Under the Dark: A Systematical Study of Stealthy Mining Pools (Ab)use in the Wild

Zhenrui Zhang[†]
Tsinghua University; QI-ANXIN Technology Research Institute
Beijing, China
zhang-zr21@mails.tsinghua.edu.cn

Geng Hong[†]
Fudan University
Shanghai, China
ghong@fudan.edu.cn

Xiang Li
Tsinghua University
Beijing, China
x-l19@mails.tsinghua.edu.cn

Zhuoqun Fu
Tsinghua University
Beijing, China
fzq20@mails.tsinghua.edu.cn

Jia Zhang[‡]
Tsinghua University; Zhongguancun Laboratory
Beijing, China
zhangjia2017@tsinghua.edu.cn

Mingxuan Liu
Tsinghua University; Zhongguancun Laboratory
Beijing, China
liumx18@mails.tsinghua.edu.cn

Chuhan Wang
Tsinghua University
Beijing, China
wch22@mails.tsinghua.edu.cn

Jianjun Chen
Tsinghua University; Zhongguancun Laboratory
Beijing, China
jianjun@tsinghua.edu.cn

Baojun Liu
Tsinghua University; Zhongguancun Laboratory
Beijing, China
lbj@tsinghua.edu.cn

Haixin Duan
Tsinghua University; Quancheng Laboratory
Beijing, China
duanhx@tsinghua.edu.cn

Chao Zhang
Tsinghua University
Beijing, China
chaoz@tsinghua.edu.cn

Min Yang
Fudan University
Shanghai, China
m_yang@fudan.edu.cn

ABSTRACT

Cryptocurrency mining is a crucial operation in blockchains, and miners often join mining pools to increase their chances of earning rewards. However, the energy-intensive nature of PoW cryptocurrency mining has led to its ban in New York State of the United States, China, and India. As a result, mining pools, serving as a central hub for mining activities, have become prime targets for regulatory enforcement. Furthermore, cryptojacking malware refers to self-owned stealthy mining pools to evade detection techniques and conceal profit wallet addresses. However, no systematic research has been conducted to analyze it, largely due to a lack of full understanding of the protocol implementation, usage, and port distribution of the stealth mining pool.

To the best of our knowledge, we carry out the first large-scale and longitudinal measurement research of stealthy mining pools to fill this gap. We report 7,629 stealthy mining pools among 59 countries. Further, we study the inner mechanisms of stealthy mining pools. By examining the 19,601 stealthy mining pool domains and IPs, our analysis reveals that stealthy mining pools carefully craft their domain semantics, protocol support, and lifespan to provide underground, user-friendly, and robust mining services. What's worse, we uncover a strong correlation between stealthy mining pools and malware, with 23.3% of them being labeled as malicious. Besides, we evaluate the tricks used to evade state-of-the-art mining detection, including migrating domain name resolution methods, leveraging the botnet, and enabling TLS encryption. Finally, we conduct a qualitative study to evaluate the profit gains of malicious cryptomining activities through the stealthy pool from an insider perspective. Our results show that criminals have the potential to earn more than *1 million* USD per year, boasting an average ROI of 2,750%. We have informed the relevant ISPs about uncovered stealthy mining pools and have received their acknowledgments.

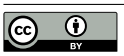
CCS CONCEPTS

• **Security and privacy → Malware and its mitigation**; **Network security**.

KEYWORDS

Cryptocurrency Mining; Cryptojacking; Malware; Botnet

ACM Reference Format:
Zhenrui Zhang[2], Geng Hong[2], Xiang Li, Zhuoqun Fu, Jia Zhang[3], Mingxuan Liu, Chuhan Wang, Jianjun Chen, Baojun Liu, Haixin Duan, Chao Zhang, and Min Yang. 2023. Under the Dark: A Systematical Study of Stealthy Mining Pools (Ab)use in the Wild . In *Proceedings of the 2023 ACM*



CCS '23, November 26–30, 2023, Copenhagen, Denmark

ACM ISBN 979-8-4007-0050-7/23/11.
https://doi.org/10.1145/3576915.3616677

[†]Both authors contributed equally to this research.
[‡]Corresponding author.

SIGSAC Conference on Computer and Communications Security (CCS '23), November 26–30, 2023, Copenhagen, Denmark. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3576915.3616677

1 INTRODUCTION

Cryptocurrency mining (cryptomining) is a crucial operation in blockchains that employ the Proof of Work (PoW) as the consensus algorithm. Miners participate in this process by solving complex hash-based puzzles and are rewarded with cryptocurrency. In order to increase their chances of winning rewards, miners frequently join mining pools by connecting their mining hardware (CPU, GPU, or ASIC) to online mining pool servers. However, the energy-intensive nature of PoW-based cryptocurrency mining has resulted in its prohibition in several regions and countries, such as New York State of the United States [30], China [13], and India [16]. As a result, mining pools, serving as a central hub for mining activities, have emerged as the primary targets for regulatory enforcement.

Furthermore, recently, criminals started abusing victims' resources to mine cryptocurrency by infiltrating victim hosts and deploying cryptomining malware, which is known as *Cryptojacking*. Cryptojacking malware soared nearly fourfold in Q3 2022 and is considered one of the most serious cybersecurity threats, according to a public report [34]. To evade detection methods, such as denylists [41], employed by security vendors, and to conceal profit wallet addresses, cryptomining malware has started to utilize self-owned stealthy mining pools [25, 31]. Previous work [52] observed that botnet malware mines cryptocurrency through underground mining infrastructures rather than public mining pools [21, 23, 29].

Stealthy mining pools, in contrast to public mining pools, are not intended of offering public services. Studying stealthy mining pools is challenging for several reasons. First, the Stratum protocol, which is the *de facto* mining protocol, lacks standardized implementation specifications. Second, not all communications based on such protocols is for mining cryptocurrency; they can also be used for other services, such as Electrum Bitcoin Wallet [3]. Third, there are no designated ports for mining pool services, making it difficult to perform large-scale scanning without prior knowledge of targeted ports. As a result, systematic research on stealthy mining pools has yet to be conducted.

In this paper, we perform a large-scale and longitudinal measurement study on the current status of stealthy mining pools by both passive analysis and active scanning, which is the first study on stealthy mining pools, to the best of our knowledge. To address the aforementioned challenges, we first collected the three most popular implementations of mining protocols from the documentation [2, 6, 11, 32], academic research [46, 48, 53, 71], and real-world mining samples [4, 8, 9, 14, 24, 37]. Then we propose a mining service discovery technique by network probing and a semantic approach to recognize the stealthy mining service. We discover the stealthy mining pool with a two-step method: a preliminary experiment for collecting candidate services' mining ports, followed by an active scan aimed at the entire IPv4 address range to achieve a comprehensive result. Finally, we find *7,629* stealthy mining pools, spanning *2,113* IPs and *17,488* domains among *59* countries.

Further, we study the inner mechanisms of stealthy mining pools. By examining the 19,601 stealthy mining pool domains and IPs, our analysis reveals that the stealthy mining pools *carefully craft their domain semantics, protocol support, and lifespan* to provide underground, user-friendly, and robust mining services. Stealthy pools tend to hide their identities in the form of domain names, which makes them less noticeable. To provide an easy-to-use mining service, around 10% of stealthy pools support requests for all three implementations on a port. What's more, 7.5% of pools are able to interact with both TLS-encrypted and non-TLS-encrypted mining requests. Besides, stealthy mining pools tend to have a shorter lifespan, with 33% of the stealthy mining pools having a lifecycle of less than one day, to avoid attracting unnecessary adversary notice, e.g., firewalls or antivirus engines.

While investigating the malicious activities related to stealthy mining pools, we find that they have a strong correlation with malware, including 23.3% of IPs and 3.3% of domains labeled as malicious. We also conduct a campaign analysis and find out 439 different campaigns, and some of them have been asserted to be the mining pools of known cryptomining botnets, which means stealthy mining pools have been popular in malware. We uncover and assess the tricks employed to evade state-of-the-art mining detection, including migrating domain name resolution methods, leveraging the botnet, and enabling TLS encryption. Our findings indicate that the third trick is the most effective evasion technique; only 9.6% of stealthy mining pools employing it being labeled by VirusTotal. Additionally, we conduct a qualitative study to evaluate the profit gains of malicious cryptomining activities through stealthy pools from an insider's perspective. Our results show that criminals have the potential to earn more than *1 million* USD per year, boasting an average ROI of 2,750%.

Contributions. We summarize the contributions as follows:

- We propose the first discovery method for stealthy mining pools and conduct a large-scale and longitudinal measurement study on the entire IPv4 range, locating 7,629 different stealthy mining pools, involving 2,113 IPs and 17,488 domains.
- We discover the unique characteristics of stealthy mining pools, revealing that stealthy mining pools carefully craft their domain semantics, protocol support, and lifespan to provide underground, user-friendly, and robust mining services.
- We uncover stealthy mining pools that collaborate with malware, analyze their campaign and survival strategies, and evaluate their profit gains from an insider's view.

2 BACKGROUND

In this section, we first focus on the process of cryptocurrency mining. Then we discuss the classification of different mining pools and introduce what is a stealthy mining pool.

2.1 Cryptocurrency Mining

Cryptocurrency mining (abbreviated as "mining" in the following) is the process of verifying transactions on a blockchain and adding them to the blockchain ledger by miners. When a new cryptocurrency block is generated or a transaction is performed, miners need to validate it and then add it to the blockchain. To achieve this work, miners must compete with each other to solve complex cryptography problems, i.e., computing the input nonce that matches a given target hash for a cryptocurrency block as Proof-of-Work (PoW).

In return for their computing efforts, miners who first finish are rewarded with a certain amount of cryptocurrency.

Mining pool. In recent years, mining has become more difficult, especially for individual miners. This is due to a number of factors, such as the growing competition among miners and the rising cost of computing resources for mining. In response to these challenges, mining pools have emerged as a new way to combine the computing resources of a group of miners. When miners join a mining pool, they simply connect their machines (mining hardware) to an online mining pool server (mining software) to share their computing power with other miners, thereby increasing their chances of finding a new block. After completing the mining task, each miner in the mining pool earns its share of the reward, depending on their contribution to the pool's computing power and rules. The mining communication between miners and mining pools used to be the HTTP-based getwork protocol, which has been replaced by the Stratum mining protocol now [65].

Stratum mining protocol. The Stratum protocol, which is a JSON-RPC-based plaintext TCP protocol, is the most common protocol used to communicate between miners and mining pools. It is originally created for Electrum Bitcoin wallet [3] to synchronize information about blocks, transfers, etc. in the Bitcoin blockchain. Although it has specialized implementations for various cryptocurrencies, the communication among them generally follows the procedures below. Figure 1 illustrates this communication process between miners and the mining pool.

• A miner sends a subscription message to the mining pool to verify its identity in case of applying for a mining job (step ①).

• After the mining pool verifies the identification, the miner enters the pool and prepares to act as mining "hardware" (step ②).

• The mining pool next generates a mining difficulty and assigns the miner a mining work (step ③).

• The miner begins the mining operation by calculating the hash using local resources and hardware (step ④). Once a result has been obtained, the miner submits it to the mining pool and awaits confirmation (step ⑤).

• The mining pool validates the result and responds with a message of success or failure (step ⑥). The mining process then repeats round by round.

