DareShark: Detecting and Measuring Security Risks of Hosting-Based Dangling Domains

Mingming Zhang, **Xiang Li**, Baojun Liu, Jianyu Lu, Yiming Zhang,
Jianjun Chen, Haixin Duan, Shuang Hao, and Xiaofeng Zheng

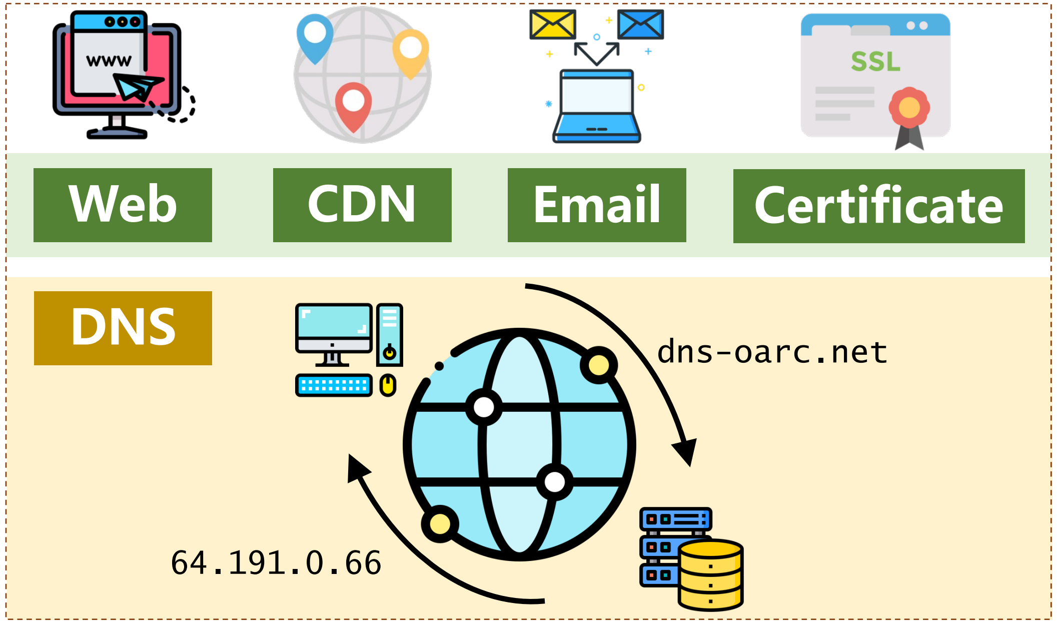
(Accepted by [ACM SIGMETRICS 2023])

Presenter: Xiang Li, Tsinghua University
February 16th, 2023

Domain Name

➢Domain name system (DNS)

- ➢Entry point of many Internet activities
- ➢Security guarantee of multiple application services
- ➢Domain names are widely registered



Web

CDN

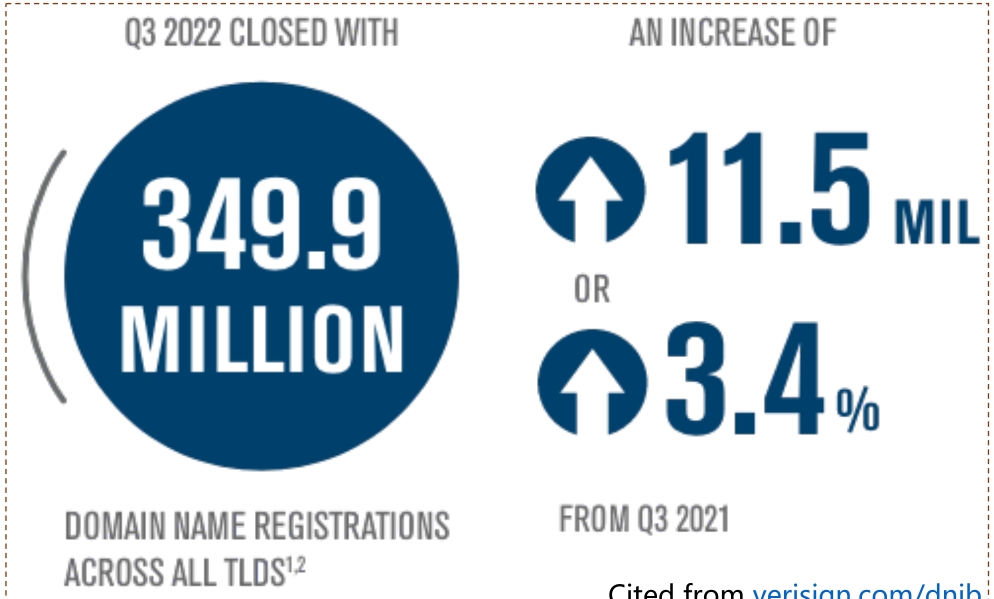
Email

Certificate

DNS

dns-oarc.net

64.191.0.66



Q3 2022 CLOSED WITH

349.9 MILLION

DOMAIN NAME REGISTRATIONS ACROSS ALL TLDS[1,2]

AN INCREASE OF

11.5 MIL

OR

3.4%

FROM Q3 2021

Cited from verisign.com/dnib

Domain Name Abuse

- **Adversaries could exploit the domains outside of their authority for malicious activities**
  - Botnet, phishing, malware distribution, etc.

Cited from bleepingcomputer.com

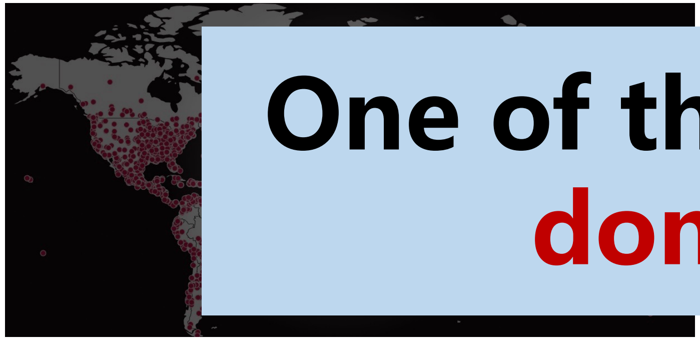
Cited from scmp.com

YOU WIN!
Malvertising
Viruses
Worms
{>_}
Fileless malware
Trojans
Rootkits
Ransomware
Spyware
Adware
Bots or botnets

Cited from norton.com

Domain Name Abuse

- ➢ **Adversaries could exploit the domains outside of their authority for malicious activities**
  - ➢ Botnet, phishing, malware distribution, etc.



Cited from bleepingcomputer.com

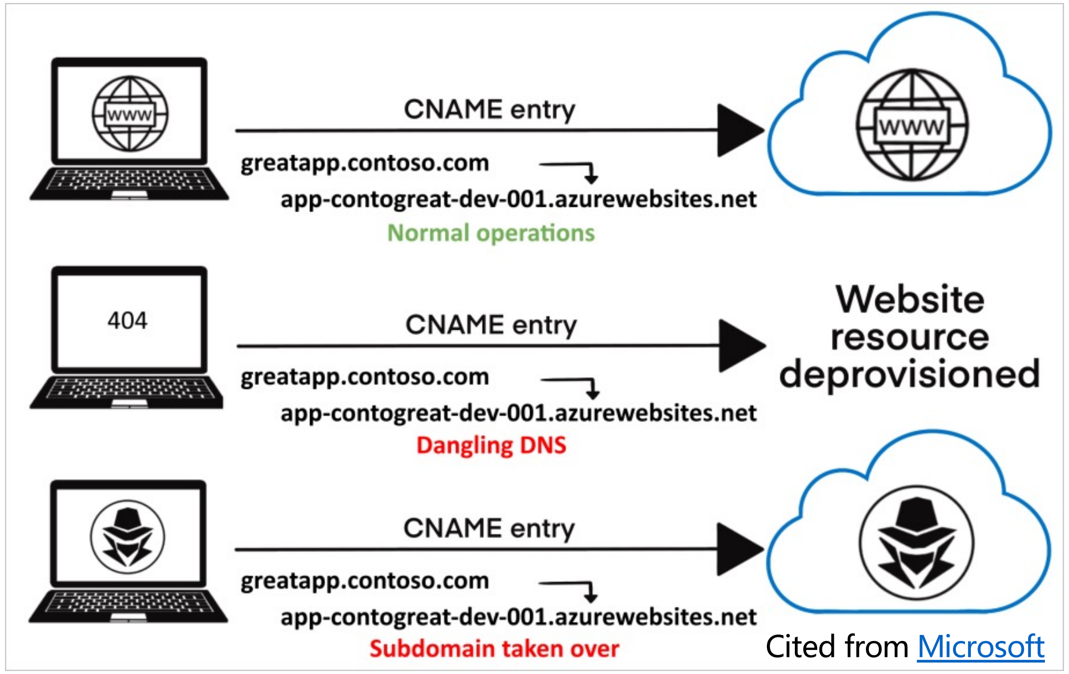
One of the powerful ways is domain takeover

Cited from scmp.com

Cited from norton.com

Domain Name Takeover

- ➢DNS Resource Records (RRs) → Use-After-Free



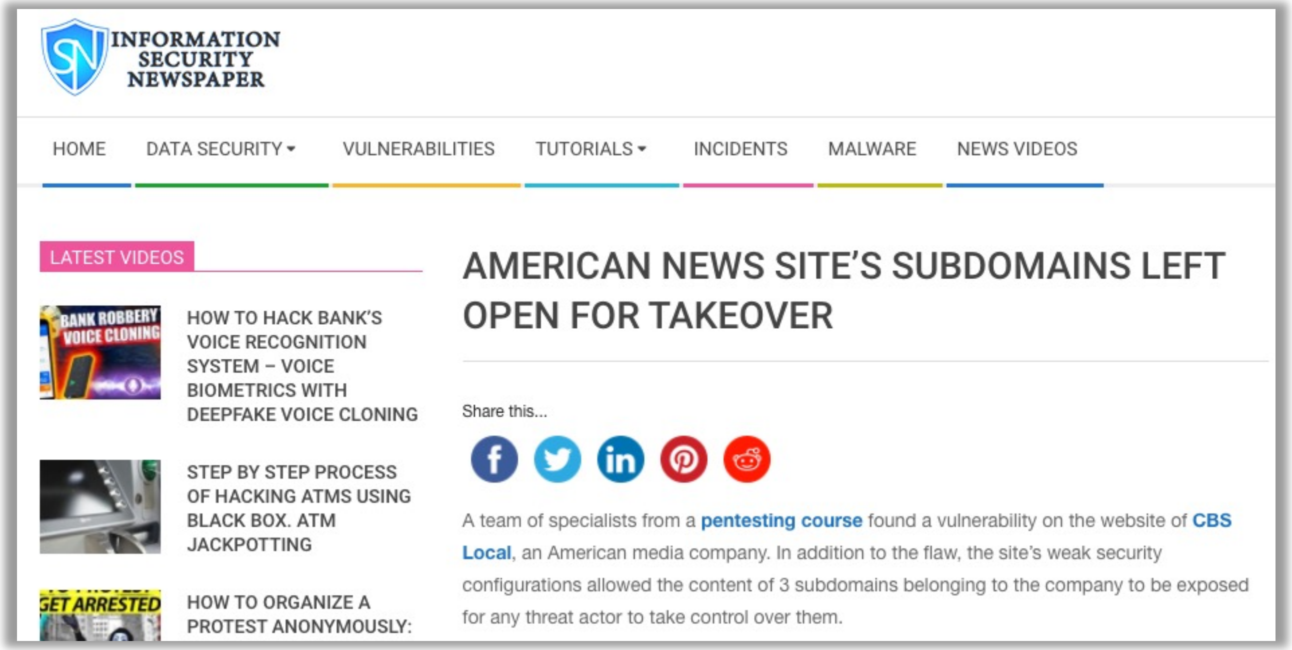
Cited from Microsoft

- ➢Security-sensitive Dangling DNS Records (Dares) → Domain Takeover
  - ➢A, CNAME, NS

All Your DNS Records Point to Us [CCS ’16]

Domain Name Takeover

- ➢Many domain-takeover incidents occur in recent years



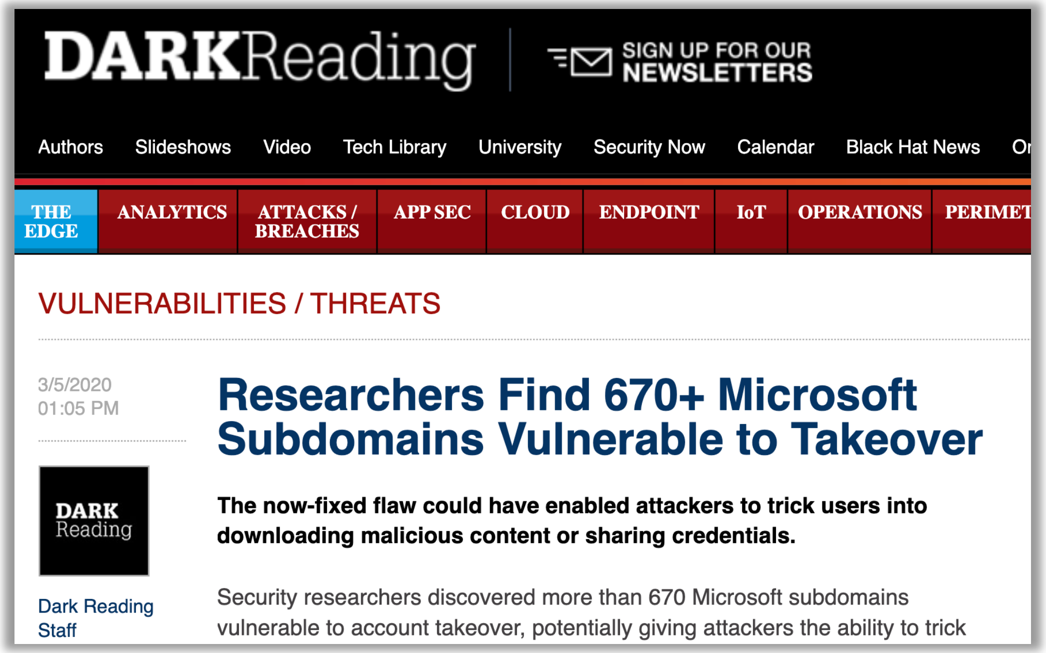
HOW TO HACK BANK'S VOICE RECOGNITION SYSTEM – VOICE BIOMETRICS WITH DEEPFAKE VOICE CLONING

STEP BY STEP PROCESS OF HACKING ATMS USING BLACK BOX. ATM JACKPOTTING

HOW TO ORGANIZE A PROTEST ANONYMOUSLY:

AMERICAN NEWS SITE'S SUBDOMAINS LEFT OPEN FOR TAKEOVER

A team of specialists from a **pentesting course** found a vulnerability on the website of **CBS Local**, an American media company. In addition to the flaw, the site's weak security configurations allowed the content of 3 subdomains belonging to the company to be exposed for any threat actor to take control over them.



