6 months. Our system has raised alarms for 5,557 new NFT projects by the time of this writing, which works as a whistle blower that pinpoints rug pull scams timely, thus mitigating the impacts.

CCS Concepts: • **Security and privacy → Intrusion/anomaly detection and malware mitigation**.

Additional Key Words and Phrases: non-fungible token; NFT; rug pull; blockchain

ACM Reference Format:
Jintao Huang, Ningyu He, Kai Ma, Jiang Xiao, and Haoyu Wang. 2023. Miracle or Mirage? A Measurement Study of NFT Rug Pulls. *Proc. ACM Meas. Anal. Comput. Syst.* 7, 3, Article 51 (December 2023), 25 pages. https://doi.org/10.1145/3626782

1 INTRODUCTION

NFTs, or non-fungible tokens, were introduced as a new type of cryptocurrency in October 2015 with the launch of the first NFT project `Etheria`. Since then, NFTs have attracted significant attention from the public and opened up new possibilities. From 2020 to 2021, the NFT market has experienced explosive growth, totally an increase of $17 billion, which is a staggering 210× compared to 2020's market value of $82 million [39]. Several NFT collections, e.g., `Mutant Ape Yacht Club`, `Azuki`, and `Bored Aple Yacht Club`, have generated significant sales volumes in NFT markets, with values reaching up to millions of USD [41]. One of the key features of NFTs is

*Haoyu Wang (haoyuwang@hust.edu.cn) is the corresponding author.

Authors' addresses: Jintao Huang, Huazhong University of Science and Technology, China; Ningyu He, Peking University, China; Kai Ma, Huazhong University of Science and Technology, China; Jiang Xiao, Huazhong University of Science and Technology, China; Haoyu Wang, Huazhong University of Science and Technology, China.

2476-1249/2023/12-ART51 $15.00
https://doi.org/10.1145/3626782

their *non-fungible* nature, which enables them to be linked to a specific digital asset, such as images, art, music, and sports highlights. This association may confer licensing rights to use the asset for a specified purpose. To enable the implementation of NFTs, Ethereum proposes two standards: `ERC-721` and `ERC-1155`. The former generates tokens that are *one-of-a-kind* and linked to a unique `token ID`, while the latter generates a set of tokens that may share the same `ID`.

Despite their popularity, NFTs, like other types of cryptocurrencies, are prone to security threats. Attackers and scammers are staring at the huge market, which are frequently reported in the media outlets [8, 38]. Among them, *rug pull* [48] is one of the most prominent type of scam, i.e., developers of an NFT project abandon it and run away with investors' funds. According to the research of `Chainanalysis` [24], cryptocurrency investors in 2021 lost over $2.8 billion to rug pulls, and NFT rug pull is on the rise. For example, the `Frosties` is a famous NFT rug pull scam, which led to the theft of over $1.3 million, after which the two founders were charged for the maximum sentence of 20 years in prison [7].

NFT rug pulls have already drawn attention from our community. For example, there are several crowd-sourcing based channels maintaining a list NFT rug pulls [23, 47]. Despite this, to the best of our knowledge, the NFT rug pull scams have not been systematically investigated or measured. Thus, there is a general *lack of an understanding* of NFT rug pulls, including 1) the concrete patterns of NFT rug pulls, 2) to what extent NFT rug pulls exist in the ecosystem, and most importantly, 3) no existing approaches can be used to mitigate or even prevent this kind of scam.

This work. In this paper, we present the first systematic study of NFT rug pulls. To understand the concrete symptoms of NFT rug pulls, we first harvest a list of 253 `confirmed rug pulls` (i.e., *initial ground truth*) revealed by our community, and based on which to perform a pilot study, highlighting the key features related to NFT rug pulls (§4). Then, we design an effective detector to measure the prevalence of NFT rug pulls in the ecosystem, by analyzing all the on-chain and off-chain data related to over 173K Ethereum NFT projects (§5). Our detector has identified 7,487 rug pulls that have taken place in the ecosystem, 30× greater than existing crowd-sourcing based collections. Next, to explore whether we can raise alarms of NFT rug pull projects before the scam happens, we have created a prediction model taking advantage of 52 kinds of features extracted from both on-chain and off-chain data (§6). Our model can raise alarms for 90% of NFT scam events within 96 hours ahead of rug pull happens. We have also implemented a prototype system to monitor NFT rug pulls in the wild since November 2022. By the time of this writing, we have successfully raised warnings for 5,557 `suspicious NFT projects`, and most of them have been confirmed to be rug pulls in later times with additional evidences.

In summary, this paper makes the following main research contributions:

- We have uncovered the key symptoms of NFT rug pulls and devised an effective approach to pinpoint NFT rug pulls that have already taken place in the wild. We have flagged 7,487 already rug pulled NFT projects in total using a strict rule-based approach, *by far the largest NFT rug pull dataset*, 30× greater than existing efforts. This can be served as the reliable benchmark for both our predictor and the future research in the community.
- We have proposed a real-time solution to mitigate the impact of NFT rug pulls. Specifically, we design a prediction model to raise alarms ahead of the scam happens, based on the signals (e.g., token transfers and trade events) released in their early stages. It can work as a *whistle blower* that pinpoints rug pull scams timely, thus mitigating the impacts.
- We have deployed our NFT rug pull warning system on Ethereum to monitor NFT transactions since November 2022. By the time of this writing, we have successfully raise alarms for 5,557 suspicious NFT projects. We further make effort to show that most of them are indeed rug pulls with additional evidences, while we can raise warnings several days earlier.

2 BACKGROUND

2.1 Ethereum & Non-Fungible Token (NFT)

Along with the prosperity of Bitcoin [2], Ethereum [35] occupies the second leading position among blockchain platforms, on which developers can deploy *smart contracts* that can be automatically executed if the condition is met. Smart contracts can interact with each other, while the interactions are recorded on-chain and is accessible for anyone, which is called *transactions*. Furthermore, transactions can not only send data across smart contracts, but also transfer tokens. For example, the official token in Ethereum is `Ether` (i.e., ETH), which is a type of cryptocurrency that can be circulated in exchanges. Within certain transactions, *events* will be emitted and recorded on-chain to provide additional details about the corresponding transaction, e.g., initiator and carrying data.

Except for official tokens, accounts (i.e., users) in Ethereum can issue tokens following the required standards. An "Ethereum Request for Comments" (ERC) [34] is a sequence of standard protocols that are used in conveying technical notes and requirements within a group of developers/users, frequently in order to issue tokens under the Ethereum blockchain. Specifically, ERC-20 [31] is the most widely adopted standard. It has six interfaces that should be implemented by developers, e.g., `transfer`, which allows the initiator to transfer a certain amount of current ERC-20 token to someone else. Note that, Ethereum only examines whether interfaces are implemented instead of the correctness of implementations. `ERC-721` [32] and `ERC-1155` [30] are two emerging standards for NFTs. Tokens under these two standards have a specific characteristic, named *non-fungibility*. Specifically, tokens under `ERC-721` are one-of-a-kind, i.e., each token corresponds to a unique ID. Tokens under `ERC-1155` also possess this characteristic, while an ID can be linked to several indistinguishable tokens. To this end, the distinguishability among *non-fungible tokens* (NFTs) can be used to bind real-world unreplicable items, such as an artwork or a piece of music [3], in one-to-one relationships. NFTs are becoming popular especially in the art world as a means for artists to monetize and sell their digital creations [4]. There are many tools that can assist in binding artworks with newly minted NFTs, e.g., `OpenSea` [11] and `Rarible` [43]. Note that, NFTs can also be traded in secondary markets [11] (i.e., where investors buy and sell securities from other investors), thus the value of an NFT is highly determined by its supply and demand relationship, or its popularity in the secondary market. After a successful trading, the ownership of the NFT will be transferred from the seller to the buyer, through the smart contract of the market. The lowest price of trade within recent period is defined as the *floor price*, which is a vital indicator to measure the value of an NFT project.

2.2 NFT Rug Pull

NFT rug pull is a type of scam that developers of an NFT project hype a project to attract investor money, and then suddenly shut down the project or disappear, taking investors' funds with them. The name comes from the idiom "to pull the rug out" from under someone, leaving the victim off-balance and scrambling [57, 58]. A successful NFT rug pull contains the following three stages.

(1) **Luring victims.** At first, scammers will lure victims by promoting their projects on social media platforms, e.g., `Twitter` and `Discord`. Meanwhile, they use kinds of tricks (e.g., posting fake comments using bots) to make an illusion that the NFTs they promoted will be in great demand, and investing in them would reap high rewards.
(2) **Pumping up the price and making a profit.** To pursue maximum profits, these rug pullers will try their best to make their projects *look* valuable and attractive, e.g., by fabricating trading volumes. After the price has been driven up, they will sell their owned NFTs, and pull as much value out of them as possible.

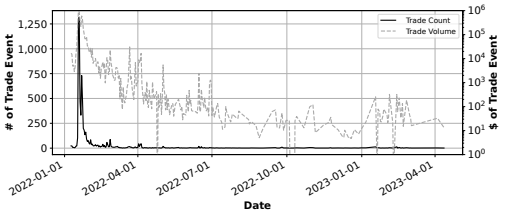


Fig. 1. The activity of AniMoon on OpenSea.

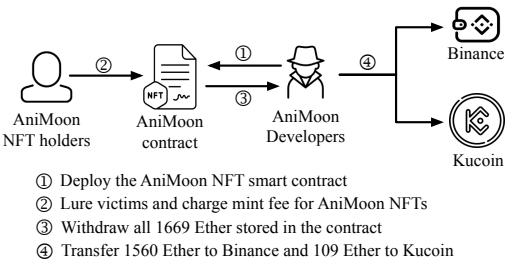


Fig. 2. The whole process of AniMoon rug pull.

(3) **Running away.** Finally, the rug pullers will run away with investors' fund, abandon the project, and usually deactivate their social media. Such a project with no endorsements will not attract any investors or collectors, leading to an extreme low price on secondary markets. All the NFT holders under this project are victims that suffer financial losses.

2.3 Motivating Example

`AniMoon` is a play-to-earn (P2E) game based on the famous `Nintendo Pokémon` cartoon series. Based on the *definition* of NFT rug pull, we label it as such since the development team abruptly abandons the project, taking investors' funds with them, and the project's social media accounts exhibit abnormal behavior following their disappearance. Fig. 1 shows its transaction numbers and volumes in the largest NFT secondary market, `OpenSea`, and Fig. 2 illustrates the whole process of how the rug pull is carried. We will briefly introduce its three stages following §2.2.

2.3.1 Luring Victims. The team attracted unsuspecting players by advertising on various social media platforms. For example, they launched an official website [18] and created a `Discord` server [16]. Two founders also advertised the project on their own `Twitter` accounts [15, 17], with 94.4K and 1.4K followers, respectively. The massive user base has successfully attracted more than 9K users (i.e., victims) to invest in the project. According to our statistics, more than 90K users are invited into their `Discord` server. As incentives, the owner promised physical rewards and cash dividends. Such a seemingly low-risk but high-returns project finally lures a large number of victims.

