Typosquatting 3.0: Characterizing Squatting in Blockchain Naming Systems

Muhammad Muzammil, Zhengyu Wu, Lalith Harisha, Brian Kondracki, Nick Nikiforakis
Stony Brook University, New York, USA
{mmuzammil, zhenwu, lharisha, bkondracki, nick}@cs.stonybrook.edu

***Abstract*—A Blockchain Name System (BNS) simplifies the process of sending cryptocurrencies by replacing complex cryptographic recipient addresses with human-readable names, making the transactions more convenient. Unfortunately, these names can be susceptible to typosquatting attacks, where attackers can take advantage of user typos by registering typographically similar BNS names. Unsuspecting users may accidentally mistype or misinterpret the intended name, resulting in an irreversible transfer of funds to an attacker's address instead of the intended recipient. In this work, we present the first large-scale, intra-BNS typosquatting study. To understand the prevalence of typosquatting within BNSs, we study three different services (Ethereum Name Service, Unstoppable Domains, and ADAHandles) spanning three blockchains (Ethereum, Polygon, and Cardano), collecting a total of 4.9M BNS names and 200M transactions—the largest dataset for BNSs to date. We describe the challenges involved in conducting name-squatting studies on these alternative naming systems, and then perform an in-depth quantitative analysis of our dataset. We find that typosquatters are indeed active on BNSs, registering more malicious domains with each passing year. Our analysis reveals that users have sent thousands of transactions to squatters and that squatters target both globally popular BNS domain names as well as the domains owned by popular Twitter/X users. Lastly, we document the complete lack of defenses against typosquatting in custodial and non-custodial wallets and propose straightforward countermeasures that can protect users without relying on third-party services.**

***Index Terms*—Web3, Typosquatting, Blockchain Naming Systems, Ethereum, Polygon, Cardano, NFTs**

I. Introduction

Since the inception of Bitcoin, there has been increased interest in the concept of cryptocurrencies, the blockchains supporting them, and the applications that they enable. Many view the distributed, trustless nature of cryptocurrencies as a welcome alternative to the increased centralization of power and control [1]. In the context of payments, cryptocurrencies offer willing parties the ability to exchange funds without the need for trusted middlemen that can arbitrarily limit transactions. In the context of the web, the so-called "second-generation" blockchains such as Ethereum promise to bring about the next iteration of the web (i.e. Web3 or Web 3.0) [2], [3]. This decentralized web allows the deployment of application logic on public blockchains where it can be vetted, ownership of digital assets that are decoupled from any specific third-party service, and the ability of users to manage their own identity by authenticating themselves using their own private keys.

Given that this concept of identity is critical in cryptocurrencies, researchers and developers soon discovered the need to build layers of abstraction on top of the public-key addresses corresponding to each user's wallet. To avoid reintroducing centralization, Blockchain Naming Systems were developed that not only enable the binding of user-friendly strings to wallet addresses (such as `vitalik.eth` to `0xd8dA6BF269[...]15D37aA96045`) but store the resolution data on blockchains where only their owner can modify them. Today, modern BNSs not only allow the easier exchange of funds between users but also enable other use cases, such as pointing to web content stored on distributed file-storage networks (e.g. on the InterPlanetary File System [4]), resulting in censorship-resistant web applications.

Security researchers have already started studying these new BNS systems, devising threat models and documenting existing types of abuse [5]–[11]. These include hoarding domains for speculation purposes, using takedown-resistant BNS names in the context of malware, domain dropcatching, and squatting trademarks and domains from the traditional web (e.g. attackers owning `google.eth`).

In this paper, we perform the first analysis of *intra-service* typosquatting on popular Blockchain Naming Systems. Instead of looking for which trademarks and domains from the traditional web are being squatted in these BNSs, we focus on attackers registering typo variations of other popular names on the same BNS. This threat model takes into account one of the original uses of BNSs (the exchange of funds between users) and highlights the disproportionate effects of a typo in a BNS, compared to typos in DNS. Whereas a typo in a DNS resolution may require additional social engineering, hosting phishing sites, the downloading of malware, and the exfiltration of sensitive user data, a single BNS typo in the context of a cryptocurrency transaction *guarantees* the loss of user funds. As Figure 1 shows, all that attackers need to do is register typosquatting variations of popular BNS domains and receive the accidental transactions sent by victim users.

We focus our work on three popular BNSs, namely the Ethereum Name Service (ENS), Unstoppable Domains (UD), and ADA Handles (ADAH). ENS and UD are built on Ethereum (with UD also supporting the minting of domains on Polygon), whereas ADAH is built on top of Cardano. We build a corpus of 4.9 million domain names registered across these BNSs and study the levels of intra-BNS squatting activity, using transaction volume as a proxy for domain popularity. We

2159-1245/24/$31.00 ©2024 IEEE
DOI 10.1109/eCrime66200.2024.00013

find tens of thousands of squatting domains across the studied BNSs, targeting as many as 37% of the most popular legitimate domains. We find that BNS users rarely register typosquatting variations of their own domains, which could route funds mistakenly sent through a typosquatting domain to their own wallet. We report on the makeup of the identified typosquatting domains and show that typosquatting registrations increase year over year, with most squatting domains being registered within 100 days of the legitimate domains they target.

Next to the domains themselves, we take advantage of the public nature of the three underlying blockchains to understand to what extent attackers have been successful in stealing funds from unsuspecting users. We find thousands of instances where a sender has sent funds to both a legitimate domain *and* a typo-variation of that same domain, with an average transaction sending $1,790 to scammers. We confirm the typosquatting phenomenon by focusing on popular cryptocurrency "influencers" on Twitter/X and characterize the 74 typosquatting domains targeting the inventor of Ethereum. Lastly, we assess the countermeasures in popular custodial and non-custodial wallets observing a *complete* lack of defenses against typosquatting.

Availability: To encourage more research in the area of BNS security, our dataset of BNS domains, and scripts to collect blockchain transactions are available here [12].

II. Blockchain Name Services (BNSs)

The fundamental concept of BNSs (resolving human-readable domains to addresses) is borrowed from the traditional Domain Name System (DNS). However, there are significant differences between the two. Central authorities like ICANN have control over DNS whereas BNSs are decentralized naming systems that operate on a blockchain network without a single point of control or failure. Moreover, BNS names can be censorship-resistant in that third parties cannot arbitrarily change the resolution address of a BNS domain name.

Ethereum Naming Service (ENS).: ENS [13] is one of the earliest and the most popular BNS, launched in May 2017. It is a decentralized naming system that, at its core, is a collection of smart contracts deployed on the Ethereum blockchain, which uses an account-based transaction model. Smart contracts are self-executing digital contracts with the terms of the agreement directly written into the code. They allow for transactions to occur on the blockchain ensuring anonymity and requiring no form of trust between the sender and the receiver. For example, the Registry contract in ENS maintains a list of all domains and subdomains, keeping track of the owners and resolvers for each domain. Resolvers are also smart contracts that respond to queries about a provided domain name, such as returning the wallet address a domain resolves to. Usually, names use the public resolver deployed by ENS, but users may develop and use their resolvers tailored to their needs. An ENS name is an NFT tied with three main entities: registrant, owner, and controller. The owner of the name is the one who can to change a resolver, expiry date,

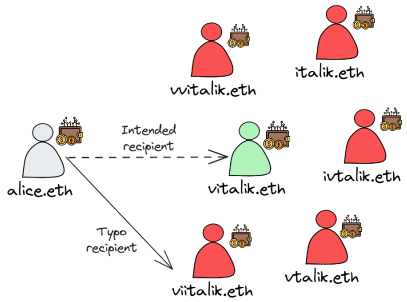

Figure 1: *Attackers can "surround" benign Web3 domains in order to capitalize from typos. Unlike traditional domain squatting, a single typo can result in the immediate and irrevocable loss of user funds.*

create or reassign subdomains, or transfer the name to another address, after which the name will change ownership to the new address. The controller may edit the records of a name. The registrant is the owner of a registration [14]. ENS makes use of a dot-separated hierarchical architecture for its names; names and subdomains registered on it have the extension `.eth`, as in `johndoe.eth` and `funds.johndoe.eth` respectively. A user can register an ENS domain for a certain period after which it will expire unless renewed.

Unstoppable Domains (UD).: UD [15] is a BNS built on the Ethereum and Polygon blockchains, both following an account-based transaction model. Similar to ENS names, UD names are also NFTs controlled by an "Owner" and an "Operator" (known as "Controller" in ENS). Both ENS and UD provide the functionality of holding a variety of records including cryptocurrency addresses, IPFS hashes, and various text records such as URLs, email, and social-media handles. It is important to note that they can be mapped to addresses for more than one cryptocurrency at a time, allowing users to receive payments in different cryptocurrencies using the same domain name. Some users (either speculators or users/developers who have not yet decided how to use their domain names) may opt not to add any resolution records to their domain names. Unlike ENS, once registered, UD domains do not expire (they can always be sold by their owners to other users). At the time of this writing, UD supports more than ten different TLDs including `.crypto`, `.nft`, and `.blockchain`.

ADA Handles (ADAH).: ADAH [16], based the Cardano blockchain, follows a UTXO-based transaction model. Unlike the previously described BNSs, ADAH do not use smart contracts for their operations. Rather, they are stored natively on the Cardano blockchain that supports tokens without the need for smart contracts. This approach circumvents potential issues affecting smart-contract-based BNSs, such as, the need to migrate names to a new smart contract after locating a bug in the current smart contract. ADA Handles are NFTs, and like any other NFT follows a specific minting policy, which is a predefined set of rules governing their creation, distribution, and management. The creators of ADAH publicize the Policy ID they use for this project. An ADA Handle resides inside a wallet address of a user, and payments directed to any handle will be routed to the wallet address in which they reside.