VULNERABILITIES / THREATS

3/5/2020
01:05 PM

Dark Reading Staff

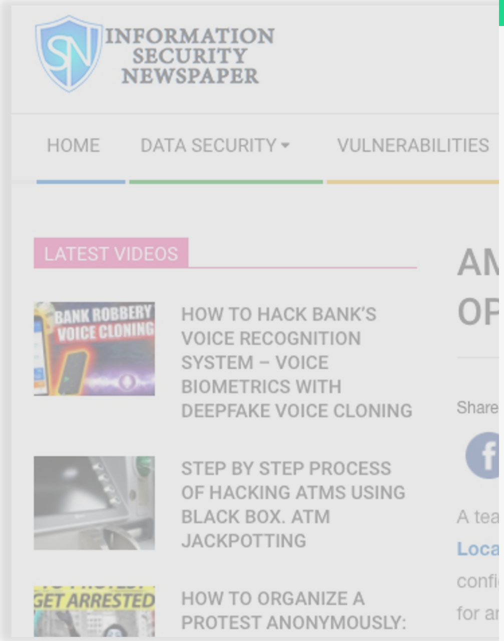
Researchers Find 670+ Microsoft Subdomains Vulnerable to Takeover

The now-fixed flaw could have enabled attackers to trick users into downloading malicious content or sharing credentials.

Security researchers discovered more than 670 Microsoft subdomains vulnerable to account takeover, potentially giving attackers the ability to trick

Domain Name Takeover

➢Many dor ecent years



INFORMATION SECURITY NEWSPAPER

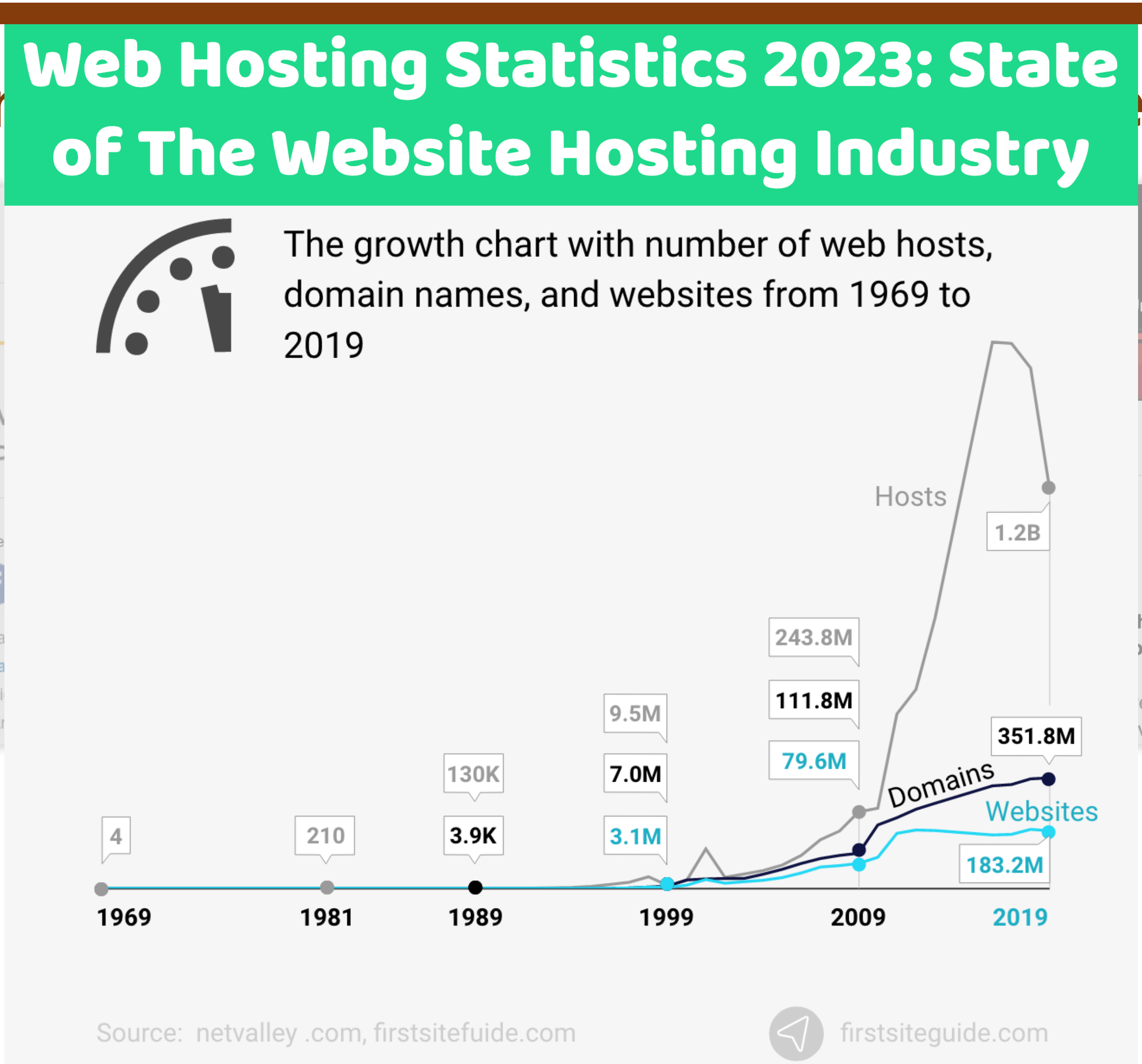
HOME DATA SECURITY ▾ VULNERABILITIES

LATEST VIDEOS

HOW TO HACK BANK'S VOICE RECOGNITION SYSTEM – VOICE BIOMETRICS WITH DEEPFAKE VOICE CLONING

STEP BY STEP PROCESS OF HACKING ATMS USING BLACK BOX. ATM JACKPOTTING

HOW TO ORGANIZE A PROTEST ANONYMOUSLY:

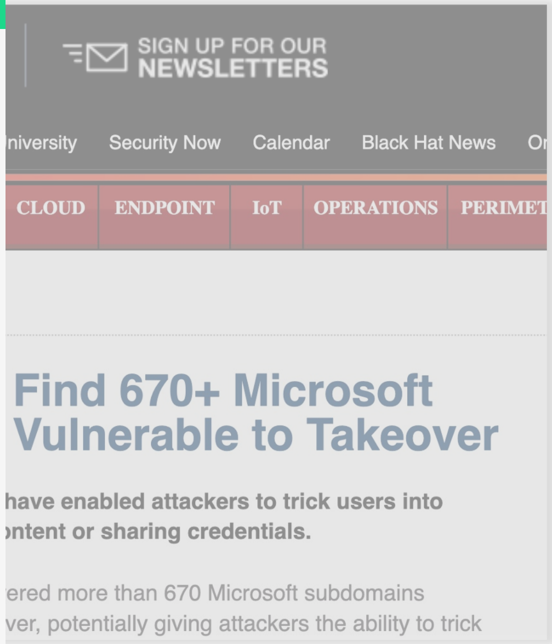


Web Hosting Statistics 2023: State of The Website Hosting Industry

The growth chart with number of web hosts, domain names, and websites from 1969 to 2019

Source: netvalley .com, firstsitefuide.com

firstsiteguide.com



SIGN UP FOR OUR NEWSLETTERS

niversity Security Now Calendar Black Hat News Or

CLOUD ENDPOINT IoT OPERATIONS PERIMET

Find 670+ Microsoft Vulnerable to Takeover

have enabled attackers to trick users into ontent or sharing credentials.

ered more than 670 Microsoft subdomains ver, potentially giving attackers the ability to trick

Domain Name Takeover

- ➢Many domain takeover incidents occur in recent years

Web Hosting Statistics 2023: State of The Website Hosting Industry

The growth chart with number of web hosts, domain names, and websites from 1969 to

INFORMATION SECURITY NEWSPAPER

HOME DATA SECURITY ▾ VULNERABILITIES TUTORIALS ▾ INCIDENTS MALWARE NEWS VIDEOS

LATEST V

STEP BY STEP PROCESS OF HACKING ATMS USING BLACK BOX. ATM JACKPOTTING

HOW TO ORGANIZE A PROTEST ANONYMOUSLY:

A team of specialists from a pentesting course found a vulnerability on the website of CBS Local, an American media company. In addition to the flaw, the site's weak security configurations allowed the content of 3 subdomains belonging to the company to be exposed for any threat actor to take control over them.

DARKReading

SIGN UP FOR OUR NEWSLETTERS

Authors Slideshows Video Tech Library University Security Now Calendar Black Hat News

PERIMET

Subdomains Vulnerable to Takeover

The now-fixed flaw could have enabled attackers to trick users into downloading malicious content or sharing credentials.

Dark Reading Staff

Security researchers discovered more than 670 Microsoft subdomains vulnerable to account takeover, potentially giving attackers the ability to trick

Source: netvalley .com, firstsitefuide.com

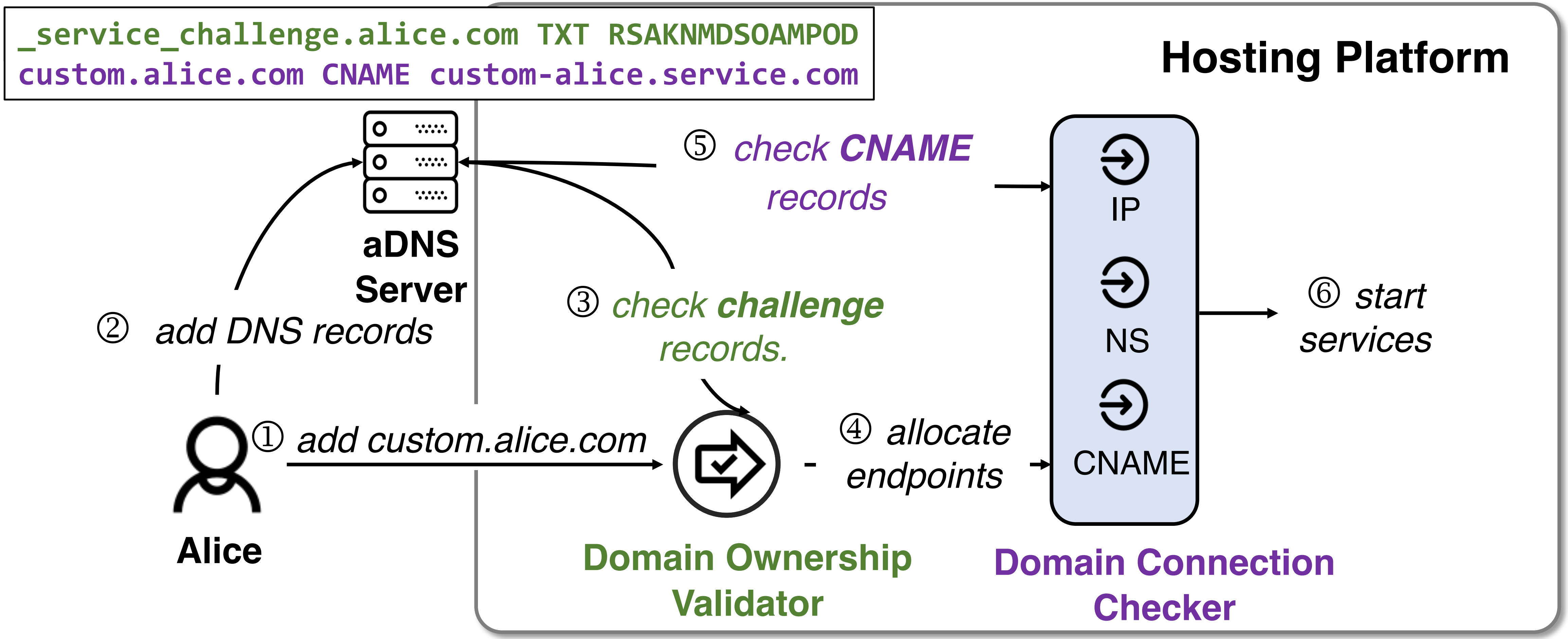
Narrowing down our vision to

hosting-based domain takeover issues!