Specifically, the Stratum v2 protocol is a next-generation implementation of the Stratum protocol that is presently available for testing by its developer, Braiins Pool [33], but is not yet widely used. Therefore, we will not include this protocol in this study. Besides the plaintext Stratum protocol over TCP, we observe that part of public mining pools [23, 29] have begun to provide services utilizing the Stratum encrypted by the TLS protocol.

2.2 Stealthy Mining Pool

Figure 2 shows three different types of mining pools that all provide services using Stratum protocol. The public mining pool announces itself with a web page under the same domain as the mining pool service, which lists the supported cryptocurrencies and related pool service ports. As a result, the domain names and IP addresses of these mining pools are easily accessible, making them susceptible to being blocked by a denylist.

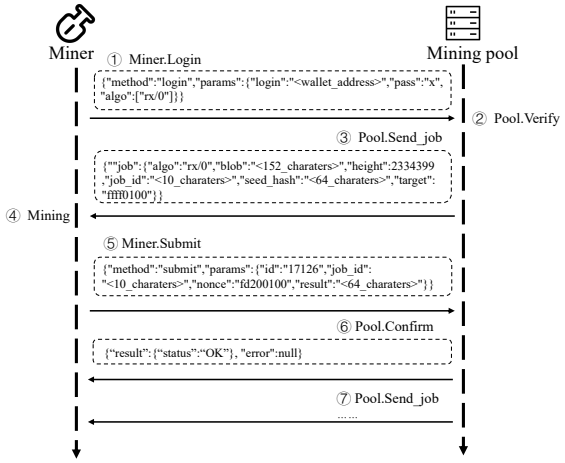


Figure 1: An example of a miner using the Stratum protocol to communicate with a mining pool.

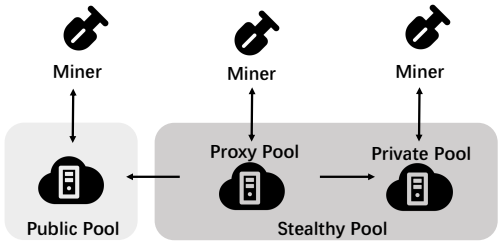


Figure 2: Different ways in which miner connects to mining pools: (a) to a public pool; (b) to a stealthy pool, including proxy pool and private pool.

Stealthy mining pool. In contrast to public mining pools, *stealthy mining pools* do not provide public services, which are classified into two types: proxy and private ones. The private mining pool, like the public one, is synchronized with the common blockchain to obtain the most recent block information, but its service addresses are not disclosed. According to publicly available statistics, it still accounts for a minor portion of pool hash rates[28]. Whereas the proxy mining pool is a kind of mining pool that behaves like a broker service between miners and upstream pools. It receives jobs from the upstream pool, decomposes them into multiple sub-tasks with lower hash rates, and distributes them to downstream miners. Then, it accepts miners' shares and aggregates them before sending them back to the upstream pool. In addition, it does not need the synchronization of block information, which makes its implementation relatively simpler. Such proxy services are commonly provided by tools such as XMRIG-PROXY, which may reduce the number of connections to the pool by up to 256 times [39]. According to public reports, attackers make extensive use of proxy pools. They are either deployed in an independent pool server [5, 25] or hosted on the same domain as the C&C server [12, 40].

In the case of illicit cryptomining, employing a stealthy mining pool has the following two advantages: First, compared to public pools, which have publicly disclosed domain names and IP addresses, stealthy pools are less likely to be detected by denylist-based approaches. Since the domain names and IP addresses of

Table 1: Sources of three Stratum protocol implementations.

Name	Doc.	Prior work	Miner
Stratum-BTC	[11, 32]	[53]	[4, 9, 14, 24]
Stratum-ETH	[2]	[48]	[9, 24]
Stratum-XMR	[6]	[46, 48, 71]	[8, 37]

stealthy mining pools are not publicly disclosed, it is difficult for regulators to actively block malicious samples before their detection. Second, the wallet address of the attacker can be disguised. Attacker's wallet address need to be encoded in the malicious sample when using public mining pools, therefore public pools can easily ban the wallet when they are reported by researchers [62]. However, with stealthy mining pools, the wallet address can be configured and updated on the server, making it impossible to obtain the attacker's wallet information from samples and/or network traffic analysis (NTA) [55], consequently increasing the difficulty of analyzing revenue and blocking cryptomining malware.

3 IDENTIFYING STEALTHY MINING POOLS

This section discusses the methodology used to identify stealthy mining pools in the wild. As illustrated in Figure 3, we first collect prominent mining protocol implementation variants, then present a mining service identification technique for detecting the mining service by network probing and a semantic approach to distinguish the stealthy mining service. Finally, we put the aforesaid strategies to the test in two steps: a preliminary experiment for gathering mining ports for candidate services, followed by an active scanning aimed at the whole IPv4 address range to produce a thorough result.

3.1 Study of Mining Protocol in the Field

The Stratum protocol, as indicated in Section 2, is the most commonly used protocol for communicating between miners and mining pools[62]. As a result, we can send a probing packet using the Stratum protocol to identify prospective mining pools. However, because the original stratum protocol[32] did not specify implementation specifics for multiple cryptocurrencies, its implementation varies today among miners and pools. To the best of our knowledge, no research has been conducted on the implementation of mining protocols, particularly the Stratum protocol. To provide a better understanding, we collected and analyzed the three most commonly used Stratum protocol implementations and summarized their sources in Table 1. We refer to them as $Stratum-BTC$, $Stratum-ETH$, and $Stratum-XMR$ which names derived from the names of the greatest market capitalization cryptocurrencies supported for mining, respectively.

Collecting and evaluating the Stratum protocol and its implementations is straightforward but non-trivial. To reach a more comprehensive result, we consult the documentation, previous academic researches, and real-world mining samples to extract the variant patterns of mining protocol. It takes two security researchers five days to obtain a full result. We start by looking for publicly available documentations of mining protocol from websites of mining pools, and README files about Stratum implementation from some open source miners, including [2, 6, 11, 32]. We then refer to related work in the field of network traffic-based cryptomining detection. We obtain the dataset from the study (e.g., [48, 71]), attempt to extract the TCP payloads that adhere to the JSONRPC format from the traffic, and manually review the protocol implementations. In order to reconstruct the original format for works (such as [46, 53]) that do not open-source a public dataset, we use the Stratum communication details specified in the study. Additionally, to investigate the traffic patterns of the real-world mining samples, we collect and install several popular miners [4, 8, 9, 14, 24, 37], initiating mining requests with the default mining pools according to their configurations, and collected traffic for analyzing the protocol types.

In the end, we summarize and focus on three implementations: $Stratum-BTC$, $Stratum-ETH$, and $Stratum-XMR$. $Stratum-BTC$ is the first implementation of Stratum, which is utilized by mining pools of Bitcoin, Litecoin, Ethereum, Zcash, and others. $Stratum-ETH$ is mainly used by Ethereum pools, and most Ethereum miners and pools support both $Stratum-BTC$ and $Stratum-ETH$. $Stratum-XMR$ is designed for Monero cryptocurrency mining, which is exclusively deployed by Monero pools and the only protocol used between Monero miners and pools. Details of the three implementations can be found in Table 2.

3.2 Methodology of Mining Pool Discovery

To discover potential mining pools in the wild, we leverage an active probing method based on the Stratum protocol implementations we summarized in Section 3.1, including the *Request Construction* and *Response Analysis* phrases that sends Stratum probing packets and analyzes responses, respectively.

Request construction. As described in Section 2, prior to starting the mining process, miners must submit a handshake JSONRPC request to prove the identification, such as subscription or login. Therefore, for each type of implementation collected from section 3.1, we construct the handshake JSONRPC request packet to determine whether a server is an active mining pool. In addition, the Stratum handshake happens following the establishment of a TCP or TLS connection, thus we also encrypt the handshake packets with TLS. Table 2 shows request and response examples of the three Stratum protocol implementations we summarized and observed in the wild. Specifically, for $Stratum-BTC$, miners should use the JSONRPC method *mining.subscribe* to subscribe to the pool server before any other connections. The initial login request for $Stratum-ETH$ is method *eth_submitLogin*, where the miner registers its ETH wallet address through *params*. Miners submit a *login* request to the $Stratum-XMR$ pool with a Monero wallet address in *params["login"]*.

Response analysis. We divide responses into three types depending on the target server. First, we discard a server directly if there is no response or the content is not in the JSON format. Otherwise, we store the responses for further analysis. Second, if the target server is a mining pool server, it will return two forms of responses after handshaking: a *success* response or an *error* response. We list examples of success and error responses in Table 2. A more detailed list of signatures can be found in Appendix E. Generally, according to our observation, the success and error responses embedded semantics. In the case of a success response, the mining pool negotiates the mining difficulty, mining algorithms, current block height,

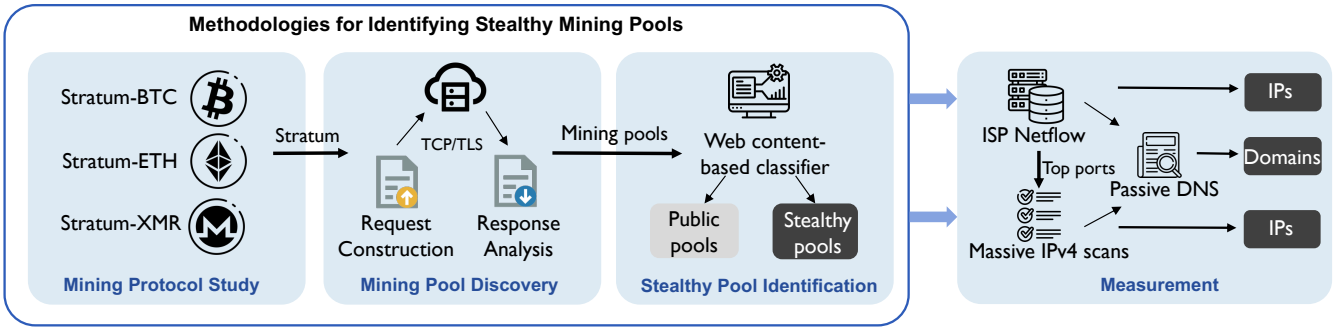


Figure 3: Methodology overview of identifying stealthy mining pools in the wild.

Table 2: Examples of Stratum implementations' requests and responses.