Like UD domains, ADAH do not expire and can be owned indefinitely by a user until sold. The use of ADA Handles is growing [17] but, being a relatively new naming system on a less popular blockchain, the absolute number of registered domains is smaller than the previous two BNSs. ADA Handles also depart from the traditional hierarchical naming of domain names, instead prepending domains with a dollar sign (e.g. `$johndoe`).

III. Motivation and Challenges

In this section, we introduce the idea and importance of studying typosquatting in BNSs along with similarities and differences with typosquatting in traditional DNS.

Typosquatting 3.0.: A traditional typosquatting attack in DNS targets users who mistype a domain name in the URL fields of their browsers. In doing so, they give the opportunity to attackers to control the IP address of that resolution and monetize the user's mistake. That monetization typically comes from redirecting users to phishing sites, affiliate-abuse scams, exploit kits, and social-engineering-based attacks. The attacker has to typically convince the user to either provide sensitive information to the landing page (e.g. for phishing sites and survey scams) or willingly accept the download of malware. Even in the worst-case of a resulting malware infection, attackers still need to somehow monetize the infection, either by exfiltrating sensitive data which they can then sell, or relying on ransomware and botnet activities. In short, a typosquatting attack in traditional DNS *may* result in loss of funds, with users having multiple chances to stop the attack before that happens. Even in rarer forms of typosquatting (e.g. users mistyping a recipient's email address) there are still no guarantees that the missent email will contain any sensitive information.

Contrastingly, when a user is trying to send cryptocurrency funds to another user and mistypes the latter's BNS domain name (depicted in Figure 1), the loss of funds is direct, immediate, and irrevocable. All that attackers need to do is register typo-variations of popular BNS domains and resolve these domains to their own wallet addresses. There is no need to serve phishing sites, host malware, or in any way try to further engage with victims.

Measurement Challenges.: Traditional domain squatting has been extensively studied in past research [18]–[24]. It is therefore tempting to assume that all prior methods used by researchers are applicable to studying squatting in BNSs. We argue that effectively studying squatting in BNSs is, in fact, more complicated than studying traditional DNS-based domain squatting for the following reasons:

Non-availability of ground truth. Every domain-squatting study starts with identifying a list of potential targets for attackers. In past studies, given that typosquatting abuse always occurred in the context of navigating the web, popularity of websites was used as a proxy for popularity of domain names. This typically means selecting the domain names of the most popular websites (e.g. top Alexa [25] or top Tranco [26]) and mutating them to arrive at possible squatting domains. In BNSs however, domain names are predominantly used in receiving payments with most domain names merely pointing to the wallet address of their owners. As such, looking just at on-chain data related to these domain names, there is no clear way to differentiate popular domains (i.e. domains that squatters are likely to target) from unpopular ones.

Aliasing of domains to wallets. Without external sources that can inform the design of a popularity metric for BNS domain names, one reasonable on-chain source of data are transactions. That is, all else being equal, if $addr_1$ has received more incoming transactions than $addr_2$, then it is reasonable to assume that $domain_1$ (i.e. the domain which resolves to that address) is more popular than $domain_2$. The issue with this metric is that it cannot account for multiple domain names resolving to the same wallet address. If $domain_a$ and $domain_b$ both resolve to $addr_1$, there is no on-chain data to conclude which of the two domains is responsible for the most transactions to $addr_1$.

One objective source of off-chain resolution data are software wallets responsible for conducting BNS resolutions. Despite reaching out to multiple wallet vendors with millions of users, those that responded to us informed us that they do not log resolution data and in turn rely on larger platforms for resolutions (e.g. MetaMask relies on Infura for resolutions [27]). At the time of this writing, we have not been able to get into contact with these larger platforms and hence must device our own popularity metrics relying on publicly available on-chain data for the majority of our analysis.

IV. Data Collection

This section describes our data-collection and analysis methodology (shown in Figure 2) and presents an overview of the datasets used in this work.

A. Collecting Names, Addresses, and Dates

Given the public nature of the blockchains supporting the evaluated BNSs, all registration/domain-management events are scattered throughout different blockchain blocks. Because of the volume of data available in these chains and the impracticality of linearly searching all these blocks, various third parties (known as ingestion services) mine these blockchains and extract data and insights which they then make available through blockchain explorers. These blockchain explorers can then be used either manually or programatically via traditional web APIs. As much as possible, we rely on these APIs to obtain the BNS domain names, on which the rest of our analysis is built.

Ethereum Name Service (ENS): Past work analyzing ENS has shown how difficult it is to extract a complete list of registered ENS domain names directly from the Ethereum blockchain (with or without the use of third-party blockchain explorers) [6]. This is mostly due to the use of the *namehash* algorithm, allowing ENS to store domains of all lengths as fixed-length (256 bit) cryptographic hashes. Depending on the smart contract mined by researchers, the extraction can be as straightforward as locating the actual domain name in the

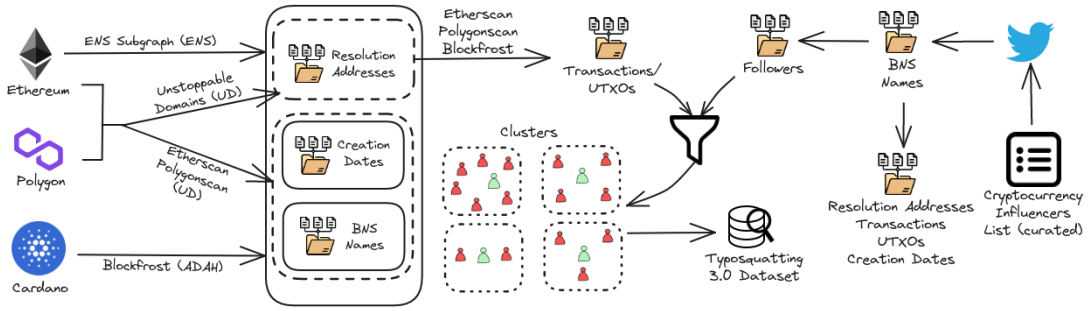

Figure 2: *High-level view of our data collection pipeline and how our analysis interfaces with different APIs and third-party services.*

payload data of a domain-registration event, or as complicated as constructing hashes following the namehash algorithm and then comparing these hashes against the ones controlled by the smart contract.

In 2020, Ethereum developers released "The Graph" a decentralized indexing protocol for organizing blockchain data and making that data accessible using GraphQL (a data-query and data-manipulation language used in APIs) [28]. Different organizations can run different "subgraphs" for indexing different blockchain data. ENS supports its own subgraph [29] which sources events from relevant ENS contracts. In this paper, we leverage this resource to extract ENS domain names, resulting in a dataset that is *significantly* larger than the one that prior work was able to extract directly from the Ethereum blockchain.

Unstoppable Domains (UD): Due to the lack of a GraphQL endpoint for UD, we adopt an alternative strategy to harvest names. Third parties take advantage of the public nature of popular blockchains to index them, and offer access to their indices via blockchain explorers, such as, Etherscan [30] and Polygonscan [31]. UD names, unlike ENS names, are minted on two different blockchains: Ethereum and Polygon. Etherscan and Polygonscan facilitate data retrieval from these respective blockchains, allowing us to query the relevant smart contracts. We focus on querying the smart contracts listed by UD for both Ethereum and Polygon. UD lists separate registry addresses for Ethereum and Polygon minted names, which we utilize to crawl domain names using Etherscan and Polygonscan respectively. Additionally, we leverage another registry smart contract listed on Etherscan that contains a substantial amount of UD names with the `.crypto` TLD, all minted on the Ethereum blockchain. Using the two block explorers, we query each event log on every block that involved the use of the three mentioned smart contracts. As with any blockchain transaction, the registration of a UD name will also appear as an event log, and will contain the name itself as an argument to the smart contract function. We supplement this data with additional data from the API offered by Unstoppable Domains itself [32] to extract the resolution addresses of each domain name.

ADA Handles (ADAH): To collect ADAH data, we make use of the Blockfrost API [33], an open-source API project that aids in accessing and processing information on the Cardano blockchain, such as addresses, blocks, assets, and transactions. Using the publicly available policy ID for ADAH, we use the "asset" endpoint of the Blockfrost API (used to query information regarding NFTs) to collect ADA Handles minted on-chain along with their creation dates and the wallet addresses in which they are located.

B. Obtaining Transactions

For both ENS and UD, we use Etherscan APIs to collect transactions on the Ethereum blockchain. We provide all the Polygon and Ethereum resolution addresses as inputs to the API calls, upon which they output both incoming and outgoing transactions that each particular resolution address was involved in. Each transaction contains a sender address, receiver address, amount of ETH sent, the transaction hash, and the timestamp. For UD minted on the Polygon blockchain, we use Polygonscan APIs which is queried exactly as the Etherscan APIs. To collect transactions for ADAH, we use the Blockfrost API to first collect the transaction hashes for each Cardano address in our dataset in which at least one ADA handle resides. For each identified transaction hash, we then collect all associated UTXOs. In the case where we have multiple input addresses in a UTXO for one output address, we count the number of transactions as the number of input addresses and the value as the amount that goes into the output address divided by the number of input addresses.

C. Dataset Overview

Table I provides a summary of the number of domain names that we collected for each BNS along with the total number of resolution addresses (i.e. the domain names can be resolved to wallet addresses by the BNS) and transactions. We also report the number of names that have a resolution address. Drawing parallels with traditional DNS, a domain name may be absent from the zone records of an authoritative server for a specific TLD (i.e. it cannot be resolved to an IP address) but it is owned by someone and unavailable for registration. We note how many ENS names have an "ETH" resolution address set and how many UD names have an "ETH" or a "MATIC" address set. Since ADAH resolution addresses are the addresses of the wallets that they reside in, all ADAH had resolution addresses.