2.3.2 Pumping up the price and making a profit. As illustrated in Fig. 1, the invested money takes a rocket. The smart contract was deployed at Jan 7th, 2022. Only 11 days later, the trading volume and the number of transactions have reached up to $944K and $1.2K, respectively. To gain a profit from players, `AniMoon` requires a certain amount of Ethers when minting an NFT. From Jan 8th to Jan 18th, 9,999 NFTs were minted, and the `AniMoon` team has earned 1,669 Ether, worth around $6.3M at that time. In addition, once an NFT is traded on `OpenSea`, the initiators of the project will receive a dividend, named *creator earnings* [40]. According to our collected transactions, the `AniMoon` team received nearly $470K as the *creator earnings* from the trading of its NFTs on `OpenSea`.

2.3.3 Running away and shutting down the project. The rug pull occurs on Jan 18th, 2022, the 11th day since project launching. Over $6.3M worth of Ether was transferred out from the NFT smart contract. All mint fees were withdrawn from the smart contract to developer controlled addresses. Interestingly, these addresses are hard-coded in the smart contract's `withdraw` function, which suggests that it is a premeditated scam. Then, 1,560 Ether was transferred to `Binance` exchange and 109 Ether was transferred to `KuCoin` exchange. In a nutshell, only after 20 days of launching, the volume of `AniMoon` has dramatically went down to $66K, only around 7% of its peak. At May 15th,

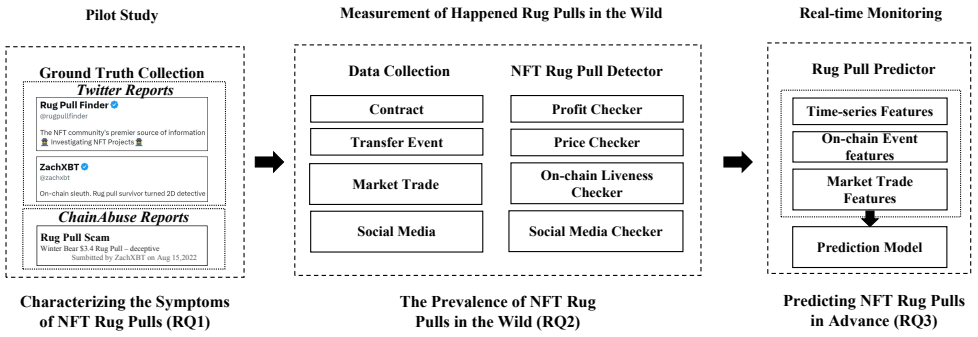


Fig. 3. An Overview of Our Study.

2022, its `Twitter` account has been suspended, the official website was no longer being maintained, and the `Discord` server was also removed. At last, all `AniMoon` NFTs are worthless.

3 STUDY DESIGN

In this work, we perform a series of *progressive* studies to gain a deep understanding NFT rug pulls. Specifically, we aim to answer the following research questions (RQs):

- RQ1 *What are the characteristics of NFT rug pulls? Can we summarize concrete patterns to depict them?* Although NFT rug pulls are emerging in the ecosystem, no existing efforts have specified the concrete rules to distinguish whether an NFT project is a rug pull scam or not. Formulating the precise symptoms of confirmed rug pulls is the key to identify them.
- RQ2 *Can we identify rug pulls in an automatic way? how many NFT rug pulls have happened in the wild?* We seek to design a strict rule-based detector to find the most reliable NFT `rug pulls` as the expansion of our initial ground truth.
- RQ3 *Can we raise early warning of NFT rug pulls in advance?* Our previous two RQs rely on reactive methods to identify `confirmed rug pulls` and flag `happened rug pulls` after the scam has performed. However, an NFT rug pull usually involves several key steps in the whole process. Proactively identifying the `predicted rug pulls` in an early stage can help eliminate the potential risks they exposed.

Fig. 3 shows the overall process of our study. To understand the concrete symptoms of NFT rug pulls, we first harvest the known NFT rug pulls revealed by our community, resulting a dataset of 253 `confirmed rug pulls`. Based on this dataset, we perform a pilot study to identify key symptoms of NFT rug pulls (§ 4). Taking advantage of the summarized symptoms, we design an effective detector to flag *happened but undisclosed* rug pull scams in the wild by analyzing all the on-chain and off-chain data related to over 173K NFT projects (§ 5). We have flagged 7,487 `happened rug pulls` using the most reliable method. Although it is the lower-bound, it is adequate for us to train the NFT alert system. Then, to raise warnings of NFT rug pulls in advance (before they run away with investors' money), we devise a proactive method based on various features for real-time monitoring, to raise warnings of the potential NFT rug pull projects in their early stages (§ 6) and successfully report 5,557 suspicious NFT projects.

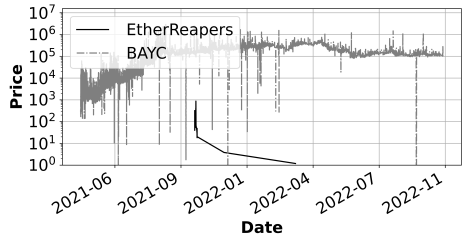


Fig. 4. The trading price of EtherReapers and BAYC in secondary markets.

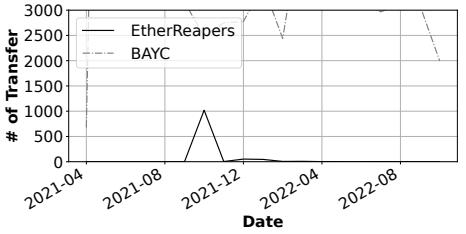


Fig. 5. The number of transfer events of EtherReapers and BAYC.

4 SYMPTOMS OF CONFIRMED RUG PULLS (RQ1)

According to the introduction in §2.2, the status of social media is the most eye-catching feature of rug pull. However, note that, it is only the necessary but insufficient condition for a rug pull, since benevolent projects could also lack or maintain abnormal social media accounts. Therefore, we need to delve into `confirmed rug pulls` to search for more comprehensive and specific symptoms for the expansion of our initial ground truth.

4.1 Initial Confirmed Rug Pulls

Some entities in the community, e.g., security companies and researchers, are proactively engaged in reporting NFT rug pull scams. Their reports can be used as the initial source of our pilot study. Specifically, we first harvest the labelled rug pull scams through `Chainabuse` [23], a well known crowd-sourcing based threat intelligence platform, and two security research accounts on Twitter, i.e., `Rug Pull Finder` [47] and `ZachXBT` [59], both of which constantly share reliable news of NFT scams. After removing duplicates, we have collected 324 reports related to NFT rug pulls. For further verification of these crowd-sourcing reports, we search the reported NFT projects in Google, intending to identify more evidences (e.g., scam accusation posts) supporting they are real scams. To the end, 253 NFT projects are labelled as `confirmed rug pulls` by at least two different sources, which are regarded as our initial ground truth.

4.2 The Symptoms of NFT Rug Pulls

Based on the collected 253 rug pulls, we next summarize their common **symptoms**. To highlight their characteristics, we further select top-300 NFT projects from `OpenSea` top list [14] for comparison. According to the general process (see §2.2), we depict their symptoms from four aspects.

4.2.1 Profit Analysis. Intuitively, making a profit is the underlying incentive of NFT rug pulls. Thus, a successful rug pulled project should be profitable. Our pilot study on the 253 confirmed NFT rug pull projects shows that all of them have made huge profits. In general, there are two ways of earning money for them. On the one hand, they can capture profit by requiring mint fee, and 251 of them belong to this category. Each of them has withdrawn 175 Ether from contracts on average. The most aggressive one is `Apes In Space` [19], the rug puller has withdrawn around 2,645.82 Ether (roughly $10.3M) through a single transaction. On the other hand, 236 of them have been traded in secondary markets, which is also a way to make money. To be specific, the rug pullers can receive a sum of creator fee for each trade up to 10% of the trading price. Furthermore, the price in secondary markets could be pumped by the rug pullers. The owners could sell tokens with

the high prices and gain profits by house wins. To sum up, all the 253 projects have gained a huge profit, which is a must-meet condition for a successful NFT rug pull.

S1 *The creators of successful rug pulled projects should gain a profit. This can be measured using both on-chain transactions and off-chain trade information in secondary markets.*

4.2.2 Secondary Markets Price Analysis. As discussed in §2.1, secondary markets play an important role in the life cycle of an NFT. For comparison, Fig. 4 shows the trading price of a rug pulled project `EtherReaper` [36] and a normal project `BAYC` [20]. Apparently, `EtherReaper` suffers an one-time dramatic rise and fall, while the price of `BAYC` is relatively stable although with some erratic fluctuation. Also, the price of `EtherReaper` will not be effectively recovered after the sharp drop. Furthermore, we can observe that `EtherReaper` has almost no trades after the dramatic decline in trading price, indicating that investors have lost confidence of it.

This symptom is not unusual. For those 236 ground truth projects that once have traded in secondary markets, 135 of them (57.2%) remain "silent" after the rug pull, i.e., the NFT holders (victims) cannot sell any tokens after the scam. In addition, for the remaining 101 projects, all of them have extremely low prices (i.e., under 1% of its top price), and the price does not rebound anymore. As for the top-300 projects, even if the price may also fall (see Fig. 4) with some erratic fluctuation, it will eventually rebound to a normal level in the end. Such a huge gap between these two types of projects is highly associated with the floor price, which will be extremely low if liquidities are lost due to some reasons. Intuitively, all top-300 projects still have active trades in the secondary market. Thus, the evolution of trading activities can be regarded as a signal to flag rug pulled projects. We thus conclude our second finding:

S2 *Compared with normal NFT projects, most of rug pulled projects will eventually expire in secondary market after the scam, causing the extremely low liquidity and low token price, and the price could be never rebound.*

4.2.3 On-chain Token Transfer Analysis. Intuitively, the liveness of on-chain activities of rug pulled projects will decrease to an extremely low level after the rug pull. Fig. 5 depicts the number of transfer events of `EtherReaper` and `BAYC`. For the rug pulled project, we can see a summit appears and then disappears after the rug pull scam happened. This follows the second stage mentioned in §2.2. However, this metric is quite stable for `BAYC`. Therefore, we compare token transfer events for rug pulled projects and top-300 projects. For the former ones, the average number of on-chain transfer within the first month since their launching is 5,819.35. However, the number drops to 51.42 (0.8% of the peak value) in the following month. As for top-300 projects, the number of on-chain transfer events is fluctuate within a reasonable range (usually 20% to 80% to its peak value). Therefore, we can observe a huge gap on on-chain liveness between these two types of projects. We can conclude our third finding:

S3 *The number of on-chain token transfer events will decrease to an extremely low level after the rug pull, which is a good indicator to identify the happened rug pulls.*

4.2.4 Social Media Analysis. We can observe from the case in §2.3 that a rug pull is highly associated with the status of its corresponding official social media. To be specific, for 253 ground truth, 97% of them have maintained social media accounts, e.g., `Twitter` accounts, `Discord` groups, `Instagram` pages, and official websites. All of them have abandoned their social media accounts after the scam, e.g., suspending `Twitter` and taking down their official websites. However, as for those normal projects, all of them have active social media events at the time of writing. Therefore, we can observe a huge gap, which can be used to distinguish rug pulls from innocent ones. Hence, we can summarize the finding:

Table 1. The statistics of our dataset (till Nov.1st, 2022).

