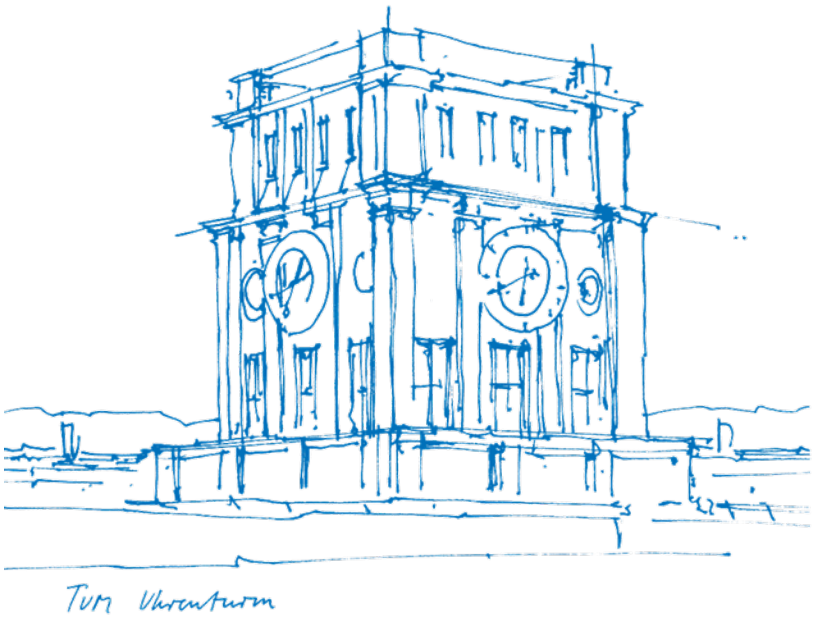
Chair of Network Architectures and Services
TUM Department of Informatics
Technical University of Munich (TUM)

Push Away Your Privacy: Precise User Tracking Based on TLS Client Certificate Authentication

Matthias Wachs, Quirin Scheitle, and Georg Carle

TMA'17, Dublin, June 21, 2017



TLS 1.2 handshake does not encrypt certificates

Known for a long time. . .

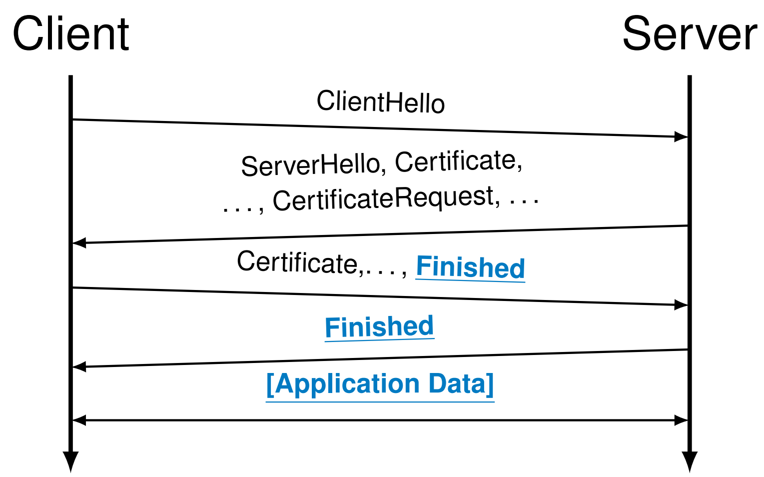


Figure: TLS 1.2 handshake, Unencrypted Data, **[Encrypted Data]**

Server Certificates

- Eavesdroppers can learn the specific websites that a user visits (not just the server's IP address)

Client Certificates

- Used by VPNs, governments, . . .
- Person names, company names, . . . → private data!

TLS 1.2 Client Certificate Authentication (CCA)

Where is CCA used?

