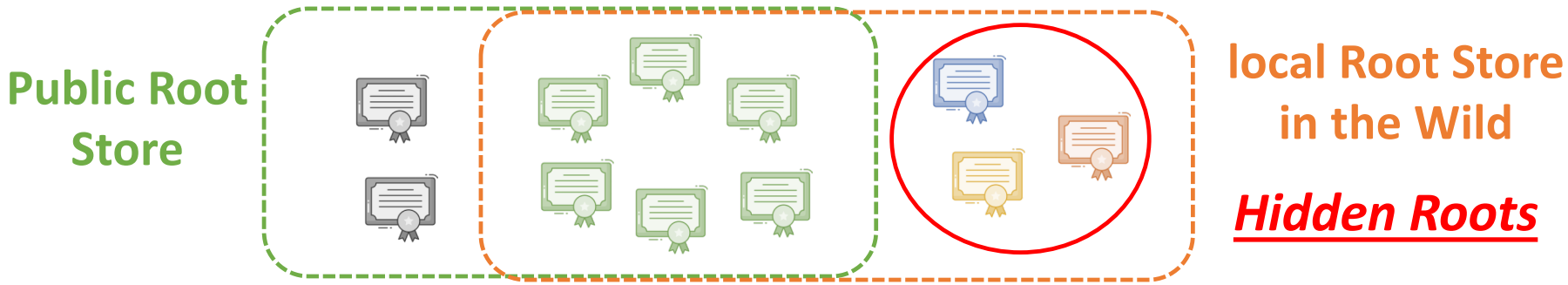
Rusted Anchors: A National Client-Side View of Hidden Root CAs in the Web PKI Ecosystem

Yiming Zhang, Baojun Liu, Chaoyi Lu, Zhou Li,
Haixin Duan, Jiachen Li, Zaifeng Zhang

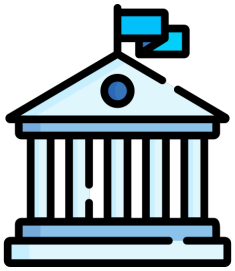
Summary



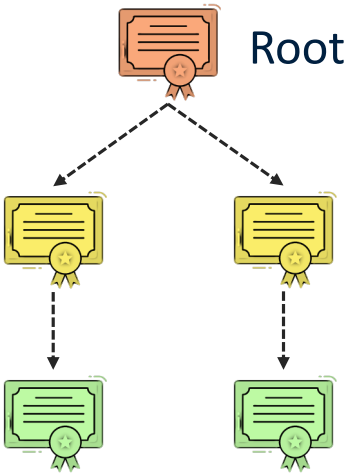
- Provided the first client-side, nation-wide measurement study of hidden root CAs in Web PKI ecosystem.
- Collected 1.17 million hidden roots from volunteers' local root stores and characterized 5,005 organizations that hold them.
- Uncovered the massive and dynamic hidden CA ecosystem, revealed the serious flawed implementations, which could be critical for Web PKI Security.

Background

Trust Model in Web PKI



Certificate Authority



Root Certificates as the trust anchors

The connection will only be considered **authenticated** if the **root** is **trusted**.

Trust Anchors - Root Stores

- Public root stores

Mozilla Microsoft Apple

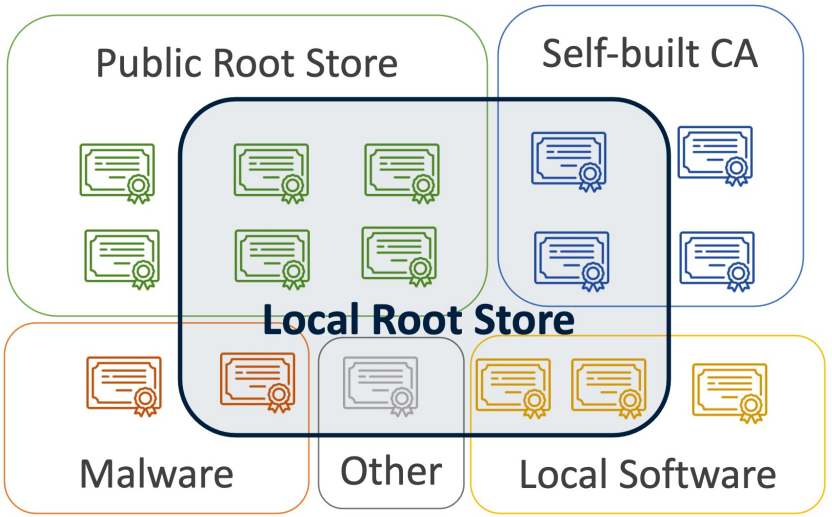
~ 590 Root Certs*

Pre-configured on local operation systems/browsers

Being rigorously reviewed and supervised

- Root stores in the wild

Local root store could be modified by several parties…



Overview of real-world CA ecosystem

** according to the snapshot statistics in Sept. 2020*

Security Incidents of Unregulated Root Stores

- Modifications of local trust lists could pose serious security problems.