Items	ERC-721	ERC-1155	Total
Contract	141,999	18,262	160,261
Transfer Event	157,930,255	17,808,803	175,739,058
mint	105,024,562	4,889,667	109,914,229
burn	1,615,223	739,805	2,355,028
swap	51,290,470	12,179,331	63,469,801
Market Trade	22,443,587		-
OpenSea	21,569,392		-
LooksRare	334,660		-
X2Y2	539,535		-
Social Media	60,021	14,968	75,008

S4 *The social media accounts of rug pulled projects will be abandoned, e.g., suspended (by the scammers themselves or by the platform) or silent for a long period, after the scam conducted.*

Answer to RQ1 *The NFT rug pulls share typical symptoms in aspects including profit capturing, price fluctuation, transfer evolution, and social media liveness. The symptoms offer opportunities for us to uncover more undisclosed rug pulls in the wild, as the expansion of the initial ground truth dataset with limited number of samples.*

5 DETECTING HAPPENED RUG PULLS (RQ2)

Based on the summarized symptoms of NFT rug pulls, we further want to identify happened but undisclosed rug pulls in the wild. Thus, we first collect all the on-chain transactions and off-line data related to Ethereum NFT projects, and then we design an NFT rug pull detector to flag rug pull scams that have already taken place, which we name as `happened rug pulls`. Note that, we aim to get the most reliable result of `happened rug pulls`, thus we have enforced a strict rule-based method in this section. Although the number of `happened rug pulls` we identified might be a lower-bound, they are sufficient to measure the overall landscape (detailed in this section) and based on which to further build the alert system (§6).

5.1 Data Collection

As aforementioned, NFT rug pull scams have typical symptoms which can be reflected by their on-chain transactions, secondary market trades and their social media activities. Thus, we make effort to create a comprehensive dataset that contains such information, which is detailed in Table 1.

5.1.1 On-chain Event & Contract. As mentioned in §2.1, under both `ERC-721` and `ERC-1155` standards, invoking the function `transfer` will not only initiate a transaction, but also emit an event which will also be recorded on-chain. Collecting these events can help us better understand the money flow. Thus, we have deployed a client node `Geth` [37], and synchronized all blocks until Nov. 1st, 2022 from its very beginning (corresponding to the first 15,871,480 blocks). Consequently, through these transactions, we have collected 157M and 17M pieces of `ERC-721` and `ERC-1155` token transfer events, respectively. These transfer events can be further divided into three categories, i.e., *mint*, *burn*, and *swap*. Specifically, they can be distinguished by the value of specific fields. The `from` field of a mint event is the *null* address, while the *to* field is the dead address for burn events. The remaining ones can be categorized into swap events, indicating ownership transfer of NFTs.

Among these NFT transfer events, we were able to identify 160,261 NFT contract addresses that were involved in these transfers (until Nov.1st, 2022). Additionally, we categorized these addresses into 141,999 ERC-721 smart contracts and 18,262 ERC-1155 smart contracts based on their distinctive function signatures. Furthermore, we extracted essential metadata from the API offered by ChainBase [25] for all these contracts. The metadata includes information such as the contract name, launch timestamp, and creator address. It provides valuable insights into the characteristics of these contracts.

5.1.2 Market Trade. Different with *token transfers*, referring to ownership transfer of NFTs between accounts, a *trade* is only related to smart contracts owned by secondary markets and has nothing to do with the project contract. To achieve this objective, we have design a *proprietary parser* to analyze the trades in the secondary market. We focus on all transactions interacted to smart contracts of the three dominant NFT markets, i.e., OpenSea [11], X2Y2 [13], and LooksRare [9], accounting for 83.87% of total history trading volumes in Ethereum [5] for ERC-721 and ERC-1155 NFTs. We have conducted a comprehensive manual analysis on their deployed smart contracts and filtered out all related transactions to extract trading information. Specifically, we firstly filtered out all external or public functions in these smart contracts to find ones that directly handle trading requests. Then, we collected all events emitted by these functions. Finally, we extracted all necessary information from these collected events, i.e., token id, token contract, and seller/buyer's address. Because trading NFTs often allows different cryptocurrencies, we use Ethplorer [6] to convert them to USD according to historically daily interest rate. Over 22M pieces of trades are collected.

5.1.3 Social Media. Our pilot study in §4 shows that as the first step of rug pulls, developers always promote their projects via social media to attract as many unsuspecting users as possible. Moreover, the shut down of social media signals the end of the NFT rug pull to some extent. We target Twitter, Discord, Instagram, and official NFT websites, to collect indicators of suspending or closing of such social media accounts. Specifically, we get access to the API from ChainBase [25], to get related links about those social media platforms. For example, project maintainers can make formal official announcements on Twitter. To examine the status of the corresponding Twitter accounts, we queried the official APIs [51] based on the user names crawled from ChainBase. In total, 60,021 ERC-721 and 14,968 ERC-1155 projects are linked to Twitter/Discord/Instagram, or their own official websites.

5.2 Pre-processing

Our objective is to identify the most reliable NFT rug pulls that have taken place using strict rules, which we believe represent a lower-bound of the rug pull scam. This will allow us to further train a reliable predictor in §6. To achieve this, we initially exclude some ambiguous projects that lie on the border of being classified as a rug pull.

First, we set the time constraint. Rug pulls may exhibit vague characteristics in the immediate aftermath of their launch. To provide ample time for projects to demonstrate clearer indications of rug pulls, we only consider projects that have been launched at least one month before our data collection. Consequently, we focus our detection efforts on the 122,249 projects that were launched before Oct. 1st, 2022 (as detailed in §3). *Second, we check the social media presence.* Basically, projects with a social media presence usually have a more comprehensive operational framework, making it easier to assess whether they may exhibit fraudulent characteristics or not. For this reason, we concentrate our detection on the 75,008 projects that possess at least one social media account. We acknowledge that this approach may overlook some rug pull scams. However, in this section, our primary objective is to expand the ground truth of rug pulls in the most reliable manner possible. In summary, we apply our detector to 70,712 projects that satisfy both of the above requirements.

5.3 Rug Pull Detector

Based on the typical symptoms we summarized in §4.2, we next measure the prevalence of NFT rug pulls in the wild. To this end, we have designed and implemented a detector, which takes all the data of a project as inputs, and outputs a report to show whether the examined project has been rug pulled or not. The detector is composed of four independent components, each of which is responsible for examining a specific feature, e.g., price fluctuation and profitability. A component will issue an alert if its corresponding rule is met. If all these four independent components raise alarm simultaneously, the project will be labeled as rug pulled.

5.3.1 Profit Checker. As stated in **S1**, a successful NFT rug pull project can usually grab profit in two ways, i.e., through on-chain Ether transfer or secondary market trades. Therefore, we design our profit checker from these two aspects. Specifically, we check on-chain logs to examine if there exist transactions sent Ether to the project. This indicates the contract creator behind the project receives profits directly. Otherwise, we firstly check whether a project can gather profits from secondary markets by analyzing the rewarding mechanism of each market. For example, if an NFT is traded successfully on `OpenSea`, the project will receive a sum of *creator fee* that can be withdrawn later. If the project can meet the requirement for one of the two above, then we assume this projects is profitable. For the sake of covering all possible profit-making activities, in the profit checker, we only focus on whether the project has profitable behavior, without considering how much profit it actually makes. In addition, even if the Ether is not directly transferred to the contract, or the owner of the contract, we also assume it is profitable, as there exist projects gain profit via *middleman* (see §5.4.4). This checker could help identify all the rug pull candidates for a further verification by other three checkers we implemented.

5.3.2 Price Checker. According to the solid line in Fig. 4 and **S2** in §4, the trading prices of rug pulled projects in secondary markets have two obvious characteristics. (*i*) the price will dramatically decline due to rug pull occurs after being pumped to the summit; (*ii*) the price of rug pulled projects will not rebound. Therefore, we introduce two metrics to quantify these two characteristics, i.e., *drawdown* and *recovery*, respectively.

Specifically, for each project, according to its trading history in secondary markets, we construct a chronological trading price sequence $\mathcal{P}$. $\mathcal{P}$ consists of a series of p_{time}^{token}, where $time$ and $token$ refer to the timestamp and the token identifier, respectively, and p is the trading price in USD. For example, $10^{3}_{1640995200}$ refers to the NFT indexed by 3 traded by \$10 at the midnight of Jan. 1st, 2022. Also note that, for any two adjacent items, the later one strictly happens after the former one, though they may correspond to different NFTs under the same project. For each p_i in $\mathcal{P}$, its $drawdown$ is defined as:

$$drawdown_i = max(\frac{p_i - p_j}{p_i}), \text{where } j > i$$

where p_j corresponds a trade happens strictly after the given p_i. In other words, the $drawdown_i$ is calculated by p_i and the *lowest trading price* after it. To this end, we can construct a *drawdown sequence*, whose length equals to $|\mathcal{P}| - 1$. Heuristically, if any element in the drawdown sequence is greater than 0.99, we can conclude that there is a dramatic price plummeting. For example, for `EtherReapers` depicted in Fig. 4, the highest trading price happened on its first trade, which corresponds a $drawdown$ around 0.9963. Note that, we build the chronological trading price sequence by involving all trade logs, regardless of the price variation for each NFT. This is primarily due to the low liquidity of an NFT project. Take `EtherReapers` as an example again, no tokens are sold for twice in this project. Thus, if we create separate price sequences for individual tokens, the price checker will be obviously invalid, as each sequence would contain only one element.

Moreover, we can also observe a long tail in Fig.4, which reflects the *unrecoverability* of all NFTs of a rug pulled project as mentioned in **S2**. Thus, for $drawdown_i$ whose value is greater than 0.99, we further examine if its p_j can be traded with a much higher price. Let us assume the p_j is related to the token t, i.e., p_j^t. Its $recovery_j$ can be defined as:

$$recovery_j = max(\frac{p_k^t - p_j^t}{p_j^t}), \text{where } k > j$$

, where p_k^t is the price that corresponds to the following trades on the same NFT. Similarly, we heuristically set the threshold of $recovery_j$ as 0.01, indicating the highest trading price after a local minimum p_j can only be 1% higher for the same token. Take `EtherReapers` as an instance again, we have identified 60 $drawdown$ greater than 0.99 in total. For all p_j^t in these $drawdown$, the highest one is 0.9986, related to the NFT whose tokenID is `0x01b9`[1]. However, as for all related tokens that result in the $drawdown$ over 0.99, there are no other trades of these tokens. In other words, after a sharp drop in prices, the price of all NFTs cannot be recovered at all. If the project does not have any trades after the last trade for over one month, for the lack of project's market activities, we also label that after the last trade, the price takes a total drawdown and could not recover. Similarly, if t is not traded at all after p_j^t, we set the $recovery_j$ as 0 due to the same reason. Note that we only apply this rule for projects that once had trades in secondary market.