Protocol	Handshake request	Success response	Error response
Stratum-BTC	{"id": 1,"method": "mining.subscribe","params": []}	{"id": 1, "result": [[["mining.set_difficulty", "<Difficulty>"],["mining.notify", "<Subscribe_Addr>"]], "<ExtraNonce1>", 4], "error": null}	{"id":1,"result":false, "error":[20,"Not supported",null]}
Stratum-ETH	{"id": 1,"jsonrpc": "2.0","method": "eth_submitLogin", "params":[<Wallet_Addr>]}	{ "id": 1, "jsonrpc": "2.0", "result": true }	{ "id": 1,"jsonrpc": "2.0", "result": null,"error": { code: -1, message: "Invalid login"}}
Stratum-XMR	{"id":1,"jsonrpc":"2.0", "method":"login","params":{ "login":"<Wallet_Addr>","pass":"x",}}	{"id":1,"jsonrpc":"2.0","result":{"id":"<ID>" ,"job":{"algo":"rx/0","blob":"<Blob>", "height":<Blockchain_Height>,"job_id":"<Job_Id>", "seed_hash":"<Seed_Hash>","target": "Target_Difficulty"},"status":"OK"},"error":null}	{ "id": 1, "jsonrpc": "2.0","error": { "code": -1, "message": "Invalid address" } }

and seed hash with the mining client. This information is then included in the response packets, which can be used as signatures to determine if the server is an active mining pool." In contrast, if the client receives an error response, the server will notify the client of the error type, such as *"Not supported"*, *"Invalid login"*, *"Invalid address"*, etc. For example, if we submit a $Stratum-BTC$ request to a $Stratum-ETH$ server, we will receive an error response with the message *"Invalid login"*. If we send our $Stratum-XMR$ request with a Monero wallet to an Ergo pool that requires an Ergo wallet address, the response will state *"Invalid address"*. Third, certain servers may return JSON payloads in response, but as these payloads do not adhere to any of the success or error semantics, we classify these servers as non-mining services and eliminate them from our study targets. Responses of non-mining servers can be found in Appendix A.

To discover potential mining pools precisely, we propose an error message spreading-based method by analyzing the responses as follows.

Discovery steps. The key observation is that mining pools have a similar implementation following the protocol specification in Table 1. Even if the wallets used by different mining pools differ, the error messages remain the same. Therefore, we design and implement Algorithm 1 to discover active mining pools by following the procedures below.

• First, we initiate three sets: *(i)* an empty set M , which will contain all mining pools; *(ii)* a set N including all the candidate servers with JSON responses; *(iii)* a signature set S that contains all the keys of key-value pairs from the success responses of each type of implementation as listed in Table 2.

• Second, we begin the first detection cycle using a signature-based detector whose input is a JSON-formatted probing response from N and output indicates whether the server is a mining service. If the keys of the input response message match the signature set of one of the implementations from S, we place the server address into M and label it as the corresponding type of implementation, then delete it from the set N.

• Third, for each pool in the mining pool set M, we collect their error codes and error messages as another set with key-value signatures. The second detection round starts by extracting all error codes and messages sent by the remaining servers in set N. If a server's error code and error message match the signature set, the server is moved from the set N to the set M.

• Finally, servers in the set M of mining pools are discovered as candidate mining pools that will be processed in Section 3.3.

3.3 Stealthy Mining Pool Identification

With the new methodology from Section 3.2, we are able to discover active mining pools. However, according to [52], not only stealthy but also public mining services are in the candidate list. The insight to distinguish the public mining pools from the stealthy ones is that the public mining pools usually promote their services from their websites. Specifically, public mining pools often announce

Algorithm 1 Discover mining pools.

Input: N, S ← Set of servers with *JSON* responses, Set of *success signatures*
Output: M ← Set of mining pools

for each candidate C in N **do**
  $flag$ ← True
  for each key k in S **do**
    if $k \notin C.Response$ **then**
      $flag$ ← False
    end if
  end for
  if $flag$ = True **then**
    Move C from N to M
  end if
end for
for each pool P in M **do**
  $E(P)$ ← Set of *error signatures* of P
end for
for each candidate C in N **do**
  $E(C)$ ← Set of *error signatures* of C
  for each pool P in M **do**
    if $E(C) = E(P)$ **then**
      Move C from N to M
    end if
  end for
end for

their service publicly with a web page containing mining-related keywords [21, 23, 29]. As a result, we build and implement a web content-based classifier to automatically distinguish between public and stealthy pools.

Ground truth collection. We collect a ground-truth dataset that includes both public mining pools and non-mining pools. The public mining pool list is collected from a mining pool statistics website [27], which contains 124 website URLs, while the non-mining pool list is compiled from the top 500 Tranco [64] website list.

Web content crawling. For the URLs from both ground-truth dataset and prediction dataset we will collect by massive scans, dynamic web-content crawling is accomplished by instructing a headless browser via the Selenium framework [43]. The HTML file of each website's homepage is crawled and saved. Then we extract all the texts in the HTML DOM contents including the title, keywords, description, and texts in the body.

Feature extraction and training. In the collected text from our crawlers, we employ NLP methods including tokenization, stop word removal, and word frequency counting for preprocessing. Next, we utilize tf-idf [66] to determine the most significant mining-related words. As shown in Table 3, 11 keywords are used as features for our classifier. We choose four different classification algorithms including SVM, KNN, GNB and RF. Results show SVM outperforms all other classifiers by using 5-fold Cross Validation (CV) for evaluation. Our model has a recall of 98.4%, a precision of 99.2%, and an F-1 score of 98.8%.

Table 3: Mining-related keywords used as features.

Keyword	Average tf-idf in mining pools	Average tf-idf in non-mining pools
pool	0.35	0
mining	0.31	0
miner	0.26	0
algorithm	0.21	0
hash	0.15	0
payout	0.14	0
hashrate	0.08	0
coin	0.05	0
payouts	0.05	0
price	0.04	0.01
block	0.03	0

3.4 Measurement and Result

In this part, we discuss how we measured stealthy mining pools using the aforementioned methodologies and the corresponding findings. First, we probe the IPs and ports from the Netflow data for candidate mining pools and determine the most commonly used mining pool ports. Then, we conduct an entire IPv4 network space scanning to retrieve a comprehensive measurement result. Lastly, we show our results about 2,113 and 17,488 identified stealthy mining pool IPs and domains.

Mining port discovery. To probe potential mining pools, we need the mining port to send constructed requests. However, the Stratum protocol lacks standard ports, making it challenging to perform a scan of all IPv4 addresses without prior knowledge [54]. To figure out the port distribution of mining pools, we utilize ISP Netflow data [1] from one of our partner providers.

The Netflow is collected at the border routers of ISP network and used as a traffic monitor. Our Netflow data set is collected at a 1:1000 sampling ratio from April 2022 to October 2022, and contains an average of *21.4* million unique (IP, port) tuples per day. We extract every unique (IP, port) pair as the probing target. For each target, we probe them with six packets, including three variations of the Stratum protocol, with or without TLS encryption. The probing experiment runs for around 6 months, from May 24, 2022 to October 31, 2022, and it takes us 10 hours to send about 128 million handshaking packets per day. Over 100 million unique (IP, port) pairs are probed during the experiment time.

Our probing result shows Stratum supports 425 distinct ports in total. To conduct the following massive scans while taking into account resource consumption and ethical concerns, we then identify the most popular mining ports by analyzing their active duration in ISP Netflow. Specifically, we calculate the number of active days for each (IP, port) tuple by subtracting the earliest and last occurrence dates in Netflow. Following that, for each port, we add the number of active dates from different IP addresses to get the total active days. Table 4 shows the top 10 ports for stealthy mining pools that are frequently observed in our Netflow dataset. The ports used by the stealthy mining pool exhibit a long-tailed distribution, as depicted in Figure 4. The top 32 ports account for more than 80% of cryptocurrency mining activities, which used for further large-scale scanning.

Table 4: Top 10 popular ports for stealthy mining pools.

Rank	Port Number	# Total active days	Percentage
1	443	7,120	19.2
2	80	5,632	15.2
3	5555	2,908	7.9
4	3333	2,408	6.5
5	8080	1,110	3.0
6	6688	855	2.3
7	13782	802	2.2
8	8888	672	1.8
9	14333	668	1.8
10	1800	539	1.5

Massive scans. To extend our mining pool IPs, apart from the potential mining pool IP obtained from the Netflow data, we use *Masscan* [17] to scan the whole IPv4 network space for the top 80% of most commonly identified mining ports within Netflow data to collect candidate server IPs. Then, we probe these candidate IPs with the Stratum handshake request to identify mining pools (see Section 3.2). Our massive scan lasts from Nov, 03 to Nov, 26 2022 with 8 server, each scanning at a rate of 100,00 packets per second. We exclude private networks, reserved networks and networks that do not allow to be scanned using the list [18] provided by *Masscan*.
Stealthy pool identification. We discovered 2,617 distinct mining pool IP addresses from either Netflow probing or massive scans. To select stealthy mining pools from public mining pools, we create a prediction dataset containing 5,600 URLs by these IPs as the input to the classifier in Section 3.3. Specifically, we first reverse the mining pool IPs to domains utilizing a large-scale passive DNS (PDNS) dataset from a public DNS resolver. The PDNS dataset contains all history records of domain resolutions collected by this resolver from April 2018 to November 2022. Then we extract their eTLD+1 [42] domains with and without the "www" prefix as the target web page URLs. Finally 5,600 distinct eTLD+1 domains are extracted from the 19,434 domains collected from the PDNS dataset.

By following the web content crawling procedure in Section 3.3, 654 out of the 5,600 URLs respond with a valid HTML file. Our classifier predicts 112 of them to be public mining pools, while the remaining are not public mining services. We further manually examined the positive ones and confirmed that they all are mining pools. To assess false negative cases, we validate the negative results with ground truth data, results show that only one public mining pool *dook.xyz* is overlooked because its homepage lacks any available text. The outcome demonstrates that our classifier is capable of distinguishing between public and stealthy mining services.

In particular, if an URL is identified as a public mining pool, we will regard the IP addresses resolved by its FQDN domains as public pool IPs. Besides, all the domains that resolve to public pool IPs are labeled as public pool domains. Therefore, When multiple domains are hosted on the same IP, if the IP is labeled as public, then all the hosted domains are labels as public too.
Result. Table 5 summarizes our key results. By scanning over 100 million (IP, port) pairs in Netflow and 32 ports in IPv4 address space, we found a total of 7,629 stealthy mining pool services, including 2,113 IPs and 17,488 domains.

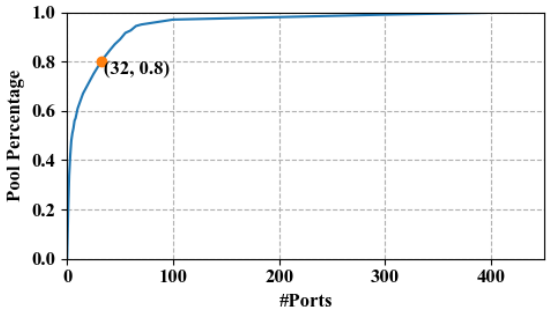


Figure 4: The CDF of top ports supporting mining protocols.

Table 5: Summary of large-scale and longitudinal scanning results.

Source	# Pool Services	# IPs	# Domains	Period (2022)
Netflow	5,534	1,221	13,317	05/24 - 10/31
IPv4	6,467	1,820	5,218	11/03 - 11/26
Total	**7,629**	**2,113**	**17,488**	-

3.5 Ethical Considerations

In our experiments, we take great ethical considerations into the Netflow and PDNS dataset analysis and active network scanning.
Passive datasets. The ISP Netflow data contains only IP and TCP packet header information, without payloads. Our experiments are done under the ISP operator's supervision, and when the daily Netflow data is uploaded to the private server, we only obtain the IP address and port of each flow, represented as a four tuple. All the stored Netflow data will be deleted upon completion of the scan.