- **Network authentication**: 802.1x EAP
- **VPN**: OpenVPN, F5 EdgeConnect, ...
- **Web**: HTTPS
- **IoT**: MQTT
- **Remote device management,** for example MobileIron
- **Apple Push Notification Service (APNs)**

Apple Statistics:

- 1 billion active devices (2016)
- 800 million iTunes accounts (2014)

Push Notification Services

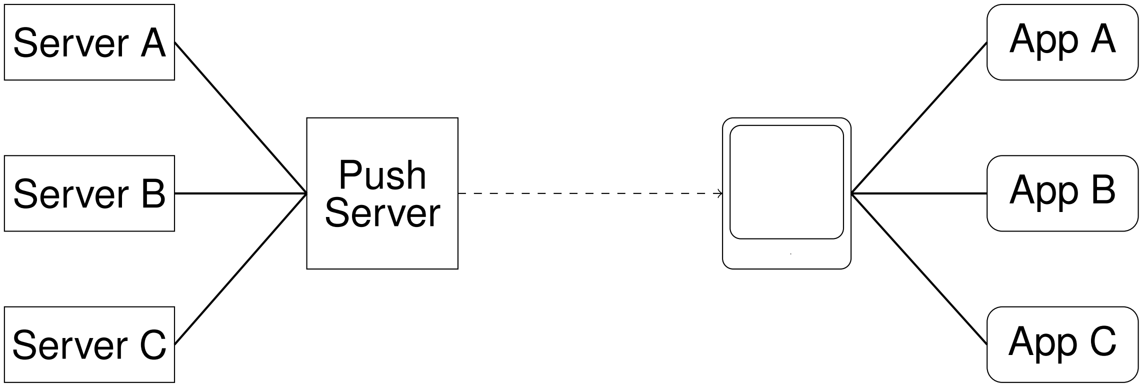


Figure: Push Service Architecture: Messages brokered to Apps through the Push Notification Service.

Resource efficient notification of (mobile) applications:

- **Apple's APNs:** OS, MacOs, iTunes
- **Google's FCM:** Android, Chrome
- **Microsoft's WNS:** Windows, Windows Phone

Paradigms:

- Tightly integrated with operating system
- Always connected to backend

Apple Push Notification Service (APNs)

Maybe the biggest user of TLS CCA?

APNs integral part of iOS and macOS – “always on”

APNs uses Client Certificates for login:

- Generated at device setup
- Unique cryptographic material (CN, public key, fingerprint)
- CN different for mobile and desktop devices

```
Serial Number: ab:12:34:56:78:9a:bc:de:f0:12
Issuer: C=US, O=Apple Inc., OU=Apple iPhone, CN=Apple iPhone Device CA
Validity Not Before: Apr  8 12:34:56 2015 GMT
Validity Not After : Apr  8 12:34:56 2016 GMT
Subject: CN=12345678-1234-1234-1234-123456789ABC
Key ...
(all data redacted)
```

Precise User[1] Tracking in APNs

Several appearances of same device easily linkable

2 of 4 Attacker Types Considered in this Work

- ~~Apple or someone infiltrating Apple: better means available~~
- ~~Local adversary: Can use MAC addresses and more~~
- Regional adversary: Access to one or several large networks
- Global adversary: Access to several core networks

Regional Adversary – Validation at Internet Uplink

- Can a regional adversary track users?

Global Adversary – Validation through Global Path Measurements

- How well can a global adversary leverage APNs to track users?

1: APNs CCA certificates are bound to devices. However, these devices are typically private and carried by a user at most times, which allows inferences into user tracking.