5.3.3 On-chain Liveness Checker. The liveness of an NFT project, summarized by **S3**, can be reflected by the number of on-chain transactions. Specifically, we denote N_1 and N_2 as the number of transfers happened in the first month once the first token was minted and the period of the last month, respectively. Based on them, we define the *on-chain liveness* as follows:

$$liveness = \frac{N_1 - N_2}{N_1}$$

, where the *liveness* should be greater than 0.99, meaning that its liveness is neglectable compared to the first month, which is a typical symptom after suffering rug pulls. For example, the number of transfer logs related to `EtherReaper` of the last month month is only 0.008 of its first month, which means the *liveness* of `EtherReaper` is 0.992. Note that this checker is not suitable for projects that have launched within the last month, as N_1 will obviously be equal to N_2 and the *liveness* will always be zero. Hence, at the beginning of this section, we mentioned that we only include the four checkers for projects launched one month before our data collection.

5.3.4 Social Media Checker. The status of social media services is an effective evidence of NFT rug pulls. As mentioned in §5.1.3, we have tried our best to collect all the social media accounts related to the NFT projects through APIs provided by `ChainBase` [25]. We next check the status of their social media accounts. For `Twitter` account, we check if it is suspended, deleted, or inactive for a long time (at least one month). Furthermore, we check whether the invitation link is expired for `Discord`, whether the account can still be found or be suspended for `Instagram`, and whether the server is down for their official websites. If any of the above requirement is satisfied for a specific project, we consider it fulfills a critical symptom of rug pull. Note that the shut down of social media is only *a necessary, but not a sufficient condition*. For example, consider a project `Parallel Alpha` which discards its original `Twitter` account (`parallelnft`) and registers a new one (`ParallelTCG`), potentially can be flagged by this module. This illustrates why we still require additional stringent checks to confirm such scams.

[1]https://etherscan.io/tx/0x1bfc75...

5.4 Rug Pulls in the Wild

5.4.1 Accuracy of our detector. We first evaluate the accuracy of our detector by applying it to our ground truth (253 rug pulls) and our randomly selected NFT projects (300 normal NFT projects).

For the 253 confirmed rug pulls, 243 of them meet the two requirements in our pre-processing (see §5.2), since 10 projects lack social media information. Then we feed all the 243 projects to our detector, and observe that 232 of 243 rug pull projects can be identified correctly, i.e., 11 false negatives. For the 11 false negatives, we found that they were not raised alarms by the price checker. Specifically, the *drawdown rates* of them are roughly 90% but not 99% used in our checker. For the 300 normal NFT projects, 278 of them are qualified after the pre-processing step (§5.2). When feeding them to our detector, although certain checker raises alarms for some projects, no project is simultaneously alarmed by the four checkers, i.e., we observe *no false positives*.

Selection of Thresholds. To further investigate the impact of threshold selection, we conduct single-factor ROC analyses by changing the threshold for *drawdown rate*, *recovery rate* and *liveness rate*, which are shown in Fig. 7 (Appendix 2). As we can see, for all these three metrics, lowering the threshold significantly increases the false positive rate and makes little contribution to the true positive rate. According to the *Cannikin Law* [1], though an alarm can only be raised when four checkers detect abnormal behaviors simultaneously, we still conservatively choose the threshold to eliminate possible false positives.

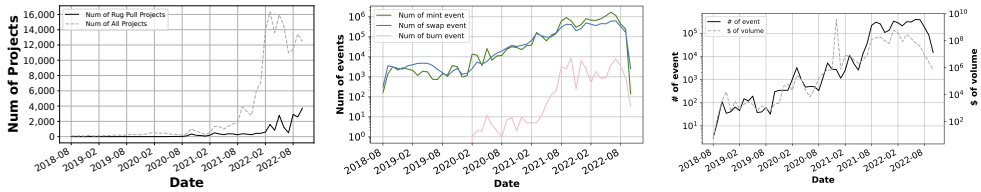
We admit that rule-based detection shares inherent limitations, e.g., rely heavily on existing knowledge. As aforementioned, even if some false negatives remain, we aim to identify the most reliable NFT rug pulls (i.e., flag no false positives), thus we have enforced most strict rules and thresholds.

5.4.2 Rug Pulls in the wild. We have applied our detector to all the 70,712 NFT projects after the pre-processing step, and it flags 7,487 rug pull projects in total (i.e., happened rug pulls), with 7,019 `ERC-721` projects and 468 `ERC-1155` projects, accounting for 5.08% and 1.70% of existing NFT projects during our study, respectively. We further sampled 50 `ERC-721` projects and 50 `ERC-1155` projects for manual verification, and we did not observe any false positives. We disclose our detection result to an anonymous leading blockchain security company, which further confirmed our findings. It suggests the reliability of our approach.

Overall Statistics. These rug pull projects correspond to 17.9M pieces of transfer events recorded on-chain, and 3.6M pieces of trade events on secondary markets, reaching up to $6.2B in terms of trading volume. Although the number is impressing, many previous studys [63, 70, 75, 79–81, 83] have revealed that scammers tend to create the mirage of a prosperity project by manipulating the trade events, i.e., wash trading. According to our statistics, they totally mint around 11.3M `ERC-721` tokens and 93K kinds of `ERC-1155` tokens. They are created by 6,778 addresses, which can be regarded as the real culprit that conducts rug pulls. Most rug pullers (92.36%) create only one project, while the other 518 ones operate multiple rug pull projects. Astonishingly, the address `0x453e`[2] has created 12 projects, all of which have been rug pulled.

5.4.3 Trending of Detected Rug Pulls. Fig. 6 shows an overview of these 7,487 happened rug pulls. As shown in Fig. 6(a), the trend of NFT rug pull is consistent with the rising of the NFT ecosystem, which becomes popular and grows rapidly at the beginning of 2022. At the peak, there are over 50 projects that are created within a single day but finally being rug pulled. Fig. 6(b) shows the number of emitted events of rug pulled projects per month. As we can see, both the *mint* and *swap* events are constantly upward till August, 2022, which indicates that rug pulls have become more and more rampant. In 2022, on average, there are over 10^4 pieces of mint events and swap events,

[2]https://etherscan.io/address/0x453e23...



(a) The number of daily transactions. (b) The number of daily events. (c) Daily trading volumes.

Fig. 6. General overview of detected rug pulls.

respectively, emitted by these rug pull projects per day. Interestingly, the number of mint events is higher than the one of swap events, which means that each NFT token will only be swapped less than once on average. This finding is consistent with the low liquidity of rug pulled projects in the NFT ecosystem. Moreover, as we mentioned in §5.1.1, a *burn* event indicates that an NFT is intentionally destroyed by its owner. Its number is also went upward from 2021. However, we can observe a sudden decrease for all these three events after August. This is because we only present the data until the end of October, while some projects are rug pulled in the following month. Fig. 6(c) shows the statistics of daily trades in terms of number and volumes. In 2020, the number of daily trade is 53.32 on average, while in 2022, it is 8,217.91. As can be seen, rug pulls have become more rampant in secondary markets that our community should pay more attention on.

5.4.4 Funds Drained by Rug Pulls. Beyond mint fee and creator fee outlined in §4.2.1, our manual investigation suggests other ways to gain funds of NFT rug pulls while diving deeper. Next, we provide a measurement of the methods that exploited by rug pulls to gather profit.

Mint Fee. Minting a token will generate a transfer event whose `from` is a null address. Taking advantage of this characteristic, we extract such transfer events whose `from` is a null address and `value` is not zero. Then, we will observe if the withdraw function is called, where the owner transfers all balance out. In total, we have found 4,821 rug pulled NFT contracts (64%) that have profited by mint fee. Table 4 in Appendix 3 shows the top 5 scam projects in terms of mint fee Ether. As we can see, rug pullers can gain huge profits (over thousands of Ethers) for each project.

Creator Fee. To identify the creator fees drained by rug pullers, we have to extract the money flow from secondary markets to contracts of projects. Note that, not all secondary markets explicitly transfer creator fee via transactions. We thus only focus on `Wywern` of `OpenSea` (see §5.1.2). At last, we found that 2,140 projects (28.6%) have once received creator fee, which is worth over $12M, as part of its revenue. Table 5 in Appendix 3 shows the top 5 scam projects that gain the most creator fee, which sums up to $4.9M, accounting for 40.8% of the total.

Middleman Reselling. Middleman reselling means that all the NFTs of a project are minted to a single account, who then resells them on secondary markets. For potential buyers, they may think that the person they are dealing with is a whale account who holds bunch of NFTs. However, the seemingly *whale account* has high likelihood to be a collusion address controlled by the rug pullers. To identify such behaviors in rug pulled projects, we extract all mint events and calculate the distribution of minted NFTs. As a result, we find that 865 rug pulled NFT projects (11.6%) mint all tokens to a single account. We first exclude NFTs of 273 projects have not been involved in any on-chain transactions at all. We find that 70 projects of the remaining 592 projects have direct Ether transfer events from some user addresses, but lack any trades on secondary markets. Interestingly, the lack of trades does not mean that these projects are not profitable, and we presume that they

```
1   totalSupply = 8390
2   maxTokens = 50000
3
4   function mint(address recipient, uint256 seed) external override whenNotPaused {
5       require(admins[_msgSender()], "Only admins can call this");
6       require(minted + 1 <= maxTokens, "All tokens minted");
7       minted++;
8       generate(recipient, minted, seed);
9       ...
10      _safeMint(recipient, minted);
11  }
```

Listing 1. An example of a hidden mint backdoor.

did this deliberately. Thus, we regard these 70 rug pull projects exploited the middleman reselling. Table 6 in Appendix 3 shows the top 5 of them in terms of the amount of profits.