Alert (TA15-051A)

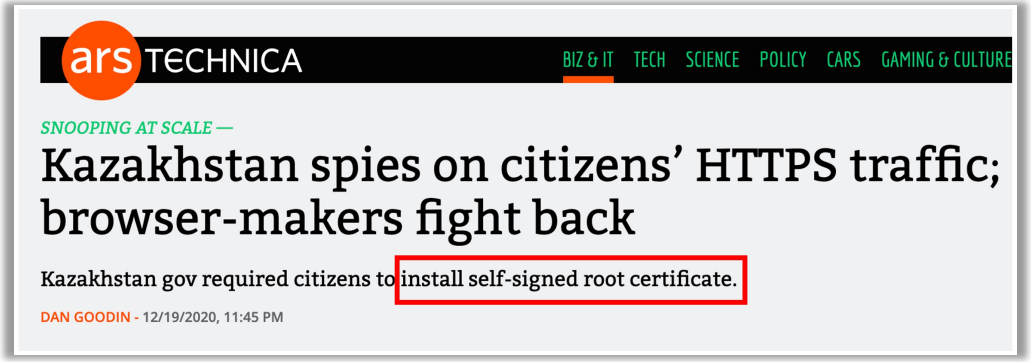
Lenovo Superfish Adware Vulnerable to HTTPS Spoofing

Original release date: February 20, 2015 | Last revised: September 30, 2016

Print Tweet Send Share

Systems Affected

Lenovo consumer PCs that have Superfish VisualDiscovery installed.



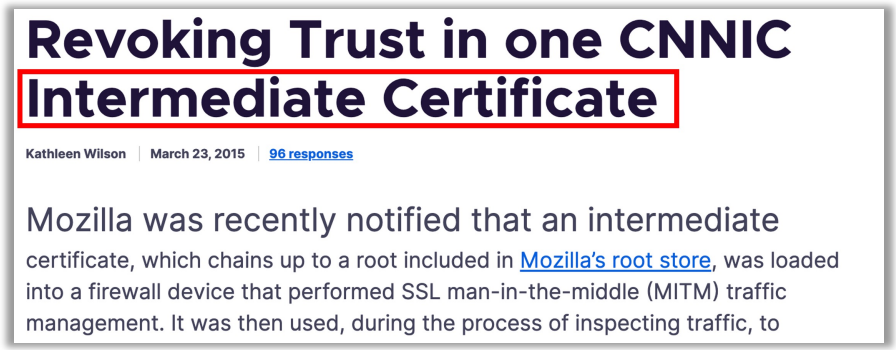
ars TECHNICA BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE

SNOOPING AT SCALE —

Kazakhstan spies on citizens' HTTPS traffic; browser-makers fight back

Kazakhstan gov required citizens to install self-signed root certificate.

DAN GOODIN - 12/19/2020, 11:45 PM



Revoking Trust in one CNNIC Intermediate Certificate

Kathleen Wilson | March 23, 2015 | 96 responses

Mozilla was recently notified that an intermediate certificate, which chains up to a root included in Mozilla's root store, was loaded into a firewall device that performed SSL man-in-the-middle (MITM) traffic management. It was then used, during the process of inspecting traffic, to

Home / Security / News

NEWS

Worse than Superfish? Comodo-affiliated PrivDog compromises web security too

By Lucian Constantin

PCWorld | FEB 23, 2015 12:37 PM PST

Research Goal: Explore Hidden Root Ecosystem

Hidden root CAs: root certificates absent from public root programs while have been imported into local root stores (gained trust from web clients).

Previous studies: all lack the client-side view of local root stores.

