Analysis of the HTTPS Certificate Ecosystem

Zakir Durumeric, James Kasten,
Michael Bailey, J. Alex Halderman

University of Michigan

HTTPS and TLS

How does HTTPS and the CA ecosystem fit into our daily lives?

Nearly all secure web communication relies on HTTPS
- online banking, e-mail, e-commerce transactions, etc.

HTTPS provides confidentiality, integrity, and authentication

HTTPS is dependent on a supporting PKI – thousands of certificate authorities we rely on to vouch for sites' identities

The supporting PKI is opaque – we blindly rely on these CAs

There has been much previous work including including the EFF's SSL Observatory, Holz et al. at IMC, and Akhawe et al. at WWW

Talk Outline

1. **HTTPS Background**
2. Data Collection Methodology
3. Identifying Trusted Authorities
4. Worrisome Trends and Observations
5. How can we make the HTTPS ecosystem more secure?

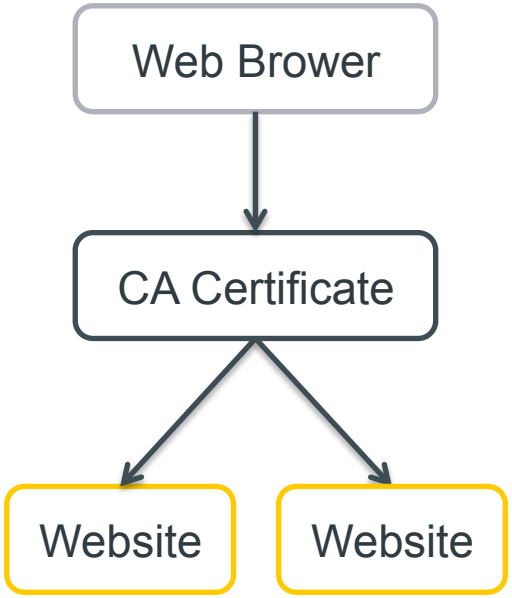
Certificate Authorities

How do certificates provide authentication?

Web browsers trust certificate authorities to investigate and vouch for the identities of trusted websites

CAs vouch for a website's identity by signing digital certificates with a browser trusted certificate and key

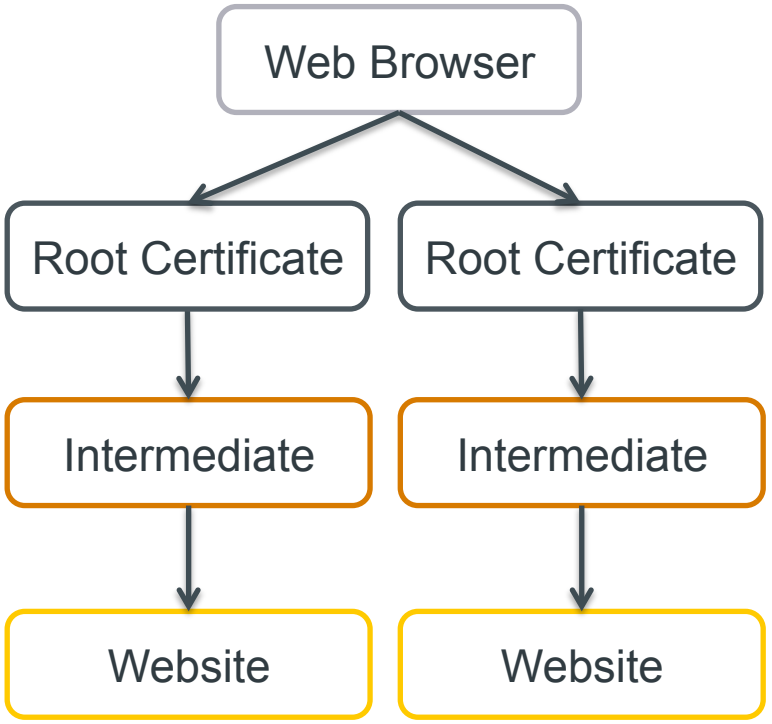
Web browsers store a list of these trusted authorities' certificates known as roots



Intermediate Authorities

How do intermediate authorities fit into the picture?