We use the described data collection methodology to collect the largest dataset for BNSs to date. We obtained 97% ENS

BNS	# Names Collected	# Resolvable Names	# Resolution Addresses	# Transactions Collected
ENS	3,047,188	2,214,012	699,548	140,183,178
UD	1,707,017	1,156,697	393,290	44,819,348
ADAH	198,121	198,121	64,612	15,487,746
Total	*4,962,326*	*3,568,830*	*1,157,450*	*200,490,272*

Table I: *Overview of the extracted domains along with the number of resolution addresses and their corresponding transactions.*

Model	Variation
Duplication	jjohndoe.eth
Addition	johndoew.eth
Removal	johndo.eth
Swapping	johnode.eth
Substitution	nohndoe.eth
Hyphenation	john-doe.eth
Pluralization	johndoes.eth

Table II: *Typosquatting models and variations for* `johndoe.eth`

names from the ENS Subgraph with only 96,667 returning empty API responses. Because of our use of the Subgraph (as opposed to trying to identify ENS domain names through blockchain explorers), we manage to collect 23% more `.eth` names than Xia et al. [6] until block 13,170,000, which was the cut off for their data collection. Due to API call failures for resolving UD, we were not able to collect resolution address records for approximately 10K UD registered names, which means that we managed to collect over 99% of UD names that have ever been registered, with 100% recovery rate for ADAH.

Overall, we were able to collect almost 5 million domain names registered across the three evaluated BNSs which we analyze in the rest of this paper.

D. Establishing Ground Truth

To investigate the prevalence of typosquatting activity using our datasets, it is important to first identify the legitimate names that attackers are likely to be targeting. However, as discussed in Section III, establishing this ground truth for BNS names is non-trivial because of the unique aspects of BNS domain names, compared to traditional DNS domain names.

We rely on two heuristics to establish ground truth (i.e. a set of legitimate domains that attackers are likely to target via squatting): i) we assume that a resolution address with a significant volume of transactions is owned by a benign user (or at least a user that attackers will target), and ii) a large number of domain names all pointing to the same resolution address constitutes suspicious behavior, suggesting either domain speculators, or squatters who are expecting to capitalize typos when resolving BNS domain names. By combining these two heuristics, we can classify all resolution addresses using the following formula: w_T/w_D, where w_T represents the number of transactions linked to each domain's resolved wallet address, and w_D denotes the number of domain names it is linked to. Note that this sorting is based on the transactions of the cryptocurrency resolution address, and not the owner address. Our metric allows us to prioritize the analysis of popular domain names (using the number of transactions as a proxy for their popularity) while penalizing wallets that contain multiple domain names (by uniformly distributing the total number of transactions of that address to all N domains).

E. Cryptocurrency Exchange Addresses and Token Contracts

Through our analysis of the sorted domain names, we discovered a number of BNS names that point to the known wallets of central cryptocurrency exchanges (like Coinbase). Given the popularity of these wallets and their volume of transactions, these domain names ranked near the top of our list. Yet, we can clearly infer that the popularity of these wallet addresses has little to do with the domain names (whether these domain names are managed by the exchanges themselves, or by ENS users who misconfigured their resolution settings is outside the scope of this work). Similarly, we located a number of domain names that resolved to addresses of token contracts as opposed to regular wallet addresses. Given the diversity of these token contracts, we cannot make any definitive general claims about how these domain names are used in the context of transactions.

As a result, we filter any addresses (along with the domain names resolving to these addresses) that belong to known central-exchange wallets and token contracts, as labeled by Etherscan and Polygonscan. Unfortunately, we could not identify a trustworthy source of labeled exchange wallets for Cardano hence we did not filter our ADAH addresses. This lack of filtering is somewhat mitigated by the lower popularity of ADAH which, as shown in Table I, has an order of magnitude fewer domain names registered.

F. Identifying Typosquatting Domains

We define a typosquatting name of a particular BNS name as a one belonging to the categories listed in Table II. These typo models have all been established by prior work in traditional typosquatting [18]–[24].

Given a list of target domain names (which constitute the top N domains after we sort the total list of domain names given our popularity metric), we generate all typosquatting variations using the aforementioned typo models and identify which ones have already been registered. We cluster these typosquatting domains together, according to the domain name that they target. When domain names are short, the likelihood of accidental squatting increases. For example, assuming that `abc.eth` is a popular ENS domain, is the less popular `abv.eth` a targeted squatting attack, or a legitimate domain that happens to be syntactically close to the first one? By manually investigating pairs of syntactically-close domain names, we empirically set a lower analysis limit to 5 characters. That is, popular domains that are shorter than that threshold are excluded from the rest of our analysis. Finally, to further reduce false

BNS	# Legitimate Names	# Legitimate Names having atleast one typo	Total Typosquatting Names
ENS	10,711	3,920 (37%)	25,396
UD	12,026	701 (6%)	1,137
ADAH	931	191 (21%)	396

Table III: *Overview for legitimate-to-typosquatting name clusters*

positives, we discard any discovered squatting domains whose registration date predates the corresponding target domain. While these are within a single-character distance of the target domains and may very well end up receiving funds intended for other domains, they do not conform to our threat model of attackers registering domains for the express purpose of squatting existing popular ones.

V. Analysis

In this section, we discuss our findings we extract through comprehensively analyzing our datasets for each of the three BNSs.

A. Typosquatting Analysis

Given that both the Ethereum Name Service (ENS) and Unstoppable Domains (UD) have attracted millions of domain registrations, we choose the top 10K wallets of our ordered list, as the ones that could be targeted by squatters. We refer the domains pointing to these wallets as "legitimate" domains, only to differentiate them from typosquatting domains. Given that vast majority of these domains are not attached to known companies and individuals, we cannot be certain that they are associated with legitimate activity. Since ADA Handles (ADAH) have an order of magnitude fewer registrations, we also pick an order of magnitude smaller set of targets (i.e. top 1K domains).

Table III provides a high-level overview of the sets of legitimate domains that we selected across all BNSs, along with the number of typosquatting domains targeting them. For both ENS and UD, we observe that there are more legitimate domains than wallet addresses, since a few wallets "hold" more than one domains. For ADAH we observe the opposite where some domains were filtered out of our analysis due to their short length. In terms of typosquatting activity, ENS is clearly more targeted than the other two BNSs which is intuitively correct given that it is the oldest and most popular BNS of the three.

Number of Typosquatting Registrations.: Figure 3 shows the number of legitimate names with one or more typosquatting registrations against them. Our analysis reveals that a significant number of names are being targeted more than once. The ENS name targeted the most was `mickey.eth` with 102 typosquatting names. Similarly for UD, `cryptoden.blockchain` attracted 19 typosquatting names, and `$xn--bs8h` for ADAH with 16 typosquatting names. The ADAH domain is clearly an internationalized domain name (IDN) which corresponds to an emoji when viewed in UTF-8. We assume that not all users can type IDNs in their native format hence we consider their ASCII-based Punycode representations as in scope for squatting.

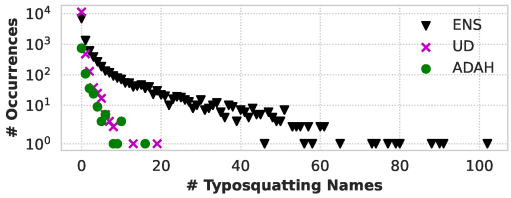

Figure 3: *Frequency of the number of typosquatting registrations against legititmate names*

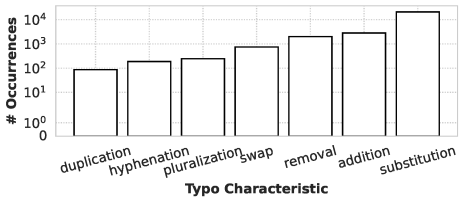

Figure 4: *Frequency of the different misspelling strategies used by typosquatters to form spelling variations of legitimate names*

Typosquatting Characteristics and Volume of Activity.: Applying different typo models to the same target domain results in tens if not hundreds of possible typosquatting domains. In terms of domains that squatters have chosen to register, Figure 4 presents the models that are most popularly used across all studied BNSs. We observe that the character-substitution model is the one most commonly employed by attackers, followed by character additions, removals, and swaps. In 2015, Agten et al. reported that, in traditional domain squatting, attackers appear to prefer character omission, over introducing additional characters [20]. We also report that 24% of legitimate domains are altered in the first character to produce the typosquatting domain.

Figure 5 shows the fluctuation in prices of Ethereum, Matic, and ADA, along with the number of typosquatting names registered over the years. The trends show that all three BNSs have attracted a rapidly increasing number of typosquatting names each year, with a sharp increase in the registration of names after 2021. In 2021, the prices of all three cryptocurrencies reached their maximum exchange rates, which could be a possible explanation for this drastic increase in registrations of typosquatting names. There is no observable slowdown in registrations since 2021 and we expect sustained registrations as cryptocurrencies further gain in popularity.