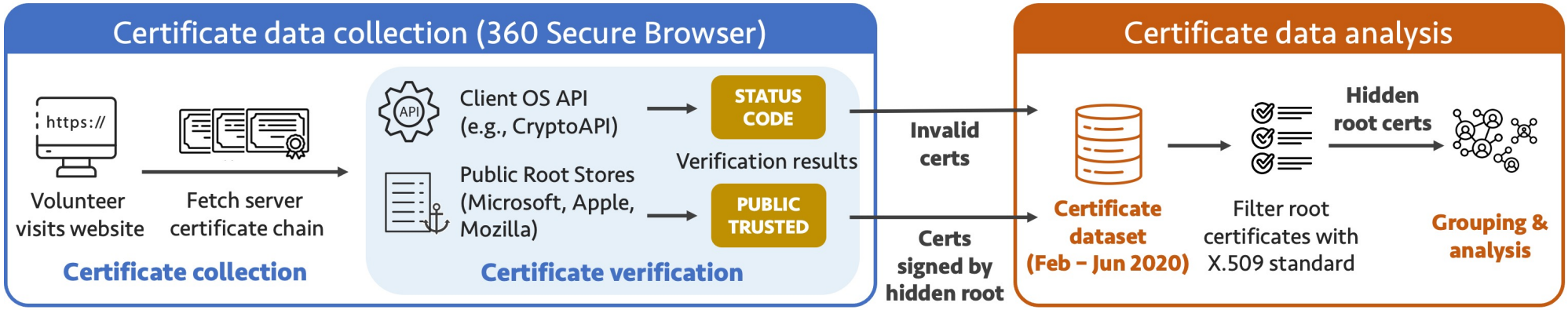
We give the first client-side, nation-wide view of hidden root CAs in the Web PKI ecosystem

Research Questions

Scale	Impact	Operators	Security Risks
Hidden root CAs CA groups	Web clients HTTPS Traffic	Organizations own hidden roots	Implementation flaws Malicious behaviors

Methodology

Methodology Overview



- Collaborate with **360 Secure Browser** (over 100 million monthly active users in China mainland)

STATUS CODE Verified by operating system API CryptpAPI on Windows

COMMON_NAME_INVALID, DATE_INVALID
AUTHORITY_INVALID, REVOKED, …

PUBLIC TRUSTED Examine whether the root includes in public root programs (Mozilla, Windows, Apple)

Check in backstage by matching the public key
Unset when the root is excluded

- Chains with **invalid verification results** or **link to hidden roots** would be collected

Identify Hidden Root Certificates

- **Collecting Period:** Feb 1, 2020-Jun 30, 2020
- **Collected data**: (anonymized)client ID, time of collection, hashes and PEM encoding of certificates, public flags, verification results.

  Chains may be disordered or incomplete
- **Filter root certificates**

Requirements of root cert

```
basicConstraints: CA = True
keyUsage: KeyCertSign
subject == issuer (non empty)
subjectKeyIdentifier ==
authoritykeyIdentifier
```


```
{
   "client_id": 1579968000000, "host": example.com,
   "collected_time": 2020-02-01T00:00:00Z,
   "leaf_sha1": "35 FE 12 … 83 5A FB",
   "parents_sha1": ["A3 A1 … D7 3A", "A8 98 … 54 36"],
   "status_code": [AUTHORITY_INVALID, WEAK_KEY, DATE_INVALID],
   "public_trusted": False
   "certs_pem": [(leaf_cert_pem), (parent1_pem), (parent2_pem)]
}
```


Example of collected data

TBSCertificate	
issuer	countryName = US, organizationName = cPanel, Inc., commonName = cPanel, Inc. Certification Authority
validity	notBefore = 20210311000000Z, notAfter = 20210610000000Z
subject	commonName = sigsac.hosting.acm.org
subjectPublicKeyInfo	subjectPublicKey = D4 06 E9 CE … 8E 33 45 C1 (2048 bits)
extensions	subjectKeyIdentifier: keyIdentifier = C0 73 03 … 52 A0 9F
	authorityKeyIdentifier: keyIdentifier = 7E 03 5A … 6A C7 65
	basicConstraints: cA = FALSE
	keyUsage = digitalSignature, keyEncipherment
Signature = 5E AB F4 D0 CD C6 … 9C CE F0 D1 79 26 (2048 bits)	

Part of an X.509 Version 3 certificate

Classify Hidden Root Groups

- **Target**: inferring certificate ownership

Public knowledge base not suitable

- **Our approach**: leverage subject distinguished names for grouping hidden roots

```
subject:
    commonName = GlobalSign Root CA
    organization = GlobalSign nv-sa
    organizationUnit = Root CA
    countryName = BE
    ……
```