Hidden Mint Backdoor. Considering the substantial funds that a rug puller can acquire through mint fees, we make the additional assumption that the rug puller might employ even more aggressive methods to exploit these fees. Consequently, our manual auditing uncovers a noteworthy backdoor—specifically, a hidden mint backdoor. Hidden mint refers to the behavior where NFT projects privately mint tokens whose amount is beyond the maximum predefined number, allowing extra profits drained by the rug pullers. Note, only one field is associated with the supply of NFTs in the standard, named `totalSupply` [33], which will return the maximum allowable minted number of tokens. However, take a famous NFT rug pull project, `Cat&Mouse` [22], as an example, whose code snippet is shown in Listing 1. First, `totalSupply` at L1 declares that at most 8,390 NFTs can be minted and circulated. Nevertheless, we have observed 8,879 available NFTs circulated on-chain, which is resulted from the implementation of `mint` (L4). Specifically, L5 asserts that only the designated administrators have the mint permission. At L6, it examines whether a mint action is allowed by comparing to the upper limitation, which is declared by `maxTokens` (set as 50,000) instead of `totalSupply`. It means that only the designated administrators are allowed to overmint NFTs, which will extremely devalue minted NFTs and gather more funds. To pinpoint how many NFT rug pulls have exploited hidden mint, we firstly fetch the declared value of `totalSupply` of all NFT contracts from APIs provided by `Alchemy` [10]. Then, we trace all the transactions initiated from each project, and calculate whether the amount of NFTs in circulation is greater than the value of `totalSupply`. Because the non-fungibility of `ERC-721` tokens, simply calculating circulated amount is feasible. However, since the IDs of `ERC-1155` NFTs are usually shared by multiple identical tokens, the same method is not applicable. As a result, we found 104 `ERC-721` projects have minted tokens covertly.

Answer to RQ2 *NFT rug pulls are prevalent in the ecosystem, and our detector has identified 7,487 happened rug pulls in the wild, which is 30× higher than the reported ground truth in our community. Scammers are taking advantages of kinds of tricks to grab profit.*

6 EARLY WARNING OF NFT RUG PULLS (RQ3)

Our previous exploration has uncovered that NFT rug pulls are prevalent in the ecosystem. However, our previous efforts are *postmortem analysis* that relies on reactive methods to flag NFT rug pulls after the scam has happened. It is urgent for our community to raise warnings of such scams at their early stage. In this section, we seek to explore whether we can identify suspicious NFT projects before rug pull happens based on the initial indicators extracted from transaction and trade events.

6.1 Dataset and Pre-processing

For all the labeled 7,487 NFT rug pulls in §5.4.2, we regard them as the ground truth, denoted as D_{pos}. We further choose the top-1000 projects listed in the OpenSea top collection [41] as normal projects, denoted as D_{neg}. Among these 1,000 cases, some of them are not based on Ethereum, thus 933 projects remained. Considering that the current cases in D_{neg} are generally more active and have higher transaction volumes than those in D_{pos}, we need more cases in the wild to avoid overfitting issue. Therefore, to ensure that the new chosen projects are not rug pulls, we randomly sampled 2,000 NFT projects beyond the top-1000 list, which have no signs of rug pulls by the time of our study, i.e., they are active in social media and secondary markets, they receive quite a lot of discussion in the community, and they show no other scam behaviors we mentioned before. We further add them into D_{neg} . Finally, we have 2,933 projects in D_{neg} and 7,487 projects in D_{pos}.

To train a model for raising early warnings, we need some fine-grained timestamps of each project. Therefore, we have to define some notations based on the life-cycle of an NFT project:

- T_{RP}: the exact timestamp when the rug pull happens if the given project is rug pulled;
- T_A: a timestamp ahead of T_{RP}, at which we try to raise an alarm for suspicious rug pulls;
- P_{FE}: a period of time ranging from the launching of the project to T_A.

Note that, for cases in D_{neg}, both T_{RP} and T_A are set as the data of our collection, i.e., P_{FE} would last until the end of data collection. For cases in D_{pos}, the first challenge is how to determine T_{RP}. Based on a series of NFT rug pull security reports (collected during the initial 253 ground truth samples) ordered by significance, We heuristically design a set of rules, as follows:

(1) If the project has been withdrawn more than once, we will set the T_{RP} as the moment of the withdrawal with the highest amount of Ethers. Even though a project may experience multiple withdrawals, by our observations, most rug pulls happen after the biggest one. The rule successfully determines T_{RP} for 5,004 projects;
(2) Otherwise, we set the T_{RP} as the moment its social media accounts last update, if it was not deleted or suspended. The rule successfully determines T_{RP} for 1,675 projects;
(3) Otherwise, if there is a drawdown greater than 0.99 (defined in §5.3), we will set the T_{RP} as the moment when the last p_j in the drawdown occurred. This indicates that the investors finally lose confidence on this project. The rule successfully determines T_{RP} for 3 projects;
(4) Otherwise, we will set the T_{RP} as the moment of the last market trade record in secondary markets. The rule successfully determines T_{RP} for 239 projects.

Not all 7,487 can be successfully labelled, for we only label the most reliable results to help train the model. If T_{RP} cannot be determined by these four strategies, to promise the accuracy of our prediction model, we will remove the case from D_{pos}. Consequently, 6,921 cases remained, and they all have a firm and determined T_{RP}. For the 253 initial collected ground truth samples, our labelled T_{RP} is inline with their original security reports, which indicates the reliability of our method.

Recall that our purpose is to raise warnings of rug pulls before T_{RP}. Thus, we decide to use a *time slicing window* to evaluate how early we can accurately predict the rug pull will happen ahead of T_{RP}, i.e., the length of $T_{RP} - T_A$. Specifically, because the T_{RP} for these 6,921 cases of D_{pos} has fixed already, we move T_A to extend the length of $T_{RP} - T_A$ to give possibilities for investors to transfer assets out as much as possible. For all cases in either D_{pos} or D_{neg}, we set different time slicing windows in different rounds of training, as shown in the first column of Table 2. After labeling on the life-cycle of cases, we merge D_{pos} and D_{neg}, and divide them as 80% training data and 20% testing data, using cross validation to evaluate the models.

6.2 Extracting Features

To effectively alert NFT rug pulls in the early stage, we need build an effective classifier. Thus, we have extracted a comprehensive set of features, which are divided into three categories as follows.

6.2.1 Time-series Features. Time-series features are composed of temporal metrics that can be used to flag *oncoming* rug pull projects. The distribution of certain activities on the life-cycle of a project is a good indicator. Intuitively, due to the dramatic price pump of rug pulled projects at the early stage, the distribution of certain types of activities (e.g., mint event and swap event) will be concentrated to the first or the second half of the project's life. However, for those healthy and innocent ones in D_{neg}, it is completely a different story. Moreover, some trading price related activities, e.g., the timestamp of the floor price occurs, also differ between the rug pulled projects and normal ones. To depict such a distinction, we introduce P_{act}, which reflects the distribution of one certain type of activities along the timeline. Specifically, it is defined as:

$$P_{act} = \frac{\frac{1}{N_{act}} \sum_{i=1}^{N_{act}} (T_{act}^{i} - T_{act}^{1})}{T_{act}^{N_{act}} - T_{act}^{1}}$$

, where *act* can be replaced by a specific type of activities mentioned above, N_{act} is the total number of this type of activities, and T_{act}^{i} denotes the timestamp of the i-th of this type of activity. For example, for those normal cases in D_{neg}, the P_{swap} will be roughly close to 0.5. This is because these projects typically operate normally for a long period and have long-term and sustained liveness on-chain, resulting in an even distribution of swap events. However, it is totally different for cases in D_{pos}. To be specific, projects are typically rug pulled when the projects suffer frequent pump at the beginning of the projects. Therefore, all swap events will be concentrated at the end of P_{FE}, which leads to P_{swap} as 1. Moreover, the pump will also push all top prices occur at the end of P_{FE}, while it tends to be random for normal projects. In a nutshell, P_{act} can measure the degree of concentration of distribution of a certain activity. We have extracted 8 features in total, including certain type of events and prices, which are shown in Table 3 in Appendix 1.

6.2.2 Features of On-chain Events. On-chain events are emitted by on-chain token transfers, which can reflect the liveness of the corresponding projects. Thus, we pay attention to events from two perspectives, i.e., the total number (denoted as N_{event}), and the number of involved participants (counted by addresses, denoted as A_{event}). For example, as discussed in §5.4.2, N_{mint} is often higher than N_{swap} for rug pulled projects because of the inborn poor liquidity of rug pulls. But for non-rug-pull projects, it is always the opposite. In addition, A_{mint} refers to how many accounts are participated in minting, which could be extremely low. For example, NFTs may be all minted to a middleman (see §5.4.4). Furthermore, we think the ratio of certain type of events to all emitted events can also reflect some characteristics. To depict the above features, we thus defined another two features, denoted as RN_{event} and RA_{event}. For example, swap events for rug pulled projects only account for a small part of all emitted events, leading to an extremely low RN_{swap}. Totally, we have extracted 14 features, related to three types of events (see Table 3 in Appendix 1).

6.2.3 Features of Secondary Market Trades. Similarly, the liveness of projects in secondary markets can also be reflected by features extracted from initiated trades. Due to rug pulled projects would always manipulate the mechanisms of secondary markets, they will definitely display different traits, which can be further utilized in P_{FE} to distinguish cases. Specifically, we mainly focus on three metrics, i.e., N, U, and V, referring to the total number, involved users, and price or trading volumes. For example, N_{trade} stands for the total number of trades, and $V_{average\_price}$ refers to the average price of all trades. Moreover, we still think the ratio can reflect some characteristics that can be utilized. We add a prefix R for those three metrics. For example, $RU_{highest\_24h}$ means the

Table 2. The prediction results. *A*, *P*, *R*, *F*1 refer to *precision*, *accuracy*, *recall*, and *F1 score*.

$T_{RP} - T_A$	**Logistic Regression**				**SVM**				**Random Forest**			
(hour)	*A*	*P*	*R*	*F*1	*A*	*P*	*R*	*F*1	*A*	*P*	*R*	*F*1
0	0.96	0.97	0.98	0.97	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98
1	0.96	0.97	0.97	0.97	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98
2	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98
4	0.95	0.97	0.97	0.97	0.95	0.96	0.97	0.97	0.97	0.97	0.99	0.98
8	0.96	0.96	0.98	0.97	0.96	0.96	0.98	0.97	0.97	0.97	0.98	0.98
12	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98	0.97	0.97	0.99	0.98
16	0.96	0.96	0.99	0.97	0.96	0.96	0.99	0.97	0.96	0.96	0.99	0.97
24	0.96	0.97	0.98	0.97	0.97	0.97	0.99	0.98	0.96	0.96	0.99	0.98
36	0.96	0.96	0.98	0.97	0.96	0.95	0.99	0.97	0.96	0.96	0.99	0.97
48	0.95	0.96	0.98	0.97	0.96	0.96	0.98	0.97	0.96	0.95	0.98	0.97
60	0.96	0.97	0.97	0.97	0.97	0.96	0.99	0.98	0.97	0.97	0.99	0.98
72	0.96	0.96	0.98	0.97	0.97	0.96	0.99	0.97	0.95	0.95	0.98	0.96
84	0.96	0.95	0.98	0.97	0.96	0.95	0.99	0.97	0.96	0.96	0.99	0.97
96	0.95	0.95	0.97	0.96	0.96	0.95	0.98	0.97	0.95	0.95	0.97	0.96

ratio of the users involved in one day that has the highest trading volume to all involved users in history. In total, we have extracted 30 features shown in Table 3 in Appendix 1.