Passive Capturing

Methodology

Analysis of $>$ 2 weeks of TLS CCA traffic at Internet uplink:

- APNs TCP ports (443, 5223, 2195, 2196)
- `pcap` Filter on certificate handshake

Stored information:

- Timestamp
- Connection 5-tupel (Source & Destination IP address, Port, Protocol TCP)
- Certificates & TLS Extensions

Working with Human Subjects

Ethical Considerations

Strict regulations by IRB:

- Documented measurement process
- Isolated measurement infrastructure
- Access only for permitted staff
- Raw data must not leave infrastructure

Our self-restrictions:

- No attempt to identify users
- No publication of identifiable data

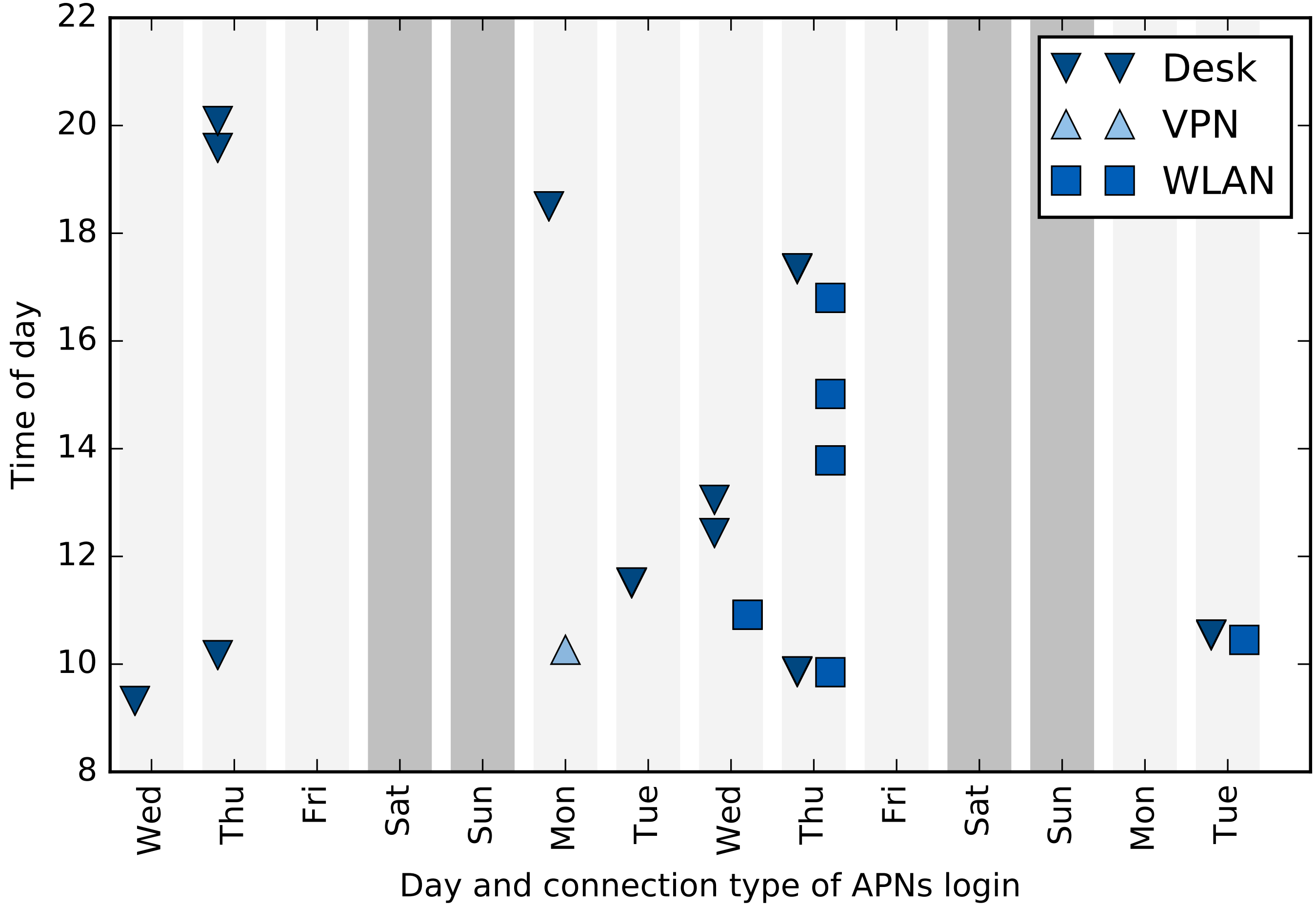
APNs by far the biggest user of CCA

#Certs	Issuer Distinguished Name
56128	/C=US/O=Apple Inc./OU=Apple iPhone/CN=Apple iPhone Device CA
334	/CN=Layer Client CA/C=US/L=San Francisco/O=Layer, Inc/ST=CA
221	/CN=AnyDesk Client
76	/C=KR/ST=Kyunggido/L=Suwon/O=Samsung Electronics (*redacted*)
52	/CN=Ricoh Remote Service (*redacted*)

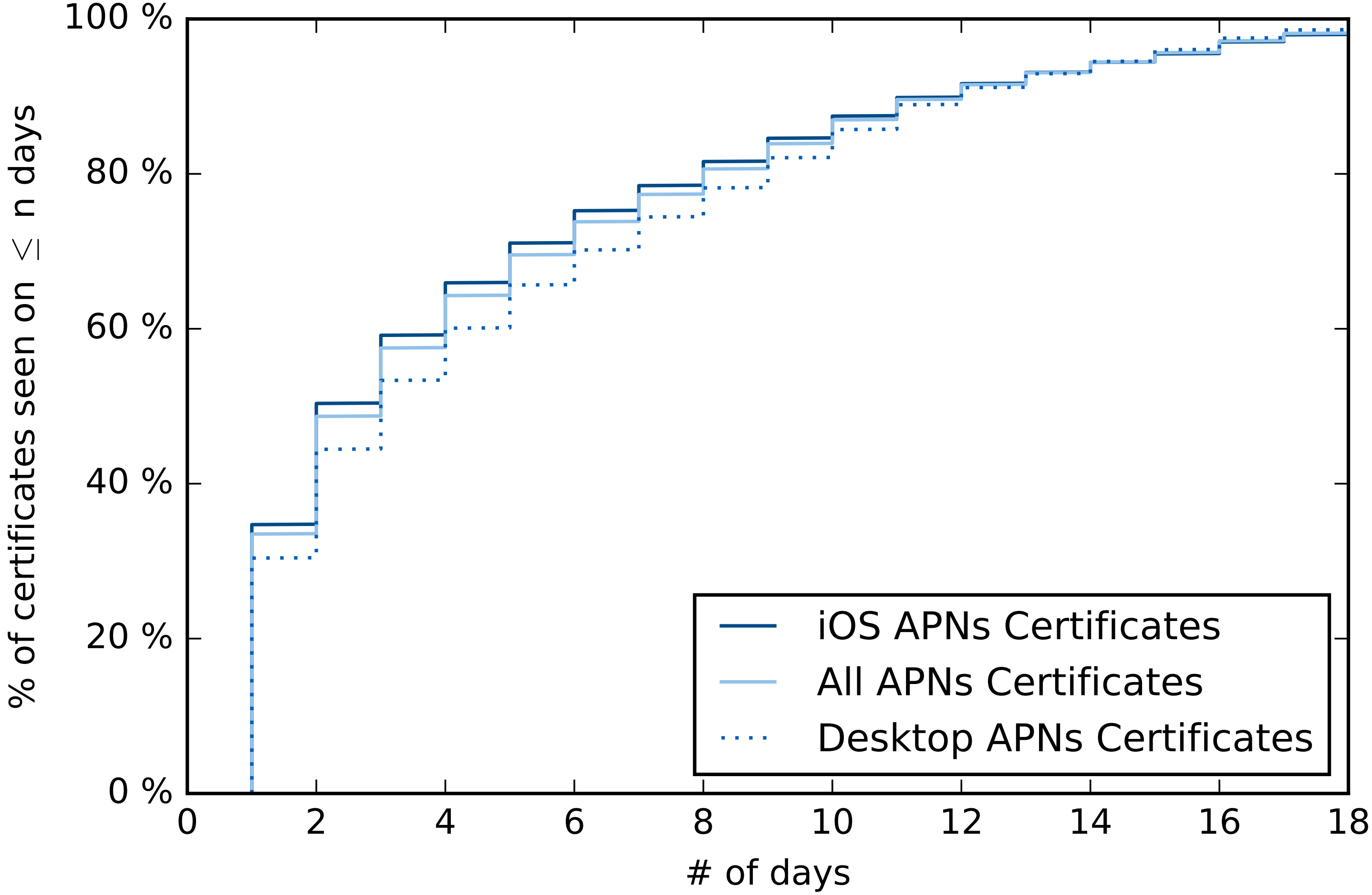
Case Study - how well can we track a single user?