`subject` template example:
`commonName = whistle.[0-9]*`

Drain algorithm to identify templated subject

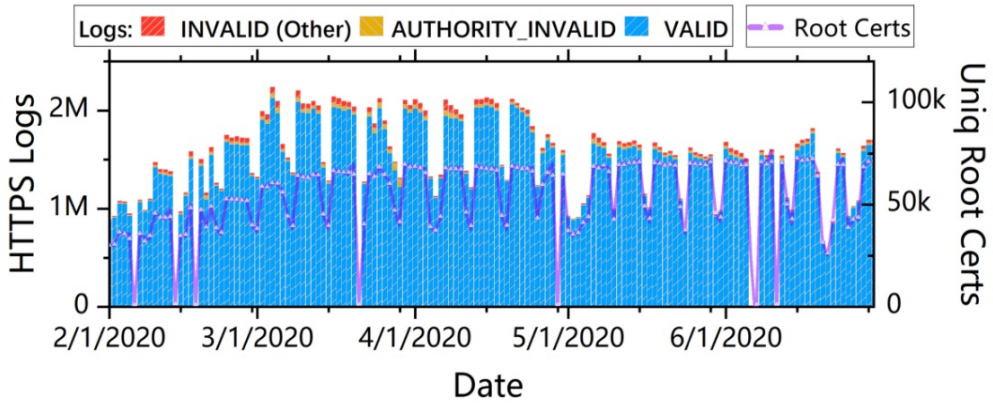
- **Grouping criteria**: identical (non-empty) or templated subject fields.

Scale and Impact Analysis

Overview of Collected Dataset

Type of Root Cert	Filtering Condition	# Distinct Root Certs	# Cert Chains	# Leaf Certs	# FQDNs
Hidden CAs	PUBLIC_TRUSTED=FALSE, Trusted by at least one client	1,175,145 (98.24%)	222,977,356	59,817,585	1,333,931
	PUBLIC_TRUSTED=FALSE, Rejected by all clients	21,010 (1.76%)	263,109	112,946	15,566
Public Trusted CAs	PUBLIC_TRUSTED=TRUE, STATUS_CODE has bits set	615	241,541,342	3,647,095	1,871,131

- 1.19 million hidden roots collected in 5 months, with 98.24% having been implanted into local root stores.
- 222 million HTTPS traffic, 1.33 million FQDNs, and 5.07 million web clients have been affected.



Daily count of hidden roots and associated HTTPS web visits

Hidden CA Group Overview

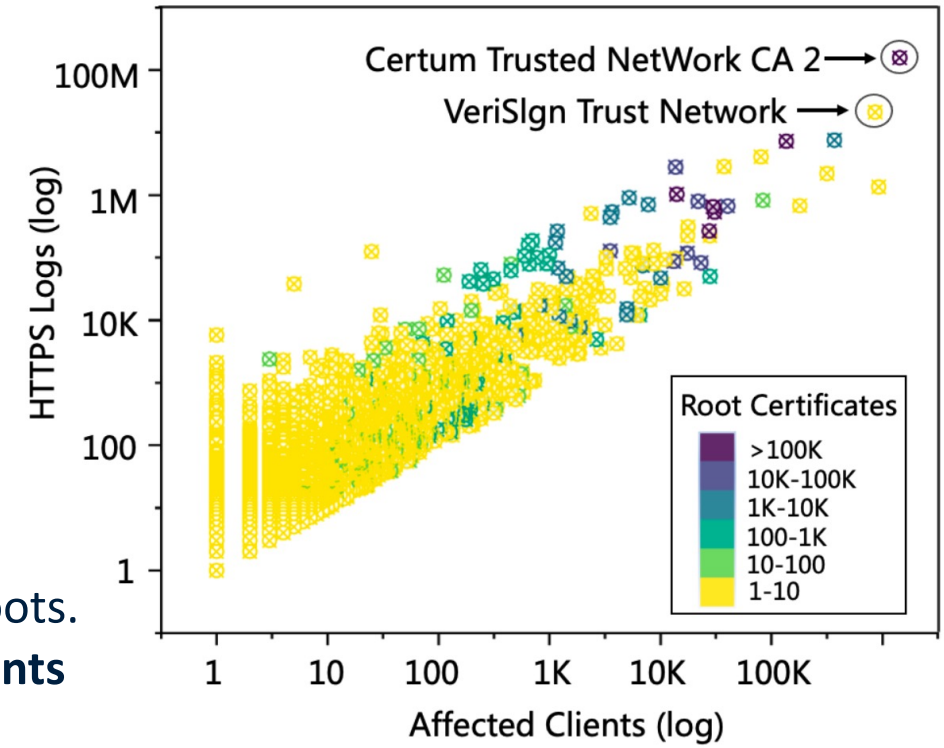
- Found **5,005** groups on 1.17 million hidden roots (imported by clients).
- Root certificates in **one group** are considered from **the same organization**.

Distribution

Long tail distribution: 4,362 groups (87.2%) holds only one cert.

Top Groups

The largest group holds 254k roots.
Top 100 groups: over **1,000 clients** and **5,000 HTTPS visits**.