6.3 Model Training & Predicting Result

By following the previous work in the research community [85], we use three basic and widely-adopted machine learning classification algorithms, i.e., logistic regression [73], SVM [68], and random forest [65]. Moreover, we directly adopt the default values of parameters in their `scikit-learn` implementation [49]. Four metrics are calculated to quantify the result, i.e., *precision*, *accuracy*, *recall*, and *F1 score*. Table 2 shows the results of metrics on different time slicing windows, i.e., $T_{RP} - T_A$. Generally speaking, all the three methods can obtain a quite good result (e.g., roughly 90% precision, recall, and F1) when the slicing window is set up to 96 hours (4 days earlier before the scam happens), which indicates the validity of our prediction method.

6.4 Real-time Monitoring

We next deploy our alert model to real-world settings, for real-time monitoring.

6.4.1 Real-world Settings. We have implemented a prototype system and launched it since Nov, 2022, and conducted a prediction once a day, predicting the projects that launched in the slicing windows from three months ago until the time of prediction. During in-the-wild study, we found that the four-day model (96h, see Table 2) is accurate enough to avoid premature predictions and timely enough to allow users to take emergency actions. Therefore, we will set the $T_{RP} - T_A$ as 96 hours, and predict once a day at midnight if the given project will be rug pulled in the next following days. All features will be extracted and calculated from the birth of each project to the day of the prediction. As all the three models on different algorithms (§6.3) can effectively predict rug pulls, thus we will label it as an oncoming rug pull project if any of these three models raises alarms.

6.4.2 Result Overview & Verification. Our system raises alarms everyday, ranging from 450 to 2,000 reports, since its deployment. Till May. 1st, 2023, the number of alarms is 1,564 everyday on average, and we have predicted 5,557 projects that have the possibility to be rug pulled in total

after removing duplication. Although it is hard for us to verify whether these projects will be rug pulled, we have observed some additional evidences that can support our prediction.

First, we apply our rug pull detector (which is used to flag happened rug pulls) proposed in §5 to the predicted NFT projects, which shows strong evidences to support of prediction. Our key idea is that, as our prediction model raises alarms for suspicious NFT projects that have not yet been rug pulled, thus intuitively, our rug pull detector in §5 cannot flag such NFT projects at the very early stage when our prediction model raises alert. However, if our predictions are correct, i.e., the corresponding NFT project will be rug pulled in several days later, our rug pull detector in §5 can indeed flag such NFT projects at the later stage when the rug pulls have been conducted. In this way, we can verify the effectiveness of our prediction. *The results are as our expected.* For the predicted suspicious NFT projects, our rug pull detector (in §5) only flags 1.01% of them as happened rug pull at the very early stage when our prediction model first raises alert, which means that most of them have not been rug pulled by the time of our alert. However, when we investigate these projects one month later by feeding them to our rug pull detector again, we observe that 77.3% of them have been flagged as rug pulls. This suggests that most of them have indeed been rug pulled after our prediction. Recall that our detector is quite conservative (i.e., with strict rules), which represents a lower bound and thus potentially underestimates the true real-world occurrences. Thus, we have strong evidences for the claim of the effectiveness of our predictor in real-world scenarios.

Second, we have been collaborating with a leading blockchain security company (anonymized due to double-blind requirement) to raise warnings for users. We have reported over 50 popular (but suspicious) NFT projects with large number of transactions, and all of them have been confirmed. For example, `Wen Sandwich` [55] is a popular NFT projects that holds total volume for 902 Ethers on OpenSea. However, our system raises alarm of this project since Feb. 13th, 2023. Four days later on Feb. 17th, 2023, the project withdrew all its Ethers (over 194 `Ether`) from its account. Since April, its official Twitter account has been abandoned [56]. Our system alarms it at the very early stage which can prevent more damage to the investors and the NFT community.

Third, beyond using our rug pull detectors, we have further manually investigated our prediction results to see if they have actually been rug pulled later. We sampled 200 projects for manual verification, based on their transaction behaviors (e.g., drain the money out) and the investors' complaints. We observed that over 90% of them have been rug pulled afterwards by the time of this writing. For example, `Donald Trump Yacht Club` [28] is a collection of 9,706 NFTs, which launched on Dec 12th, 2022. However, it has withdrawn all the `Ether` in the contract on Dec. 21st, 2022 [29], and its `twitter` account has been abandoned [27], as well as its floor price drops to 0.002 Ether later [26]. However, we start to raise alarms of this project since Dec. 19th, 2022, 2 days before its withdrawn. Note that some contracts are not detected by our detector but identified by our manual examination. Recall that, our detector follows a conservative approach, which detects the lower bound of rug pulls. Therefore, it may potentially underestimate certain cases. To further investigate the impact of our threshold selection, we conduct the ROC analysis for our detection on prediction results. However, even if we lower the *drawdown rate*, *recovery rate* or *liveness rate* to 0.9, the true positive rate remains the same (77.3%). Recall that, our four checkers together contribute to the results, which means changing one of them will not influence much on the results.

6.4.3 Assessment on Financial Impact. We further assess the potential financial losses that our system can mitigate. Note that the measurement process in this part is a roughly approximation, given that both victims and rug pullers might engage in the project dynamically.

First, our model can effectively prevent investors from incurring significant financial losses on mint fees after our prediction. According to our calculation, 1,461 projects still had ongoing mint fee process after being first predicted as potential rug pull. The total mint fee accumulates to 10.78K

Ether. The project that accrued the highest mint fee even after being predicted as potential rug pull was UINTS [52], accumulating 104.24 Ether. Now, the floor price of this project stands at under 0.0001 Ether on OpenSea[53] now. Furthermore, its official twitter account has stopped updating since May 5th, 2023 [54] (also its public mint day), providing strong indications that this project indeed constitutes a rug pull. Through the implementation of our model, the risk of such substantial losses in minting fees subsequent to our predictions can be effectively eliminated.

Second, our model can be served as a valuable indicator for investors to prevent them falling into the trap and mitigating the huge earnings of creator fees by scammers. After first being predicted as potential rug pulls, 4,147 projects continued to perform trades in secondary markets, for a total of 314,690 instances of trade events and over $507M history trading volumes, bringing huge creator fees flowing to the rug pullers. Therefore, we can prevent the investors from falling into the trap after the alarms. The project that exhibited the most activity in the secondary market after being predicted as rug pulls was RPlanet [45], with 826 trades and $25M history trading volumes. Note that the project has been rug pulled, evidenced by the profit grabbing of mint fee of 694.64 Ether [12], abandonment of its Twitter account by the creator team since March 2023 [46], coupled with a decline in the floor price to 0.002 Ether [42]. By deploying our model, we can mitigate such a significant volume of token trades in the secondary market, thereby safeguarding potential new victims from investing in tokens and huge creator fee drained by the rug pullers.

Answer to RQ3 *We have implemented a real-time alert system that predicts suspicious NFT projects before rug pull happens based on the initial indicators extracted from transaction and trade events. We have predicted 5,557 suspicious NFT projects and most of them have been confirmed with strong evidences. Our system can work as a whistle blower that pinpoints rug pull scams timely, thus mitigating the impacts.*

7 DISCUSSION

7.1 Lessons Learned

Our work reveals that rug pull events are rampant in the NFT ecosystem. Therefore, we summarize some suggestions for investors, developers and secondary markets managers.

For investors. Several best practices can be given for NFT investors. First, only open-source projects that are audited by prestigious security companies are acceptable. Otherwise, rug pulls may occur due to the hidden backdoors (see §5.4.4). Second, pay special attentions to the NFT projects that require a sum of mint fees, as withdrawing mint fee is one of the mainstream profiting way for rug pullers. Third, do not be fooled by the seemingly prosperity of NFT projects. If a popular NFT can be bought by a lower price than market price, pay attention to examine if it is a "mirage".

For creator team. For a better NFT ecosystem, developers should avoid rug pulls happen, which can be generally divided in twofold. On the one hand, developers should strictly follow the best practices of implementing ERC standards. For example, use totalSupply instead of other self-defined variables to limit the circulation of available NFTs. Developers should also open-source the implementation or even ask for code auditing and bug bounty to eliminate investors concern. On the other hand, the team should pay attention to behaviors that may result in misunderstanding of investors, e.g., requiring a bunch of mint fees, issuing all NFTs to a single account. In general, building a project's reputation and gaining users' confidence heavily require efforts from developers.

For secondary markets managers. First, all circulated NFTs should be critically reviewed to avoid scams which bring in financial losses for both trading platforms and holders. Second, creator fee is a huge part of profits for rug pullers, which urges the disclosure and traceability of all trading

transactions to track attackers if rug pulls happened. Last but not least, the predicting method proposed in this paper is proven to be effective and efficient. It is reasonable and practical to integrate such predicting methods into secondary markets to raise alarms for investors in advance.

7.2 Threats of Validity

Our study carries certain limitations. First, our data of secondary markets might be incomplete. As we discussed in §5.1.2, we only collected trades from the top three secondary markets in terms of trading volumes. However, these three have accounted for over 83% of total trading volumes in Ethereum for `ERC-721` and `ERC-1155` NFTs, which means including other markets will not influence the final results significantly. Second, our rule-based detector in §5 is quite straightforward and conservative, which heavily depends on the symptoms we summarized from the pilot study. However, to get a reliable results, we make them quite conservative for both rules and our detector. To this end, we can guarantee that we can identify the lower bound of rug pulled NFT projects. Thirdly, as our predictor aims to predict projects with malicious-intention before a rug pull occurs, it remains challenging to precisely determine when and how such projects will execute a rug pull in the real world. Also, it is hard for us to assess the exact number of financial loss that the system can prevent. Nevertheless, we have tried our best to show the reliability and effectiveness of our alert system in real-world scenarios.

7.3 The Collapse of the NFT Market

Recently, some media reports [44, 50] suggest that the NFT market starts collapsing, with the trading volume declined after its prosperity in 2021 and 2022, and many NFTs are now worth nothing. Although the price may drop, our approach is still effective in distinguishing them. Even if some projects may suffer price decline due to the depression of NFT ecosystem, only the price checker would be affected, while other three checkers will work as before. However, a project will be labelled as being rug pulled by the detector only if four checkers simultaneously raise alarms. Similarly, our prediction models consider 52 kinds of features, whose robustness is difficult to be affected by a few features. Additionally, the price decline resulting from the collapse of the environment is gradual, whereas the price decrease due to a rug pull is extremely sudden. For instance, the floor price of the `Bored Ape Yacht Club` occasionally experiences a drop [21]. Nevertheless, the floor price is still approximately half of what it was three months ago. The price checker continues to be unable to trigger an alert for this project, and the other three checkers remain unaffected.