Defensive Registrations.: Like in traditional domain names, users of BNSs may proactively register typosquatting variations of their own domain names to protect themselves from the types of attacks discussed in this paper. In the aforementioned 2015 paper, Agten et al. had discovered that 156 of the 500 evaluated top Alexa domains (i.e. 31.2%) were engaging in at least one defensive registration of typosquatting domains [20]. In our case, we note a substantially lower rate of defensive registrations, with just 13 ENS legitimate domains (out of the 10,711 selected domains) owned by the same wallet address as at least one typographic variation of itself. We were not able to identify *any* defensive registrations

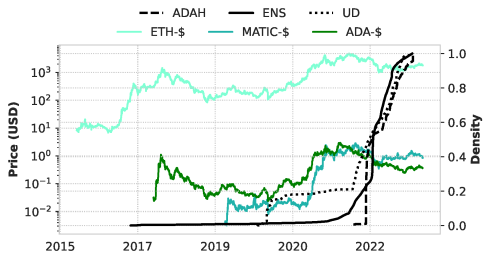

Figure 5: *Correlation between the number of typosquatting names registered each year with the price of each cryptocurrency*

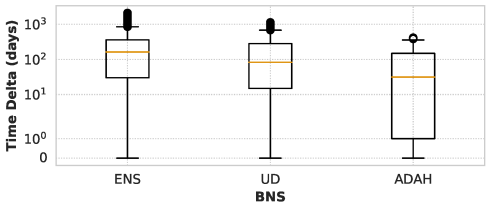

Figure 6: *Time difference (in days) between the registration of the legitimate and typosquatting names*

in UD and ADAH. One possible reason behind this lack of defensive registrations is the current absence of large commercial interest in these BNSs. Companies with billions of dollars in revenue can afford the recurring registrations costs of defensive registrations whereas individuals (even if they are aware of the issue) may not be able to sustain these costs.

Trends of Typosquatting Registrations.: Next, we look at the registration-time deltas between the legitimate name and typosquatting names, the results of which are shown in Figure 6. We note that the median time difference is ≈100 days, which could be because of the time it takes for a legitimate name to become popular enough to attract squatters. However, it is important to note that the number of typos registered at the very same day of the original name is non zero. One potential explanation for this behavior is that speculators observe the registration of a domain name and predict that it may eventually become popular, thereby immediately registering typo variations of it. One source for these predictions could be social media, where accounts with large followings obtain one or more BNS domain names. Since they are already popular, attackers can immediately register variations of their domain names. We analyze the phenomenon of BNS squatting in relation to social media popularity in Section VI.

Behavior of Typosquatters.: Furthermore, we explore the behavior of squatters, specifically on how they choose their targets. We analyze the number of times a squatter (as identified by their wallet address) has registered a typosquatting name for the same legitimate name, and the number of unique names a typosquatter has registered a typosquatting name for. Figure 7 shows our results. The attacker that has targeted the most typosquatting names from ENS has targeted a total of 46 legitimate names, registering a total of 90 typosquatting names, and holds 1,368 domains at the time of our study.

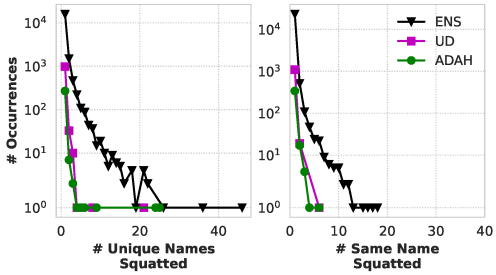

Figure 7: *Number of unique names targeted by the same squatter (Left). Number of times a given name is targeted by a squatter (Right)*

Figure 8 is focused on the top 50 ENS attackers, in terms of their registration activity. There we see not just that attackers register tens of squatting domains related to the legitimate ones in our study but that they hold hundreds (and occasionally thousands) of other domains in the same wallets.

Of all the typosquatting wallets, 71% of ENS, 87% of UD, and 66% of ADAH wallets have registered at least one other domain. Since a wallet address can register both UD and ENS names, we also observe that 25% wallets that had registered UD names, had also registered ENS names, and 7% vice versa, indicating that squatters are registering multiple domains across different BNSs as well. The leftmost subplot in Figure 9 shows that squatters are indeed registering multiple other domains, indicating suspicious behavior, with the maximum number of domains registered by a single wallet being 3,099 for ENS, 6,963 for UD, and 1,477 for ADAH. To further understand the motivation behind this mass registration behavior, we compute the average cosine similarity score of each typosquatting name with all other domains registered by its address (one-to-many), and the average cosine similarity of all names registered by the squatter (many-to-many). Cosine similarity values range from 0 to 1, with a higher score depicting a high semantic similarity. These scores are calculated by using word embeddings obtained using the `bert-large-uncased` model [34], which is currently state-of-the-art for natural language processing tasks. Using manual inspection, we decide on thresholds to categorize whether two domains are semantically similar in meaning or not, where a score of $x < 0.5$ indicates that the domains are not similar, $0.5 \leq x < 0.75$ indicates moderate similarity, and $x \geq 0.75$ indicates high similarity. Figure 9's middle and rightmost subplots show our results, where the average cosine similarity trend suggests that names registered by typosquatters are more often than not different in their meaning. From this analysis, we notice that there are three squatting styles among attackers: *i)* registering a single typosquatting domain, *ii)* registering multiple domains for different targets (most popular style of attack in our data), and *iii)* registering multiple domains with similar targets. Examples of each type are shown in Appendix A.

B. Transaction Analysis

Next, we investigate how successful attackers have been with their typosquatting registrations in BNS. Our aim is to

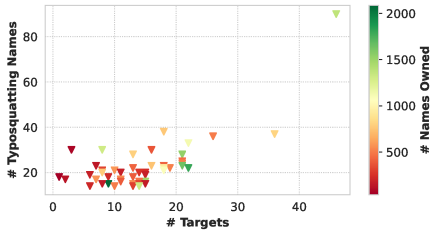

Figure 8: *Number of targets vs. number of typosquatting names for Top 50 ENS attackers*

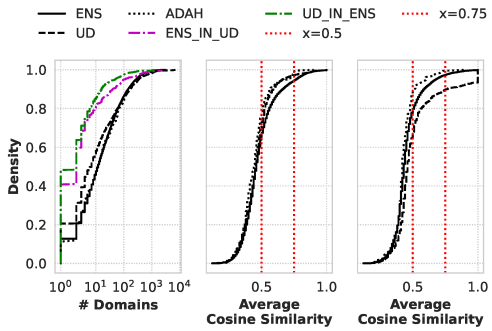

Figure 9: *Distributions for the number of other domains owned by typosquatting wallets (left), the average cosine similarity of typosquatting names and other domains owned by typosquatting wallets (middle), and the average cosine similarity of all domains owned by typosquatting wallets (right)*

identify the number of times users have mistakenly sent funds to an attacker-controlled wallet through a typosquatting BNS name. We note that inferring the intent of each transaction is difficult, since on-chain data just reveal the transactions themselves and not which domain the sender used (aliasing issue discussed in Section III), let alone which BNS domain a user intended to type. This lack of data can lead to false positives when performing our analysis, as funds going into an attacker-controlled wallet address might not necessarily be through the typosquatting BNS name, but rather through another BNS name resolving to the same wallet address.

To counter these possible false positives, we present two different analyses. First, we report on all transactions that these typosquatting wallets received, with the caveat that some of these transactions may have occurred through domains other than the identified typosquatting ones (i.e. domains that point to the same squatter wallets). Second, we search for senders who have sent funds to both the legitimate *and* the squatting domain, indicating that a mistype was involved in one of their regular transactions. To this end, we extract the transactions of the addresses that resolve from the legitimate name and the corresponding typosquatting names and find the senders that have initiated transactions to both addresses. We report on how many custodial and non-custodial common senders have sent assets to both, the amount of funds that were sent to the squatting address, and the time difference between the transaction to the legitimate address and the squatting ones. We note that the results of this analysis are a lower bound of

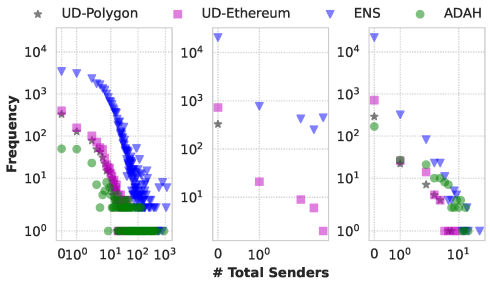

Figure 10: *Frequency of the total number of victims attracted by a unique typosquatting name (left), number of common senders (victims) that have sent assets to a pair of legitimate and typosquatting name through a custodial address (middle)/non-custodial address (right)*

the transactions that typosquatting domains attracted since we ignore transactions from senders who did not attempt to send any funds to the corresponding legitimate domain.

On-chain data also lacks relevant insights needed to distinguish whether a transaction was initiated by entering a BNS name or the actual wallet address. At the time of our study, only Coinbase, among all custodial wallets, offers a name resolution service, supporting ENS and UD domains. This suggests that transactions processed via custodial wallets other than Coinbase most likely occurred through direct wallet address rather than BNS name resolution. We exploit this characteristic as a mechanism to filter out such transactions, which accounted to 16% of all transactions to typosquatting domains.

Victims.: Figure 10 provides three different types of analysis we perform to report the number of transactions towards the typosquatting domains in our dataset. The leftmost figure highlights the frequency of the total number of senders/victims attracted by a typosquatting domain. There are 6 typosquatting ENS domains held by a single wallet address which have attracted the highest number of victims, i.e. 1,141. At the time of this study, this wallet holds a total of 112 ENS domains that are all either hexadecimal or decimal numbers. Examples that have been flagged as typosquatting domains are `0x1861.eth`, `0x1862.eth`, and `0x1863.eth`, which are targetting the domain `0x186d.eth` (linked with a wallet that has over 2,000 transactions, and only holds that domain). Similarly, 796 senders have sent assets to 8 different typosquatting names in our dataset that are linked to the same wallet. This wallet owns a total of 99 domains that are all decimal numbers, examples include `93813.eth` and `93815.eth` targetting `95813.eth` (over 5,600 transactions and linked with only one other domain).