What is hosting-based domain takeover?

Public Hosting Service

➢Domain hosting procedures



```
_service_challenge.alice.com TXT RSAKNMDSOAMPOD
custom.alice.com CNAME custom-alice.service.com
```

Hosting Platform

aDNS Server

Alice

① add custom.alice.com

② add DNS records

③ check challenge records.

④ allocate endpoints

⑤ check CNAME records

⑥ start services

IP

NS

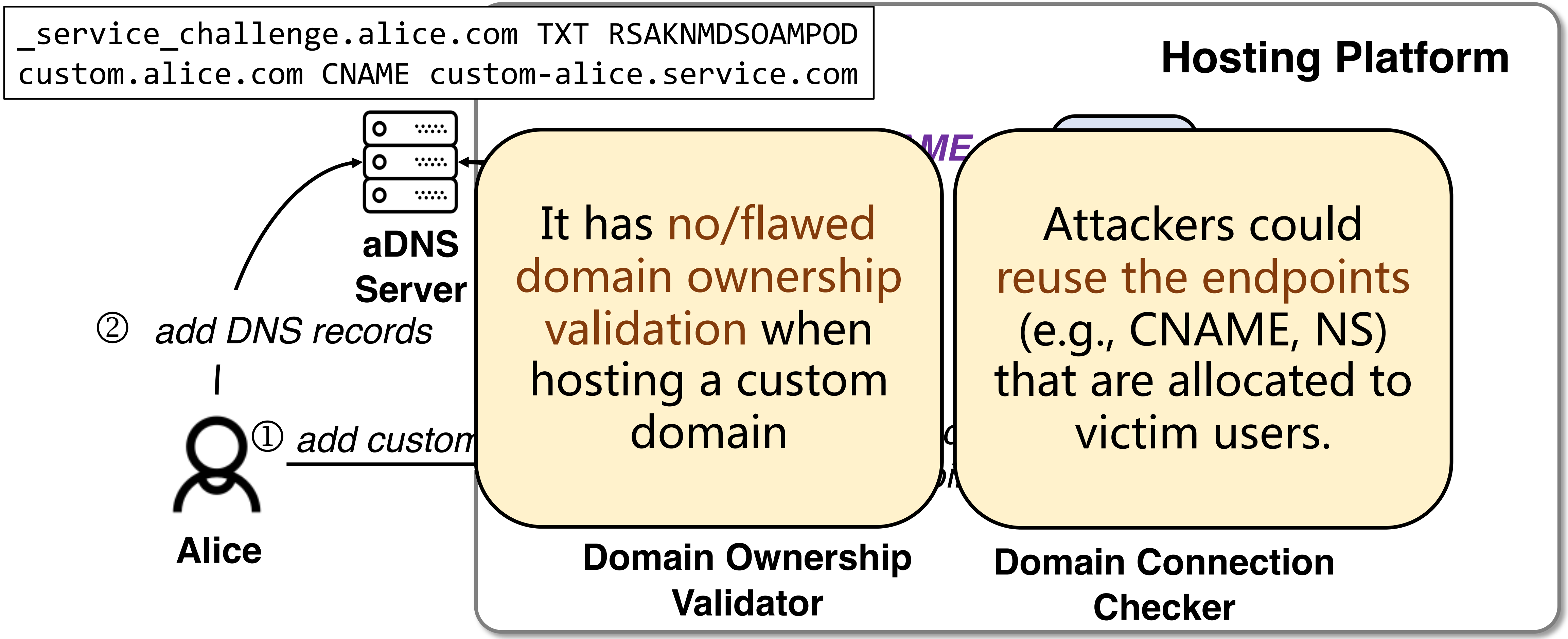
CNAME

Domain Ownership Validator

Domain Connection Checker

Vulnerable Hosting Service

- ➢**However, a hosting service might be vulnerable if:**



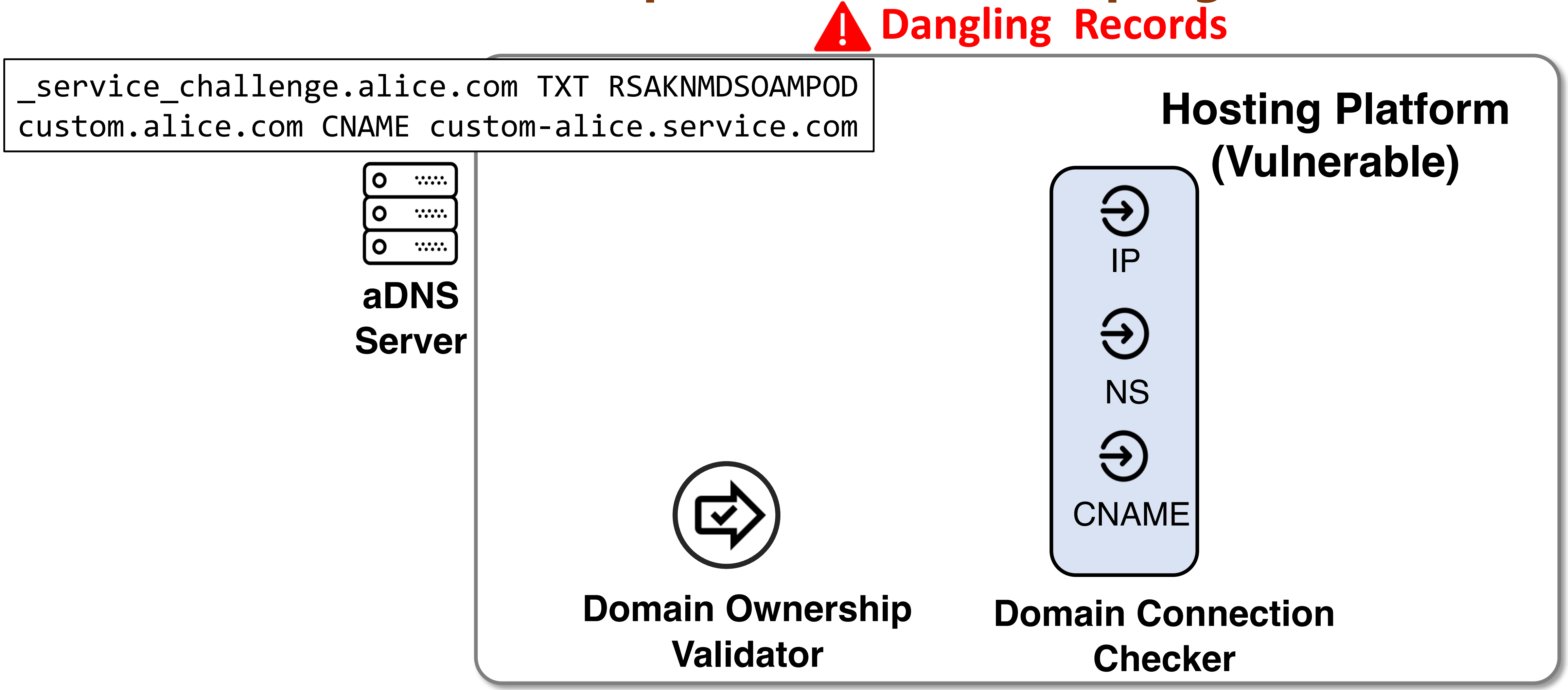
```
_service_challenge.alice.com TXT RSAKNMDSOAMPOD
custom.alice.com CNAME custom-alice.service.com
```

It has no/flawed domain ownership validation when hosting a custom domain

Attackers could reuse the endpoints (e.g., CNAME, NS) that are allocated to victim users.

Hosting-based Domain Takeover

- When Alice’s service expires, she doesn’t purge RRs



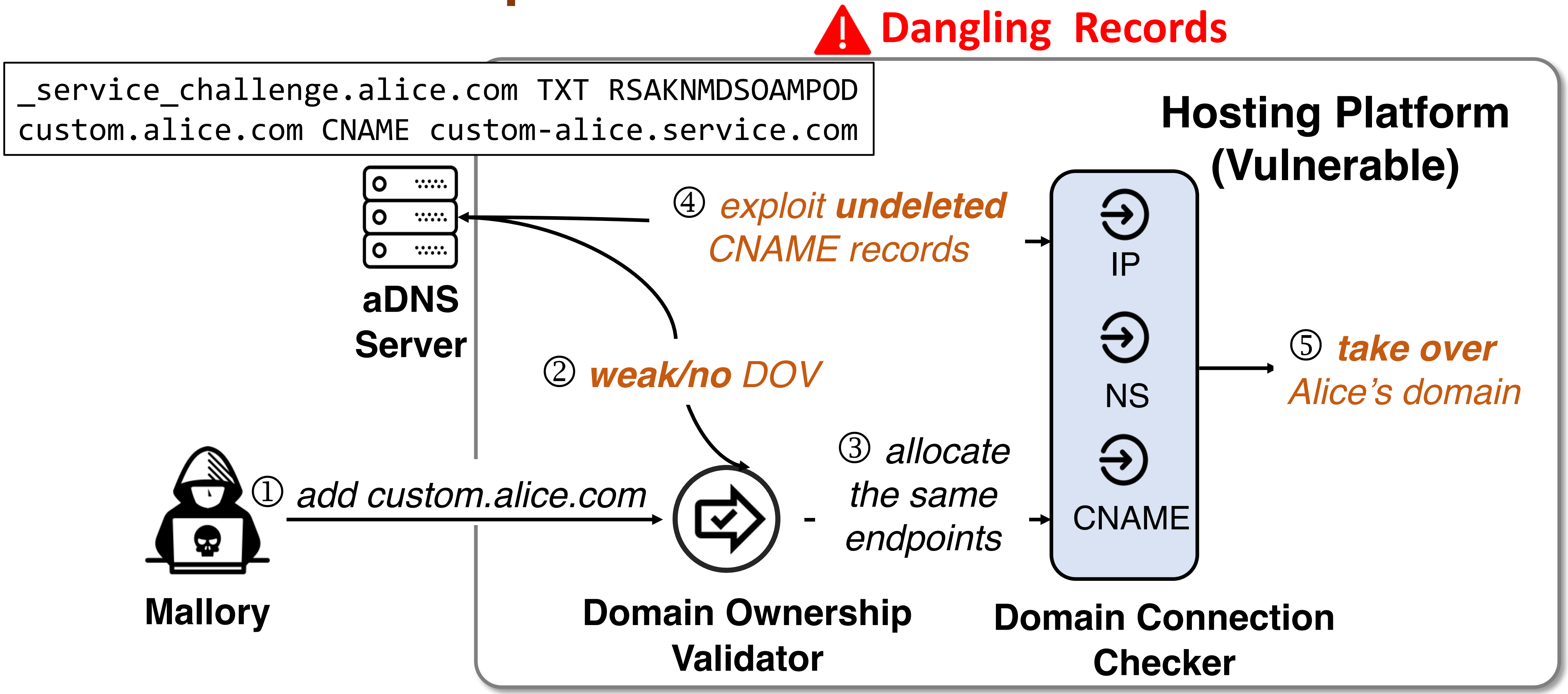
⚠ Dangling Records

```
_service_challenge.alice.com TXT RSAKNMDSOAMPOD
custom.alice.com CNAME custom-alice.service.com
```

aDNS Server
Hosting Platform (Vulnerable)
IP
NS
CNAME
Domain Ownership Validator
Domain Connection Checker

Hosting-based Domain Takeover

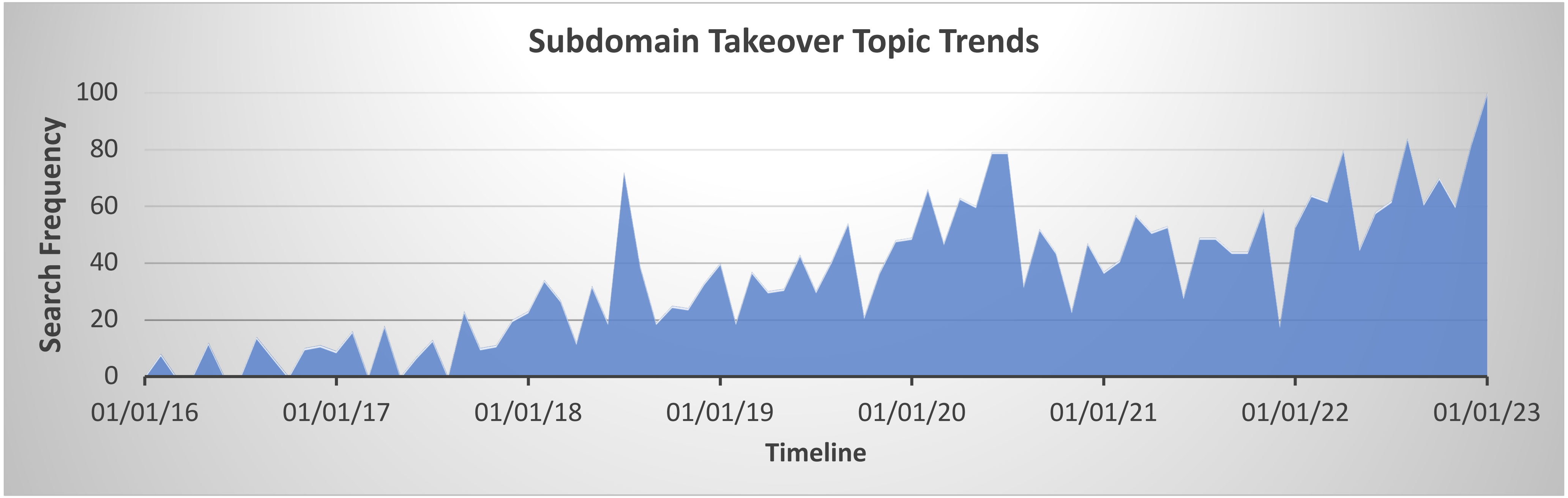
➢ **Domain takeover procedures**



```
_service_challenge.alice.com TXT RSAKNMDSOAMPOD
custom.alice.com CNAME custom-alice.service.com
```

⚠ Dangling Records
Hosting Platform (Vulnerable)
aDNS Server
Mallory
① add custom.alice.com
② weak/no DOV
③ allocate the same endpoints
④ exploit undeleted CNAME records
⑤ take over Alice's domain
IP
NS
CNAME
Domain Ownership Validator
Domain Connection Checker

Why domain takeover occurs ceaselessly?