Informed Consent

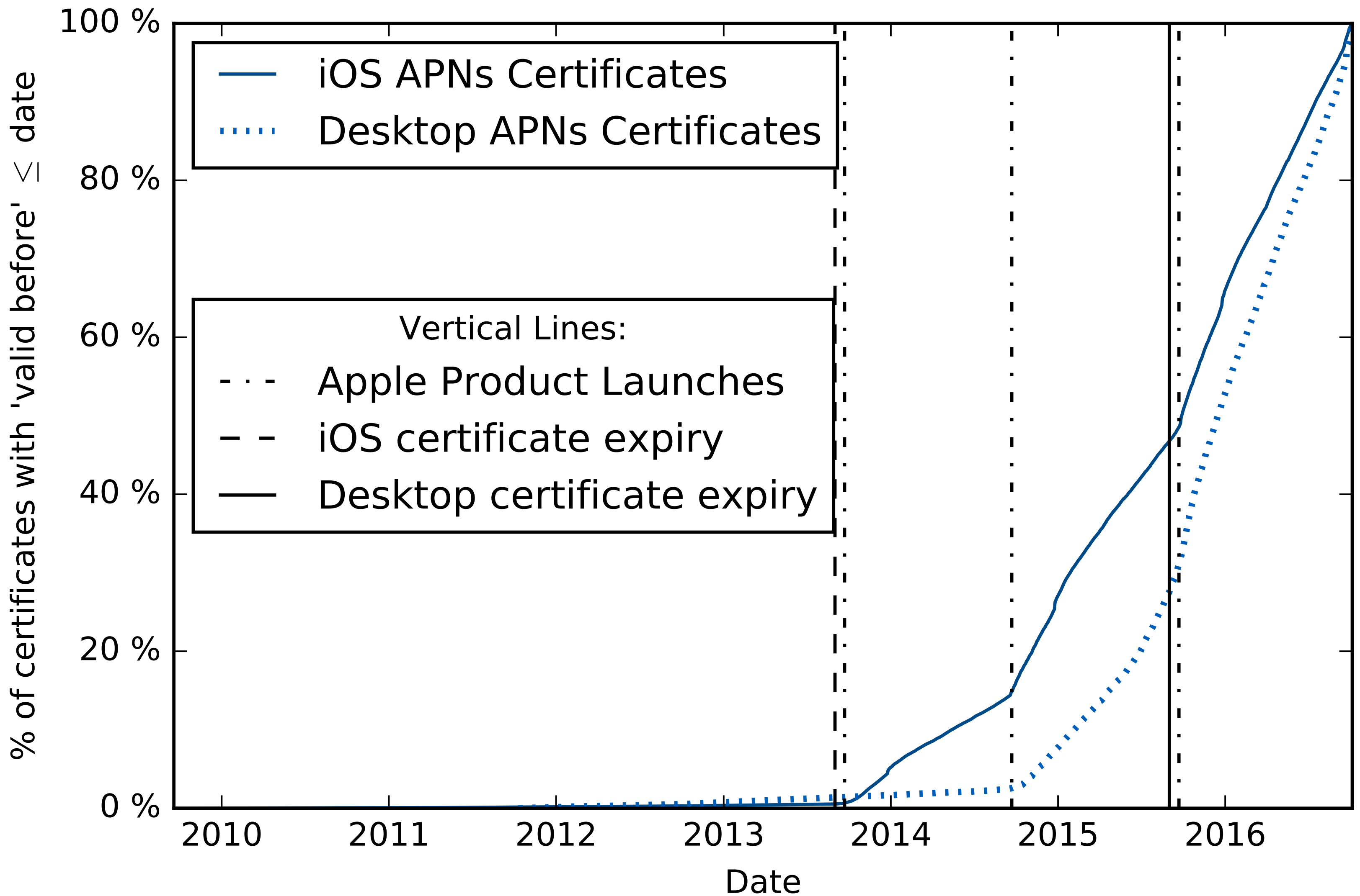
Note: We are tracking a device. As mobile devices are typically closely carried, they allow conclusions about users.