Since it is plausible that the transactions were initiated through a domain name not flagged as a typosquatting model but owned by the same wallet, we also report our findings using the common senders approach in the middle and rightmost plots of Figure 10. In total, we find 483 pairs of legitimate and typosquatting ENS names, 100 for ADAH, 51 for UD on the Ethereum blockchain, and 38 for UD on the Polygon blockchain that have received transactions sourced through a

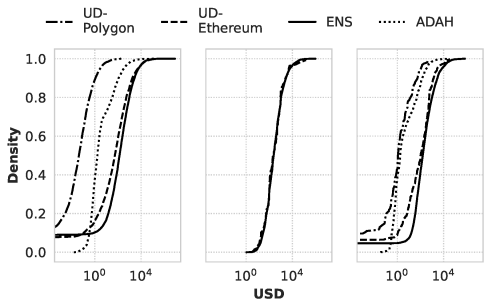

Figure 11: *Amount in USD sent to addresses of typosquatting names through the set of all senders (left), common custodial-senders (middle), and common non-custodial senders (right)*

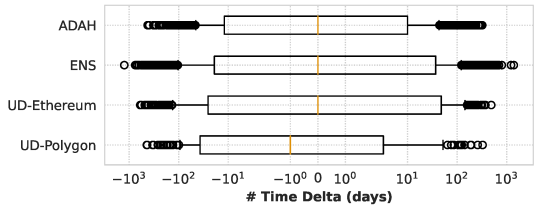

Figure 12: *Distribution of time (in days) between a transaction to a legitimate name vs a transaction to a typosquatting name through a common custodial or non-custodial wallet address*

non-custodial wallet. In terms of Coinbase (custodial-wallet) sourced transactions, we find 1,899 such pairs for ENS, 38 for UD on Ethereum blockchain and 0 for UD on the Polygon blockchain. The top three pairs with most number of transactions from common senders all belonged to ENS with the top pair being `qwerky.eth` and `qw3rky.eth` (target) receiving transactions from 28 common senders.

Funds sent to squatters.: Figure 11 shows the funds in US dollars that were sent by the set of all senders (left), as well as common senders through custodial and non-custodial wallets to the wallet addresses of the typosquatting names (middle and right). To convert the amount of each cryptocurrency into USD, we utilize the transaction timestamps along with the closing conversion rate for the day of the transaction. Our analysis reveals that the registration of typosquatting names have been successful for attackers in the context of extracting funds from users. From all transactions sent to typosquatting domains, the average amount of funds sent was $1,790 ($127 median). The average loss through senders that had sent transactions to both the legitimate and typosquatting names in custodial wallets was $1,999 ($241 median) dollars compared to $1,862 ($95 median) in non-custodial wallets. This finding supports our distinction between traditional typosquatting and BNS-based typosquatting where a single typo can now cost users thousands of dollars in lost funds.

Differences in transaction times.: Lastly, we calculate the time difference between the transactions of a common sender to a legitimate name vs. its typosquatting equivalent. Figure 12 shows that the median time difference is approximately one day and the interquartile range of all transactions is less than $\pm$100 days around the origin. Both positive and negative differences are intuitive since users may mistype either in their first transaction, or a later one.

VI. Case Studies

In this section, we complement the high-level results of Section V with an analysis of BNS typosquatting activity against popular Twitter/X users, as well as against the inventor of Ethereum.

A. Squatting cryptocurrency influencers

It is common for BNS users to publicize their names on their Twitter/X profiles for branding (e.g. associating themselves with a specific technology) or simply to enable their followers to send them cryptocurrency payments. From an attacker's point of view, these Twitter-associated BNS names are ideal typosquatting targets, particularly when their legitimate owners are influencers with millions of followers and thereby a large pool of prospective victims. From our point of view, these influencer-associated BNS names give us an opportunity to verify our earlier observations since Twitter allows us to objectively establish account popularity (and thereby likelihood of targeting specific BNS domains) without restricting ourselves to on-chain data.

Data Collection.: We perform our analysis of typosquatting against cryptocurrency influencers as follows: we start by collecting popular influencer accounts in the cryptocurrency space on Twitter. We curated two such lists, one of 187 popular influencers in the cryptocurrency space from the top 3 Google search results in January 2023 for the search query "Top 100 Cryptocurrency Influencers on Twitter". Since ADAH were not sufficiently represented in the resulting list, we searched for the keywords "adahandle" and "cardano" on Twitter and selected 120 additional legitimate accounts.

Given this initial list of influencer accounts, we use the Twitter API to access the accounts that they follow, and the accounts that their "followers" follow in search for other popular accounts that advertise BNS domains in their profiles.

In total, we manage to collect 14,910 ENS names, 733 UD names, and 1,849 ADAH names from Twitter accounts. In terms of their footprint on social media, ENS is significantly more popular than the other two, while UD is the least. We perform another step to narrow down and select the BNS names that were found in the profiles of users with large numbers of followers (and thereby large numbers of potential victims for squatters). We create two lists, one comprising of a total of 1,000 ENS and UD domains belonging to popular accounts, and the other consisting of 100 ADAHs. Our decision to combine the ENS and UD domains list was inspired by the fact that very few UD names in our dataset belonged to accounts with significant numbers of followers. Figure 14 (Appendix A) shows the profile of the most popular user on Twitter among those who advertise their ENS domains with 13.1 million followers. Figure 15 (Appendix A) shows the number of followers of users collected before and after we filter out unpopular accounts. We obtain all relevant addresses and transactions using the same methodology as the one in our Section IV analysis.

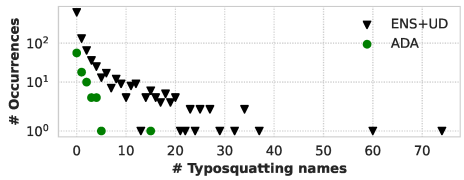

Figure 13: *Frequency of the number of typosquatting registrations against legitimate names popularized on Twitter/X*

Analysis.: Our findings from analyzing this Twitter dataset suggest that typosquatting activity for such BNS names is popular among attackers. Specifically, we indentified 86 typosquatting names related to ADAH and 2,116 tied to ENS and UD names. Of all legitimate ENS and UD names, 38% were targetted at least once; this percentage was marginally higher (40%) for ADAH. Figure 13 shows the frequency of the number of typosquatting names targeting legitimate names, showing a similar trend as that of Figure 3. For example, the ADAHs `$bullion` and `$villion` are registered against the legitimate name `$billion`. A total of 17 names target the ENS name `hayden.eth` (inventor of the Uniswap protocol), and a few examples of typosquatting names include `hayd3n.eth`, `haydin.eth`, and `hyden.eth`. UD domains in the Twitter dataset are overall targeted less than the ones in Section V, with some attack instances against popular accounts. For instance, `metascan.nft` is the BNS name owned by the popular Web3 service, MetaScan, and has attracted 3 typosquatting names: `metascam.nft`, `metacan.nft`, and `m3tascan.nft`.

The three most targeted names in our Twitter dataset from each BNS were `vitalik.eth` (further analysis in Section VI-B), `metascan.nft`, and `$crypto`. We hypothesise that since such influencers are aware that their BNS are well-known, they could be more careful and might have defensive registrations for their domains. However, we find that only 16 ENS users and 2 ADAH users have one defensive registration each against their BNS names. We report that 85 wallet addresses hold more than one typosquatting names that target the legitimate names in our dataset. A single wallet holding 9 domains that are typos of popular domains in our Twitter dataset holds in total over 800 domains.

After performing the same transaction analysis of typosquatting names on this dataset, we find that Twitter-targeting squatters have been just as successful as the ones studied in Section V. When all transactions to typosquatting names are considered, an average of $2,277 per transaction ($140 median) is transferred to typosquatting names. When focusing just on transactions from senders who have sent funds to both legitimate and typosquatting domains, an average of $1,192 ($196 median) per transaction is transferred through Coinbase and $174 ($0.5 median) through non-custodial wallets to typosquatting domains. We also find that 297 pairs of legitimate and typosquatting ENS and UD names have received at least one transaction through a common custodial sender (Coinbase), and 55 through a common, non-custodial wallet. For ADAH, 8 pairs received funds through the same non-custodial sender.

Typosquatting Name (Number of names linked to the same wallet)	Funds (in USD) sent to squatter
fitalik.eth (14)	33,310
vitalak.eth (17)	24,144
v1talik.eth (286)	7,109
italik.eth (103)	3,773
vytalik.eth (115)	1,523

Table IV: *Typosquatting names that had Coinbase as a common sender from `vitalik.eth`, and the funds sent to them.*

B. Squatting against Vitalik Buterin

In this case study, we report on the typosquatting activity against Vitalik Buterin, the inventor of Ethereum. Vitalik Buterin owns `vitalik.eth` and advertises it on his Twitter profile. Given the association between Buterin and Ethereum, `vitalik.eth` is the most squatted domain in our Twitter dataset, with a total of 74 typosquatting names registered against it resolving to 66 unique wallet addresses.

Typosquatting Name Strategies.: Among the typosquatting names registered, the most popular typo model was that of character substitutions, with 40 typosquatting instances. A few examples of this are `vitqlik.eth`, `vlitalik.eth` and, `vigalik.eth`. Other strategies involved substitutions that convert the name into one that sounds similar (known as soundsquatting [35]) including `witalik.eth`, `vetalik.eth` and, `vitalyk.eth`. Less popular strategies included the removal of characters (`italik.eth`), swapping two characters (`vitalki.eth`), pluralization (`vitaliks.eth`), character duplication (`vitalikk.eth`), and hyphenation (`v-italik.eth`).

Temporal Analysis.: Typos targeting `vitalik.eth` are constantly being registered starting from just four days after the registration of the original domain (`vitalikb.eth` was the first typosquatting domain registered in 2017). The name that was registered most recently in our dataset is `v-italik.eth` in December 2022. The number of typosquatting names registered each year is also increasing, consistent with the overall typosquatting growth observations in Section V. Namely, 41 typosquatting domains targeting `vitalik.eth` were registered in 2022, compared to just 18 in 2021.