Size and impact of certificate groups

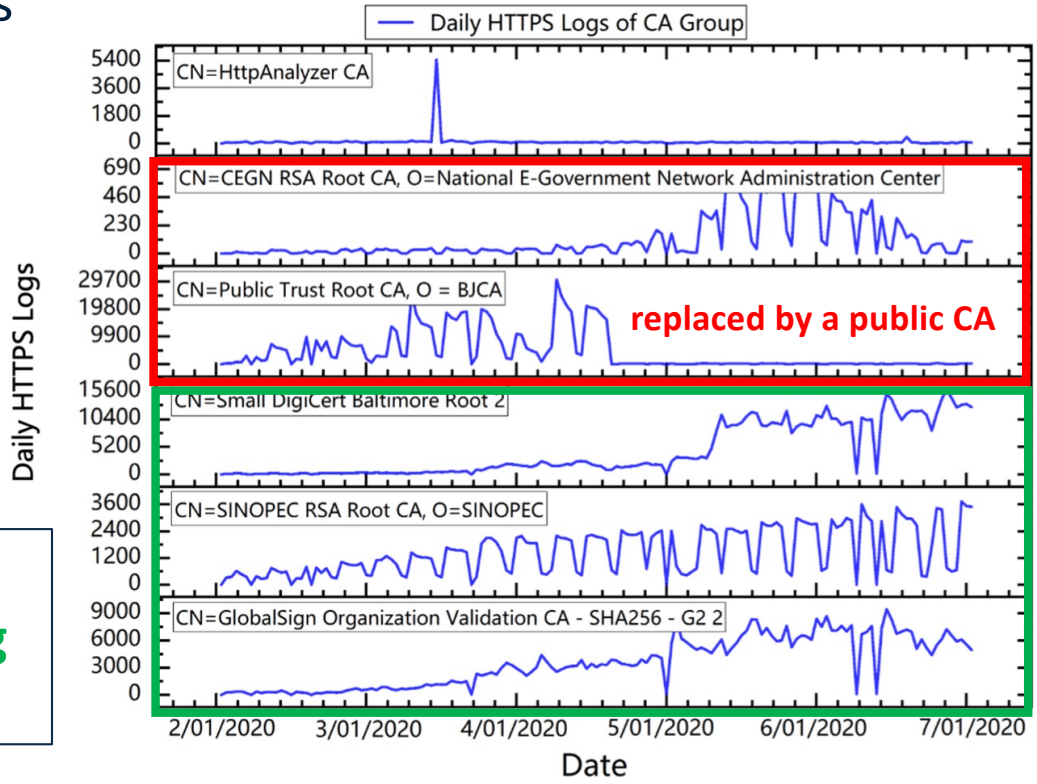
Dymanically Updating Ecosystem

- **Active time**: days when certificate chains linked to the roots when captured in our datasets.

Top 100 groups: **146 days** on average.

- **Stability:** the **coefficient of variation** of daily HTTPS visits on active days

The ecosystem of hidden root CAs is **dynamic and updating**, as new **emerging** CAs and **retiring** ones are both observed.



Temporary, retiring, and emerging group examples

Categories of Hidden CA Groups

- **Infer CA group categories**: 3 security researchers manually inspect the semantics in `subject` fields, and search keywords/hashes when subject is not identifiable.

Categories	# clusters (groups)	# hidden root certs	# affected connections	# affected clients	Invalid (Authority)	Invalid (Other)	Example of hidden root certificate
Enterprise Self-built	24	48	2,071,344	199,743 (3.94%)	35.54%	75.66%	CN = SZSE ROOT CA, O = Shenzhen Stock Exchange
Digital Authentication	13	18	3,261,905	539,711 (10.65%)	28.37%	96.66%	CN = CFCA ACS CA, O = China Financial Certificate Authentication
Government Self-built	13	16	314,351	62,032 (1.22%)	30.46%	89.67%	O = National E-Government Network Administration Center
Fake Authentications	11	817,532	192,901,548	**2,798,985 (55.21%)**	**0.00%**	0.25%	CN = VeriSlgn Class 3 Public Primary Certification Authority - G4
Packet Filter	11	15,587	3,622,177	73,725 (1.45%)	13.57%	14.39%	CN = NetFilterSDK 2
Proxy/VPN	10	90,131	3,050,138	1,029,648 (20.31%)	2.26%	4.27%	CN = koolproxy.com, O = KoolProxy inc
Security Software	2	7,187	509,645	4,719 (0.09%)	0.01%	0.32%	O = Beijing SkyGuard Network Technology Co., Ltd
Parent Control	1	7,554	70,8129	7,787 (0.15%)	0.00%	0.57%	CN = UniAccessAgentFW 2
Unknown	15	207,957	14,048,377	289,198 (5.07%)	2.89%	4.73%	CN = VRV NDF RootCA 2

- Three major sources of hidden Cas are identified: **Self-built CAs** (50 groups), **Fake root CAs** (11 groups) and **Local software CAs** (24 groups).