What % of certificates is traceable?



Can we derive device types from certificate data?



Is global tracking feasible?

Methodology

Research Question: How many networks does an attacker have to eavesdrop on to observe a significant share of APNs logins?

- We identify APNs backend infrastructure and conduct distributed traceroute measurements towards it
  - Measurements confirm that clients resolve one of *[1-50]-courier.push.apple.com*
  - We globally resolve *[1-50]-courier.push.apple.com* using 1000 RIPE Atlas probes each
  - We find 69 /24 subnets and pick one random observed IP address in each of the 69 subnets
  - Using 1000 RIPE Atlas probes per measurement, we conduct traceroute measurements towards all 69 IP addresses
- We map transit router's IP addresses to ISPs and IXPs
- We count what % of routes traverses a certain ISP or IXP

Is global tracking feasible?

Eavesdropping capabilities on just 10 networks allows to follow APNs messages of over 80% of users globally or nationally

Rank	Global		Germany	
	IXP/AS	Σ% Paths	IXP/AS	Σ% Paths
1	AS3356 (L3)	25%	IXP DE-CIX	30%
2	AS1299 (Telia)	40%	AS3320 (DTAG)	52%
3	AS174 (Cogent)	54%	IXP E-CIX	61%
4	AS7922 (Comcast)	61%	AS6830 (Liberty)	69%
5	AS12322 (Free)	67%	AS31334 (VF/Kabel D)	75%
6	AS6830 (Liberty)	71%	AS1273 (C&W)	78%
7	AS4637 (Telstra)	75%	AS3356 (L3)	81%
8	AS6453 (Tata)	78%	AS34419 (VF Group)	84%
9	AS2828 (XO)	81%	AS680 (DFN)	86%
10	AS3320 (DTAG)	84%	AS6805 (Telefonica)	88%

Note: % is based on RIPE Atlas probe distribution as a proxy for APNs user distribution.

Responsible Disclosure

We informed Apple's product security team before publication:

- Contact with OpenPGP secured mail
- Very quick response
- Several phone calls, continuous contact
- Several engineers in calls and working on resolution

Impact:

- MacOS & iOS fixed with January 2017 security patches
- APNs Backend patched
- iTunes on Windows patched a bit later (SChannel is complicated . . .)

Discussion: The Value of Internet Measurements

It has been known and criticized for a while that TLS1.2 does not encrypt certificates, which may have specific adverse impact for client certificates. Anyhow . . .

- Discussions eroded . . .
- Draft RFCs expired . . .
- Apple decided to use CCA for APNs . . .

Lack of taking the issue seriously?

We believe that Internet measurements can overcome inertia in security improvements by . . .

- Quantifying impact and scale of a problem with hard evidence
- Benefitting issue prioritization
- Providing means to track patching progress

What now?