There are no privacy data evaluated in the PDNS dataset. Regarding our strategy, we only use this dataset to get history domain resolution records of a given mining pool IP address. Any sensitive data, like client IP address or DNS query time, is not accessible. Besides, the PDNS dataset is stored in our partner's server, and we only get the query API for obtaining IP history resolutions.
Active Scanning. We performed active scans in both Netflow data and the IPv4 address space. We run the scanning application on a dedicated server, and we've taken several measures to minimize the harm to the network. First, we sent legitimate packets at no more than half the machine's bandwidth to ensure no impact on the local network. Second, only one probe packet is received per target port at a time, thus the load on the target machine is very low. Third, we made it clear in the probe packet and a website (with the same scanning source IP) about our research intentions. Throughout the experiment, we did not receive any complaints from any organizations or individuals.

4 CHARACTERISTICS OF STEALTHY MINING POOLS

This section dives into the inner workings of stealthy mining pools, analyzing their domain semantics, protocol support, and lifespan in detail. Our investigation of the 19,601 stealthy mining pool domains

and IPs indicates how they function as underground, user-friendly, and robust mining services.

Landscape By checking the ISP information of IP address, we find that our discovered stealthy mining pools distribute across 59 countries. Specifically, they are present with a long-tail distribution such that the top five countries – the United States (30.15%), China (22.39%), Germany (11.31%), Singapore (6.29%), and France (3.36%) – account for more than 70% of all stealthy pools. One notable finding is that the vast majority (94.7%) of mining pools in China are located in Hong Kong. As a result of the Chinese government's strict regulations on the mining industry, operating mining servers in mainland China becomes increasingly dangerous. Instead, by relocating the mining service to Hong Kong, the operators can avoid regulatory problems while still providing rapid access to their underground users.

4.1 Domain Semantics of Stealthy Pools

The domain name of a service often reflects its purpose, and it has been reported that mining blockers may determine mining behaviors based on domain semantic information [26]. This raises the question of how the domain names of stealthy mining pools are constructed, and whether they reveal the underlying mining service.

We evaluated the difference in mining-related semantics between domain names of public and stealthy pools to determine the extent to which stealthy mining pools are related to mining activity. A public mining pool domain typically follows the form <coin>.<region>.<SLD>, like *eth.usa.antpool.com*, of the three parts, <coin> and <SLD> usually contain mining-related semantics. Inspired by this naming pattern, We sample and investigate 100 public mining pool domains, as collected in Section 3.4, and summarize the mining semantics patterns, which consist of eight mining-related keywords. The mining semantic pattern can be categorized into two types: *(i)* mining-related activities and infrastructure: including *pool*, *mine*, *mining* and *hash*;*(ii)* popular cryptocurrencies and its abbreviations, including *xmr*, *monero*, *eth* and *btc*.

We examine the 17,488 stealthy pool domains and 2,442 public pool domains collected from Section 3.4. For each keyword of mining semantics, we count the number of times it appears in the domains. Figure 5 depicts the percentages of mining semantics found in domain names. It has been discovered that the mining semantics of stealthy pools are significantly lower than public pools. More than half of domain names of public pools contain the word "pool", while only 3.7% of all contain "pool" for stealthy pools. Overall, 93.0 % of public domains and 5.9% of stealthy domains show relevance to mining semantics.

We further make use of the *Pearson's Chi-square test of independence*[63] to see if there is a statistically significant difference in the frequency of mining-related words in public pools vs stealthy pools. Our null hypothesis H_0 is that the distribution of mining-related semantics does not have a significant difference. The result shows the p-value of the test is less than 0.001, so we can reject the null hypothesis H_0 and conclude that stealthy mining pools is significantly different and contain fewer mining-related semantics in domain names compared to public pools. This result indicates

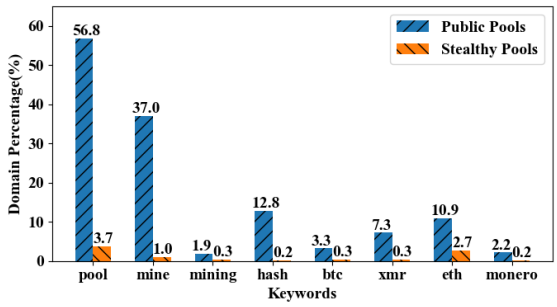


Figure 5: Percentages of mining-related keywords appeared in mining pool domains.

that stealthy pools disguise their mining-related semantics in the form of domain names, making them less noticeable.

4.2 Protocol Support of Stealthy Pools

Even though supporting multiple implementations inevitably requires additional software development efforts, nowadays, stealthy pools still tend to provide services that support various mining protocols. In this section, we investigate the number of implementations supported by stealthy pools to demonstrate the extent of effort they have made to provide user-friendly services.

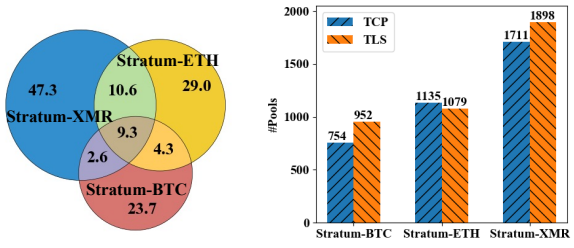
As shown in Figure 6, for the three implementations $Stratum-BTC$, $Stratum-ETH$, and $Stratum-XMR$, the numbers of supporting pools are 1,806 (23.67%), 2,214 (29.02%), and 3,609 (47.31%), respectively. Since Stratum-XMR is a protocol dedicated to Monero mining, nearly half of the stealthy mining pools offer Monero mining services.

In addition, We find that one pool often hosts mining services that support multiple different implementations at the same time. 9.3% of stealthy pools support requests for all three implementations $Stratum-BTC$, $Stratum-ETH$, and $Stratum-XMR$ at the same time, and 7.5% of pools can respond to both TLS-encrypted and non-TLS-encrypted Stratum requests. Each mining pool supports 2.8 services, considering TLS support and different Stratum implementations. Supporting multiple mining protocol implementations, stealthy mining pools make it easier for clients to conduct mining operations by enabling them to change the mining currency or TLS settings without updating the pool configurations.

4.3 Lifespan of Stealthy Pool Domains

Public mining pools need to provide stable services for their users, therefore, a public pool domain name can typically survive for a long period. Nevertheless, apart from attracting mining clients' favorites, a longer lifespan may lead to drawing unnecessary adversary notice, e.g, firewalls or antivirus engines. This raises the question of how long a stealthy mining pool domain lasts.

By using the PDNS dataset, we find that domains of stealthy pools have a much shorter lifespan than public pools. Specifically, the lifespan of a stealthy mining pool domain is estimated by combining the first and last times its DNS record was resolved, which represent the start and end of its lifetime. This serves as a lower-bound estimation, as the domain could have been active before



(a) Implementations support (%)

(b) Stratum over TLS

Figure 6: Protocol support of mining pools.

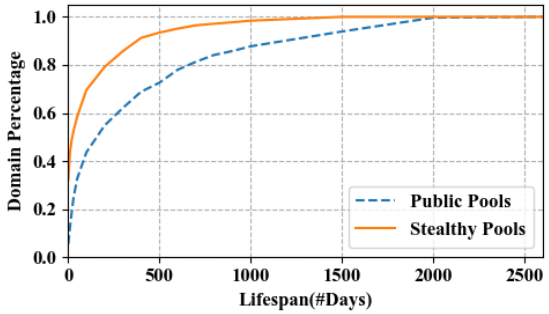


Figure 7: The CDF of public and stealthy mining pools' lifespan.

the first resolution or after the last. Figure 7 shows the CDF of lifespan distribution of public mining pools and stealthy mining pools respectively, generated from 17,488 domains of stealthy pools and 2,442 domains of public pools. Nearly half of the mining pool domains have a lifecycle of fewer than 10 days, including 33% of them are active for less than one day. In contrast, more than half of public mining pools have a lifecycle of more than one year.

Compared to public mining pools, the lifecycle of stealthy mining pools is significantly shorter. Since stealthy pools do not provide public services, maintaining a domain name for a long period of time does not seem very necessary. Besides, when the stealthy pool is used for malicious purposes, the short lifecycle helps to escape the denylists and maintain a low profile.

5 CRYPTOMINING CAMPAIGNS

In this section, we investigate the malicious behavior of cryptomining campaigns abusing stealthy mining pools. We achieve this goal by implementing a four-step process. First, we recognize the malicious activities associated with stealthy mining pools by utilizing threat intelligence (TI). Next, we group and identify the cryptomining campaigns by analyzing their underlying infrastructure, including IP addresses, top-level domains, and public keys. Then, we reveal and evaluate the strategies for evading detection and spreading samples. Finally, we conduct a qualitative study to demonstrate the profit gains of these cryptojacking activities from an insider's view.

Table 6: Statistics of stealthy mining pools' malicious activities reported by VirusTotal. Note that comm. is the abbreviation for communication.

IP				Domain	
host	refer	download	comm.	malicious	suspicious
20.6%	2.60%	2.40%	14.50%	3.30%	1.30%

5.1 Malicious Activities

There have been many anecdotal reports and academic papers describing that malicious campaigns utilize stealthy pools as a covert channel for cryptocurrency mining [25, 31, 52]. However, it is unclear how many stealthy pools have been used to facilitate these malicious activities. In this section, by correlating the mining infrastructures with state-of-the-art threat intelligence, we aim to shed light on the extent of this issue.

Specifically, we use VirusTotal [35] as the threat intelligence source for identifying malicious activities, which is a publicly available open threat intelligence platform that synthesizes data from more than 70 anti-virus engines. The intelligence report for stealthy mining pool can be categorized into two folds, the IP analysis report and the domain analysis report. IP analysis report of VirusTotal includes four types of malicious behavior: 1) *hosting*, which refers to malware or malicious URLs hosted on this IP address; 2) *download*, which refers to malware samples downloaded from URLs associated with this IP address; 3) *communication* means the IP under study has performed communication with malware samples through their execution in a sandboxed virtual environment; and 4) *referred* means domains are witnessed embedding in malware samples as strings. As for the domain report, the VirusTotal labels the domain as *malicious*, *suspicious*, or *undetected*, according to the result generated by its security vendors.

Results. By examining the VirusTotal labels of the stealthy mining pools, surprisingly, we found 23.3% IP of mining pools are labeled as malicious by at least one security vendor. As shown in Table 6, 20.6%, 2.6%, 2.4%, and 14.5% mining pools are labeled as pools hosting, referred, downloaded, or communicated to at least one malware sample, respectively. The strong correlation between stealthy mining pools and malicious activities indicates that such mining infrastructure has widely facilitated the malware gaining profit. Interestingly, as for the domain report, we find only 3.3% of the domains were classified as malicious and 1.3% of the domains were marked as suspicious. The huge gap between IP and domain mainly attributes to the short lifespan of domains as we've discussed in Section 4.3.

5.2 Malicious Campaign Analysis

In contrast to public mining pools that open to a wide range of miners, stealthy mining pools are privately owned and many (23.3% of them) are witnessed in malicious scenarios. In this section, we aim to investigate the distribution of malicious mining services and examine their characteristics.

We define a campaign to be a group of malicious stealthy mining pools correlated by some indicators like shared IP addresses, eTLD+1, and public keys of TLS certificates. Similar definitions are

Table 7: Top 10 malware campaigns abusing stealthy mining pool, ranked by request volumes in passive DNS.