(15 of the top 100 groups could not be identified, marked as “Unknown”)

Self-built Certificate Authorities

Categories	# roots	# conns	# clients	%error	Example
Enterprise	48	2,071,344	199,743	75.66%	CN=SZSE ROOT CA, O=Shenzhen Stock Exchange
Digital Authentication	18	3,261,905	539,711	96.66%	CN=CFCA ACS CA, O=China Finnacial Certificate Authentication
Government	16	314,351	62,031	89.67%	O=National E-Government Network Administration Center

- **Widely used**: organizations as government agencies and enterprises still use self-built CAs to issue website certificates (for 3,311 FQDNs we observed).
- **Incorrectly implemented**: over 75% chains linked to self-built CAs received verification errors. The most common error is `WEAK_SIGNATURE_ALGORITHM` (78.3%).
- **Publicly deployed**: active scanning experiments find 2,439 FQDNs (73.6%) are publicly accessible websites that still linked to self-built CAs.

Fake Certificate Authorities

- **Fake root CAs**: **impersonate trusted CAs** with deceptive subject names. None of their public keys are located in the official lists of public trusted CAs.

Examples of fake Certificate Authorities

Subject Common Name	# hidden roots	# connections	# FQDNs
Certum Trusted NetWork CA 2	254,414	158.54M	1,137,121
VeriSIgn Class 3 Public Primary Certificate Authority - G4	2	21.20M	210
GlobalSign Root CA	1,419	7.61M	6,023
GlobalSignature Certificates CA 2	1	2.85M	74,555

Wide impact range:

11 groups, 817K hidden roots, 192M HTTPS connections

High trusted rate:

0.0001% connnections receive `AUTHORITY_INVALID` error.

- **Uncover behind owners**: match certificates in:

Threat Intelligences

Sandbox Logs of Malware

Find 44 fake roots associated with:
Trojan, CoinMiner, Adware…

Local Software

- **Local software**: imported at software installation, for (benign) purposes as virus detection, download acceleration and ad blocking.

Categories	# groups	# roots	# conns	# clients	% error	Example
Packet Filter	11	15,587	3,622,177	73,725	14.39%	CN=NetFilterSDK 2
Proxy/VPN	10	90,131	3,050,138	1,029,648	4.27%	CN=koolproxy.com, O=KoolProxy inc
Security Software	2	7,187	509,645	4,719	0.32%	O=Beijing SkyGuard Network Technology Co., Ltd
Parent Control	1	7,554	708,129	7,787	0.57%	CN=UniAccessAgentFW 2

- **Traffic owned by local software**: shows a **much lower percentage** (3.58% in this study) than previous works (commonly over 50% for HTTPS interceptions).

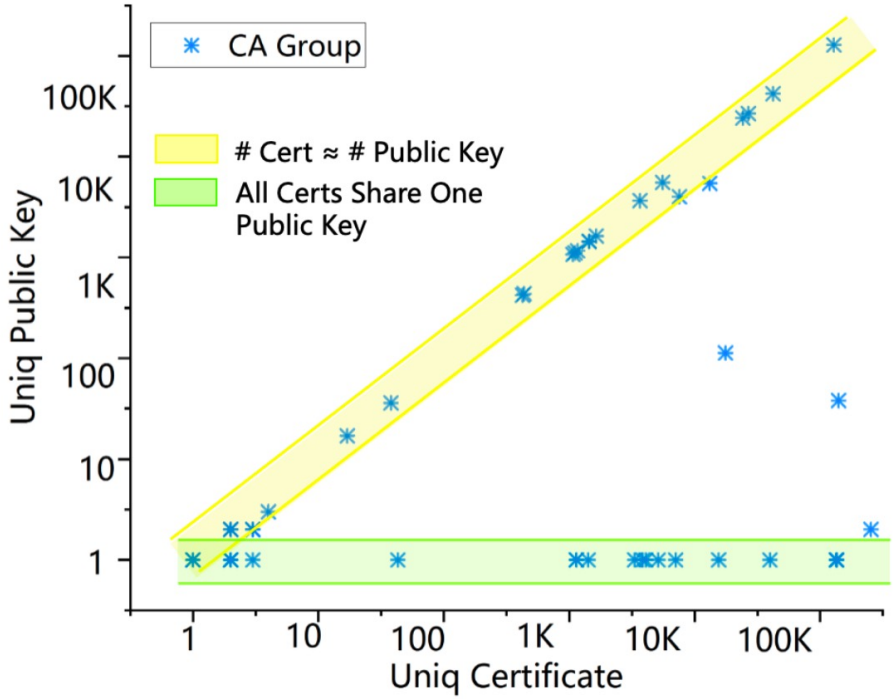
  From our (client-side) perspective, "user-informed" or "regulated" interceptions of local sofware may not predominate as commonly expected.