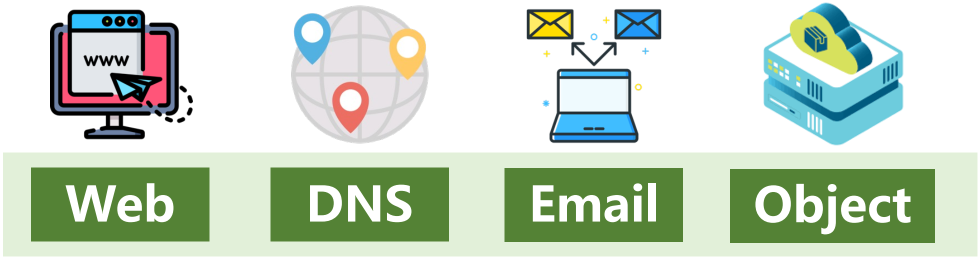
“Domain takeover incidents are still on the rise, increasing by 25% from 2020 to 2021.”[1]

[1] https://blog.detectify.com/ 2022/03/22/subdomain-takeover-on-the-rise-detectify-research/.

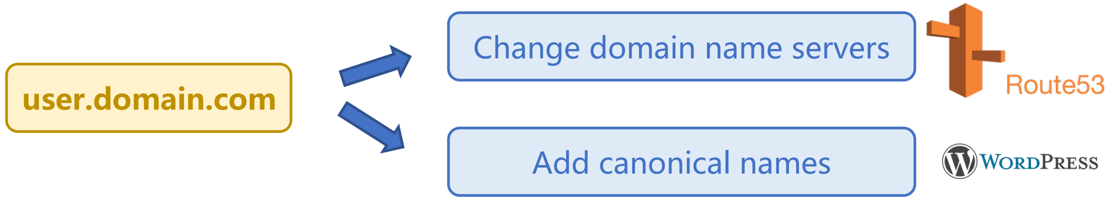
Motivation

1. A generic method for discovering third-party hosting services is needed

- Various hosting service types



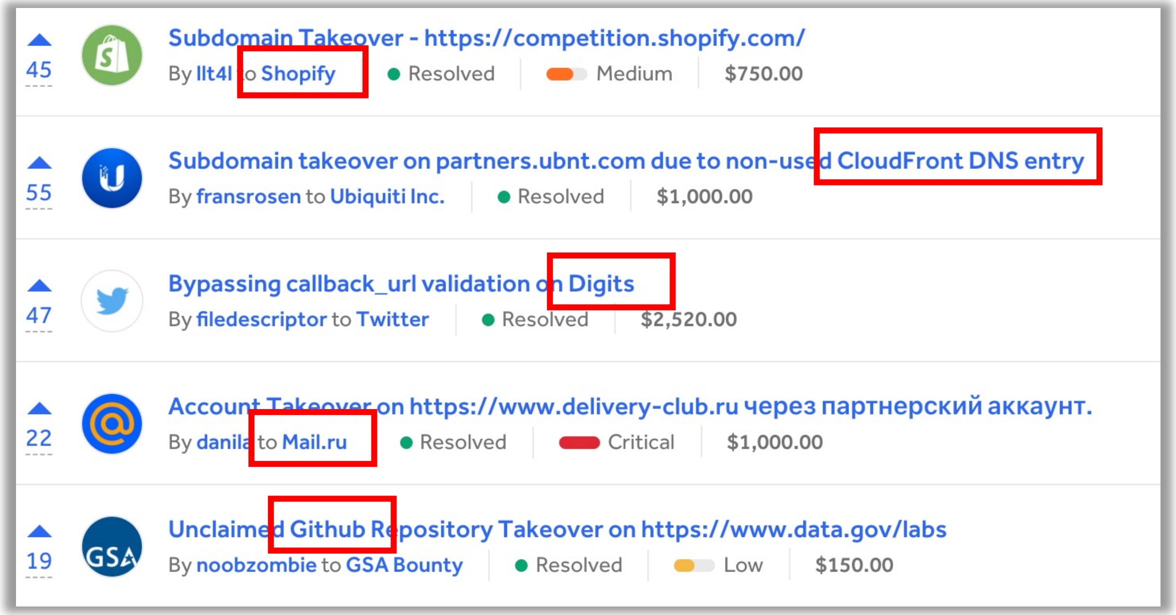
- Various domain hosting strategies


user.domain.com

Change domain name servers

Add canonical names

- Ad-hoc hacktivity reports on HackerOne


45 Subdomain Takeover - https://competition.shopify.com/
By llt4l to Shopify ● Resolved Medium $750.00

55 Subdomain takeover on partners.ubnt.com due to non-used CloudFront DNS entry
By fransrosen to Ubiquiti Inc. ● Resolved $1,000.00

47 Bypassing callback_url validation on Digits
By filedescriptor to Twitter ● Resolved $2,520.00

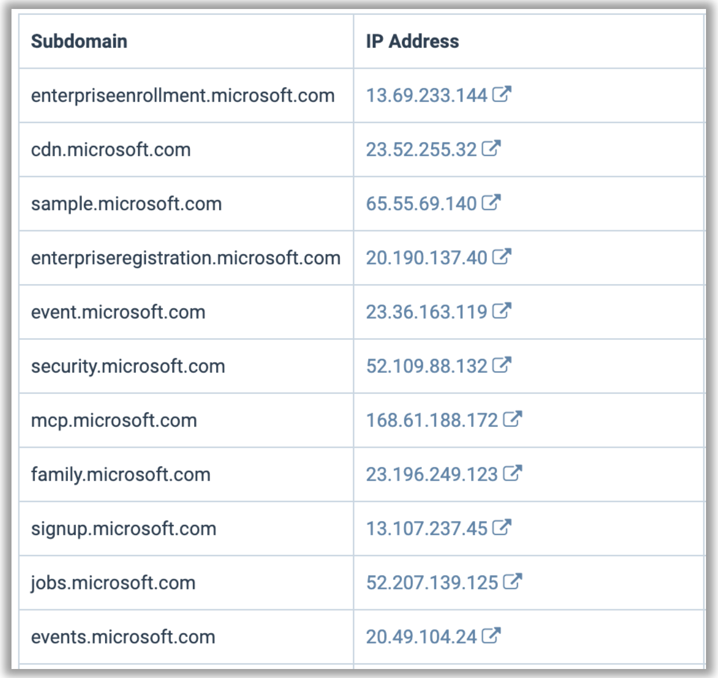
22 Account Takeover on https://www.delivery-club.ru через партнерский аккаунт.
By danila to Mail.ru ● Resolved Critical $1,000.00

19 Unclaimed Github Repository Takeover on https://www.data.gov/labs
By noobzombie to GSA Bounty ● Resolved Low $150.00

Motivation

2. An efficient detection system is absent for quickly digging out vulnerable domains in the wild

- **Large companies have thousands of subdomains, with DNS chains changing frequently**



Subdomain	IP Address
enterpriseenrollment.microsoft.com	13.69.233.144
cdn.microsoft.com	23.52.255.32
sample.microsoft.com	65.55.69.140
enterpriseregistration.microsoft.com	20.190.137.40
event.microsoft.com	23.36.163.119
security.microsoft.com	52.109.88.132
mcp.microsoft.com	168.61.188.172
family.microsoft.com	23.196.249.123
signup.microsoft.com	13.107.237.45
jobs.microsoft.com	52.207.139.125
events.microsoft.com	20.49.104.24

How to timely detect vulnerable domains among them?

Previous work: active DNS resolution

[Daiping 2016, Eihal 2020 , Marco 2021]

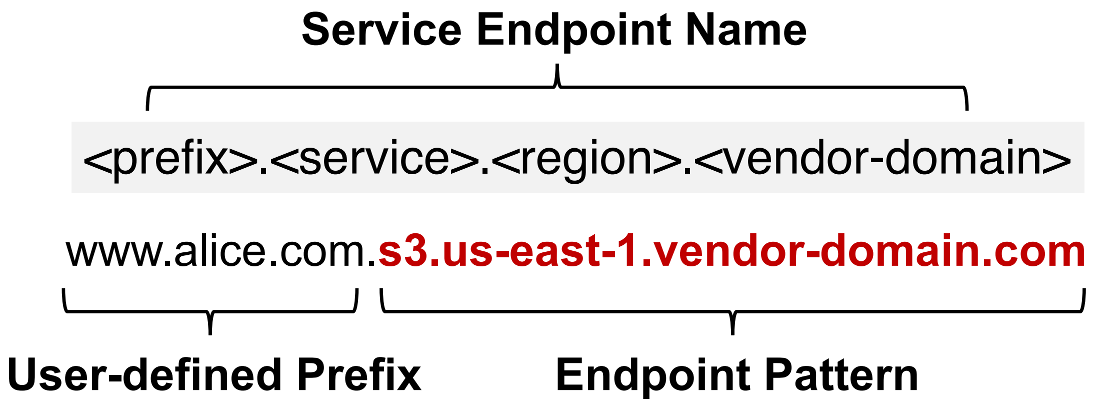
Can we discover more hosting services and detect vulnerable domains timely?

The domain characteristics of hosting services and the DNS chains of domains are logged in DNS traffic.

Empirical Observations

O1. Similar endpoint naming conventions

- **Service Endpoint Patterns**



Service Endpoint Name

<prefix>.<service>.<region>.<vendor-domain>

www.alice.com.**s3.us-east-1.vendor-domain.com**

User-defined Prefix

Endpoint Pattern

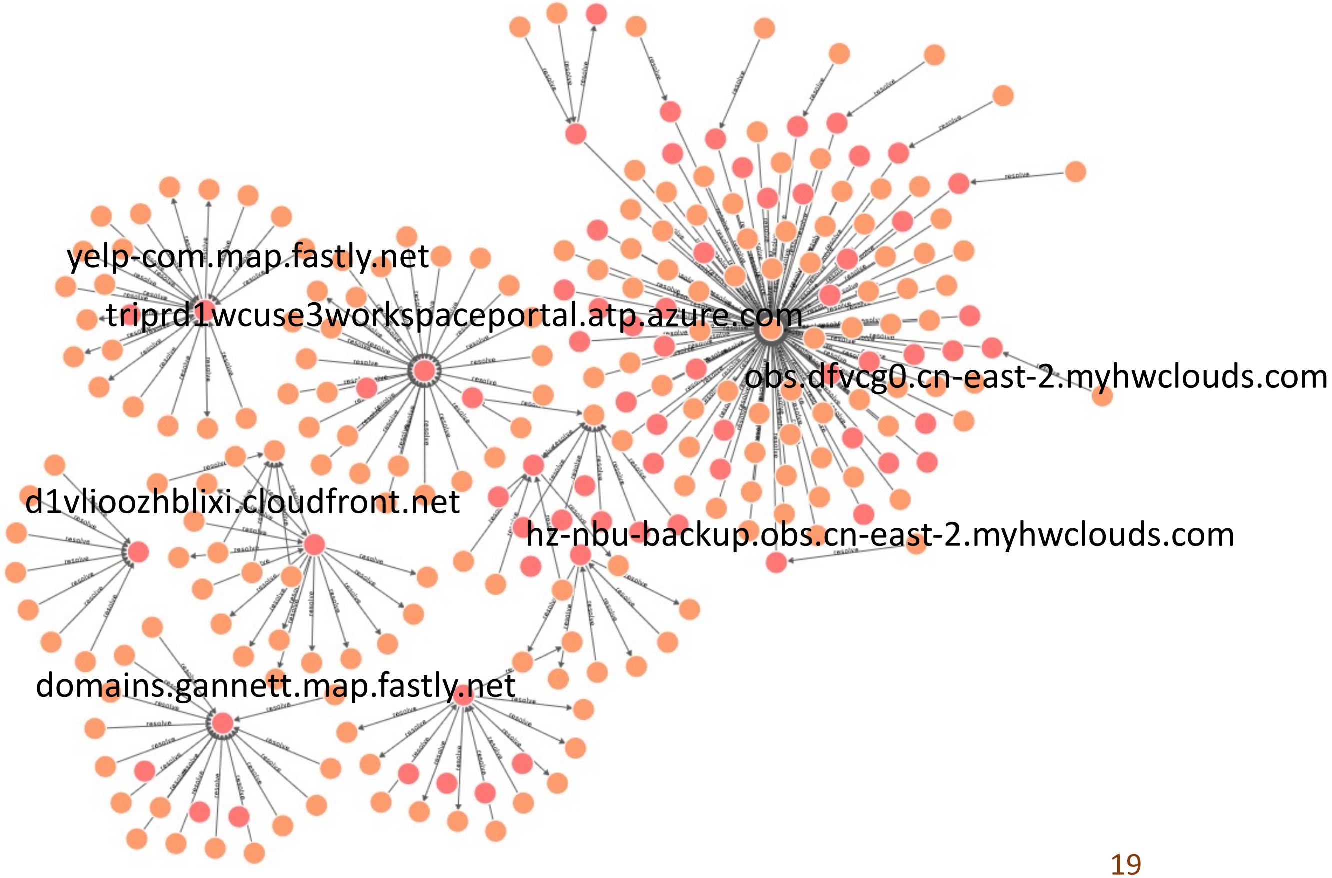
Empirical Observations

O2. High domain dependency number

- One service apex domain may serve thousands of customers’ domains

```
custom1.com   CNAME   prefix1.service.com
custom2.com   CNAME   prefix2.service.com
                 ...
customN.com   CNAME   prefixN.service.com
```

$$DN(\text{“}service.com\text{”}) = N$$


yelp-com.map.fastly.net
triprd1wcuse3workspaceportal.atp.azure.com
obs.dfvcg0.cn-east-2.myhwclouds.com
d1vlioozhblixi.cloudfront.net
hz-nbu-backup.obs.cn-east-2.myhwclouds.com
domains.gannett.map.fastly.net

Our solution

Automate the approach to discovering services and vulnerable domains using passive DNS traffic.