Transaction Analysis.: Lastly, we perform a transaction analysis, finding that only custodial wallets were used to send funds to *both* `vitalik.eth` and a typosquatting name. We find 108 transactions where funds have been transferred to typosquatting names using only Coinbase as the source wallet. Table IV shows the domains these funds were sent to and the amount of funds that has been sent. It highlights not only that the various typosquatting techniques utilized by attackers have been successful, but also that the typosquatters of `vitalik.eth` have also registered hundreds of other names either hoping to sell them for profit in the future, or capitalizing on typos for multiple accounts. As in all cases involving custodial wallets, we cannot know exactly how many individual users were behind these transactions.

BNS	Custodial	Non-Custodial		
ENS	Coinbase	Metamask v10.32.0	Bitcoin.com v8.5.1	Alpha Wallet v3.65
UD	Coinbase	Atomic Wallet v1.11.5	Bitcoin.com v8.5.1	Alpha Wallet v3.65
ADAH	N/A	Eternl v1.10.10	Nami Wallet v3.5.0	Typhon Wallet v2.5.6

Table V: *Digital wallets used in cold/warm typos experiments*

VII. Discussion

In this section we describe the overall limitations of our study and assess the presence of typosquatting countermeasures in wallets and exchanges. Based on our results, we propose some directions for future work in this area.

A. Limitations

The limitations of this study are grounded on the two challenges listed in Section III, i.e., the lack of ground truth regarding domain popularity and the aliasing of multiple domains to the same wallet address. These limitations make it difficult for us to categorically state that each and every transaction sent to a typosquatting variation of a legitimate domain name was the result of a typo. There is insufficient on-chain information to differentiate between a user consciously sending funds to a domain other than a typosquatting domain name, if these two are resolving to the same wallet address. At the same time, we argue that even if some of the identified typosquatting transactions are false positives, this paper sheds light to the overall typosquatting problem in BNSs. Without such a study, centralized exchanges and wallet providers can never try to solve an issue that they do not know they have.

We approached these fundamental limitations through a series of conservative filters, focusing on the senders that have sent funds to pairs of legitimate/typosquatting domain names and thereby excluding one-off transactions that could very well have been the result of typos. Similarly, regarding centralized exchanges, we focused just on Coinbase-originating transactions since that is the only exchange that supports ENS/UD resolutions and thereby the only exchange where transactions could have realistically been the result of a typo. Lastly, our Twitter/X case study confirms the overall problem of typosquatting since we see similar levels of typosquatting activity even when we change our approach for identifying legitimate domains, i.e., domains that attackers target with typosquatting registrations.

B. Defenses and future work

On the traditional web, public DNS servers will resolve all registered domains (both legitimate as well as squatting ones) but some browsing software may warn users about potential typos, requiring confirmation before actually navigating to the website hosted on a suspicious domain [36]. One may wonder whether similar countermeasures exist in the systems that support BNS resolutions, namely wallet software for non-custodial wallets and the web applications operated by centralized exchanges.

Experiments.: To assess whether these defenses exist, we perform the following two experiments on a centralized exchange that supports BNS resolution as well as on popular wallet software that resolves domains for all three evaluated BNSs:

Cold typos: Attempt to send funds to a squatting domain without any other related interactions. Do exchanges and wallet software operate global squatting-related blocklists?

Warm typos: Attempt to send funds to a squatting wallet *after* having sent funds to the legitimate (i.e. squatted) domain. Do exchanges and wallet software operate local (user-specific) squatting-related blocklists?

We use the wallets and custodial exchanges shown in Table V. Coinbase is, to our knowledge, the only centralized exchange that supports the resolution of ENS and UD domains at the time of this writing. No exchanges currently support ADA Handles. In terms of non-custodial wallets, we select popular wallets advertised by each respective BNS. For example, the Metamask wallet is installed by more than 10 million users, just on the Google Chrome store. For our cold-typos experiment, we simply send a minimal amount of cryptocurrency to typosquatting names of popular names from each BNS and note the presence of any warnings or errors. For our warm-typos experiment, we first send funds to a legitimate address resolved through a popular BNS domain and then send funds to a typosquatting variant of the same domain.

Ethical considerations:: We are familiar with the ethical issues surrounding the sending of funds to attackers, as part of research experiments. For all experiments, we sent the smallest amount of cryptocurrency that the exchange/wallet software would allow us. For example, while Coinbase has a minimum transaction amount of 0.001 ETH, we were able to submit significantly lower transactions through our non-custodial wallets.

In this, we follow prior security studies where authors make modest payments to attackers and underground economies as a way of shedding light to the studied malicious ecosystems. In past work researchers have, among others, sent small amounts of funds to cyber criminals in order to understand how they construct fake Twitter accounts [37], operate CAPTCHA farms [38], attempt to compromise users [39], and create artificial backlinks [40]. As with this prior work, we argue that these small payments are acceptable if the value of the gained insights outweigh the funds sent, and no alternative method was available.

For this paper, we sent a total of $12.52 spread over multiple attackers, a vanishingly-small amount considering the actual typo-payments that these attackers are stealing from victims (Section V-B). These were unavoidable since we lack access to the logic of centralized exchanges and resolution-services of software wallets, meaning that unless we try to send a small payment, we cannot know if there is some resolution-level/transaction-level mechanism that will stop that payment from going through.

Results.: All variations were allowed to go through without any warnings. The only exceptions are Eternl and Bitcoin.com wallets which show a warning after *every* name resolution (for both legitimate and typosquatting domains) reminding users that it is their responsibility to confirm the addresses that they are sending funds to, as all transactions are irreversible. Figure 16 (Appendix A) illustrates the screen views from Metamask when sending funds to a typosquatting domain. These results highlight that there is ample room for deploying global and local defenses to protect cryptocurrency users. Even without any trusted third parties (i.e. along the overall ethos of cryptocurrencies and trustless P2P payments), software wallets can keep local lists of the domains that their users have resolved and sent funds to, calculate typosquatting variations using well-known typo models, and warn users if they ever try to send funds to these variations in the future.

VIII. Related Work

To the best of our knowledge, this is the first study that investigates the issue of intra-squatting in modern BNSs, i.e., users of a BNS targeting other users of the same BNS. In this section, we briefly describe the limited related work in the BNS space, as well as how our work compares to squatting research in traditional DNS domain names.

The work that is most closely related to ours was the ENS measurement study by Xia et al. [6] in which the authors present a systematic analysis of the Ethereum Name Service, its growth, and the different types of attackers that it has attracted. In terms of squatting, the authors focus on *inter-system* squatting, i.e., evaluating to what extent attackers are squatting on popular domains and trademarks that do not belong to them (e.g. `google.eth` and `facebook.eth`). The threat models behind these types of squatting domains are distinctly different from the ones we explored in this paper with attackers hosting malware or intending to sell the squatting domains back to their original trademark holders, as opposed to monetizing typos occurring during the exchange of funds. Similar to that work, Patsakis et al. [7] describe possible squatting attacks on decentralized blockchain-based naming services focusing on NameCoin and EmerCoin, two blockchain-based name services that behave like traditional DNS, resolving domain names to IP addresses. The authors evaluate the level of trademark squatting in these two services, but do not explore intra-chain squatting, which is the focus of this paper. Kalodner et al. [8] explore the prevelance of name-hoarding in NameCoin's `.bit` TLD domains, where users purchase domains with the intent of speculative resale at higher prices. They develop techniques to analyze the transfers of these domains from one user to another and find that at the time of their study, transfers of these domains was rare. Muzammil et. al [5] investigated dropcatching attacks on ENS domains, where attackers can re-register expired ENS domains to attract transactions that were meant for their previous owners. Apart from BNS attacks, various types of scam activities in the Web3 space has caused significant amounts of financial losses [41]–[52].

In traditional DNS, cybersquatting (individuals registering trademarks not belonging to them) and typosquatting (registering typo-variations of existing popular domains) can be traced back to the 1990s [53], [54]. Due to their popularity and level of abuse, these phenomena attracted a large body of research attempting to understand how typos are constructed, how fast users get compromised, whether keyboard layouts affect typosquatting, and how typosquatting domains are abused by their owners [19], [20], [22]–[24], [55]–[62]. Researchers also identified and studied other types of squatting including *homograph domains* (malicious domains that abuse visually-similar characters with their victim domains [63]), *soundsquatting* (malicious domains that sound like popular domains [35], [64]), *bitsquatting* (malicious domains that are one-bit-flip variations of popular domains [21], [65]), and *combosquatting* (malicious domains that include popular trademarks [18], [66]).

The main difference between the BNS squatting we studied in this paper and all types of traditional domain squatting is the complexity and likelihood of successfully exploiting users through a squatting attack. As we argued in Section III, BNS squatting is a *significantly* stronger attack vector where a single typo in a wallet software or online exchange translates to the immediate and irrevocable loss of funds, without attackers needing to further social-engineer users, infect them with malware, or exfiltrate sensitive information from their machines.

IX. Conclusion

In this paper, we drew attention to the issue of typosquatting in Blockchain-based Naming Systems (BNSs) and performed the first study of intra-chain squatting across three BNSs. We were able to build a corpus of 4.9 million domains hosted across these services which we used to define legitimate domains and identify squatting domains targeting them. Through a set of conservative filters aimed at overcoming the inherent limitations of the available on-chain data on these domains, we discovered tens of thousands of squatting domains, targeting as many as 37% of the benign domains in their corresponding BNS. Among others, we observed an increasing number of typosquatting registrations in BNSs per year, with defensive domain registrations being almost entirely absent. In terms of transactions, we focused on senders who have sent funds to pairs of legitimate/typosquattting domains, observing thousands of such pairs with transactions involving substantial amounts of cryptocurrencies, often the equivalent of hundreds to thousands of dollars. Lastly, we confirmed the typosquatting phenomenon on Twitter data and observed that modern software wallets and online exchanges do not currently support any mechanisms to protect their users from typosquatting.