Implementation Flaws

Irregular Key Usage and Management

- **Direct signing with root certificates**: Over 97% chains are signed directly by root certificates. 41.4% of self-built CAs issued chains are directly signed.

- **Public key sharing**: Prevalent in the hidden CA ecosystem, with 144 groups (22.4%) suffering from this threat (including two security software).



Count of hidden roots and public keys in each group

Chaotic and Improper Validity Period Settings

- **Baseline of validity**: 6 months to 16 years according to security standards. Less than 40 years for all roots in public programs with an median value of 20 years.

Unreasonable settings:

Created in the year 1899

Experated in the year 9999

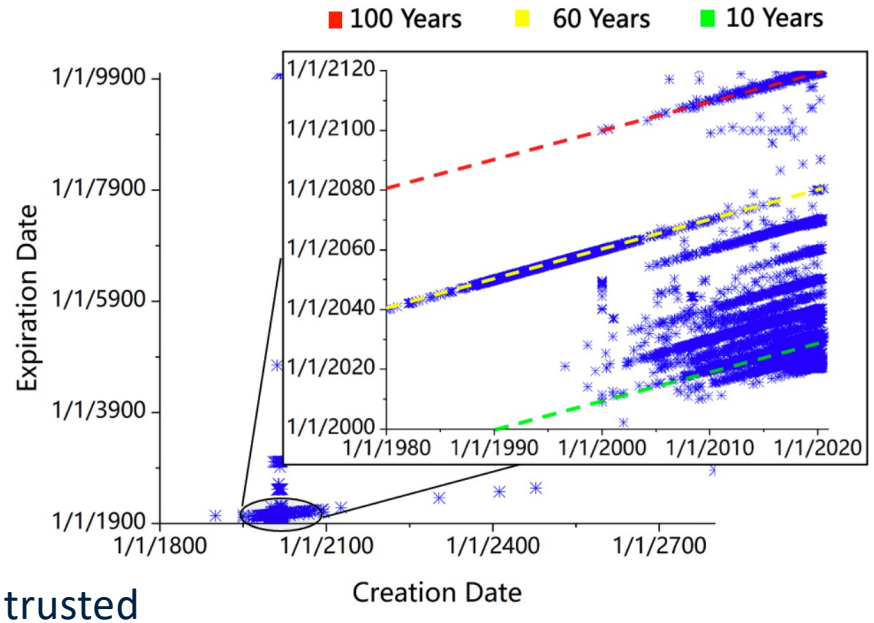
Long validity periods:

79% are valid for over 60 years

317 are valid for over 100 years

Incomplete revocation:

23 roots that should be revoked are still active, trusted by over 34K clients and affecting 264K HTTPS traffic.



Creation dates and expiration dates of hidden roots

Non-compliant Certificate Content

- **Identify content in-compliance** by **Zlint**, a certificate linter with 266 lints specified by X.509 standards and CA/Browser Forum.

Zlint **ERROR** messages of hiddend root certificates

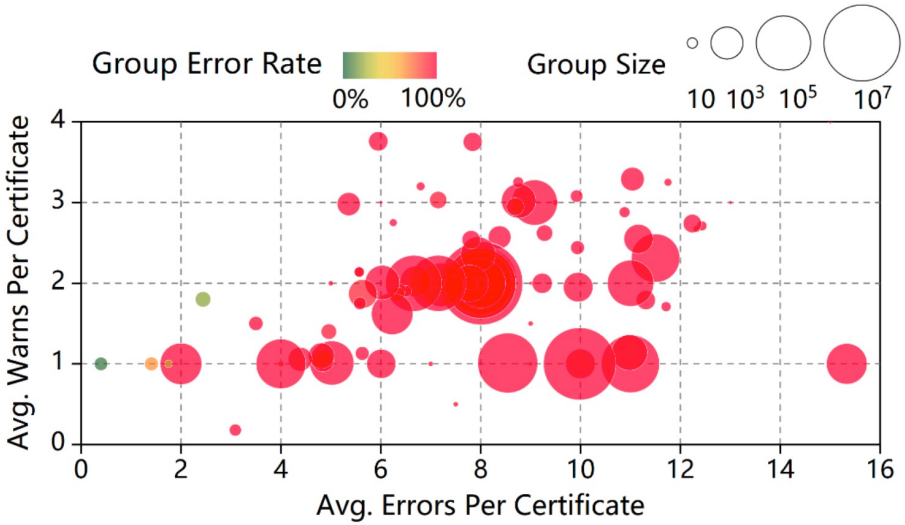
Error Type	# Lints	# Cert Errors	Example
Missing extensions	15	954,453 (79.46%)	Missing key identifier
Invalid values	49	121,745 (10.14%)	Negative serial number
Missing fields	3	89,763 (7.47%)	Missing CA organization
Vulnerable values	6	35,228 (2.93%)	RSA key < 2048 bits

- **Problematic content of root certificates:**

  1,201,189 **ERROR** reported of 73 lints, over 87% of roots violating at least one basic requirement.
  Most roots miss critical extensions such as key identifiers (could lead verification errors).
  Vulnerable settings are also prevalent, e.g., weak public keys.

