



Network Measurement Methods for Locating and Examining Censorship Devices

ACM CoNEXT 2022

Ram Sundara Raman*, Mona Wang*, Jakub Dalek, Jonathan Mayer, Roya Ensafi

6 December 2022



SNOOPING AT SCALE —

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



SIGN IN NPR SHOP

NEWS CULTURE MUSIC PODCASTS & SHOWS SEARCH

TECHNOLOGY



Russia is restricting social media. Here's what we know

The Washington Post
Democracy Dies in Darkness

MIDDLE EAST

Sanctions and censorship are making the Internet in Iran less accessible, analysts say



Large-scale censorship and surveillance events

Enabled by advanced network software and hardware

Netsweeper

- Citizen Lab Identified an “**Alternative Lifestyles**” blocklist curated by Netsweeper was used by several countries such as UAE to block LGBTQ content.
- After advocacy based on Citizen Lab’s findings, Netsweeper claims they have **removed the option** to block based on this category.

VICE

Canadian Internet Filtering Company Says It's Stopped 'Alternative Lifestyles' Censorship

The UAE was found to be blocking LGBTQ content using a pre-set category in Netsweeper's software. Amid pressure from rights groups, the company says it's disabled that category.

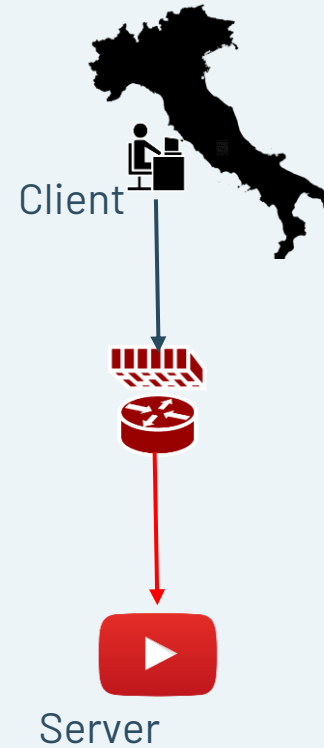
By [Jordan Pearson](#)

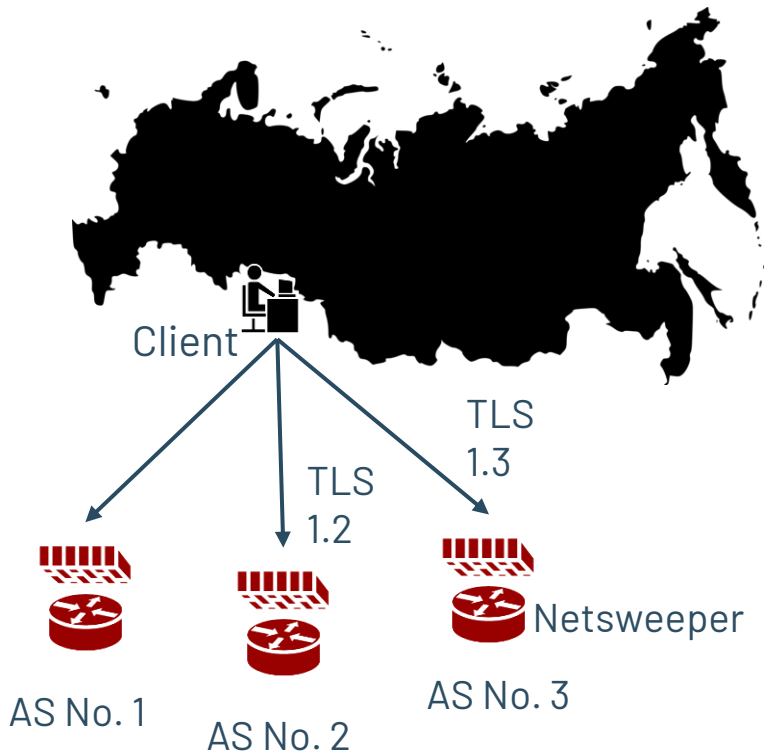
Jan 21 2019, 12:25pm [Share](#) [Tweet](#) [Snap](#)



What and When?

- Censorship Measurement Platforms





Who, Where and How?

- Specific censorship systems
 - Great Firewall of China
 - Iran's national firewall
 - Russia's TSPU system