Root authorities delegate the ability to sign certificates to intermediate authorities

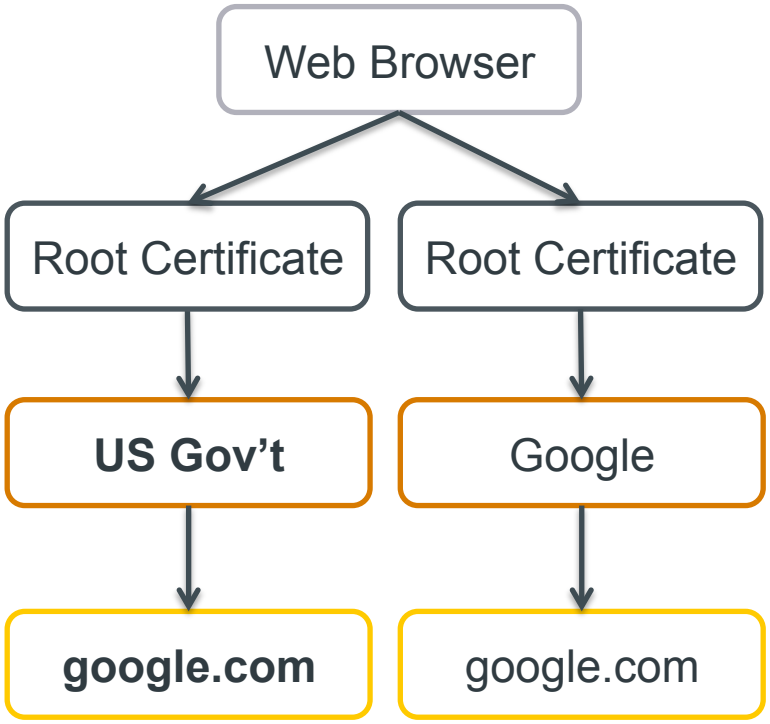


Intermediate Authorities

How do intermediate authorities fit into the picture?

Root authorities delegate the ability to sign certificates to intermediate authorities

In all but a handful of cases, intermediates can sign for certificates for *any* domain



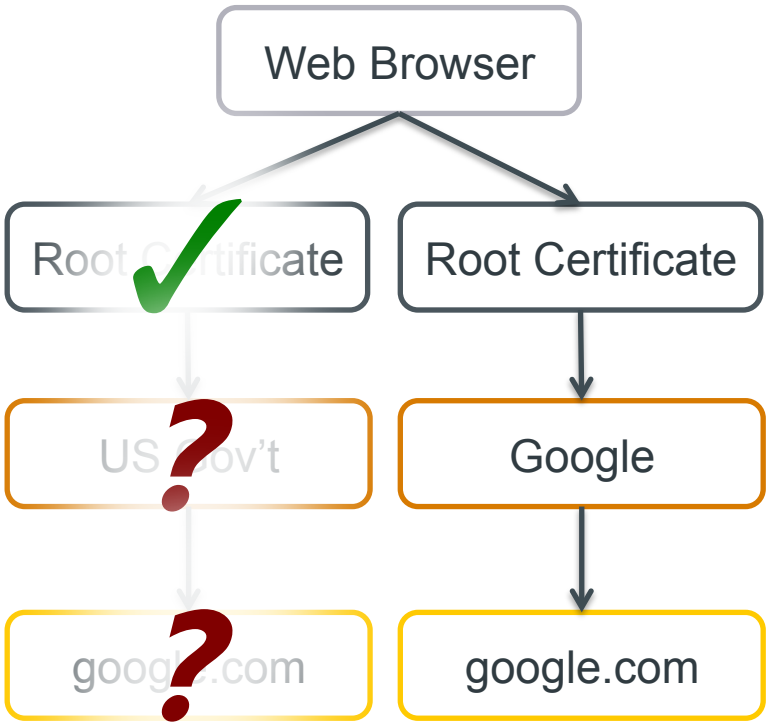
Intermediate Authorities

How do intermediate authorities fit into the picture?