Camp.	# Req.	# IP	# Domain	# Sample	Time first	Malware family	Most resolved domain	Supported Stratum			# Strategies
								BTC	ETH	XMR	
C1	6.45 B	7	15	268	04/2021	WannaMine[7]	o.aunt***.com	○	○	●	3
C2	6.30 B	5	2	288	08/2019	Outlaw[10]	debian-pack***.center	○	○	●	3
C3	3.97 B	1	3	3	05/2021	-	apache***.top	○	○	●	2
C4	3.11 B	15	34	33	11/2016	8220 Gang[12]	rx.the***.win	○	○	●	3
C5	1.0 B	67	131	2,130	08/2014	8220 Gang[12]	xmr-rx0.pwn***.pw	●	●	●	3
C6	769 M	1	2	2	12/2018	-	donate.xmr***.pro	○	○	◐	2
C7	580 M	8	2	8	08/2020	-	btc.my1***.com	◐	◐	◐	1
C8	65.6 M	2	24	45	05/2018	-	a.be***.website	○	○	●	3
C9	63.9 M	1	28	6	05/2021	-	zq2021.zao***.com	○	◐	○	2
C10	35.1 M	1	7	100	07/2022	-	poole.laofu***.com	○	○	◐	1

Camp.: Campaign. Req.: Request. For the visible features, we use "○" when we observe neither TCP-based nor TLS-based Stratum is supported.
"◐" when only TCP-based Stratum is supported, and "●" when both TCP-based and TLS-based Stratum are supported. #Strategies means encountered strategies.
Samples and corresponding *Malware family* are retrive from the IP report in Section 5.1.
Time first means the first seen time of the campaign, We take the earlier one from either the first time recorded by the PDNS or the date when the samples are first seen.

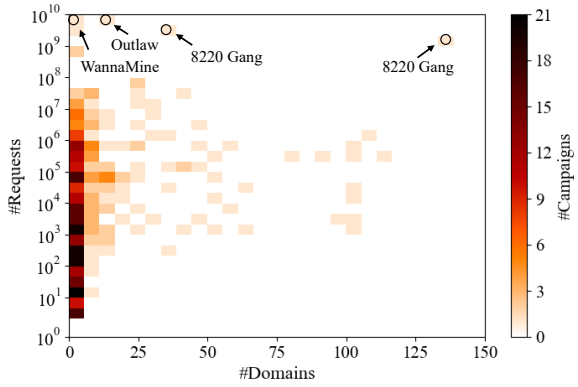


Figure 8: Distribution of identified campaigns in terms of the number of domains and counts of requests.

widely used in current research, such as [44, 60]. We utilize the following indicators to further categorize malicious mining pools into campaigns.

• **Common IP addresses**. Miscreants set up multiple mining pools on the same physical or virtual machine. Thus, if different domain addresses are hosted on the same IP address, they will be considered to belong to the same campaign.

• **Common eTLD+1**. eTLD+1 is commonly referred to as registrable domain, which suggests that they are controlled by the same registrant [69]. Thus, if domains with the same eTLD+1 are found in separate clusters, then they are merged into one campaign.

• **Common public key of TLS certificate**. Stealthy mining pools leverage TLS to encrypt their communications, and the miscreants won't share their private keys with each other in most cases. If two different mining pools share the same public key of TLS certificate, we consider they are in the same campaign, then group them into the same campaign.

Using these features, we identified a total of 439 campaigns involving 880 IPs and 4,503 domains. Figure 8 shows the distribution of domain counts and client request counts for the identified campaigns. Note that the counts of client requests are generated from the PDNS dataset.

Table 7 shows an overview of the top 10 campaigns ranked by request counts in PDNS. C1,C2,C4,C5 are all from known cryptomining campaigns, where C1 belongs to WannaMine [7], C2 belongs to Outlaw [10], and C4,C5 are all from 8220 Gang [20]. Although C4, C5 belong to the same malware campaign, they are not related in any identifiers we've adopted, thus they belong to two different mining pool infrastructures of 8220 Gang, which shows the robustness of 8220 Gang's mining topology.

In terms of protocols, we found that nine of the top 10 campaigns support *Stratum-XMR*, which is used for Monero mining. This also confirms previous studies [62] that Monero is the preferred currency by criminals for malicious cryptomining since it is friendly to CPU mining.

Case Study. Take C5 as an example. It is one of the largest campaigns for the known cryptomining malware family 8220 Gang, which has grown rapidly since 2021 according to public reports [12]. Our analysis finds that the earliest activity of C5 dates back to 2014, indicating it has been active for a long time. Interestingly, we find that some of its domains do not hide the mining-related semantics like most other stealthy pools as we've discussed in Section 4.1. these domains follow the pattern *<coin>-<algo>-(tls).pwn***.pw* like *xmr-rx0-tls.pwn***.pw*. We can deduce from this naming pattern that at least 10 cryptocurrency mining services are provided. We also discover that C5 uses TLS extensively, with over 100,000 requests to its Monero TLS pool. We further refer to the threat intelligence labels for 2,130 C5 samples and categorize them into three types: (i) CoinMiner; (ii) Tsunami botnet [19]; and (iii) Port scanner, which shows its effort to spread aggressively.

5.3 Surviving Strategies

As mining pool detection techniques continue to improve, it's crucial for stealthy mining to take actions to keep underground and survive. In this section, we uncover three tricks used by stealthy mining pools as countermeasures. Our investigation revealed that these strategies, including migrating domain name resolution method, leveraging known botnets, and enabling Transport Layer Security (TLS) encryption, can greatly increase the surviving rate.

Table 8: Evolution of resolution strategies for campaigns.

Year	CNAME	Public IP	Self-owned IP
2022	2.10%	4.50%	98.20%
2021	4.60%	8.30%	93.20%
2020	4.30%	5.40%	92.00%
2019	11.90%	6.70%	85.70%
2018	9.60%	3.60%	84.10%

We collect the history records for each year from PDNS at the timestamp of October 31.
The campaign is considered to have adopted the resolution strategy if it was used by one of the campaign's domains.

5.3.1 Migrating domain names resolution method. Before stealthy mining pools are massively adopted by attackers, previous research [62] has suggested that attackers try to escape denylist-based detection by creating CNAME domain aliases, i.e., domains they hold in the form of CNAMEs that point to domains in public mining pools. Similarly, we observe from the PDNS dataset that some stealthy pool domains used to resolve its mining pool domains to public mining pools by configuring A records as public mining pool IP addresses. In both cases, the mining pool address is not under the control of the attacker and may be blocked when the attacker's wallet address is reported to some responsible mining pool [62].

All the stealthy pool domains we collected are A records pointing to stealthy pool IPs. We further examined the PDNS resolution history of domains from the campaigns, and found instances where they were CNAME aliases or A records pointing to public pools in the past five years. Table 8 shows the evolution of the resolution strategies. The proportion of campaigns holding their own mining pool IPs has been increasing year by year, reaching 98.2% in 2022.

Take the campaign C30 for example. In 2020 its mining pool domain *xmrs.wuli***.nl* first pointed to a Monero mining pool *pool.supp***.com* by way of CNAME, but it was altered to point directly to *pool-ca.supp***.com*'s IP 192.******.114 later in this year. Then it resolves the mining pool domain to its self-owned IP address 192.***.***.90 from 2022. During this evolutionary process, the stealthiness of the campaign's mining pools gradually grew, suggesting that criminals will continue to change their mining strategies in order to maximize profits.

To assess the effectiveness of this strategy, we submit the domain and IP list of public and stealthy pools to a security vendor we partner with and compared the detection rates of public and stealthy mining pools. Table 9 reveals that the detection rate of public pools (82.5%) is much higher than stealthy pools (23.3%), indicating that the stealthy mining pool can hugely hide the mining behavior. Besides, we find setting up the stealthy mining pool with a self-owned IP address effectively decreases the detection rate. Specifically, when considering the three different ways of domain resolution, the detection rate drops to 20.6% of IPs and 4.0% of domains while using self-owned IP.

5.3.2 Leveraging the botnets. Anecdotal reports suggest that mining malware campaigns have begun to employ botnets to propagate samples in order to make malware spread more easily [12]. We investigated the malicious labels linked with the malware campaigns obtained in section 5.2 to assess the rate of usage of this method.

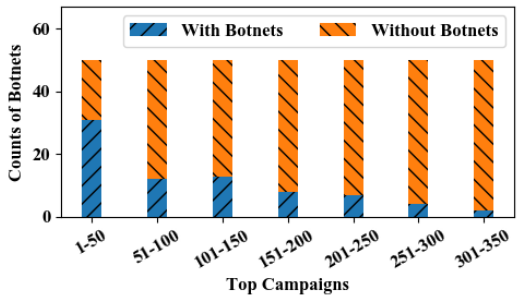


Figure 9: Usage of botnets in malware campaigns.

We found a total of 18.9% (2,041/10,824) samples from 77 campaigns associated with commodity botnets, including 11 different botnet services like Tsunami, Virut, and Graftor [15, 19, 36]. Interestingly, as shown in Figure 9. we find that the more active the campaign is, the higher the percentage of botnet recruiting. Among the top 100 campaigns sorted by the count of DNS requests, 43% of them have used the botnet to spread the mining malware. This rate decreases to 21% among the top 100 to 200 campaigns and further to 5.4% in the remaining campaigns. This difference suggests that campaigns that use botnets to spread malware are more prosperous.

5.3.3 Enabling TLS encryption. According to Section 3.4, we probe the targets with and without TLS encryption. Among the mining pools we collected, TLS has a deployment rate of 51.5%, which means more than half of the stealthy mining pools tend to communicate with mining via encrypted traffic. By checking the IP report in Section 5.1, we find that among all the TLS pools, only 9.6% are labeled as malicious, which is much lower than the overall malicious rate (23.3%). As for mining campaigns, we found that 238 out of 439 campaigns have deployed at least one mining pool service that supports TLS connection. In contrast to the 2021 study [67] that stated there is no usage of SSL/TLS by cryptomining malware, we have observed a high adoption rate (54.2%) of TLS encryption in cryptomining campaigns through protocol scanning. This suggests that these cryptomining campaigns are evolving rapidly.

We further scanned 757 TLS-enabled stealthy mining pools that were still active on November 10, 2022, and find out that 66.5% of these server certificates are self-signed certificates and 7.1% of them are expired TLS certificates. Appendix B contains examples of self-signed certificates. We can conclude that stealthy mining pools deploying TLS simply make use of the encryption capabilities of TLS without caring about the security of the entire session. This is due to the fact that, on the one hand, self-signed certificates are cost-free and quick and easy to generate, and on the other hand, a previous study has proved that encrypted mining traffic is sufficient to escape deep packet inspection (DPI) based detection [49].

5.4 Revenue Estimation

Estimating the revenues of cryptocurrency mining campaigns via stealthy mining pools can be challenging due to a lack of information on controlled miners and criminal wallet addresses. However,

Table 9: The detection rate of public pools and stealthy pools with different domain resolution strategies. We consider a domain or IP detected if it is flagged as malicious. CNAME and Public IP mean the stealthy pool adopted this strategy in the past.

Category	Public pool	Stealthy pool			
		CNAME	Public IP	Self-owned IP	Total
Domains	91.6%(2237/2442)	75.5%(34/45)	64.8%(79/122)	4.0%(691/17321)	4.6%(804/17488)
IPs	82.5%(416/504)	96.3%(26/27)	69.6%(55/79)	20.6%(435/2113)	23.3%(516/2219)

Table 10: Overview of the stealthy mining pool domains we have taken over.