Acknowledgements We thank the anonymous reviewers for their helpful feedback. This work was supported by the National Science Foundation (NSF) under grants CNS-2211575, CNS-2126654, and CNS-1941617.

REFERENCES

[1] Andreas M Antonopoulos and SH El Hariry. *The Internet of Money*, volume 1. Merkle Bloom LLC Columbia, MD, 2016.

[2] Shermin Voshmgir. *Token Economy: How the Web3 reinvents the internet*, volume 2. Token Kitchen, 2020.

[3] Ethereum. What is Web3 and why is it important? https://ethereum.org/en/web3/.

[4] Juan Benet. IPFS-content addressed, versioned, p2p file system. *arXiv preprint arXiv:1407.3561*, 2014.

[5] Muhammad Muzammil, ZhengYu Wu, Aruna Balasubramanian, and Nick Nikiforakis. Panning for gold.eth: Understanding and Analyzing ENS Domain Dropcatching. In *Proceedings of the Internet Measurement Conference (IMC)*, 2024.

[6] Pengcheng Xia, Haoyu Wang, Zhou Yu, Xinyu Liu, Xiapu Luo, Guoai Xu, and Gareth Tyson. Challenges in decentralized name management: the case of ens. In *Proceedings of the 22nd ACM Internet Measurement Conference*, pages 65–82, 2022.

[7] Constantinos Patsakis, Fran Casino, Nikolaos Lykousas, and Vasilios Katos. Unravelling ariadne's thread: Exploring the threats of decentralised dns. *IEEE Access*, 8:118559–118571, 2020.

[8] Harry A Kalodner, Miles Carlsten, Paul M Ellenbogen, Joseph Bonneau, and Arvind Narayanan. An empirical study of namecoin and lessons for decentralized namespace design. In *Workshop on the Economics of Information Security*, 2015.

[9] Fran Casino, Nikolaos Lykousas, Vasilios Katos, and Constantinos Patsakis. Unearthing malicious campaigns and actors from the blockchain dns ecosystem. *Computer Communications*, 179:217–230, 2021.

[10] Zecheng Li, Shang Gao, Zhe Peng, Songtao Guo, Yuanyuan Yang, and Bin Xiao. B-dns: A secure and efficient dns based on the blockchain technology. *IEEE Transactions on Network Science and Engineering*, 8(2):1674–1686, 2021.

[11] Daiki Ito, Yuta Takata, Hiroshi Kumagai, and Masaki Kamizono. Investigations of top-level domain name collisions in blockchain naming services. In *Proceedings of the ACM on Web Conference 2024*, pages 2926–2935, 2024.

[12] typosquatting3.0. https://github.com/pragseclab/typosquatting3.0, 2024.

[13] Ethereum naming service. https://ens.domains/, 2023.

[14] Terminology - ens documentation. https://docs.ens.domains/terminology, 2023.

[15] Unstoppable domains. https://unstoppabledomains.com/, 2023.

[16] Ada handle. https://adahandle.com/. https://adahandle.com/, 2023.

[17] Florian. What is adahandle? the cardano name service, Dec 2022.

[18] Panagiotis Kintis, Najmeh Miramirkhani, Charles Lever, Yizheng Chen, Rosa Romero-Gómez, Nikolaos Pitropakis, Nick Nikiforakis, and Manos Antonakakis. Hiding in plain sight: A longitudinal study of combosquatting abuse. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pages 569–586, 2017.

[19] Janos Szurdi, Balazs Kocso, Gabor Cseh, Jonathan Spring, Mark Felegyhazi, and Chris Kanich. The long "taile" of typosquatting domain names. In *23rd USENIX Security Symposium)*, pages 191–206, 2014.

[20] Pieter Agten, Wouter Joosen, Frank Piessens, and Nick Nikiforakis. Seven months' worth of mistakes: A longitudinal study of typosquatting abuse. In *Proceedings of the 22nd Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2015.

[21] Nick Nikiforakis, Steven Van Acker, Wannes Meert, Lieven Desmet, Frank Piessens, and Wouter Joosen. Bitsquatting: Exploiting bit-flips for fun, or profit? In *Proceedings of the 22nd international conference on World Wide Web*, pages 989–998, 2013.

[22] Anirban Banerjee, Dhiman Barman, Michalis Faloutsos, and Laxmi N Bhuyan. Cyber-fraud is one typo away. In *IEEE INFOCOM-The 27th Conference on Computer Communications*, pages 1939–1947. IEEE, 2008.

[23] Benjamin Edelman. Large-scale registration of domains with typographical errors. *2003a. Harvard University.,"Domain Name Typosquatter Still Generating Millions."," 2003b. Harvard University*, 2003.

[24] Tyler Moore and Benjamin Edelman. Measuring the perpetrators and funders of typosquatting. In *International Conference on Financial Cryptography and Data Security*, pages 175–191, 2010.

[25] Alexa top websites — last save before it was closed. https://www.expireddomains.net/alexa-top-websites/, 2023.

[26] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. Tranco: A research-oriented top sites ranking hardened against manipulation. *arXiv preprint arXiv:1806.01156*, 2018.

[27] Christof Ferreira Torres, Fiona Willi, and Shweta Shinde. Is Your Wallet Snitching On You? An Analysis on the Privacy Implications of Web3. 2023.

[28] Graphql — a query language for your api. https://graphql.org/, 2023.

[29] Ens subgraph. https://thegraph.com/hosted-service, 2023.

[30] Etherscan — ethereum (eth) blockchain explorer. https://etherscan.io/, 2023.

[31] Polygonscan — polygon (matic) blockchain explorer. https://polygonscan.com/, 2023.

[32] Unstoppable domains developer portal. https://docs.unstoppabledomains.com/openapi/partner/v2/, 2023.

[33] Blockfrost.io - cardano api. https://blockfrost.io/, 2023.

[34] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805, 2018.

[35] Nick Nikiforakis, Marco Balduzzi, Lieven Desmet, Frank Piessens, and Wouter Joosen. Soundsquatting: Uncovering the use of homophones in domain squatting. In *Information Security: 17th International Conference, ISC, Hong Kong, China, October 12-14, 2014. Proceedings 17*, pages 291–308. Springer, 2014.

[36] Google chrome now detects typos in your urls - the verge. https://www.theverge.com/2023/5/18/23728705/google-chrome-detects-typos-urls-accessibility, may 2023.

[37] Kurt Thomas, Damon McCoy, Chris Grier, Alek Kolcz, and Vern Paxson. {Trafficking} fraudulent accounts: The role of the underground market in twitter spam and abuse. In *22nd USENIX Security Symposium (USENIX Security)*, pages 195–210, 2013.

[38] Marti Motoyama, Kirill Levchenko, Chris Kanich, Damon McCoy, Geoffrey M Voelker, and Stefan Savage. Re:{CAPTCHAs—Understanding}{CAPTCHA-Solving} services in an economic context. In *19th USENIX Security Symposium (USENIX Security 10)*, 2010.

[39] Ariana Mirian, Joe DeBlasio, Stefan Savage, Geoffrey M Voelker, and Kurt Thomas. Hack for hire: Exploring the emerging market for account hijacking. In *The World Wide Web Conference*, pages 1279–1289, 2019.

[40] Tom Van Goethem, Najmeh Miramirkhani, Wouter Joosen, and Nick Nikiforakis. Purchased fame: Exploring the ecosystem of private blog networks. In *Proceedings of the ACM Asia Conference on Computer and Communications Security*, pages 366–378, 2019.

[41] Xigao Li, Anurag Yepuri, and Nick Nikiforakis. Double and nothing: Understanding and detecting cryptocurrency giveaway scams. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2023.

[42] Xigao Li, Amir Rahmati, and Nick Nikiforakis. Like, comment, get scammed: Characterizing comment scams on media platforms. In *Proceedings Network and Distributed System Security Symposium*. In Network and Distributed System Security Symposium (NDSS) San Diego, CA, USA, 2024.

[43] Enze Liu, George Kappos, Eric Mugnier, Luca Invernizzi, Stefan Savage, David Tao, Kurt Thomas, Geoffrey M Voelker, and Sarah Meiklejohn. Give and take: An end-to-end investigation of giveaway scam conversion rates. *arXiv preprint arXiv:2405.09757*, 2024.

[44] Kai Li, Shixuan Guan, and Darren Lee. Towards understanding and characterizing the arbitrage bot scam in the wild. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 7(3):1–29, 2023.

[45] Seung Ho Na, Sumin Cho, and Seungwon Shin. Evolving bots: The new generation of comment bots and their underlying scam campaigns in youtube. In *Proceedings of the ACM on Internet Measurement Conference*, pages 297–312, 2023.

[46] Yazan Boshmaf, Charitha Elvitigala, Husam Al Jawaheri, Primal Wijesekera, and Mashael Al Sabah. Investigating mmm ponzi scheme on bitcoin. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, pages 519–530, 2020.

[47] Hanna Kim, Jian Cui, Eugene Jang, Chanhee Lee, Yongjae Lee, Jin-Woo Chung, and Seungwon Shin. Drainclog: Detecting rogue accounts with illegally-obtained nfts using classifiers learned on graphs. *arXiv preprint arXiv:2301.13577*, 2023.

[48] Dipanjan Das, Priyanka Bose, Nicola Ruaro, Christopher Kruegel, and Giovanni Vigna. Understanding security issues in the nft ecosystem. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pages 667–681, 2022.