Our Tool: DareShark

- **A novel framework that can assist in:**
  - **Discovering vulnerable hosting services**

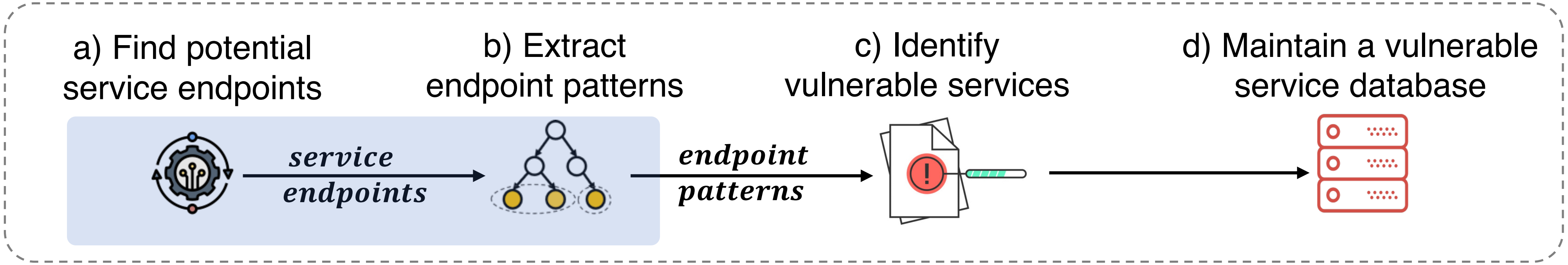
    ↳ **Expand the detection scope**

  - **Detecting hosting-based vulnerable domains efficiently**

    ↳ **Prevent potential security threats**

DareShark Workflow

Part 1. Vulnerable service discovery (offline procedure)



a) Find potential service endpoints

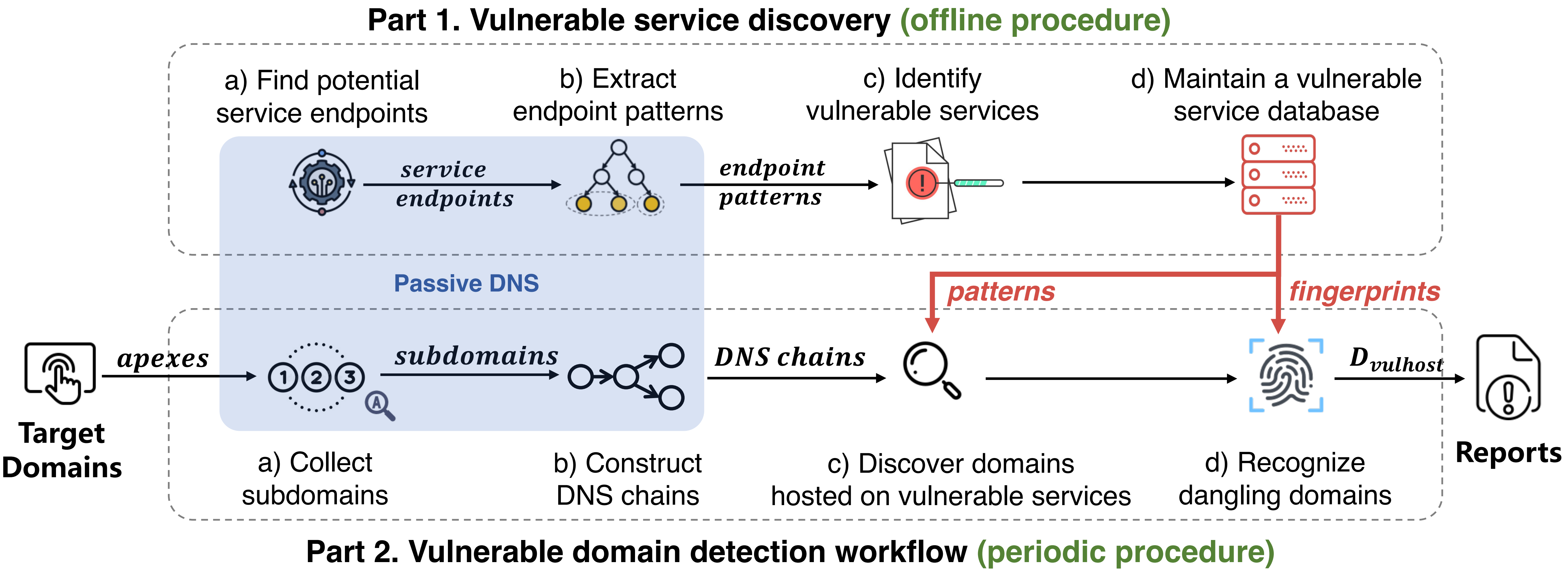
b) Extract endpoint patterns

c) Identify vulnerable services

d) Maintain a vulnerable service database

Passive DNS

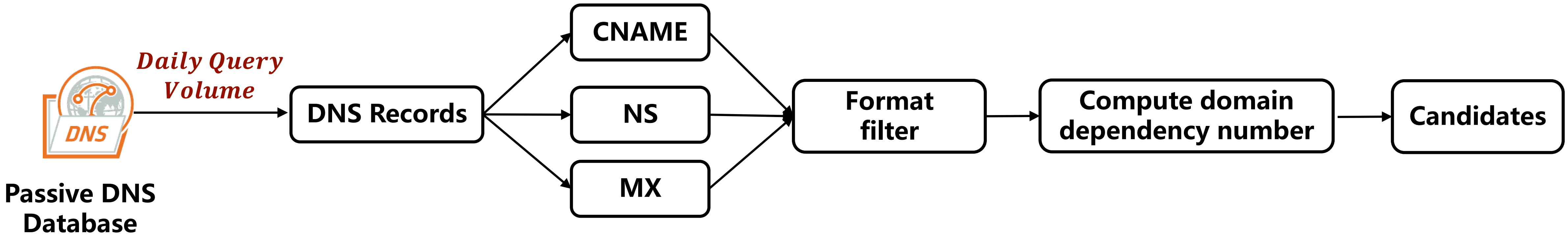
DareShark Workflow



Part 1. Vulnerable service discovery (offline procedure)
a) Find potential service endpoints
b) Extract endpoint patterns
c) Identify vulnerable services
d) Maintain a vulnerable service database
service endpoints
endpoint patterns
Passive DNS
patterns
fingerprints
Target Domains
apexes
subdomains
DNS chains
$D_{vulhost}$
Reports
a) Collect subdomains
b) Construct DNS chains
c) Discover domains hosted on vulnerable services
d) Recognize dangling domains
Part 2. Vulnerable domain detection workflow (periodic procedure)

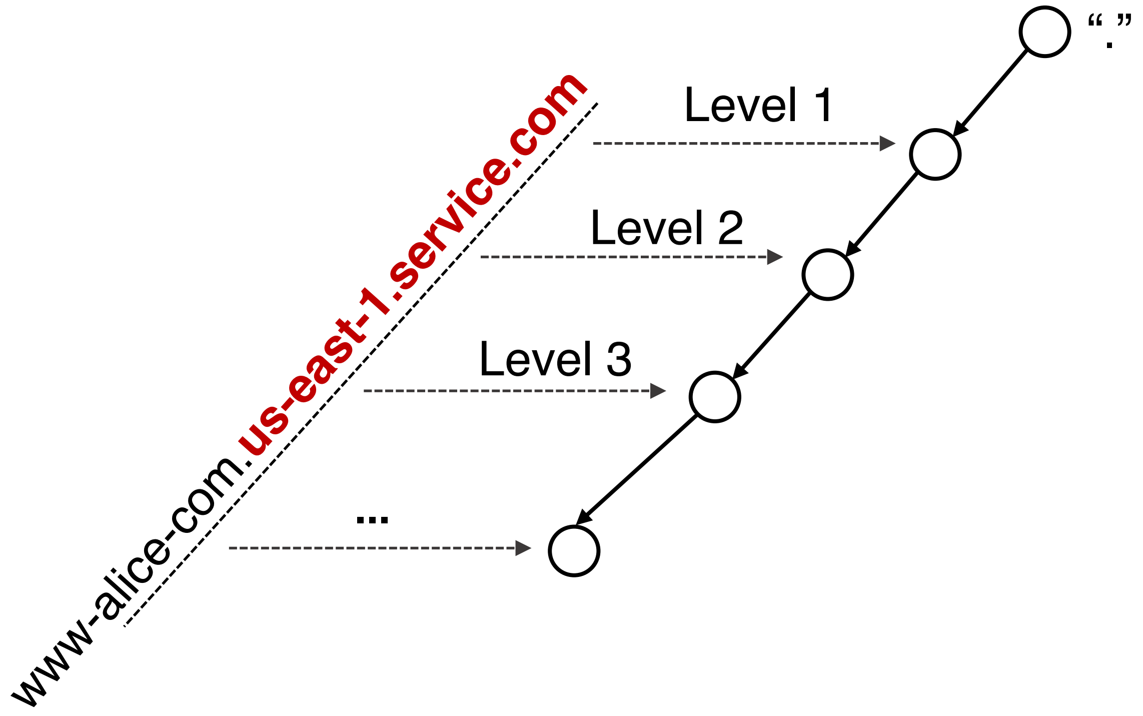
Part I: Discovering Vul. Services

- **Step 1: Finding service endpoint candidates**
  - Filtering endpoint domains by DNS resolution popularity and domain dependency.



Part I: Discovering Vul. Services

➢ Step 2: Extracting endpoint patterns via a Domain Suffix Tree



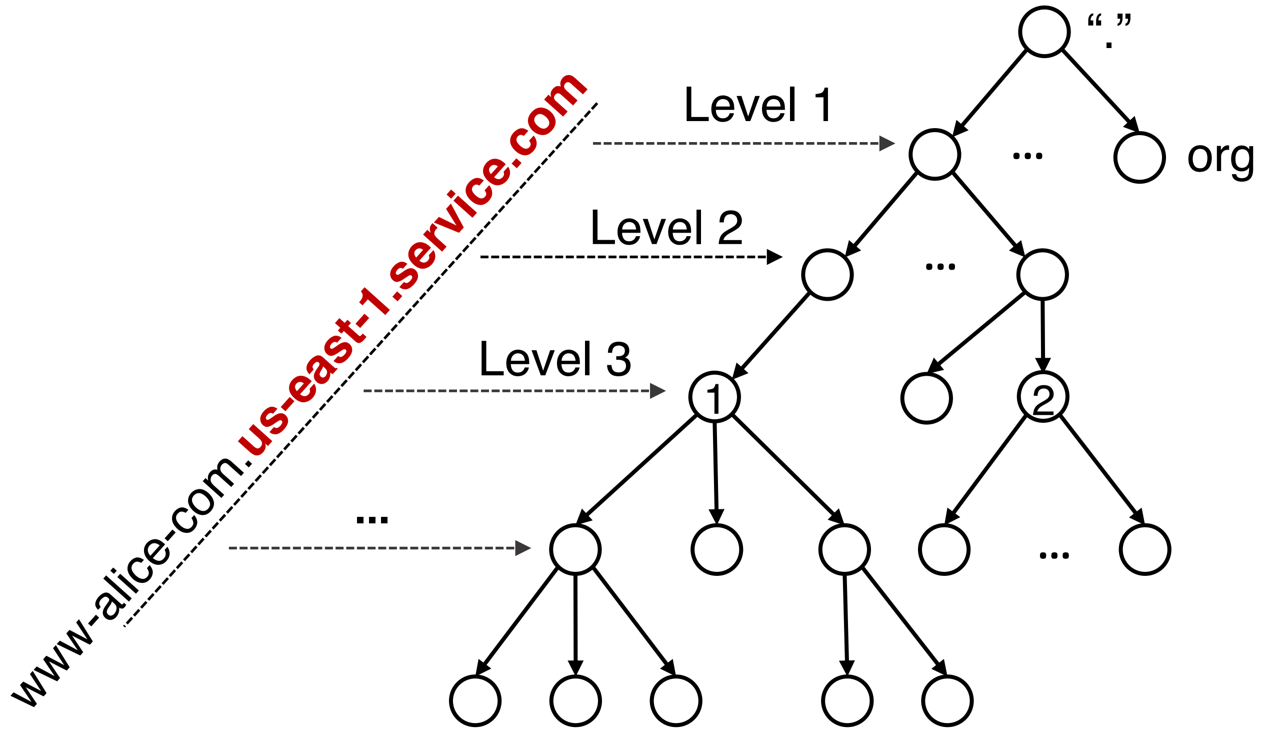
Domain Tree

Domain Tree Construction:

- The root is "." , and children nodes are eTLDs, apex domains, apex+1, apex+2, and so on

Part I: Discovering Vul. Services

➢ Step 2: Extracting endpoint patterns via a Domain Suffix Tree



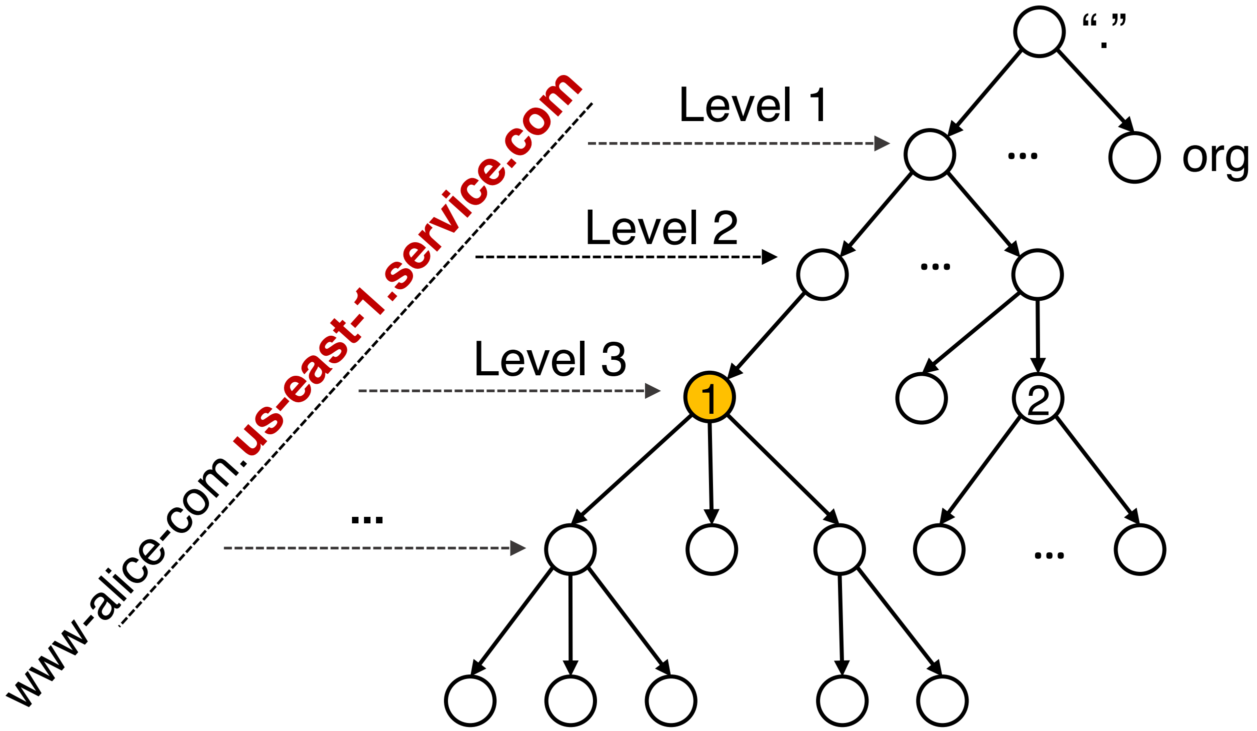
Domain Tree

Domain Tree Construction:

- The root is "." , and children nodes are eTLDs, apex domains, apex+1, apex+2, and so on

Part I: Discovering Vul. Services

- **Step 2: Extracting endpoint patterns via a Domain Suffix Tree**



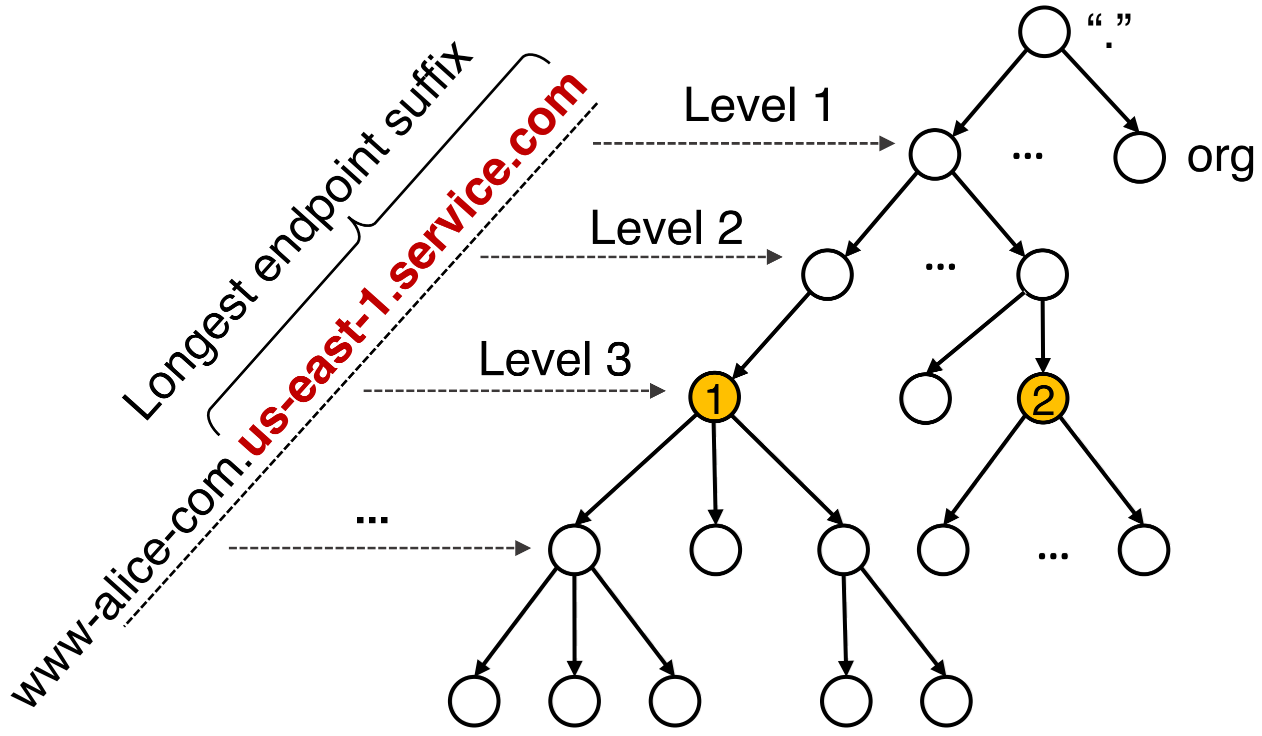
Domain Tree

Tree node attributes (Example of Node 1)

```
{
   "name"        : "us-east-1.service.com",
   "suffixLevel": 3,
   "DN"          : Dependency Number,
   "subCount"    : 3,
   "subList"     : ['a', 'b', 'c'],
   "subEntropy" : Shannon entropy of subList
}
```

Part I: Discovering Vul. Services

➢ **Step 2: Extracting endpoint patterns via a Domain Suffix Tree**



Domain Suffix Tree (DST)

Domain Tree Pruning

- Prune the tree from the bottom up, by limiting number of hosted FQDNs, subCount, and subEntropy of each node

Part I: Discovering Vul. Services

- Step 2: Extracting endpoint patterns via a **Domain Suffix Tree**

  - **Service Endpoint Examples**

Services	Endpoint Names (endpoint patterns)
Aliyun OSS	alice.storage.com.oss-cn-hongkong.aliyuncs.com
Amazon S3	a.b.c.d.s3.us-east-1.amazonaws.com ab-cd.s3.dualstack.us-gov-west-1.amazonaws.com
GitHub	abcd.github.io

Part I: Discovering Vul. Services

- **Step 3: Identifying services and checking service vulnerabilities**
  - **Narrow down the candidate list of endpoint patterns**

    e.g., remove highly randomized endpoint domains

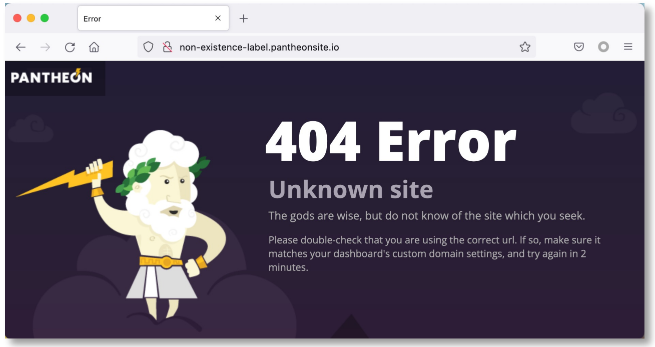
  - **Map endpoint patterns to services**

    e.g., access homepages, dig through search engines

  - **Check vulnerabilities in domain connection and domain ownership validation**

Part I: Discovering Vul. Services

➢ Step 4: Maintaining a database for vulnerable services

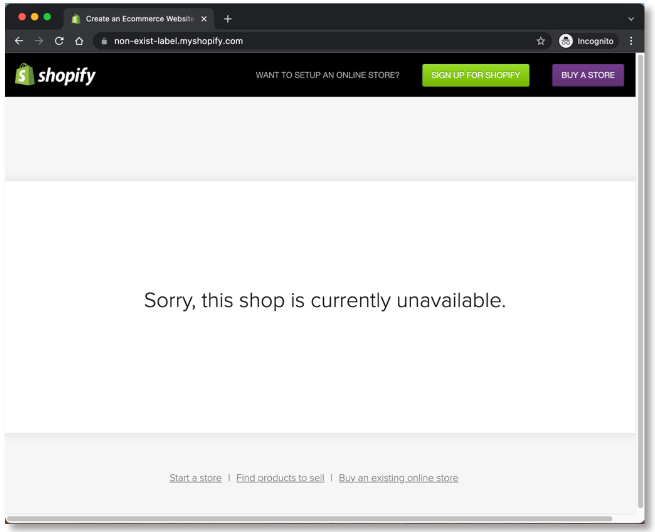


404 Error

Unknown site

The gods are wise, but do not know of the site which you seek.

Please double-check that you are using the correct url. If so, make sure it matches your dashboard's custom domain settings, and try again in 2 minutes.


Create an Ecommerce Website
non-exist-label.myshopify.com
Incognito
shopify
WANT TO SETUP AN ONLINE STORE?
SIGN UP FOR SHOPIFY
BUY A STORE
Start a store | Find products to sell | Buy an existing online store

Sorry, this shop is currently unavailable.

Vulnerable Service Fingerprints

Type	Response Example	# Banner	# Service	# Vendor
HTTP Response		**106**	**59**	**48**
Header	"404 Unknown site"	14	13	10
Body	"NoSuchBucket"	92	52	47
DNS Answer		**4**	**13**	**9**
NX-CNAME[1]	status:NXDOMAIN	1	11	7
Default Rdata[2]	127.0.0.1 nx.aicdn.com	3	2	2
Total		**110**	**64**	**51**

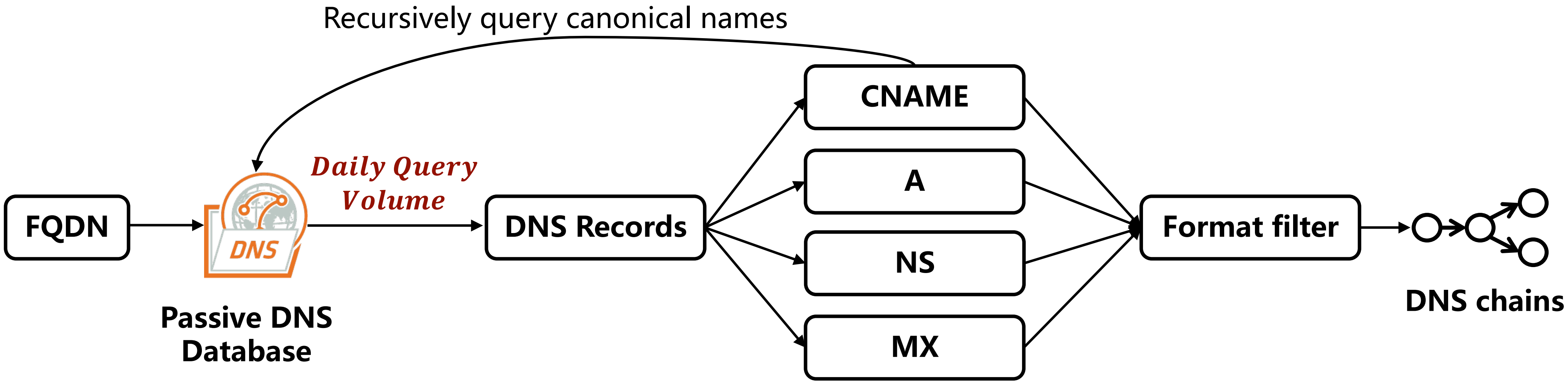
Part II: Detecting Hosting-based Dares

➢ Collecting subdomain names from passive DNS logs

- ➢ Legal format **[RFC 1034] Domain Names - Concepts And Facilities**
- ➢ Filter disposable domains created on demand