Push TLS 1.3 standardization which encrypts certificates

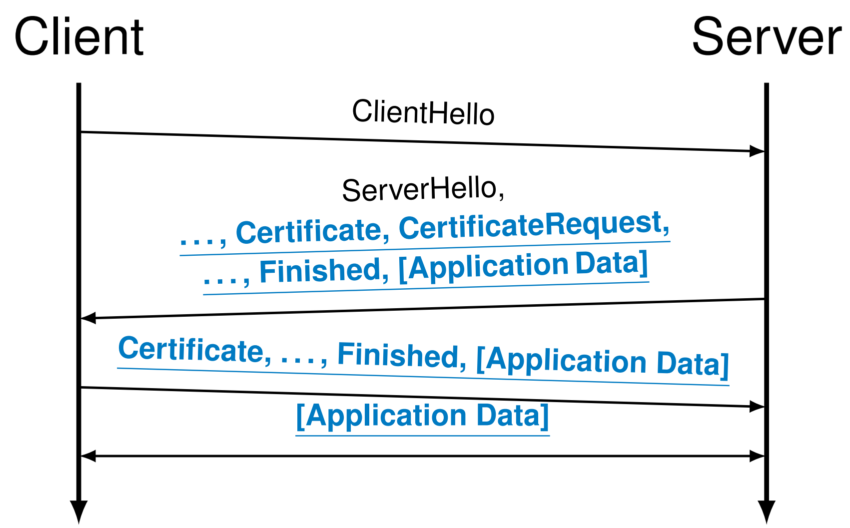


Figure: TLS 1.3 handshake, Unencrypted Data, **[Encrypted Data]**

But: ClientHello Extensions still unencrypted:

- Server Name Indication (SNI)
- Application-specific data

Reproducibility...

We aim for repeatability, replicability, and reproducibility[1]

Repeatability — same team, same experimental setup

Packing of “Reproducibility Bundle” along with camera-ready version requires detailed repetition of paper creation.

Replicability — different team, same experimental setup

We provide “artifacts” (scripts, data, documentation) so any other team can easily replicate our work.

Reproducibility — different team, different experimental setup

We provide a detailed documentation of our approach (*which pcap filter was set? what precise traceroute parameters were set?*) so other teams can reproduce our work without using our artifacts.

We ran an exercise[2] at the TMA PhD school that followed the research question and methodology of this paper. This resulted in a partial mix of replication, reproduction, and extension of our work.

1: Terms as defined by ACM: `http://www.acm.org/publications/policies/artifact-review-badging`
2: `https://github.com/quirins/tma17-ripeatlas-lab-participants/`

How to deal with Reproducibility and Private Data?

Much of the data in this work contains private and sensitive data

Passively Captured TLS handshakes and certificates

- No publication of raw data
- Cut open of analysis pipeline (for example, “not valid before” attribute of certificate)
- Anonymize output of database query with **documented** script
- Feed the anonymized data into analysis pipeline, published – figures in paper clickable:
  `https://github.com/tumi8/cca-privacy/blob/master/userstudy/userstudy.ipynb`

Active Measurement of APNs backend and traceroutes

- Public data – RIPE Atlas measurements per default public, for example
  `https://atlas.ripe.net/measurements/5719601/`
- Publish everything: Measurement scripts, RIPE Atlas IDs, raw data, analysis tools

Future Work

- Measuring uptake of APNs patch
- In-depth analysis of APNs backend infrastructure
- Controlling for AS population vs. RIPE Atlas probe count bias

Key Messages, Data, and Code

- TLS-CCA sends certificates unencrypted
- In an “always-on” mobile scenario, this can cause serious privacy issues
- We quantified this issue in the Apple Push Notification Service (APNs), Apple fixed promptly

Data and Code:

`https://github.com/tumi8/cca-privacy`

Matthias Wachs, Quirin Scheitle, and Georg Carle
Chair of Network Architectures and Services — `https://net.in.tum.de`
Technical University of Munich (TUM)