8 RELATED WORK

8.1 NFT measurement

Following the surge in popularity of NFTs in 2021, several researchers have focused on this area [60–64, 67, 69–72, 74, 75, 77, 79–84]. In 2021, Wang et al. [82] introduced the NFT ecosystem as a first step. Kugler et al. [74] proposed the use of non-fungible tokens and measured their economic impact. In 2022, White et al. [84] conducted a study on OpenSea, and found that despite sparsity in the network, communities of users are forming and power users tend to congregate in these structures. In 2023, Roy et al. [77] used machine learning to detect NFT phishing. Gupta et al. [72] conducted a security survey of the NFT ecosystem and identified various security issues. In addition, Von et al. [81] used different methods to detect wash trading behavior in the NFT ecosystem on Ethereum. However, as the most prominent type of scam, NFT rug pulls have not been systematically explored. Our research serves as the first to characterize NFT rug pulls, which is of significant importance to stakeholders in the NFT community.

8.2 Cryptocurrency rug pull

In 2021, Xia et al. [85] employed machine learning methods to identify scam tokens in Uniswap, a decentralized exchange of DeFi. Among the scam schemes, rug pulls were discussed in their work. In 2022, Mazorra et al. [76] introduced the environment of ERC-20 tokens and two types of confirmed rug pulls. Scharfman et al. [78] discussed DeFi case studies, including rug pulls, pump and dump scams, and regulatory actions involving DeFi. Cernera et al. [66] discussed three types of malicious behavior in Binance, including rug pulls. Our work builds upon previous research on rug pulls but is specific to NFT projects, as the characteristics of NFTs are distinct from those of previous work. We propose a method to detect NFT rug pulls and raise alarms ahead of the scam happens, which is different with all existing work.

9 CONCLUSION

This paper presents the first comprehensive study of NFT rug pulls. By summarizing the key symptoms of rug pull scams, we have formulated a list of concrete rules to flag rug pull projects in the NFT ecosystem, and curated a list of 7,487 rug pull projects, by far the largest dataset of NFT rug pulls. To further impede the expansion of the scam, we further design a prediction model to proactively identify the potential rug pull projects in an early stage ahead of the scam happens. This paper presents the first solution to detect and mitigate NFT rug pulls.

ACKNOWLEDGMENTS

We are deeply grateful to our shepherd Professor Vishal Misra (Columbia University) and the anonymous reviewers for their insightful suggestions. This work was supported in part by National Key R&D Program of China (2021YFB2701000, 2021YFB2700700), the Key R&D Program of Hubei Province (2021BEA164, 2023BAB017, 2023BAB079), the National Natural Science Foundation of China (grant No.62072046), the Knowledge Innovation Program of Wuhan-Basic Research, and HUST CSE-HongXin Joint Institute for Cyber Security.

REFERENCES

[1] The bucket effect. https://www.drcarrierigoni.com.au/blog/the-bucket-effect, 2019.
[2] Official bitcoin portal. https://bitcoin.org/en/, 2019.
[3] Nfts for artists: how to sell digital art in 2021. https://www.theverge.com/22300888/nfts-explainer-what-is-blockchain-crypto-art-faq, 2021.
[4] Beeple homepage. https://www.beeple-crap.com/, 2022.
[5] Dappradar marketplace. https://dappradar.com/nft/marketplaces/protocol/ethereum, 2022.
[6] Ethplorer api. https://github.com/EverexIO/Ethplorer/wiki/Ethplorer-API, 2022.
[7] Frosties nft: 2 charged in us over $1.3m rug pull. https://coingeek.com/frosties-nft-2-charged-in-us-over-1-3m-rug-pull/, 2022.
[8] Hacker steals us$1mn worth of crypto and nfts in 24 hours. https://www.cshub.com/attacks/news/hacker-steals-us1mn-worth-of-crypto-and-nfts-24-hours, 2022.
[9] Looksrare - nft marketplace. https://looksrare.org/, 2022.
[10] Nft api quickstart. https://docs.alchemy.com/reference/nft-api-quickstart, 2022.
[11] Opensea: Buy crypto collectibles, cryptokitties, decentraland, and more on ethereum. https://opensea.io/, 2022.
[12] Rplanet withdraw event. https://etherscan.io/tx/0x6ec4d2eee19e96413397a07bea0701f687fa6a3e911dde85c7396a687e0a0840, 2022.
[13] X2y2 marketplace - x2y2.io. https://x2y2.io/, 2022.
[14] Y combinator. https://www.ycombinator.com/companies/opensea, 2022.
[15] Animoon co-founder twitter acount. https://twitter.com/marc_blata, 2023.
[16] Animoon discord server. https://discord.com/channels/920250158552350731/932657204551622787, 2023.
[17] Animoon founder twitter acount. https://twitter.com/MaximAdam_, 2023.
[18] Animoon official website. https://www.animoon.io/, 2023.
[19] Apes in space etherscan page. https://etherscan.io/address/0x7a3b97a7400e44dadd929431a3640e4fc47daebd, 2023.

[20] Bayc etherscan page. https://etherscan.io/address/0x0a5b9b930fc5be638232d8b9b69cb5b46249c06e, 2023.
[21] Blue chip nft bored ape yacht club (bayc) floor price declines to two year low. https://finance.yahoo.com/news/blue-chip-nft-bored-ape-024926657.html, 2023.
[22] Cat&mouse etherscan page. https://etherscan.io/address/0xe19e0cb95a9e39cb0ecde82e8a9f9f432835a0d0, 2023.
[23] Chainabuse official. https://www.chainabuse.com/, 2023.
[24] Chainalysis reactor certification: Live course. https://academy.chainalysis.com/live-crc, 2023.
[25] Chainbase official. https://chainbase.com/, 2023.
[26] Donald trump yacht club - collection | opensea. https://opensea.io/collection/donald-trump-yacht-club, 2023.
[27] Donald trump yacht club | twitter. https://twitter.com/DonaldTrumpYC, 2023.
[28] Donald trump yacht club etherscan page. https://etherscan.io/address/0x3b1fbe997c2253cffa975c066fa3feec326337dd, 2023.
[29] Donald trump yacht club withdraw event. https://etherscan.io/tx/0x8a9d43e5c0c6dbb3aeea98c8c6fe6f5e9491085789dec091c286a95cf66
2023.
[30] Erc-1155: Multi token standard. https://eips.ethereum.org/EIPS/eip-1155, 2023.
[31] Erc-20 token standard. https://ethereum.org/en/developers/docs/standards/tokens/erc-20/, 2023.
[32] Erc-721: Non-fungible token standard. https://eips.ethereum.org/EIPS/eip-721, 2023.
[33] Erc-721 non-fungible token standard. https://eips.ethereum.org/EIPS/eip-721, 2023.
[34] Ethereum improvement proposals. https://eips.ethereum.org/erc, 2023.
[35] Ethereum official. https://ethereum.org/en/, 2023.
[36] Etherreaper etherscan page. https://etherscan.io/address/0x0a5b9b930fc5be638232d8b9b69cb5b46249c06e, 2023.
[37] go-ethereum. https://geth.ethereum.org, 2023.
[38] Moonbirds nft creator loses $1 million almost instantly after wallet hack – here's what happened. https://cryptonews.com/news/moonbirds-nft-creator-loses-1-million-almost-instantly-after-wallet-hack-heres-what-happened.html, 2023.
[39] Non-fungible token - wikipedia. https://en.wikipedia.org/wiki/Non-fungible_token/, 2023.
[40] Opensea creator earnings. https://support.opensea.io/hc/en-us/articles/14068991090067-What-are-OpenSea-s-fees-, 2023.
[41] Opensea top collection. https://opensea.io/rankings/trending, 2023.
[42] R planet - genesis collection | opensea. https://opensea.io/collection/rplanet-genesis, 2023.
[43] Rarible. https://rarible.com, 2023.
[44] Remember when nfts sold for millions of dollars? 95collectibles may now be worthless. https://markets.businessinsider.com/news/currencies/nft-market-crypto-digital-assets-investors-messari-mainnet-currency-tokens-2023-9, 2023.
[45] Rplanet etherscan page. https://etherscan.io/address/0x656d34a8309363302e46de99853f4cef30b85a1d, 2023.
[46] Rplanet twitter. https://twitter.com/RplanetNFT, 2023.
[47] Rug pull finder, providing every human an outlet to safely navigate web3 without the threat of victimization. https://www.rugpullfinder.io/, 2023.
[48] Rug pull wikipedia. https://en.wiktionary.org/wiki/rug-pull, 2023.
[49] scikit-learn: Machine learning in python. https://scikit-learn.org/stable/, 2023.
[50] Study reveals nft market has collapsed with majority of tokens now worth less than jpegs. https://technext24.com/2023/09/22/nft-market-collapse-tokens-worth/, 2023.
[51] Twitter developer official. https://developer.twitter.com/, 2023.
[52] Uinits etherscan page. https://etherscan.io/address/0x7c10c8816575e8fdfb11463dd3811cc794a1d407, 2023.
[53] Uints | opensea. https://opensea.io/collection/uints, 2023.
[54] Uints twitter. https://twitter.com/uintsnft, 2023.
[55] Wen sandwich etherscan page. https://etherscan.io/address/0x29ae5ad9d3f50810d81a562c563c1ae9d799232d, 2023.
[56] Wen sandwich twitter. https://twitter.com/wensandwichxyz, 2023.
[57] What is a rug pull? https://www.bankrate.com/investing/what-is-a-rug-pull/, 2023.
[58] What is rug pull? https://www.bitdegree.org/crypto/learn/crypto-terms/what-is-rug-pull, 2023.
[59] Zachxbt twitter. https://twitter.com/zachxbt, 2023.
[60] Hong Bao and David Roubaud. Recent development in fintech: Non-fungible token, 2021.
[61] Hong Bao and David Roubaud. Non-fungible token: A systematic review and research agenda. *Journal of Risk and Financial Management*, 15(5):215, 2022.
[62] Sangam Bhujel and Yogachandran Rahulamathavan. A survey: Security, transparency, and scalability issues of nft's and its marketplaces. *Sensors*, 22(22):8833, 2022.
[63] Gianluca Bonifazi, Francesco Cauteruccio, Enrico Corradini, Michele Marchetti, Daniele Montella, Simone Scarponi, Domenico Ursino, and Luca Virgili. Performing wash trading on nfts: Is the game worth the candle? *Big Data and Cognitive Computing*, 7(1):38, 2023.