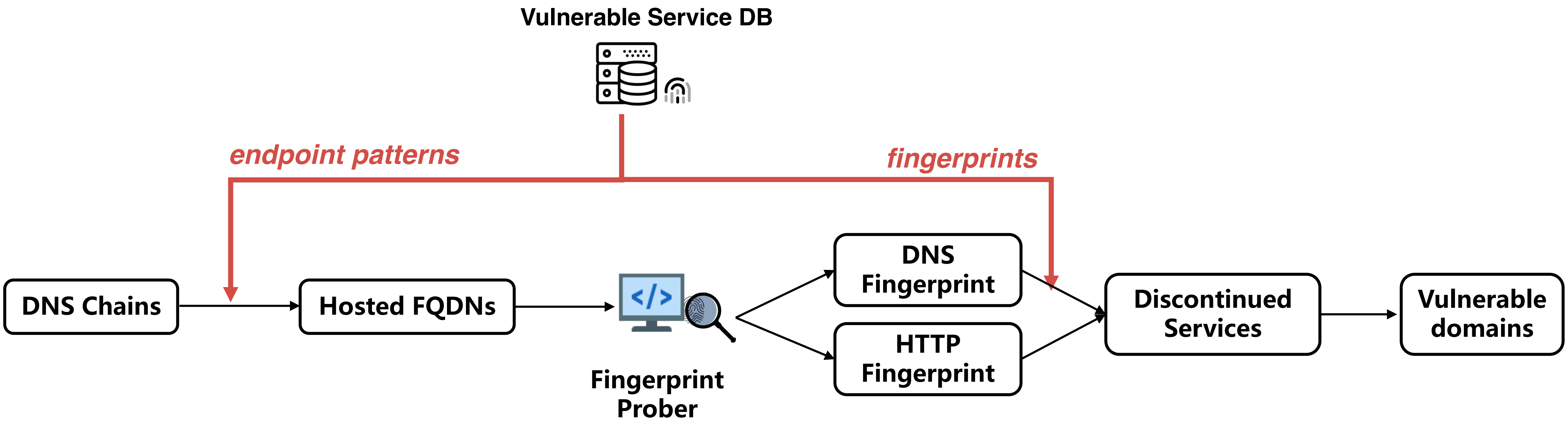
  e.g., scanning, convey “one-time signals” $Total\ Query\ Volume > 100$

➢ Reconstructing domain dependencies (DNS chains)



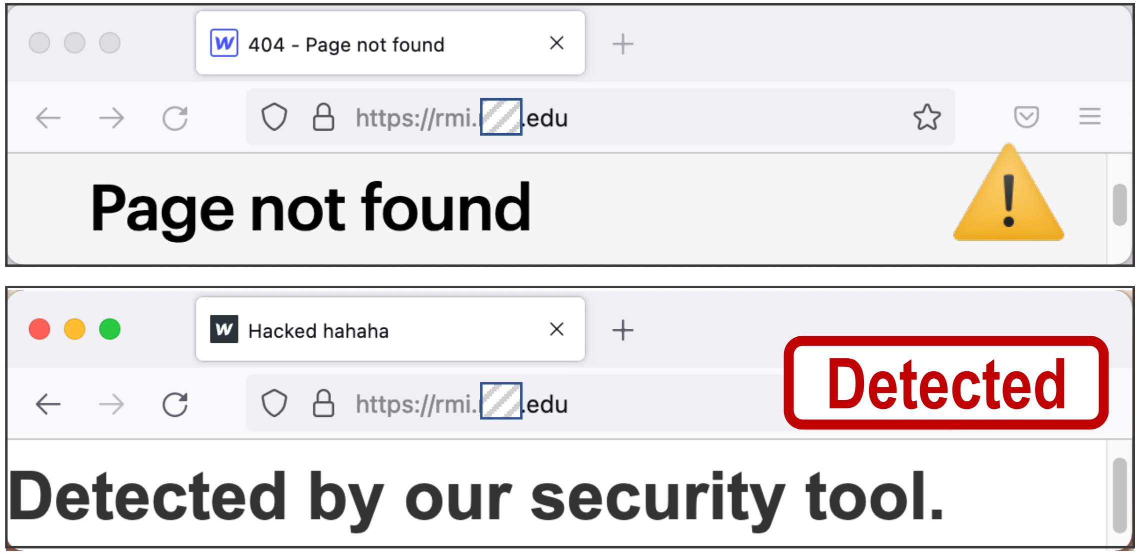
Part II: Detecting Hosting-based Dares

➢ **Probing hosted domains to inspect service status**



Part II: Detecting Hosting-based Dares

➢ **Probing hosted domains to inspect service status**



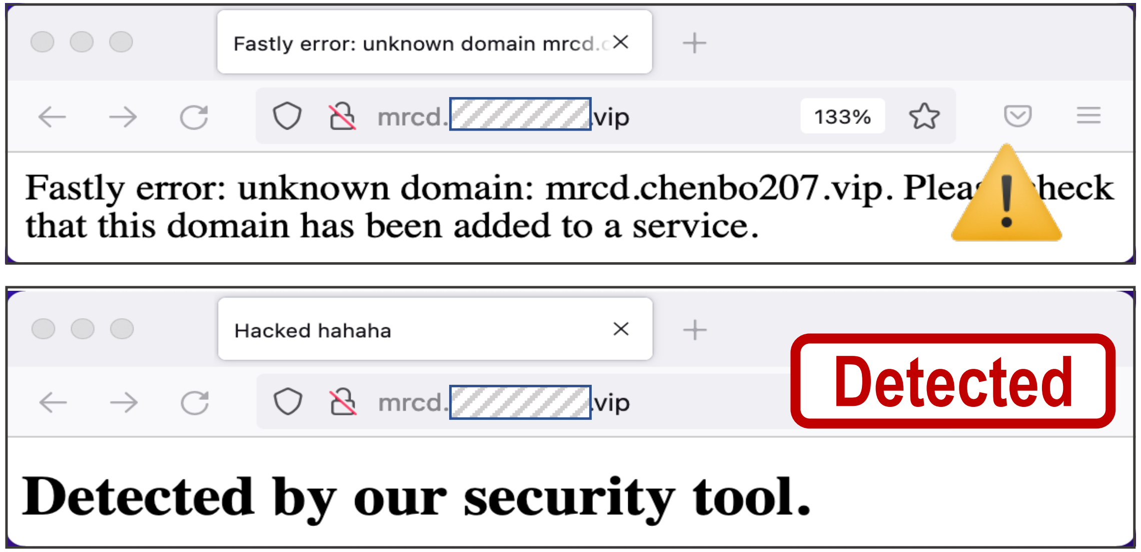
Page not found ⚠

Detected

Detected by our security tool.

(1) Webflow

rmi.xxxx.edu



Fastly error: unknown domain: mrcd.chenbo207.vip. Plea⚠heck that this domain has been added to a service.

Detected

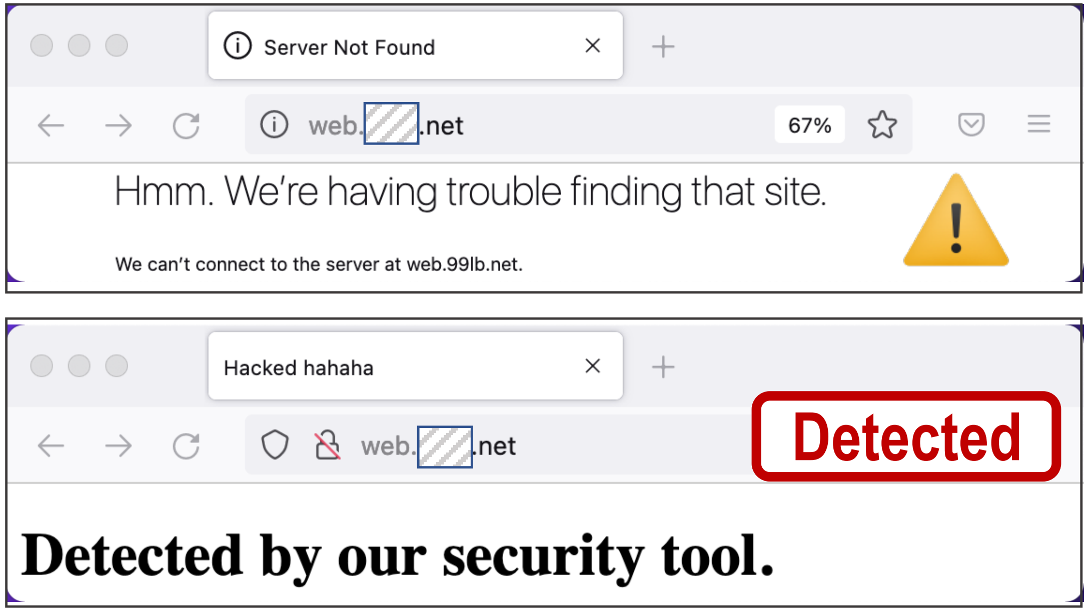
Detected by our security tool.

(2) Fastly

mrcd.xxxxxxxxx.vip

Part II: Detecting Hosting-based Dares

➢ **Probing hosted domains to inspect service status**



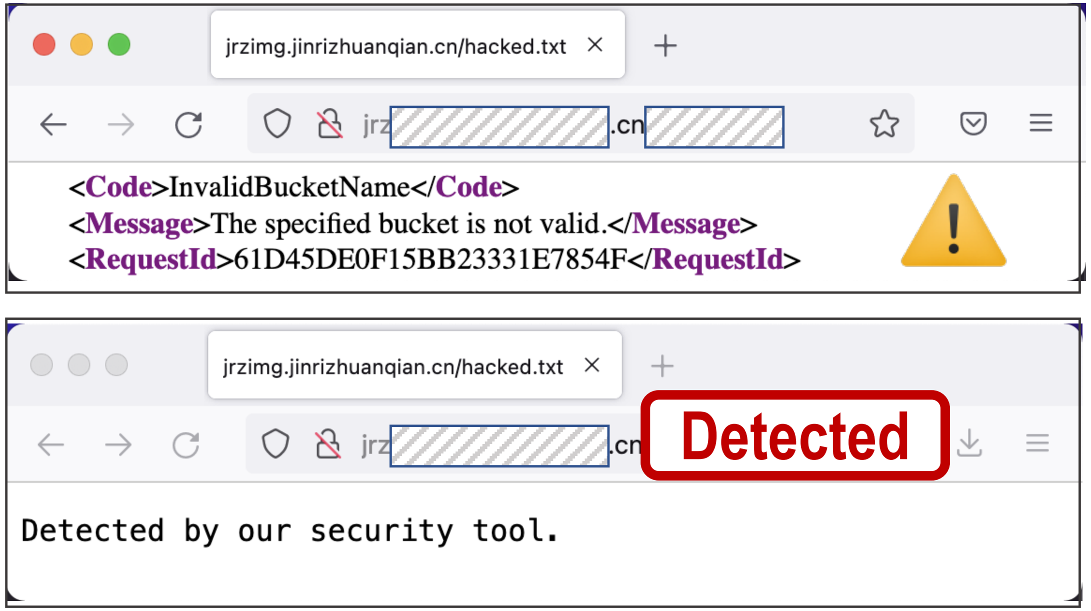
Hmm. We're having trouble finding that site.

We can't connect to the server at web.99lb.net.

Detected by our security tool.

(3) Cloudflare

web.xxxx.net


jrzimg.jinrizhuanqian.cn/hacked.txt
jrz[redacted].cn[redacted]
jrzimg.jinrizhuanqian.cn/hacked.txt
jrz[redacted].cn
Detected

```
<Code>InvalidBucketName</Code>
<Message>The specified bucket is not valid.</Message>
<RequestId>61D45DE0F15BB23331E7854F</RequestId>
```

```
Detected by our security tool.
```

(4) Alibaba Cloud

rrzxxx.xxxxxxxxxxxxx.cn

DareShark Deployment

- **Passive DNS dataset**
  - DNS response data from public DNS resolvers for **114DNS**, the largest DNS provider in China
  - **600B** DNS queries per day, covering **99.9%** of Tranco Top 1M domains
  - DNS queries originate from telecom companies (e.g., China Telecom), research institutions (e.g., MIT and NUS), and large providers (e.g., Alibaba and Google)

What did we find for hosting services?

- The current practice of hosting services is in a mass, resulting in various types of service vulnerable to domain takeover.

Vulnerable Hosting Services

- 65 services vulnerable to domain takeover threats.
- Vulnerable services comprise a variety of service types.