Pool domain	Pool IP	Pool port	Stratum protocol	Sample	Monthly observation			
					Cost ($)	# Victim	# Req.	Revenue ($)
seve-amam.jn***.com	23.*****.17	6666	BTC/ETH/XMR+TLS	-	0.721	1	81	0.709
5rx***.cn	106.*****.216	5555	BTC+TlS	-	0.498	2	925	1.38
apkxm.faye***.org	206.*****.113	80	XMR+TCP/TLS	✓	0.927	116	408,482	82.112
data.halo***.club	67.*****.14	38071	XMR+TCP/TLS	✓	0.179	1,057	532,898	747.101
lzsc***.cn	152.*****.53	8080	BTC/ETH/XMR+TLS	-	0.361	0	0	0
nhiai***.xyz	185.*****.8	8080	XMR+TLS	✓	0.167	11	9,674	7.8
neo***.cn	106.*****.216	5555	BTC+TCP	-	0.178	0	0	0
www.bs***.co	51.*****.37	7777	BTC/ETH/XMR+TCP/TLS	✓	0.584	609	11,851,143	430.619
uxi***.com	47.*****.240	443	BTC/ETH/XMR+TCP	-	0.923	0	0	0
Total	-	-	-	-	**4.538**	**1,796**	**12,803,203**	**1,269.717**

taking over the expired domain of the mining pool gives us the opportunity to evaluate the profits from an insider's view. We checked for registrable domains among the mining pool belonging to campaigns and finally obtained nine expired mining pool domains. The overview of taken-over mining pools is displayed in Table 10.

To evaluate the scales and impacts of taken-over pools, we deployed a mining pool honeypot based on the port and protocol information obtained from Section 3.4. To avoid causing any negative effects on victims, our honeypot only collects mining pool login requests and sends a response informing them that the domain has been taken over, thus the victims do not actually start mining and contribute computing power to the taken-over domains. Specifically, to further validate that the client connects to the taken-over domains to start the mining process without performing mining activities, we conduct a proof-of-concept (PoC) experiment under our controlled environment (see PoC details in Appendix C).

The honeypot was served from 2022-12-01 to 2022-12-31. During this one-month-long period, six of the nine taken-over pools received mining login requests from miners, totaling 1,796 victims from 44 countries (details of victims distribution can be seen in Appendix D).

Revenue estimation from taken-over pools. In order to understand the revenue gains by the miscreants, we estimated the profits we make by taking over the mining pool and compared it with the investment made in purchasing the domain. Specifically, we used the following revenue estimation model.

$$R_i = \sum_{j=0}^{Victims} L_j * h_j * P \quad (1)$$

Where R_i refers to the revenue that can be obtained through mining pool i. *Victims* refers to the number of victims in the pool. L_j denotes the victim's lifespan, while h_j denotes the hash rate of the victim's system. P refers to mining profitability, meaning how much USD can be obtained each day by a certain mining hash rate.

To estimate the lifecycle L_j of the victim, we assume that we start mining with the victim machine from the first time the victim sends a mining request and that the victim has been actively mining since then. The victims we take over are all Monero cryptocurrency miners. Therefore, to estimate the hash rate h_j of the victim machines, we refer to the benchmark provided by xmrig [38], the most popular client for Monero coin mining. The average hash rate is based on the mainstream CPU (Intel i5-7400 Processor) in desktop PCs, With the algorithm set to RandomX and considering the average performance of single thread and multi threads, we set the hash rate of all machines to 1000H/s.

To find the daily mining profitability, we collected historical data from BitinfoCharts [22], which provides the mining profitability for a day in USD with a hash rate of 1 kHash/s based on the daily mining difficulty and block returns.

Table 10 summarizes our costs and profits for each domain. Using these data, we estimated the profitability for one month is $1,269.717 on average. Considering the renting price of servers for deploying taken-over pools, our overall purchase cost is $44.538 for one month ($40 for renting servers and $4.538 for registering domains). By taking over the mining pool, we can earn $1,225.179 per month, with a return on investment (ROI) of 2,750%.

Revenue estimation from PDNS. Our research has shown that criminals engaging in stealthy mining activities are able to generate significant profits, with an average return on investment of 2,750%. To gain an extensive understanding of the financial impact of these activities, we employ PDNS to assess the potential victims.

Since *Stratum-XMR* is exclusively used by Monero, we focus on the 134 campaigns that only adopt *Stratum-XMR* as their mining protocol and estimate revenues based on Monero price. Specifically, for domains of Monero campaigns, we get the first and last access days of users from the PDNS dataset as the start and end of the

mining process and then aggregate the overall mining duration for each campaign. It's worth noting that the IP addresses for all users are anonymized when we access the PDNS dataset.

We take a sample of one month's history record of PDNS from November 1 to November 30. Using the revenue estimation model for taken-over pools (Equation 1), our result shows that the 134 Monero campaigns can profit around $84,836.32 per month from more than 200 thousands victims. Therefore, criminals have the ability to acquire more than 1 million USD per year.

6 DISCUSSION

6.1 Limitations

Bias of data collection. Since our mining pool port distribution comes from a single ISP's Netflow statistics, the result inevitably has some bias. However, we made the following efforts to make the data more representative: (1) Our Netflow collection lasted six months, and the sampled port distribution is closer to the real situation, i.e., it shows a long-tailed distribution. (2) We scanned the top 32 ports accounting for more than 80% of cryptocurrency mining activities in the entire IPv4 space to mitigate the limitation of the potential bias ISP Netflow (Section 3.4).

Mining pool protocols. Stratum is the *de facto* protocol for mining pool communication. However, there are no standardized implementation specifications for the Stratum, leading to specific implementation varies among different mining pools. To make the probing results convincing, we seek documentation, previous work, and miner clients to collect different implementations (Section 3.1). Through these efforts, we have been able to identify three different variations of the Stratum protocol implementation.

We also admit that, as noted in the browser-based cryptojacking study [56], our methodology cannot fully detect mining pools with obfuscated traffic (e.g., utilizing base64 encoding of the mining payload) or custom protocols. However, we believe this case only accounts for a few real-world stealthy mining pool services. First, traffic obfuscation or custom protocols require both miners and mining pools to be programmed and negotiated to support the same encoding/encryption protocol, making the mining program less portable. Second, to bypass ISP censorship, e.g., DPI, the TLS encryption of Stratum protocol can satisfy this requirement, and our experiments confirm that this obfuscation method is able to evade detection (Section 5.3.3). Finally, during our study, all public reports of cryptojacking malware utilize one of the three implementations of Stratum protocols (Section 3.1).

6.2 Responsible Disclosure

Stealthy mining pools. For the IP addresses of stealthy mining pools from cryptomining campaigns, we initiate a responsible disclosure by collecting the IP WHOIS data and extracting the email addresses for reporting network abuse. Then we send emails to report the IP address abuse we've found. By the submission of this paper, we have reported the abuse to 24 ISPs or hosting providers and received acknowledgments from two of them.

Victim miners. Our experiment in Section 5.4 found 1,796 victims of taken-over cryptomining campaigns from 44 countries. 78 different ISPs were found by the IP WHOIS data, and we have reported these issues to the related ISPs. Up to now, the ISP where we get Netflow data has confirmed the cryptomining malware activities and taken down the cryptojacking domains.

7 RELATED WORK

Mining pool. There have been a few earlier studies on the mining pool ecosystem. Miller et al. [61] discovered peer-to-peer links in Bitcoin. They inferred details about the organization of mining pools by corroborating these details with supplemental evidence found by public records on the web and DNS records. Kai et al. [58] revealed Ethereum's Network Topology and demonstrated mining pools' biased neighbor selection strategies. Cao et al. [45] explored the Monero Peer-to-Peer Network. They linked public and private mining pools to Monero P2P nodes by the nodes' degree. Some works focus on mining pool attacks Ittay Eyal [47] discussed the withholding attack on mining pools. Variants of this attack include Fork After Withholding (FAW) attack proposed by Yujin et al. [57], and the Power Adjusting Withholding (PAW) attack proposed by Shang et al. [50]. Kai el al. [59] conducted a novel attack that can disable a remote Ethereum node's txpool service. However, there is no systematic research to analyze the stealthy mining pool yet due to a lack of a full understanding of the protocol implementation, usage, and port distribution.

Cryptojacking. The first study of cryptomining malware was conducted by Huang et al. [52], they analyzed the prevalence of Bitcoin mining botnets and discovered the use of dark pools via network protocol in 26% of the samples. The most related work is Pastrana et al. [62] which conducted the largest measurement of cryptomining malware. They found some stealth techniques including using CNAMEs of public pools and mining proxies to bypass denylist-based detection. These two studies were performed by malware sample analysis, which neither covers most of the stealthy pools we've collected nor performs a large-scale and longitudinal study.

Recently Li et al. [60] studied real-world illicit cryptomining on public CI platforms. Only public mining pools are included in their crawling samples. As browser-based cryptojacking emerged, Geng et al. [51] and Konoth et al. [56] reported a systematic study on browser-based cryptojacking ecosystem. They all mentioned the extensive use of Websocket proxy servers by browser-based mining. Some works, such as [44, 68], are also devoted to browser-based cryptojacking measurement.

Many studies discovered cryptomining activity using content-agnostic traffic flow [46, 48, 67, 73]. Their approach was based mostly on the spatio-temporal properties of the Stratum protocol. In addition to network-based detection, host-based detection studies utilized sample fingerprints and unique features of hardware to find out mining activities [70, 72]. In contrast to previous works, which primarily focus on cryptojacking malware samples and behaviors, our research proposes and examines a novel mining underlying infrastructure: the stealthy mining pool. We investigate its ecosystem, characteristics, evasion techniques, and revenues.

8 CONCLUSION

The presence of stealth mining pools in the cryptocurrency mining ecosystem has become impossible to overlook. Recent reports demonstrate that underground miners and cryptojacking malware have turned to stealthy pools to bypass law enforcement activities

or security censorship. However, due to lacking a comprehensive understanding of the protocol characteristics, there is no systematic research to analyze stealth mining pools yet.

In this paper, we shed light on the (ab)use of stealthy mining pools in the wild. By performing a large-scale and longitudinal measurement study of stealthy mining pools, we report 7,629 stealthy mining pools, spanning 2,113 IPs and 17,488 domains among 59 countries. Our analysis reveals the stealthy mining pools carefully crafting their domain semantics, protocol support, and lifespan to provide underground, user-friendly, and robust mining services. What's worse, we uncover a strong correlation between the stealthy mining pool and malware. Stealthy pools are also leveraging tricks, e.g., migrating domain names resolution method to evade state-of-the-art mining detection. Finally, a qualitative study is performed to evaluate the profit gains of malicious cryptomining activities from an insider's perspective, demonstrating that criminals have the ability to acquire more than 1 million USD per year with an average ROI of 2,750%.

ACKNOWLEDGMENTS

We sincerely thank all anonymous reviewers for their valuable comments to improve the paper. This work is in part supported by the National Key Research and Development Program of China (2021YFB2701000), the National Natural Science Foundation of China (62102218, U19B2034, 62272265, 61972224, 62302101), Beijing National Research Center for Information Science and Technology (BNRist) under Grant BNR2022RC01006, Alibaba Innovative Research Program (AIR), CCF-Tencent Rhino-Bird Young Faculty Open Research Fund, and the Huawei Technologies Co., Ltd under Grant No. TC20200917004.