[64] Nicola Borri, Yukun Liu, and Aleh Tsyvinski. The economics of non-fungible tokens. *Available at SSRN*, 2022.
[65] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.
[66] Federico Cernera, Massimo La Morgia, Alessandro Mei, and Francesco Sassi. Token spammers, rug pulls, and sniperbots: An analysis of the ecosystem of tokens in ethereum and the binance smart chain (bnb). *arXiv preprint arXiv:2206.08202*, 2022.
[67] Yanto Chandra. Non-fungible token-enabled entrepreneurship: A conceptual framework. *Journal of Business Venturing Insights*, 18:e00323, 2022.
[68] Vladimir Cherkassky and Yunqian Ma. Practical selection of svm parameters and noise estimation for svm regression. *Neural networks*, 17(1):113–126, 2004.
[69] Usman W Chohan. Non-fungible tokens: Blockchains, scarcity, and value. *Critical Blockchain Research Initiative (CBRI) Working Papers*, 2021.
[70] Dipanjan Das, Priyanka Bose, Nicola Ruaro, Christopher Kruegel, and Giovanni Vigna. Understanding security issues in the nft ecosystem. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, pages 667–681, 2022.
[71] Michael Dowling. Fertile land: Pricing non-fungible tokens. *Finance Research Letters*, 44:102096, 2022.
[72] Yash Gupta and Jayanth Kumar. Identifying security risks in nft platforms. *arXiv preprint arXiv:2204.01487*, 2022.
[73] David G Kleinbaum. *Logistic regression*. Springer.
[74] Logan Kugler. Non-fungible tokens and the future of art. *Communications of the ACM*, 64(9):19–20, 2021.
[75] Massimo La Morgia, Alessandro Mei, Alberto Maria Mongardini, and Eugenio Nerio Nemmi. Nft wash trading in the ethereum blockchain. *arXiv preprint arXiv:2212.01225*, 2022.
[76] Bruno Mazorra, Victor Adan, and Vanesa Daza. Do not rug on me: Leveraging machine learning techniques for automated scam detection. *Mathematics*, 10(6):949, 2022.
[77] Sayak Saha Roy, Dipanjan Das, Priyanka Bose, Christopher Kruegel, Giovanni Vigna, and Shirin Nilizadeh. Demystifying nft promotion and phishing scams. *arXiv preprint arXiv:2301.09806*, 2023.
[78] Jason Scharfman and Jason Scharfman. Decentralized finance (defi) compliance and operations. *Cryptocurrency Compliance and Operations: Digital Assets, Blockchain and DeFi*, pages 171–186, 2022.
[79] Sven Serneels. Detecting wash trading for nonfungible tokens. *Finance Research Letters*, 52:103374, 2023.
[80] Syed Ahzam Tariq and Imtiaz Sifat. Suspicious trading in nonfungible tokens (nfts): Evidence from wash trading. *Available at SSRN 4097642*, 2022.
[81] Victor von Wachter, Johannes Rude Jensen, Ferdinand Regner, and Omri Ross. Nft wash trading: Quantifying suspicious behaviour in nft markets. *arXiv preprint arXiv:2202.03866*, 2022.
[82] Qin Wang, Rujia Li, Qi Wang, and Shiping Chen. Non-fungible token (nft): Overview, evaluation, opportunities and challenges. *arXiv preprint arXiv:2105.07447*, 2021.
[83] Xiaolin Wen, Yong Wang, Xuanwu Yue, Feida Zhu, and Min Zhu. Nftdisk: Visual detection of wash trading in nft markets. *arXiv preprint arXiv:2302.05863*, 2023.
[84] Bryan White, Aniket Mahanti, and Kalpdrum Passi. Characterizing the opensea nft marketplace. In *Companion Proceedings of the Web Conference 2022*, pages 488–496, 2022.
[85] Pengcheng Xia, Haoyu Wang, Bingyu Gao, Weihang Su, Zhou Yu, Xiapu Luo, Chao Zhang, Xusheng Xiao, and Guoai Xu. Trade or trick? detecting and characterizing scam tokens on uniswap decentralized exchange. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 5(3):1–26, 2021.

Appendix 1 FEATURES USED IN OUR MACHINE LEARNING PREDICTOR

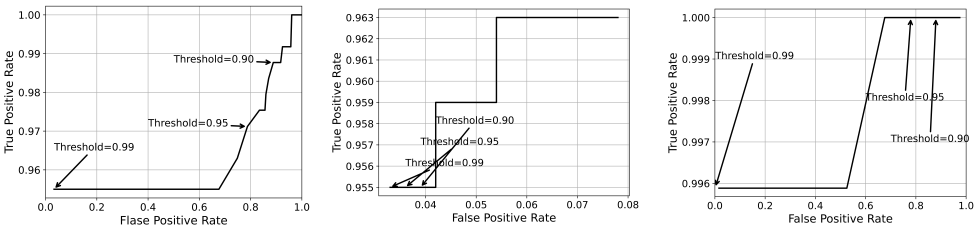
Table 3 shows the 52 kinds of features we extracted to train the predictor, including 8 kinds of *time-series features*, 14 kinds of *on-chain events features*, 30 kinds of *secondary market trades features*.

Table 3. Features are used in the NFT rug pull classifier (see §6). The feature that cannot be calculated (e.g., due to the lack of data, inevitably) will be set as -1.

	Feature	Description
Time-series	$P_{transfer}$	Average timepoint of each transfer
	P_{mint}	Average timepoint of each mint event
	P_{swap}	Average timepoint of each swap event
	P_{burn}	Average timepoint of each burn event
	P_{trade}	Average timepoint of each trade
	P_{top_price}	The timepoint of the trade with highest price
	P_{floor_price}	The timepoint of the trade with the lowest price
	$P_{highest\_24h\_trade}$	Average timepoint of the trade with the highest trade events in 24 hours
On-chain Events	$N_{Transfer}$	Total times of transfer event
	N_{mint}	The number of mint event
	N_{swap}	The number of swap event
	N_{burn}	The number of burn event
	$RN_{mint\_transfer}$	The ratio between the mint events and transfer events
	$RN_{swap\_transfer}$	The ratio between the swap events and transfer events
	$RN_{burn\_transfer}$	The ratio between the burn events and transfer events
	A_{all}	The number of addresses that have participated in transfer events
	A_{mint}	The number of addresses that have participated in mint events
	A_{swap}	The number of addresses that have participated in swap events
	A_{burn}	The number of addresses that have participated in burn events
	RA_{mint_all}	The ratio between the mint events and transfer events
	RA_{swap_all}	The ratio between the swap events and transfer events
	RA_{burn_all}	The ratio between the burn events and transfer events
Secondary Market Trades	N_{trade}	Total times of trade
	V_{volume}	The history volume of trade
	$V_{average\_price}$	The average price of trades
	$N_{beyond\_average}$	The number of trades whose price is beyond the average price
	$N_{below\_average}$	The number of trade whose price is below the average price
	$RN_{beyond\_average}$	The ratio between the "beyond average" trades and all trades
	$RN_{below\_average}$	The ratio between the "below average" trades and all trades
	V_{top_price}	The highest price of the trades.
	V_{floor_price}	The lowest price of the trades.
	U_{all}	Total users that have participated in trade
	U_{buyer}	Total users who are buyers
	U_{seller}	Total users who are sellers
	RU_{buyer_all}	The ratio between the buyers and all users
	RU_{seller_all}	The ratio between the sellers and all users
	$N_{highest\_24h\_trade}$	The number of trades of day that have the highest amount of trades events
	$RN_{highest\_24h\_trade}$	The ratio between $N_{highest\_24h\_trade}$ and all trade events
	$V_{highest\_24h\_volume}$	The history volume of the day that have highest trades events
	$RV_{highest\_24h\_volume}$	The ratio between $D_{highest\_24h\_volume}$ and total volume
	$V_{highest\_24h\_average\_price}$	The average price of the day that have highest trades events
	$RV_{highest\_24h\_average\_price}$	The ratio between $V_{highest\_24h\_average\_price}$ and average price
	$U_{highest\_24h\_user}$	The number of users of the day that have highest trades events
	$RU_{highest\_24h\_user}$	The ratio between $U_{highest\_24h\_user}$ and total number of users
	$N_{recent\_24h\_trade}$	The number of trade events of the recent day
	$RN_{recent\_24h\_trade}$	The ratio between $N_{recent\_24h\_trade}$ and all trade events
	$V_{recent\_24h\_volume}$	The history volume of the recent day
	$RV_{recent\_24h\_volume}$	The ratio between $V_{recent\_24h\_volume}$ and total volume
	$V_{recent\_24h\_average\_price}$	The average price of the recent day
	$RV_{recent\_24h\_average\_price}$	The ratio between $V_{recent\_24h\_average\_price}$ and average price
	$U_{recent\_24h\_user}$	The number of users of the recent day
	$RU_{recent\_24h\_user}$	The ratio between $U_{recent\_24h\_user}$ and total number of users

Appendix 2 ROC ANALYSIS OF OUR DETECTOR

Fig. 7 shows the ROC analysis of our detector for *drawdown rate*, *recovery rate*, *liveness rate*, respectively.



(a) ROC analysis for *drawdown rate*. (b) ROC analysis for *recovery rate*. (c) ROC analysis for *liveness rate*.

Fig. 7. ROC analysis for drawdown rate, recovery rate and liveness rate.

Appendix 3 FUNDS DRAINED BY RUG PULLS

Here are a set of lists in which the projects drain their funds using various methods. Specifically, Table 4, Table 5, and Table 6 present the top-5 rug pull projects that drain their funds through mint fee withdrawal, bonus creator fee, and middleman reselling, respectively.

Table 4. Top 5 profitable NFT projects rug pulled via mint fee.

Project	Contract Address	# Ether (USD)
Apes In Space	0x7a3b97a7400e44dadd929431a3640e4fc47daebd	2,645.82 (10.93M)
Fat Ape Club	0xf3114dd5c5b50a573e66596563d15a630ed359b4	2,573.00 (11.70M)
Bored Bunny	0x9372b371196751dd2f603729ae8d8014bbeb07f6	1,991.00 (7.08M)
MURI	0x4b61413d4392c806e6d0ff5ee91e6073c21d6430	1,872.60 (4.92M)
HULLYUniverse	0xb8b6cb37c0968f72c6d37dc3074c80ad73521024	1398.90 (3.92M)

Table 5. Top 5 NFT rug pull projects that earn most via creator fee.

Project	Contract Address	Creator Fee (USD)
MURI	0x4b61413d4392c806e6d0ff5ee91e6073c21d6430	1.486M
SkuxxVerse Pass	0x19350eb381ab2f88d274e740bd062ab5ff15542e	1.198M
hausphases	0x5be99338289909d6dbbc57bb791140ef85ccbcab	0.958M
Beyond Earth Land	0x28c6ea3f9cf9bc1a07a828fce1e7783261691b49	0.797M
Moonbirds2	0xdb7b094fdc04f51560a03a99f747044951b73727	0.490M

Table 6. Top 5 NFT projects gain the largest profit via middleman reselling.

Project	Middleman Address	# Mint Tokens	Ether (USD)
CypherHumans	0x6e40ea6202d5bc2ace21bc904c9c772c484320a1	696	103.92 (188K)
Muttniks	0xa00f56b263d3c3e016c33d9b31791b625d90ae3b	1080	64.20 (116K)
DigDragonz	0x0ce353f8bca317024e4ae6b87a0e14ca0377f476	944	46.35 (83K)
DigDragonzReborn	0xcd6d7e5a31cb3cf43734398e6506d5422072a172	1507	27.45 (50K)
DRM1	0x171ab540b9cb730626db91f648e2b09eb5363484	101	19.99 (36K)

Received August 2023; revised October 2023; accepted October 2023