Root authorities delegate the ability to sign certificates to intermediate authorities

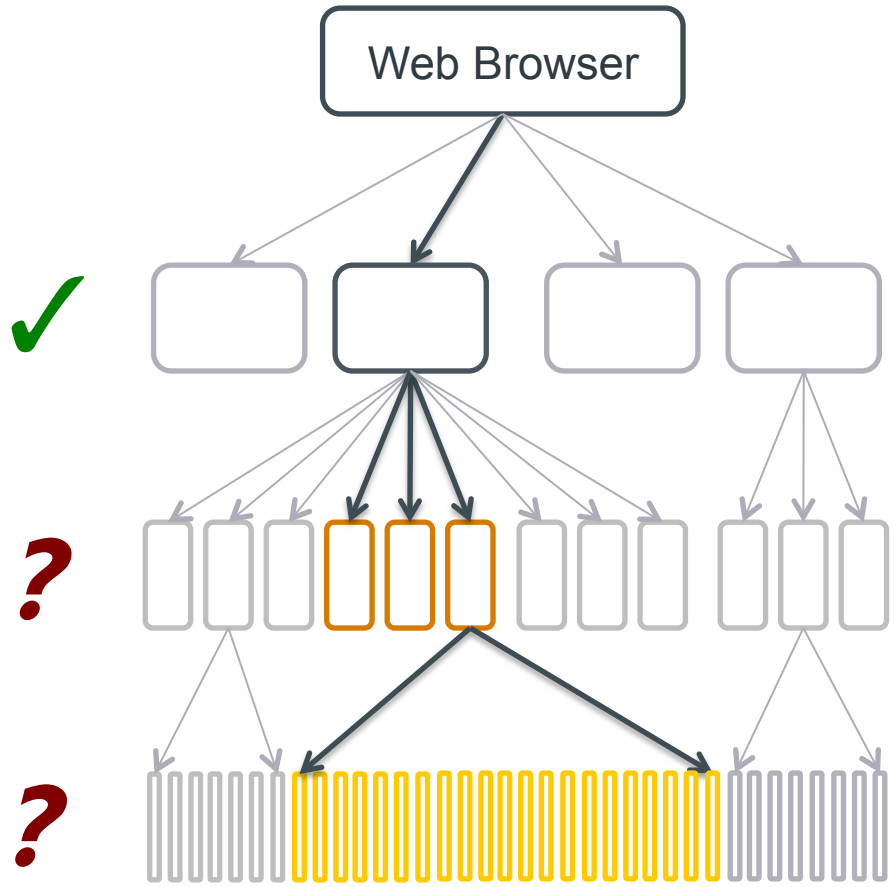
In all but a handful of cases, intermediates can sign for certificates for *any* domain

Non-roots aren't publicly known until they are found in the wild



Intermediate Authorities

How do intermediate authorities fit into the picture?



Hundreds of roots

Thousands of intermediates

Millions of leaf certificates

Talk Outline

1. HTTPS Background
2. **Data Collection Methodology**
3. Identifying Trusted Authorities
4. Worrisome Trends and Observations
5. How can we make the HTTPS ecosystem more secure?

Dataset and Methodology

How do we measure the certificate authority ecosystem?

We performed 110 scans of the IPv4 address space over an 18 month period using ZMap, OpenSSL, and libevent

Completed 1.8 billion TLS handshakes and collected certificates

We collected:

- 42 million unique certificates
- 6.9 million browser trusted

from 109 million hosts

Dataset available at https://scans.io and code at https://zmap.io

Responsible Data Collection

How do we reduce the impact of active scanning?

Reducing Scan Impact

Scan in random order and at a reduced scan rate

Signal benign nature over HTTP, DNS, and WHOIS

Honor all requests to be excluded from future scans

Excluded Networks

Correspondence with 145 individuals and organizations

Excluded 91 networks (.11% of the address space)

2 requests from ISPs account for 50% of addresses

Talk Outline

1. HTTPS Background
2. Data Collection Methodology
3. **Identifying Trusted Authorities**
4. Worrisome Trends and Observations
5. How can we make the HTTPS ecosystem more secure?

Identifying Trusted Authorities

Who do we trust to sign a certificate for any website?