REFERENCES

[1] 2012. Netflow. https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/ios-netflow/prod_white_paper0900aecd80406232.html.
[2] 2016. EthereumStratum_v1.0.0. https://github.com/sammy007/open-ethereum-pool/blob/master/docs/STRATUM.md.
[3] 2017. Electrum Protocol Specification. https://electrum.readthedocs.io/en/latest/protocol.html#electrum-protocol-specification.
[4] 2018. Equihash miner. https://github.com/nemosminer/DSTM-equihash-miner.
[5] 2018. Large Scale Monero Cryptocurrency Mining Operation using XMRig. https://unit42.paloaltonetworks.com/unit42-large-scale-monero-cryptocurrency-mining-operation-using-xmrig/.
[6] 2018. Stratum mining protocol. https://github.com/xmrig/xmrig-proxy/blob/master/doc/STRATUM.md.
[7] 2018. Wannamine. https://www.techtarget.com/searchsecurity/news/252434485/Cryptojacking-malware-using-EternalBlue-to-build-botnets.
[8] 2019. Xmr-Stak. https://github.com/fireice-uk/xmr-stak.
[9] 2020. Claymore. https://github.com/Claymore-Dual/Claymore-Dual-Miner.
[10] 2020. Outlaw. https://www.oguzhantopgul.com/2020/06/outlaw-botnet-xmrig-miner-and-shellbot.html.
[11] 2020. Stratum Implementation. https://github.com/edsonayllon/Stratum-Implementation-For-Pantheon#4-stratum-implementation.
[12] 2021. 8220 Gang. https://www.lacework.com/blog/8220-gangs-recent-use-of-custom-miner-and-botnet/.
[13] 2021. China's top regulators ban crypto trading and mining, sending bitcoin tumbling. https://www.reuters.com/world/china/china-central-bank-vows-crackdown-cryptocurrency-trading-2021-09-24/.
[14] 2021. CPUMiner. https://github.com/pooler/cpuminer.
[15] 2021. Graftor. https://en.wikipedia.org/wiki/Hupigon.
[16] 2021. India to propose law banning cryptocurrency trading, mining and possession. https://economictimes.indiatimes.com/industry/banking/finance/india-to-propose-law-banning-cryptocurrency-trading-mining-and-possession/banning-cryptocurrencies/slideshow/81526975.cms.
[17] 2021. Masscan. https://github.com/robertdavidgraham/masscan.
[18] 2021. Masscan exclude IP list. https://github.com/robertdavidgraham/masscan/blob/master/data/exclude.conf.
[19] 2021. Tsunami. https://malwiki.org/index.php?title=Tsunami.
[20] 2022. 8220 Gang Deploys a New Campaign with Upgraded Techniques. https://blog.aquasec.com/8220-gang-confluence-vulnerability-cve-2022-26134#Discovery.
[21] 2022. BTC.com. https://pool.btc.com/.
[22] 2022. Cryptocurrency statistics. https://bitinfocharts.com/.
[23] 2022. Ethermine. https://etc.ethermine.org/.
[24] 2022. Ethminer. https://github.com/ethereum-mining/ethminer.
[25] 2022. LemonDuck Cryptojacking Botnet. https://www.crowdstrike.com/blog/lemonduck-botnet-targets-docker-for-cryptomining-operations/.
[26] 2022. Minerblock. https://github.com/xd4rker/MinerBlock.
[27] 2022. MiningPoolStats. https://miningpoolstats.stream/.
[28] 2022. Monero Network Monitoring. https://web.archive.org/web/20221209071310/https://minexmr.com/pools.html.
[29] 2022. Nanopool. https://nanopool.org/.
[30] 2022. New York governor signs first-of-its-kind law cracking down on bitcoin mining. https://www.cnbc.com/2022/11/23/new-york-governor-signs-law-cracking-down-on-bitcoin-mining.html.
[31] 2022. Orchard botnet. https://blog.netlab.360.com/a-new-botnet-orchard-generates-dga-domains-with-bitcoin-transaction-information/.
[32] 2022. Stratum v1. https://www.braiins.com/stratum-v1/docs.
[33] 2022. STRATUM V2. https://www.braiins.com/stratum-v2.
[34] 2022. The state of cryptojacking in the first three quarters of 2022. https://securelist.com/cryptojacking-report-2022/107898/.
[35] 2022. Virustotal API v3 Overview. https://developers.virustotal.com/reference.
[36] 2022. Virut. https://en.wikipedia.org/wiki/Virut.
[37] 2022. XMRig. https://xmrig.com/docs/miner.
[38] 2022. XMRig benchmark. https://xmrig.com/benchmark.
[39] 2022. XMRig proxy. https://xmrig.com/docs/proxy.
[40] 2023. CrowdStrike Uncovers I2Pminer MacOS Mineware Variant. https://www.crowdstrike.com/blog/i2pminer-macos-mineware-analysis/.
[41] 2023. Maltrail cryptomining list. https://github.com/stamparm/maltrail/blob/master/trails/static/suspicious/crypto_mining.txt.
[42] 2023. Mozilla foundation's public suffix list. https://publicsuffix.org/list/public_suffix_list.dat.
[43] 2023. Selenium. https://www.selenium.dev/.
[44] Hugo L. J. Bijmans, Tim M. Booij, and Christian Doerr. 2019. Just the Tip of the Iceberg: Internet-Scale Exploitation of Routers for Cryptojacking. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019, London, UK, November 11-15, 2019*, Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz (Eds.). ACM, 449–464. https://doi.org/10.1145/3319535.3354230
[45] Tong Cao, Jiangshan Yu, Jérémie Decouchant, Xiapu Luo, and Paulo Veríssimo. 2020. Exploring the Monero Peer-to-Peer Network. In *Financial Cryptography and Data Security - 24th International Conference, FC 2020, Kota Kinabalu, Malaysia, February 10-14, 2020 Revised Selected Papers (Lecture Notes in Computer Science, Vol. 12059)*, Joseph Bonneau and Nadia Heninger (Eds.). Springer, 578–594. https://doi.org/10.1007/978-3-030-51280-4_31
[46] Maurantonio Caprolu, Simone Raponi, Gabriele Oligeri, and Roberto Di Pietro. 2021. Cryptomining makes noise: Detecting cryptojacking via Machine Learning. *Comput. Commun.* 171 (2021), 126–139. https://doi.org/10.1016/j.comcom.2021.02.016
[47] Ittay Eyal. 2015. The Miner's Dilemma. In *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*. IEEE Computer Society, 89–103. https://doi.org/10.1109/SP.2015.13
[48] Yebo Feng, Jun Li, and Devkishen Sisodia. 2022. CJ-Sniffer: Measurement and Content-Agnostic Detection of Cryptojacking Traffic. In *25th International Symposium on Research in Attacks, Intrusions and Defenses, RAID 2022, Limassol, Cyprus, October 26-28, 2022*. ACM, 482–494. https://doi.org/10.1145/3545948.3545973
[49] Sergey Frolov and Eric Wustrow. 2019. The use of TLS in Censorship Circumvention. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019, San Diego, California, USA, February 24-27, 2019*. The Internet Society. https://www.ndss-symposium.org/ndss-paper/the-use-of-tls-in-censorship-circumvention/
[50] Shang Gao, Zecheng Li, Zhe Peng, and Bin Xiao. 2019. Power Adjusting and Bribery Racing: Novel Mining Attacks in the Bitcoin System. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019, London, UK, November 11-15, 2019*, Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz (Eds.). ACM, 833–850. https://doi.org/10.1145/3319535.3354203
[51] Geng Hong, Zhemin Yang, Sen Yang, Lei Zhang, Yuhong Nan, Zhibo Zhang, Min Yang, Yuan Zhang, Zhiyun Qian, and Hai-Xin Duan. 2018. How You Get Shot in the Back: A Systematical Study about Cryptojacking in the Real World. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018, Toronto, ON, Canada, October 15-19, 2018*, David Lie, Mohammad Mannan, Michael Backes, and XiaoFeng Wang (Eds.). ACM, 1701–1713. https://doi.org/10.1145/3243734.3243840

[52] Danny Yuxing Huang, Hitesh Dharmdasani, Sarah Meiklejohn, Vacha Dave, Chris Grier, Damon McCoy, Stefan Savage, Nicholas Weaver, Alex C. Snoeren, and Kirill Levchenko. 2014. Botcoin: Monetizing Stolen Cycles. In *21st Annual Network and Distributed System Security Symposium, NDSS 2014, San Diego, California, USA, February 23-26, 2014*. The Internet Society. https://www.ndss-symposium.org/ndss2014/botcoin-monetizing-stolen-cycles

[53] Jordi Zayuelas i Muñoz, José Suárez-Varela, and Pere Barlet-Ros. 2019. Detecting cryptocurrency miners with NetFlow/IPFIX network measurements. In *5th IEEE International Symposium on Measurements & Networking, M&N 2019, Catania, Italy, July 8-10, 2019*. IEEE, 1–6. https://doi.org/10.1109/IWMN.2019.8804995

[54] Liz Izhikevich, Renata Teixeira, and Zakir Durumeric. 2021. LZR: Identifying Unexpected Internet Services. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael Bailey and Rachel Greenstadt (Eds.). USENIX Association, 3111–3128. https://www.usenix.org/conference/usenixsecurity21/presentation/izhikevich

[55] Manish R. Joshi and Theyazn Hassn Hadi. 2015. A Review of Network Traffic Analysis and Prediction Techniques. *CoRR* abs/1507.05722 (2015). arXiv:1507.05722 http://arxiv.org/abs/1507.05722

[56] Radhesh Krishnan Konoth, Emanuele Vineti, Veelasha Moonsamy, Martina Lindorfer, Christopher Kruegel, Herbert Bos, and Giovanni Vigna. 2018. MineSweeper: An In-depth Look into Drive-by Cryptocurrency Mining and Its Defense. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018, Toronto, ON, Canada, October 15-19, 2018*, David Lie, Mohammad Mannan, Michael Backes, and XiaoFeng Wang (Eds.). ACM, 1714–1730. https://doi.org/10.1145/3243734.3243858

[57] Yujin Kwon, Dohyun Kim, Yunmok Son, Eugene Y. Vasserman, and Yongdae Kim. 2017. Be Selfish and Avoid Dilemmas: Fork After Withholding (FAW) Attacks on Bitcoin. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017*, Bhavani Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu (Eds.). ACM, 195–209. https://doi.org/10.1145/3133956.3134019

[58] Kai Li, Yuzhe Tang, Jiaqi Chen, Yibo Wang, and Xianghong Liu. 2021. TopoShot: Uncovering Ethereum's Network Topology Leveraging Replacement Transactions. *CoRR* abs/2109.14794 (2021). arXiv:2109.14794 https://arxiv.org/abs/2109.14794

[59] Kai Li, Yibo Wang, and Yuzhe Tang. 2021. DETER: Denial of Ethereum Txpool sERvices. In *CCS '21: 2021 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, Republic of Korea, November 15 - 19, 2021*, Yongdae Kim, Jong Kim, Giovanni Vigna, and Elaine Shi (Eds.). ACM, 1645–1667. https://doi.org/10.1145/3460120.3485369

[60] Zhi Li, Weijie Liu, Hongbo Chen, XiaoFeng Wang, Xiaojing Liao, Luyi Xing, Mingming Zha, Hai Jin, and Deqing Zou. 2022. Robbery on DevOps: Understanding and Mitigating Illicit Cryptomining on Continuous Integration Service Platforms. In *43rd IEEE Symposium on Security and Privacy, SP 2022, San Francisco, CA, USA, May 22-26, 2022*. IEEE, 2397–2412. https://doi.org/10.1109/SP46214.2022.9833803

[61] Andrew Miller, James Litton, Andrew Pachulski, Neal Gupta, Dave Levin, Neil Spring, and Bobby Bhattacharjee. 2015. Discovering bitcoin's public topology and influential nodes. *et al* (2015).