[49] Jiahua Xu and Benjamin Livshits. The anatomy of a cryptocurrency {Pump-and-Dump} scheme. In *28th USENIX Security Symposium (USENIX Security)*, pages 1609–1625, 2019.
[50] Ross Phillips and Heidi Wilder. Tracing cryptocurrency scams: Clustering replicated advance-fee and phishing websites. In *IEEE international conference on blockchain and cryptocurrency (ICBC)*, pages 1–8. IEEE, 2020.
[51] Svetlana Abramova, Artemij Voskobojnikov, Konstantin Beznosov, and Rainer Böhme. Bits under the mattress: Understanding different risk perceptions and security behaviors of crypto-asset users. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2021.
[52] Danny Yuxing Huang, Maxwell Matthaios Aliapoulios, Vector Guo Li, Luca Invernizzi, Elie Bursztein, Kylie McRoberts, Jonathan Levin, Kirill Levchenko, Alex C Snoeren, and Damon McCoy. Tracking ransomware end-to-end. In *IEEE Symposium on Security and Privacy (SP)*, pages 618–631. IEEE, 2018.
[53] Anticybersquatting Consumer Protection Act (ACPA). https://www.govinfo.gov/content/pkg/PLAW-106publ113/html/PLAW-106publ113.htm, 1999.
[54] David Kesmodel. *The domain game: how people get rich from Internet domain names*. Xlibris Corporation, 2008.
[55] Victor Le Pochat, Tom Van Goethem, and Wouter Joosen. A smörgåsbord of typos: exploring international keyboard layout typosquatting. In *IEEE Security and Privacy Workshops (SPW)*, pages 187–192. IEEE, 2019.
[56] Mohammad Taha Khan, Xiang Huo, Zhou Li, and Chris Kanich. Every second counts: Quantifying the negative externalities of cybercrime via typosquatting. In *IEEE Symposium on Security and Privacy*, pages 135–150. IEEE, 2015.
[57] Rashid Tahir, Ali Raza, Faizan Ahmad, Jehangir Kazi, Fareed Zaffar, Chris Kanich, and Matthew Caesar. It's all in the name: Why some urls are more vulnerable to typosquatting. In *IEEE INFOCOM-IEEE Conference on Computer Communications*, pages 2618–2626. IEEE, 2018.
[58] Takashi Koide, Naoki Fukushi, Hiroki Nakano, and Daiki Chiba. Phishreplicant: A language model-based approach to detect generated squatting domain names. In *Proceedings of the 39th Annual Computer Security Applications Conference*, pages 1–13, 2023.
[59] Francesco Blefari, Angelo Furfaro, Giovambattista Ianni, Alessandro Viscomi, et al. Typoalert: a browser extension against typosquatting. In *Proc. of SEBD: 32nd Symposium on Advanced Database Systems*, 2024.
[60] Rodolfo Vieira Valentim, Idilio Drago, Marco Mellia, and Federico Cerutti. X-squatter: Ai multilingual generation of cross-language sound-squatting. *ACM Transactions on Privacy and Security*, 2024.
[61] Durjoy Majumdar, S Arjun, Pranavi Boyina, Sri Sai Priya Rayidi, Yerra Rahul Sai, and Suryakanth V Gangashetty. Beyond text: Nefarious actors harnessing llms for strategic advantage. In *International Conference on Intelligent Systems for Cybersecurity (ISCS)*, pages 1–7. IEEE, 2024.
[62] Anastasios Lepipas, Anastasia Borovykh, and Soteris Demetriou. Username squatting on online social networks: A study on x. In *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security*, pages 621–637, 2024.
[63] Tobias Holgers, David E Watson, and Steven D Gribble. Cutting through the confusion: A measurement study of homograph attacks. In *USENIX Annual Technical Conference, General Track*, pages 261–266, 2006.
[64] Rodolfo Valentim, Idilio Drago, Federico Cerutti, and Marco Mellia. Ai-based sound-squatting attack made possible. In *IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 448–453, 2022.
[65] Artem Dinaburg. Bitsquatting: Dns hijacking without exploitation. *Proceedings of BlackHat Security*, 2011.
[66] Duc-Ly Vu, Ivan Pashchenko, Fabio Massacci, Henrik Plate, and Antonino Sabetta. Typosquatting and combosquatting attacks on the python ecosystem. In *ieee european symposium on security and privacy workshops (euros&pw)*, pages 509–514, 2020.

APPENDIX A

Figure 14: *Profile of the most popular Twitter/X user advertising his ENS domain name.*

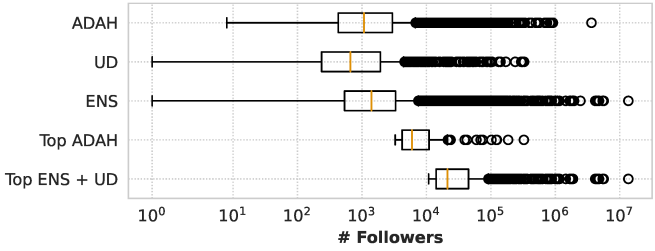

Figure 15: *Distribution of the number of followers of all the collected BNS names popularized on Twitter/X*

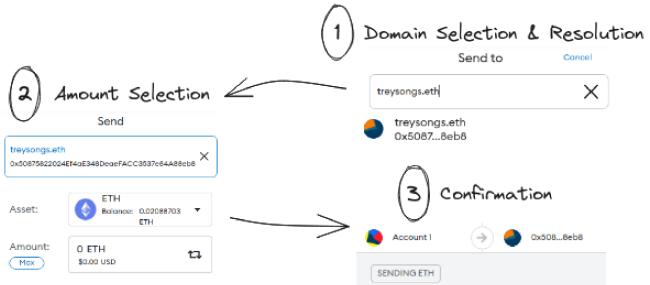

Figure 16: *Example of the "Cold Typos" experiment, showing the stages to sending assets to* `treysongs.eth` *(targeting* `treysongz.eth`*) through Metamask without warnings.*

Target	Typosquatting Name	Typosquatting Address
vitalik.eth	vitqlik.eth	0x6c30...83a90
vitalik.eth	vita1ik.eth	0x4bcb..d920
vitalik.eth	vitlaik.eth	0x90e9...84d6
vitalik.eth	vitalii.eth	0x4c81...5492
vitalik.eth	vitaiik.eth	0x0325...e09ec
vitalik.eth	vitalki.eth	0x8c26...fe68
vitalik.eth	vitalij.eth	0x5201...80a3
metascan.nft	m3tascan.nft	0xcd9f...aab9
play2earn.crypto	play2ern.crypto	0xaf81...b4bb
jjlin.eth	jhlin.eth	0xc3e0...67e7
jjlin.eth	jnlin.eth	0x3682...333a
jjlin.eth	jlin.eth	0x793b...0805
blackcoin.crypto	balckcoin.crypto	0x8e77...0fa4
$ada_astronaut	$adaastronaut	addr1q82...40t9
$cnftjunky	$cnftjunk	addr1q9t...uq6c

Table VI: *Examples of targets, typosquatting names, and typosquatting addresses of* ***Type (i)*** *typosquatting attacks (i.e. where attackers register only one typosquatting domain, and no other domains)*

Address	0x4ea8...567e	0x7afd...e4a2	0xba44...ae33	addr1q87...py2f	addr1qyd...hk6z
Sample Domains	googleeth.eth	vitalik3.eth	0ldtrafford.crypto	$-handles-	$olidity
	treysong.eth	masterkard.eth	0ldtrafford.nft	$-handle-	$occer
	ahopkin.et	maricle.eth	vitalybuterin.nft	$-handles	$tewart
	vitulik.eth	mckillip.eth	vitalik0.x	$-handle	$anchez
	vitalikbuteri.eth	bl0ckhead.eth	j0ebiden.nft	$dana.swap	$hrek
	votalik.eth	btcafe.eth	b0bmarley.nft,	$dana_	$stevenshandle
	paradigmi.eth	akward.eth	0x00001.x	$bitfinex	$steveshandle
	paradigmu.eth	beelive.eth	0x000001.x	$binance	$paulshandle
	pradigm.eth	greatfull.eth	k0bebryant.blockchain	$usbank	$tephanieshandle
	jimmyfalloneth.eth	dallasfans.eth	krypt0.blockchain	$-eternlwallet-	$stephanies
	starbuk.eth	chesleafans.et	0ptimusprime.nft	$-eternl.wallet	$stephans
	mariogotz.eth	heatfan.eth	an0nymous.blockchain	$-eternl_wallet-	$tephanshandle
	pornhb.eth	wilsonfamily.eth	0xpunk.crypto	$ada_handles-	$ex-appeal
	pornhbu.eth	taylorfamily.eth	nakamoto1.crypto	$ada_handles.	$ex_appeal
	simsung.eth	figurati.eth	stevej0bs.crypto	$adahandles	$sex.appeal
	amsang.eth	johnwaynegacy.eth	w0nderw0man.nft	$ada-handle-	$haman
	davidsiwonchi.eth	p0rntv.eth	daviddunn.nft	$walgreens	$motor-yacht
	davidsiwonchio.eth	nfasset.eth	w0lfman.nft	$rockwell	$showbiz
Total Domains Owned	116	773	5862	1475	438

Table VII: *Examples of targets, typosquatting names, and typosquatting addresses of* ***Type (ii)*** *typosquatting attacks (i.e. where attackers register multiple domains targetting dissimilar legitimate domains)*

Address	0xc900...e268	0x250a...cc2f	0xc7d9...2434	0x143a...a1d1	0x92c3...bb57
Sample Domains	w3bank.eth	98223.eth	derick2.eth	0878888888.blockchain	1818181818.wallet
	web3nk.eth	98454.eth	erick2.eth	0908888888.888	181818181818.wallet
	web4nk.eth	98534.eth		0998888888.888	1818181818.blockchain
		98393.eth		0908888888.bitcoin	181818181818.blockchain
		98441.eth		0908888888.dao	1818181818.nft
		98094.eth		0998888888.dao	181818181818.nft
Total Domains Owned	3	179	2	13	14

Table VIII: *Examples of targets, typosquatting names, and typosquatting addresses of* ***Type (iii)*** *typosquatting attacks (i.e. where attackers register >1 domains targetting similar legitimate domains)*