Identified 1,832 CA certificates belonging to 683 organizations

80% of the organizations were not commercial CAs

Organizations included religious institutions, libraries, cities, corporations, and non-profits

CA certificates were owned by organizations in 57 countries

CAs by Organization Type	
Academic Institutions	40%
Commercial Authorities	20%
Governments	12%
Corporations	12%
Other Types	16%

CAs by Owning Country	
United States	30%
Germany	21%
France	4%
Japan	3%
Other Countries	42%

The Path to Power

How are organizations obtaining CA certificates?

311 (45%) of the organizations were provided certificates by German National Research and Education Network (DFN)

A large number of root CAs have provided CA certificates to unrelated third-party organizations and governments

Largest were *GTE CyberTrust Solutions* (Verizon) and Comodo

- Provided unrestricted CA certificates to 62 organizations

Financial institutions (e.g. Visa) and some countries used CA certificates included in each brower root store

Distribution of Trust

Who do we trust on a day-to-day basis?

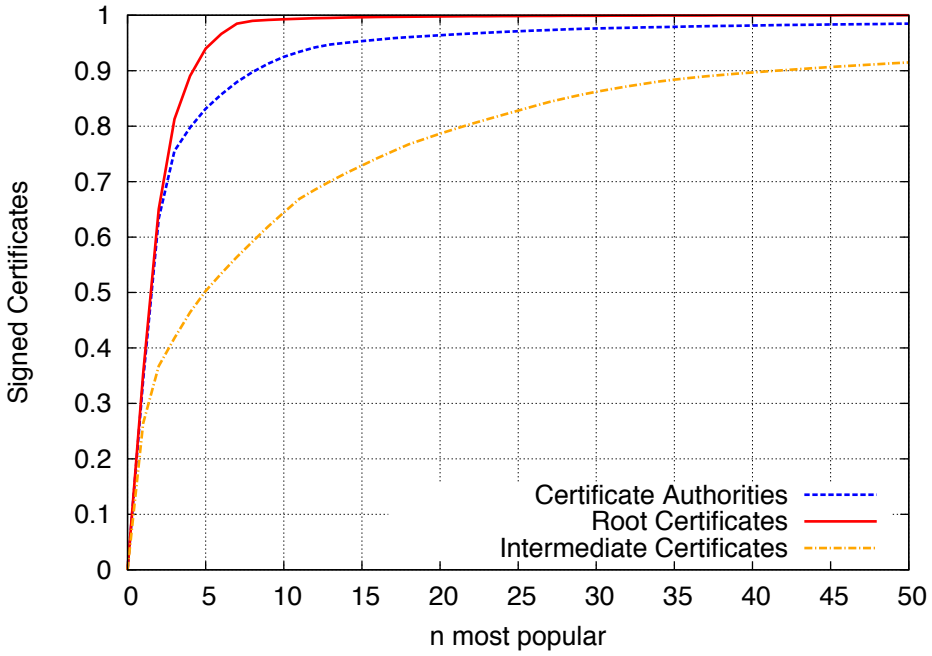
Large companies have acquired smaller CAs

75% are signed by Comodo, Symantec, and GoDaddy

90% are descendants of 4 roots

90% are signed by 40 intermediates

26% are signed by a single intermediate certificate



Talk Outline

1. HTTPS Background
2. Data Collection Methodology
3. Identifying Trusted Authorities
4. **Worrisome Trends and Observations**
5. How can we make the HTTPS ecosystem more secure?

Misaligned Objectives

CAs are providing services that harm the HTTPS ecosystem

Only 7 of the CA certificates we found had a name constraints

- All other organizations can sign for ***any*** domain

Only 40% of CA certificates had a length constraint

- All other organizations can create new CA certificates

Almost 5% of certificates are trusted for a local domain

- e.g. *mail*, *exchange*, and *intranet*
- provide no actual protection against attackers