Categories	# Vendor		# Endpoint Patterns		# Services	
	All	Vulnerable	All	Vulnerable	All	Vulnerable
Cloud Storage	7	7	130	118	12	9
CDN	25	7	247	31	44	8
Website Builder	51	40	156	105	60	44
Others	27	4	462	4	49	4
Newly Discovered	55	19	920	183	125	34
All	88	52	995	258	165	65

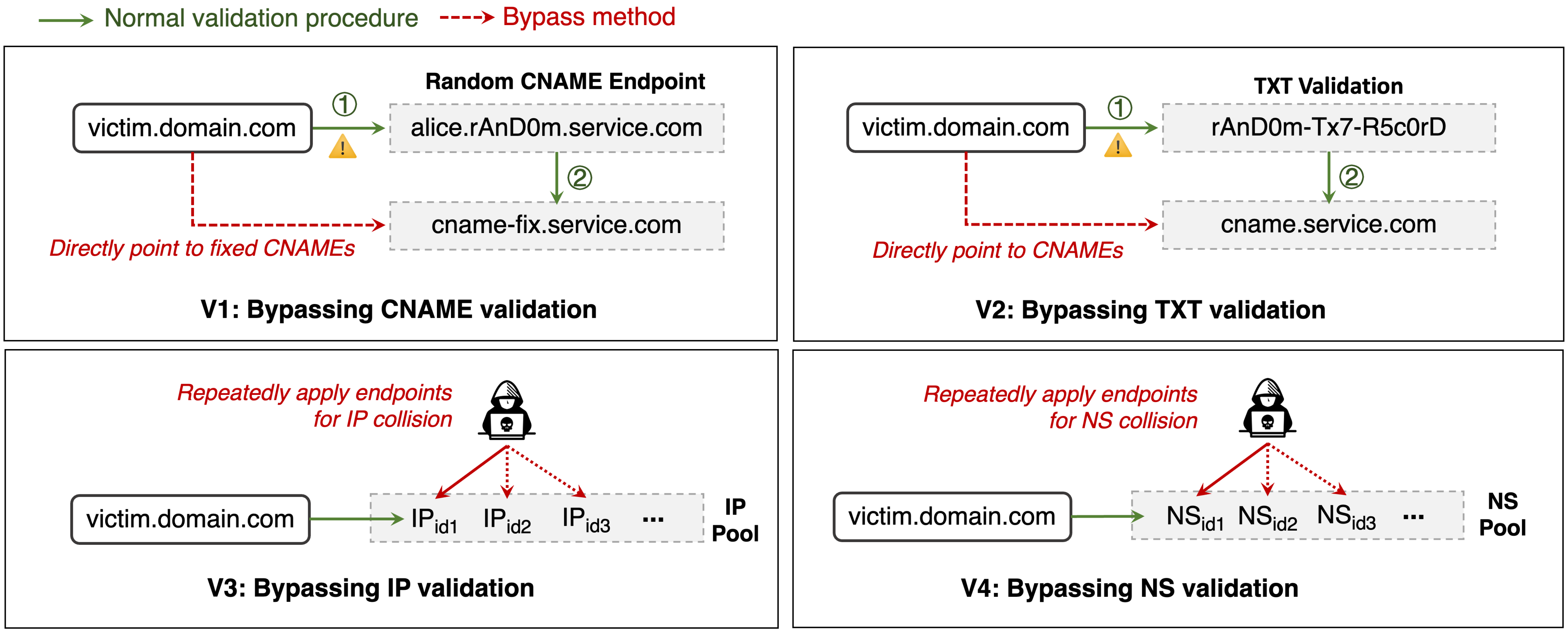
Vulnerable Hosting Services

- 7/9 domain connecting methods are exploitable

Method	Type	Connect a custom domain to...	# Services	Exploitable
CNAME	M1	Fixed canonical domains	12	●
	M2	Any canonical domains customized by any users	70	●
	M3	New canonical domains customized by new users	12	○
	M4	The canonical domains allocated from a candidate pool	5	◐
	M5	Canonical domains containing newly generated random labels	47	○
NS	M6	Fixed nameservers	1	●
	M7	The nameservers allocated from a candidate pool	5	◐
IP	M8	Fixed IPs	8	●
	M9	The IPs allocated from a candidate pool	4	◐

Vulnerable Hosting Services

➢ 4 new threat models that can bypass flawed DOV



→ Normal validation procedure ---→ Bypass method

Random CNAME Endpoint

victim.domain.com ① alice.rAnD0m.service.com ② cname-fix.service.com

Directly point to fixed CNAMEs

V1: Bypassing CNAME validation

TXT Validation

victim.domain.com ① rAnD0m-Tx7-R5c0rD ② cname.service.com

Directly point to CNAMEs

V2: Bypassing TXT validation

Repeatedly apply endpoints for IP collision

victim.domain.com IP_{id1} IP_{id2} IP_{id3} ··· IP Pool

V3: Bypassing IP validation

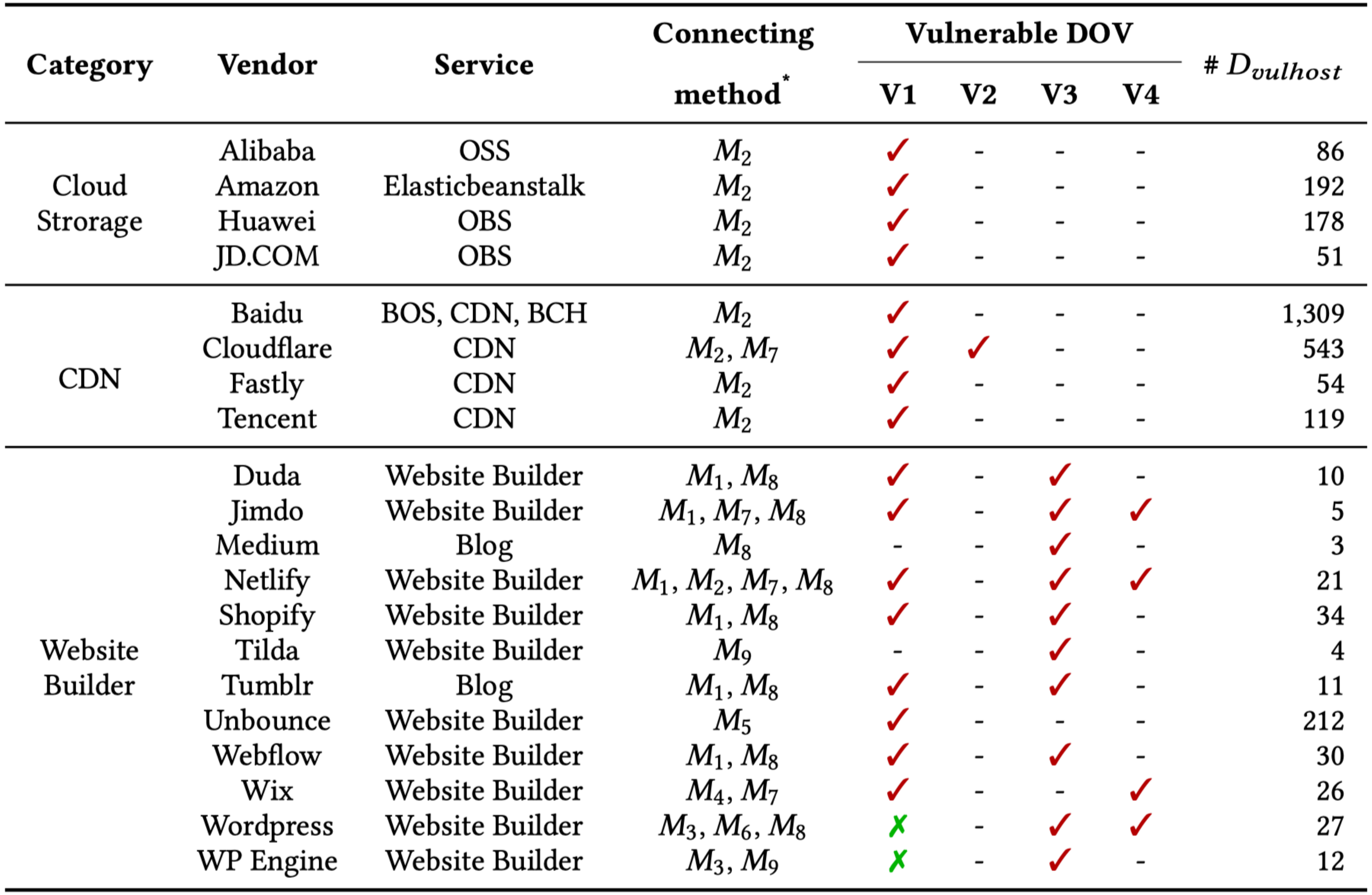
Repeatedly apply endpoints for NS collision

victim.domain.com NS_{id1} NS_{id2} NS_{id3} ··· NS Pool

V4: Bypassing NS validation

Vulnerable Hosting Services

- **Top 20 vendors with 70% market share are vulnerable**



Category	Vendor	Service	Connecting	Vulnerable DOV				# $D_{vulhost}$
			method*	V1	V2	V3	V4	
Cloud Strorage	Alibaba	OSS	M_2	✓	-	-	-	86
	Amazon	Elasticbeanstalk	M_2	✓	-	-	-	192
	Huawei	OBS	M_2	✓	-	-	-	178
	JD.COM	OBS	M_2	✓	-	-	-	51
CDN	Baidu	BOS, CDN, BCH	M_2	✓	-	-	-	1,309
	Cloudflare	CDN	M_2, M_7	✓	✓	-	-	543
	Fastly	CDN	M_2	✓	-	-	-	54
	Tencent	CDN	M_2	✓	-	-	-	119
Website Builder	Duda	Website Builder	M_1, M_8	✓	-	✓	-	10
	Jimdo	Website Builder	M_1, M_7, M_8	✓	-	✓	✓	5
	Medium	Blog	M_8	-	-	✓	-	3
	Netlify	Website Builder	M_1, M_2, M_7, M_8	✓	-	✓	✓	21
	Shopify	Website Builder	M_1, M_8	✓	-	✓	-	34
	Tilda	Website Builder	M_9	-	-	✓	-	4
	Tumblr	Blog	M_1, M_8	✓	-	✓	-	11
	Unbounce	Website Builder	M_5	✓	-	-	-	212
	Webflow	Website Builder	M_1, M_8	✓	-	✓	-	30
	Wix	Website Builder	M_4, M_7	✓	-	-	✓	26
	Wordpress	Website Builder	M_3, M_6, M_8	✗	-	✓	✓	27
	WP Engine	Website Builder	M_3, M_9	✗	-	✓	-	12

What did we find for domain takeover?

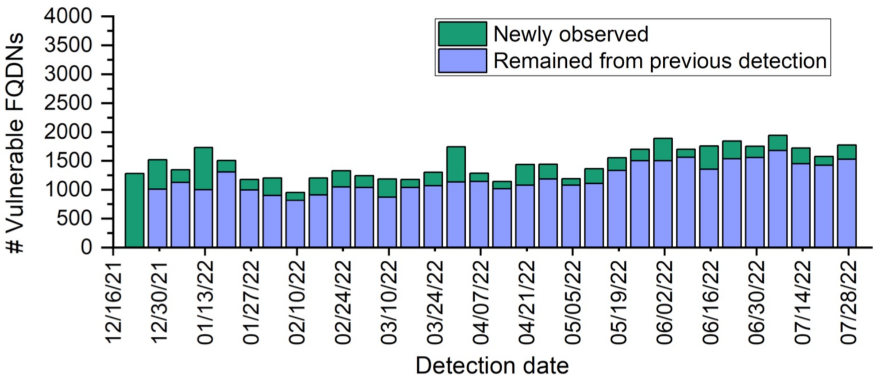
- Hosting-based domain takeover threats are still prevalent.

Measurement and Findings

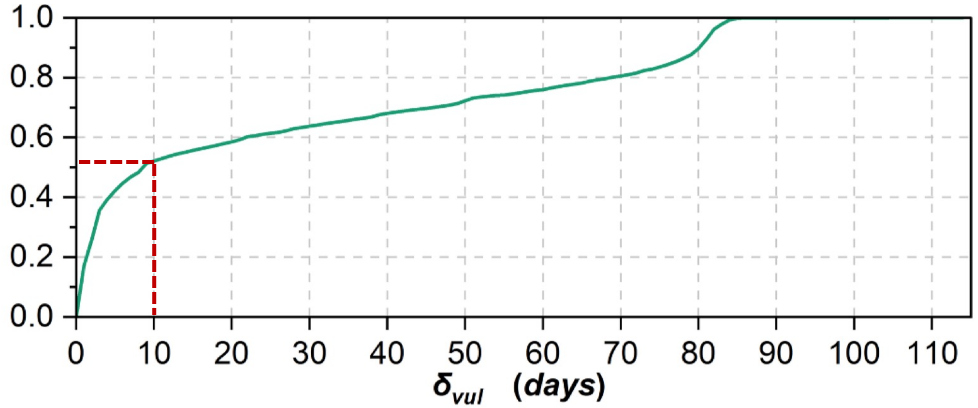
- **Detection target domains**
  - **Tranco Top 1M apex domains +9,808 .edu and 7,198 .gov apexes**
  - **We collect 11,446,359 subdomains from PDNS for all apexes.**
- **Longitudinal and periodic measurement**
  - **101 rounds (Dec. 16, 2021 – Jul. 28, 2022)**
  - **~1 day/round**

Measurement and Findings

- **114,063 (1.0%) FQDNs have been hosted on vulnerable services**
- **10,351 FQDNs are vulnerable, covering 2,096 apex domains**
  - Reputable universities (e.g., Stanford and Rice)
  - Famous companies (e.g., Baidu, Huawei, and Marriott).
- **Hosting-based domain takeover appears frequently and long-lasting**



Weekly cumulative detection results.



Distribution of vulnerable days

270 new vulnerable domains emerge per week.

Over 50% remain vulnerable for over 10 days.

Conclusion

- **DareShark: A novel and effective detection framework**
  - High efficiency and coverage
- **Comprehensive measurements**
  - 7-month longitudinal measurement on Tranco 1M apexes' subdomains
  - Detect 10,351 vulnerable domains (8x more than previous study)
- **Systematic service inspection and threat analysis**
  - Discover 65 vulnerable services and new security flaws
  - Receive vulnerability confirmation from 10 vendors, and provide solutions

aws HUAWEI HUAWEI CLOUD 腾讯云 百度智能云 duda JIMDO

Thanks for listening!

Any question?

Xiang Li, Tsinghua University

x-l19@mails.tsinghua.edu.cn