Non-compliant Certificate Content

- The implementation of **leaf certificate** signed by hidden CAs is even **more worrisome**.



Zlint **ERROR** ratio of leaf certificates in top 100 groups

Problematic content of leaf certificates:

8.14 **ERROR** and 1.93 **WARNING** messages for each leaf certificate on average.

Vulnerable settings are more common than root certificates (e.g., 1024-bit RSA keys).

85 of the top 100 hidden root groups sign every leaf certificate with implementation flaws.

The performance of hidden CAs on compliance is significantly worrisome than public trusted CAs (only 0.02% ERROR rate, 2018[1]) .

[1] Kumar, Deepak, et al. "Tracking certificate misissuance in the wild." *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018.

Conclusions

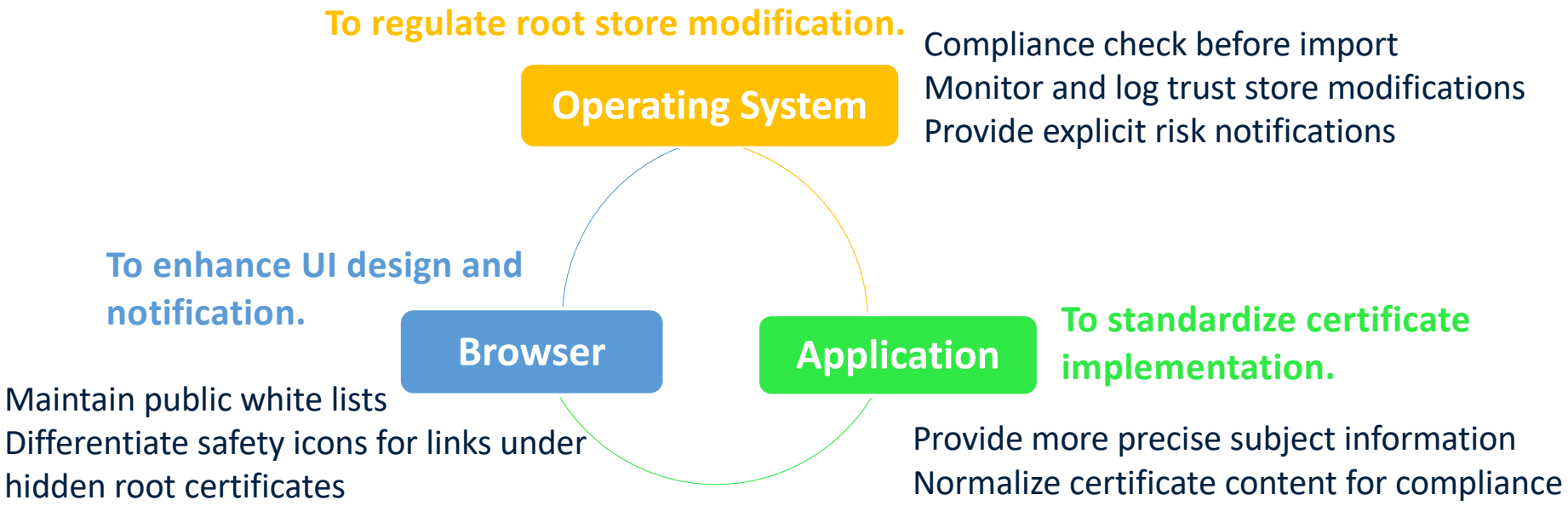
- The ***Hidden Root Ecosystem*** is **large-scale, highly active and widely incluential**, which plays an important role in Web PKI, despite being less noticed before.

- Through comprehensive measurement studies, we identified its scales, usages, and distribution resources. And **demonstrated numerouse implementation flaws**.

- The community should immediately **review the security of the local root store**, and seek a best security practice to regulate and manage hidden root CAs.

Recommendations

It’s impractical to rigorously check and block all roots outside the public trust list. However, at least **several parties** could cooperate to **mitigate the potential risks**:



Rusted Anchors: A National Client-Side View of hidden Root CAs in the Web PKI Ecosystem

Yiming Zhang

Tsinghua University

zhangyim17@mails.tsinghua.edu.cn