Community Shortsightedness

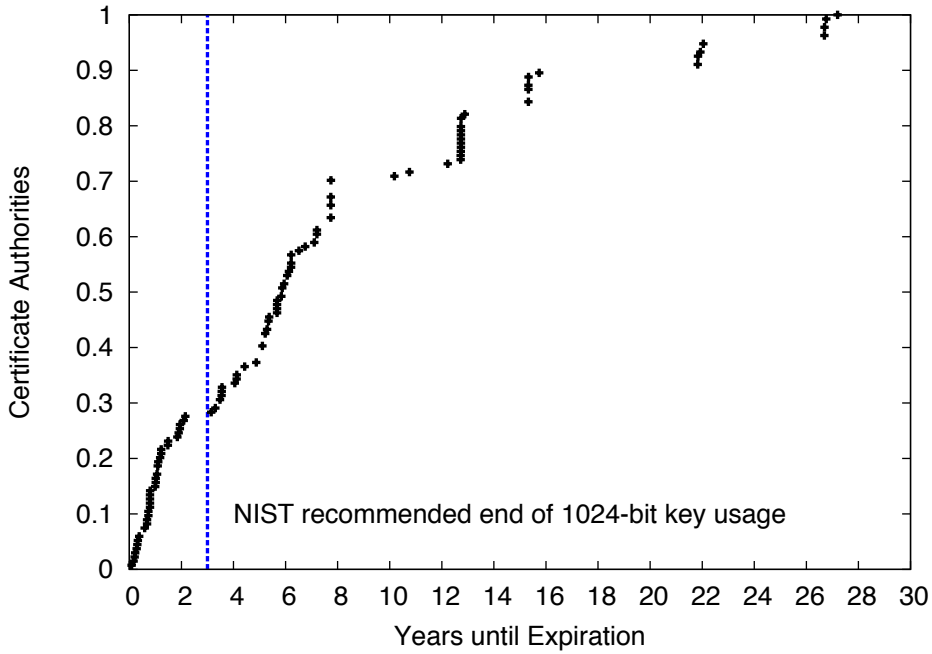
We are not considering the long-term consequences today

CA certificates are being issued for 40+ years in the future

49% of certificates have a 1024-bit key in their trust chain

In 2012, 1.4 million signed new certificates were signed using a 1024-bit root authority

15 organizations provide no avenues for revocation



70% of trusted CA using a 1024-bit key certificates expire after 2016

Client Deployment Difficulties

It remains difficult for end-users to correctly deploy HTTPS

13% of hosts serving once trusted certificates are misconfigured

20% of hosts remove expired certificates after expiration or revocation

47 of the signing certificates were not for web traffic due to end users using code signing keys, etc.

CAs by Owning Country	
Expired	6%
Not Yet Valid	.02%
Revoked	.3%
Incorrect Intermediates	7%
Unnecessary Root	42%
Optimal Configuration	45%

Impact of the Lack of Oversight

What is the real world impact of these observations?

We are making errors on a day-to-day basis

There's has been an impact to ignoring our community's guidelines such as *least privilege* and *defense in depth*

Case 1: A mis-issued CA certificate issued by Turktrust to a transit authority that was revoked after signing for *.google.com

Case 2: South Korea misissued 1,400 CA certificates that were prevented from causing harm by a root length constraint

Talk Outline

1. HTTPS Background
2. Data Collection Methodology
3. Identifying Trusted Authorities
4. Worrisome Trends and Observations
5. **How can we make the HTTPS ecosystem more secure?**

Moving Forward

How do we make the HTTPS Ecosystem more secure?

A lack of oversight has led to an unmaintainable ecosystem

- We need to consider ideas such as *Certificate Transparency*

Groups such as the CA/B Forum are on the right track, but are toothless and even their members are ignoring their policies

We need web browsers to coordinate and demand change

A lot of recent work on bugs in TLS implementations, but also important to consider how we help users to move to HTTTPS

Conclusion

Analyzed HTTPS Ecosystem by performing regular comprehensive scans of the IPv4 address space

Identified CA certificates, the organizations that control a CA certificate, and how they gained these rights

Explored several of the most worrisome trends that we observed over the past year of scanning

Discussed potential avenues for the security community to improve the HTTPS ecosystem moving forward

Questions?

Zakir Durumeric
University of Michigan

zakir@umich.edu

https://httpsecosystem.org