[62] Sergio Pastrana and Guillermo Suarez-Tangil. 2019. A First Look at the Crypto-Mining Malware Ecosystem: A Decade of Unrestricted Wealth. In *Proceedings of the Internet Measurement Conference, IMC 2019, Amsterdam, The Netherlands, October 21-23, 2019*. ACM, 73–86. https://doi.org/10.1145/3355369.3355576

[63] Karl Pearson. 1900. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50, 302 (1900), 157–175.

[64] Victor Le Pochat, Tom van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019, San Diego, California, USA, February 24-27, 2019*. The Internet Society. https://www.ndss-symposium.org/ndss-paper/tranco-a-research-oriented-top-sites-ranking-hardened-against-manipulation/

[65] Ruben Recabarren and Bogdan Carbunar. 2017. Hardening Stratum, the Bitcoin Pool Mining Protocol. *Proc. Priv. Enhancing Technol.* 2017, 3 (2017), 57. https://doi.org/10.1515/popets-2017-0028

[66] John Repplinger. 2011. G.G. Chowdhury. *Introduction to Modern Information Retrieval*. 3rd ed. London: Facet, 2010. 508p. alk. paper, $90 (ISBN 9781555707156). LC2010-013746. *Coll. Res. Libr.* 72, 2 (2011), 194–195. http://crl.acrl.org/content/72/2/194.full.pdf

[67] Michele Russo, Nedim Srndic, and Pavel Laskov. 2021. Detection of illicit cryptomining using network metadata. *EURASIP J. Inf. Secur.* 2021, 1 (2021), 11. https://doi.org/10.1186/s13635-021-00126-1

[68] Jan Rüth, Torsten Zimmermann, Konrad Wolsing, and Oliver Hohlfeld. 2018. Digging into Browser-based Crypto Mining. In *Proceedings of the Internet Measurement Conference 2018, IMC 2018, Boston, MA, USA, October 31 - November 02, 2018*. ACM, 70–76. https://dl.acm.org/citation.cfm?id=3278539

[69] Marco Squarcina, Mauro Tempesta, Lorenzo Veronese, Stefano Calzavara, and Matteo Maffei. 2021. Can I Take Your Subdomain? Exploring Same-Site Attacks in the Modern Web. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael Bailey and Rachel Greenstadt (Eds.). USENIX Association, 2917–2934. https://www.usenix.org/conference/usenixsecurity21/presentation/squarcina

[70] Rashid Tahir, Muhammad Huzaifa, Anupam Das, Mohammad Ahmad, Carl A. Gunter, Fareed Zaffar, Matthew Caesar, and Nikita Borisov. 2017. Mining on Someone Else's Dime: Mitigating Covert Mining Operations in Clouds and Enterprises. In *Research in Attacks, Intrusions, and Defenses - 20th International Symposium, RAID 2017, Atlanta, GA, USA, September 18-20, 2017, Proceedings (Lecture Notes in Computer Science, Vol. 10453)*, Marc Dacier, Michael Bailey, Michalis Polychronakis, and Manos Antonakakis (Eds.). Springer, 287–310. https://doi.org/10.1007/978-3-319-66332-6_13

[71] Ege Tekiner, Abbas Acar, and A. Selcuk Uluagac. 2022. A Lightweight IoT Cryptojacking Detection Mechanism in Heterogeneous Smart Home Networks. In *29th Annual Network and Distributed System Security Symposium, NDSS 2022, San Diego, California, USA, April 24-28, 2022*. The Internet Society. https://www.ndss-symposium.org/ndss-paper/auto-draft-196/

[72] Wenhao Wang, Benjamin Ferrell, Xiaoyang Xu, Kevin W. Hamlen, and Shuang Hao. 2018. SEISMIC: SEcure In-lined Script Monitors for Interrupting Cryptojacks. In *Computer Security - 23rd European Symposium on Research in Computer Security, ESORICS 2018, Barcelona, Spain, September 3-7, 2018, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 11099)*, Javier López, Jianying Zhou, and Miguel Soriano (Eds.). Springer, 122–142. https://doi.org/10.1007/978-3-319-98989-1_7

[73] Shize Zhang, Zhiliang Wang, Jiahai Yang, Xin Cheng, Xiaoqian Ma, Hui Zhang, Bo Wang, Zimu Li, and Jianping Wu. 2021. MineHunter: A Practical Cryptomining Traffic Detection Algorithm Based on Time Series Tracking. In *ACSAC '21: Annual Computer Security Applications Conference, Virtual Event, USA, December 6 - 10, 2021*. ACM, 1051–1063. https://doi.org/10.1145/3485832.3485835

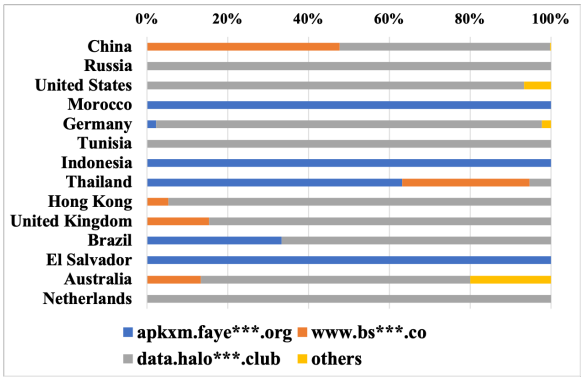


Figure 10: Geolocation distribution of victim miners from countries or regions.

Table 11: Response examples from non-mining services.

Type	Response
Echo server	{"id": 1,"method": "mining.subscribe","params": []}
Electrum	{"jsonrpc": "2.0", "error": {"code": -32601, "message": "unknown method "mining_subscribe""}, "id": 1}
Others	{"action":"create connection","nonce": "66d803b3c743cb30972fba8c83fda1aa", "success":"True"}

Table 12: Examples of self-signed TLS certificates.

Common name	Org. name	Count	# Public key
localhost	-	195	34
caocao.cao	CC	39	1
cn	-	24	1
mining.pool	Mining Pool	17	11
sslserver	-	14	1
Eth Proxy	Developer	12	1

A EXAMPLES OF NONE-MINING SERVICES

Table 11 presents the response examples from non-mining services. We categorize them into 3 types: *(i)* echo servers that return original requests; *(ii)* Bitcoin client Electrum mentioned in section 2.1; *(iii)* other services that provide a JSON format response but are not mining pools.

B SELF-SIGNED TLS CERTIFICATES

Table 12 lists examples of self-signed certificates.

C PROOF-OF-CONCEPT EXPERIMENT

To further validate that client connects to the taken-over domains can start mining while does not perform actual mining activities, we conduct a proof-of-concept (PoC) experiment under our controlled environment. Out of the nine mining pools, four Monero pools have a record of communication with known malicious samples.

In the PoC experiment, we simulated the victims by running the malicious samples in a closed virtual environment and built a mining pool environment using the open-source software XMRIG-PROXY, configuring the domain names of the mining pools associated with the samples and setting the wallet address to our own wallet address. Subsequently, we ran these samples in the virtual environment, and without modifying the content of the samples, three samples successfully started mining operations. The remaining unsuccessful sample had the corresponding mining pool *data.halo***.club:38071*. We further analyzed its network traffic and found that the sample sends HTTP GET requests to port 63145, requesting a configuration file. We successfully obtained this configuration file by accessing the previous pool IP address, i.e., 67.*****.***.14, as shown in Listing 1.

```
/*HuTaoConfig*/

[Download]
FrameworkUrl=http://fdjkgs.cn-gd.uf***.com/

[Framework]
Name=htv13.exe
Ver=20220225

[MinIng]
MineUpdate=on
MiningPool=67.***.***.14:38071
MiningPoolBackUp=67.***.***.14:38071
MiningMode=80
[Scan]
Download=http://1.1.1.1/nxc.exe
```

Listing 1: Mining pool config file of data.halo*.club**

In this configuration file, the parameter *MiningPool* specifies the mining pool address, where the pool is directly accessible via IP. We hosted the same file on port 63145 and changed the parameter *MiningPool* to our host IP and put it on port 62145, then the mining process started successfully.

D DISTRIBUTION OF THE VICTIMS

The geolocation of victims from taken-over pools is shown in Figure 10, which account for 96.7% of all victims, with China having the highest number of victims at 1,245. Among the three most requested mining pools, *apkxm.faye***.org* has a more dispersed distribution, with the largest victim country being Morocco, which accounts for 38.8% of the total number of victims of this pool, and the rest of the countries with a larger distribution are Indonesia, Thailand, El Salvador. most of the victims of pool *www.bs***.co* are located in China, accounting for 97.9%. *data.halo***.club* also controls the largest number of victims in China, with 57.1%, and it is the most widely distributed of the pools taken over, including 42 countries or regions in total.

E SIGNATURES OF ERROR RESPONSES

There is only one success response signature for each type of Stratum implementation, but error responses can vary depending on the implementation. All the error Stratum responses we collected during the scanning are listed on Table 13.

Table 13: Signatures of Stratum error responses.

Protocol	Error response
Stratum-BTC	{"error":[20,"Unknown method",null],"id":1,"result":false} {"id":1,"result":false,"error":[20,"Not supported",null]} {"id":1,"result":false, "error":[24, "Unauthorized workers", null]} {"id":1,"result":false, "error":[25, "Not subscribed", null]} {"id":1,"result":false,"error":[26,"Illegal method",null]} {"id":1,"jsonrpc":"2.0","error": {"code":27,"message":"Illegal params"}} {"error": [20,"Client Pre-authorization was not accepted (incorrect workername, check your settings).",null], "id": 1, "result": false}
Stratum-ETH	{ "id": 1,"jsonrpc": "2.0", "result": null,"error": {code: -1, message: "Invalid login"}} {"error":"Invalid params wrong zil bech32 addr","id":999,"jsonrpc":"2.0","result":false}
Stratum-XMR	{"id": 1, "jsonrpc": "2.0","error": {"code": -1, "message": "Invalid address"}} {"id":1,"jsonrpc":"2.0","error":{"code":-1,"message":"Invalid method"}} {"id":1,"error":{"code":-3,"message":"Method not found"}} {"id":1,"jsonrpc":"2.0","error":{"code":-1,"message":"Invalid address used for login"}} {"id":1,"jsonrpc":"2.0","error":{"code":-1,"message":"Invalid BTC address."}} {"id":1,"jsonrpc":"2.0","error":{"code":-1,"message":"invalid format of user_name"}} {"id":1,"jsonrpc":"2.0","error":{"code":-1,"message":"Please update your XMRig miner (XMRig/0.8.2) to v3.2.0+ to support new rx/0 Monero algo (miner will connect after several attempts)"}}