Challenges and Gaps

1

Opaque nature of censorship

2

Lack of transparency

3

Variety of devices and censorship techniques

4

Reliance on specific behaviors

5

Large manual effort does not scale

Need: **General-purpose, robust methods**

To study censorship devices

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

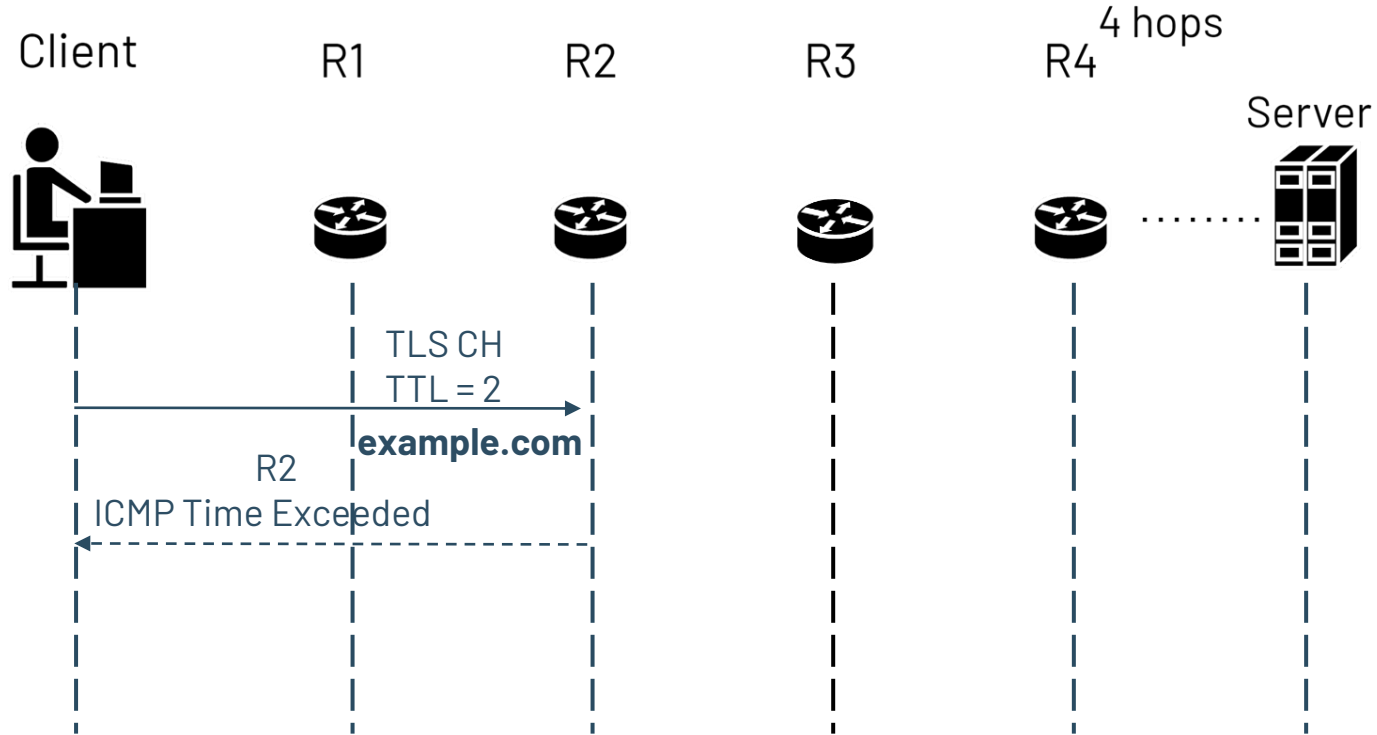
Banner grabs and Clustering

3

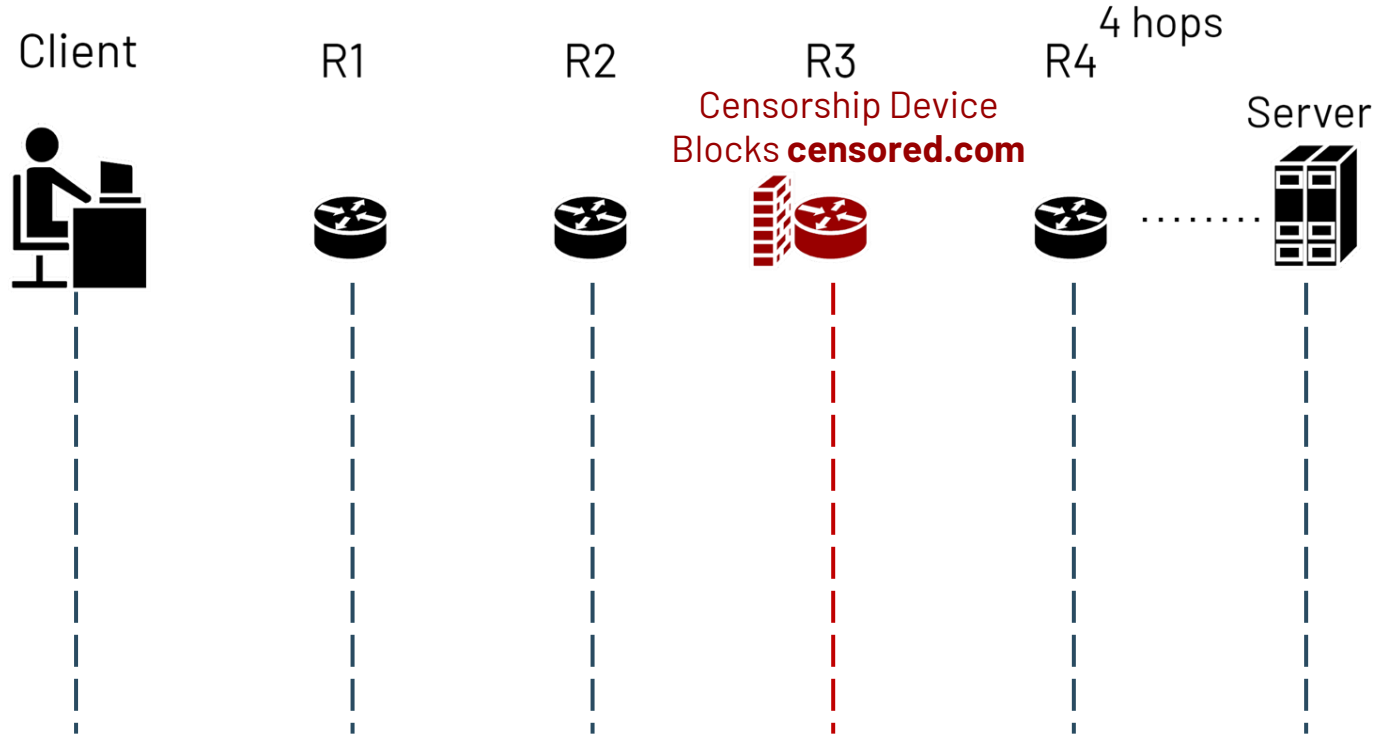
Reverse-engineer censorship triggers

Censorship Fuzzer

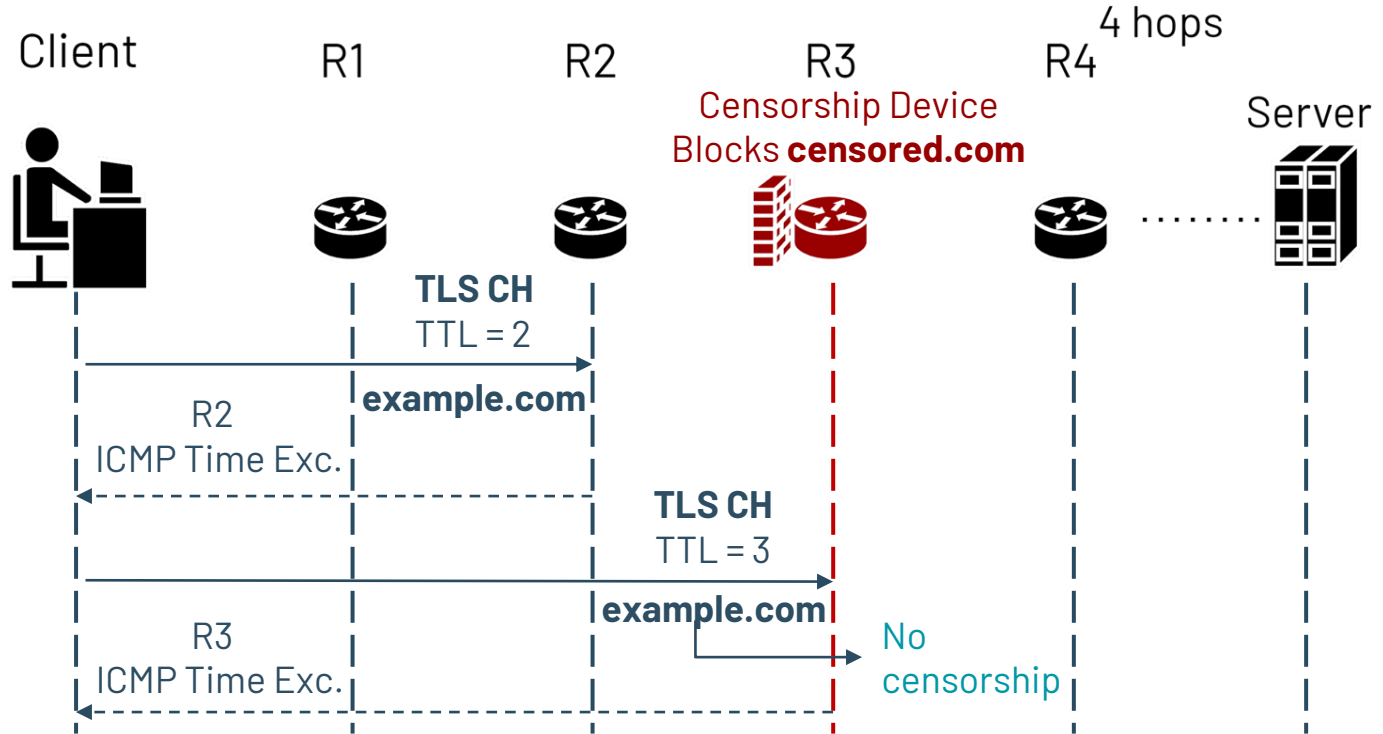
Application Traceroute



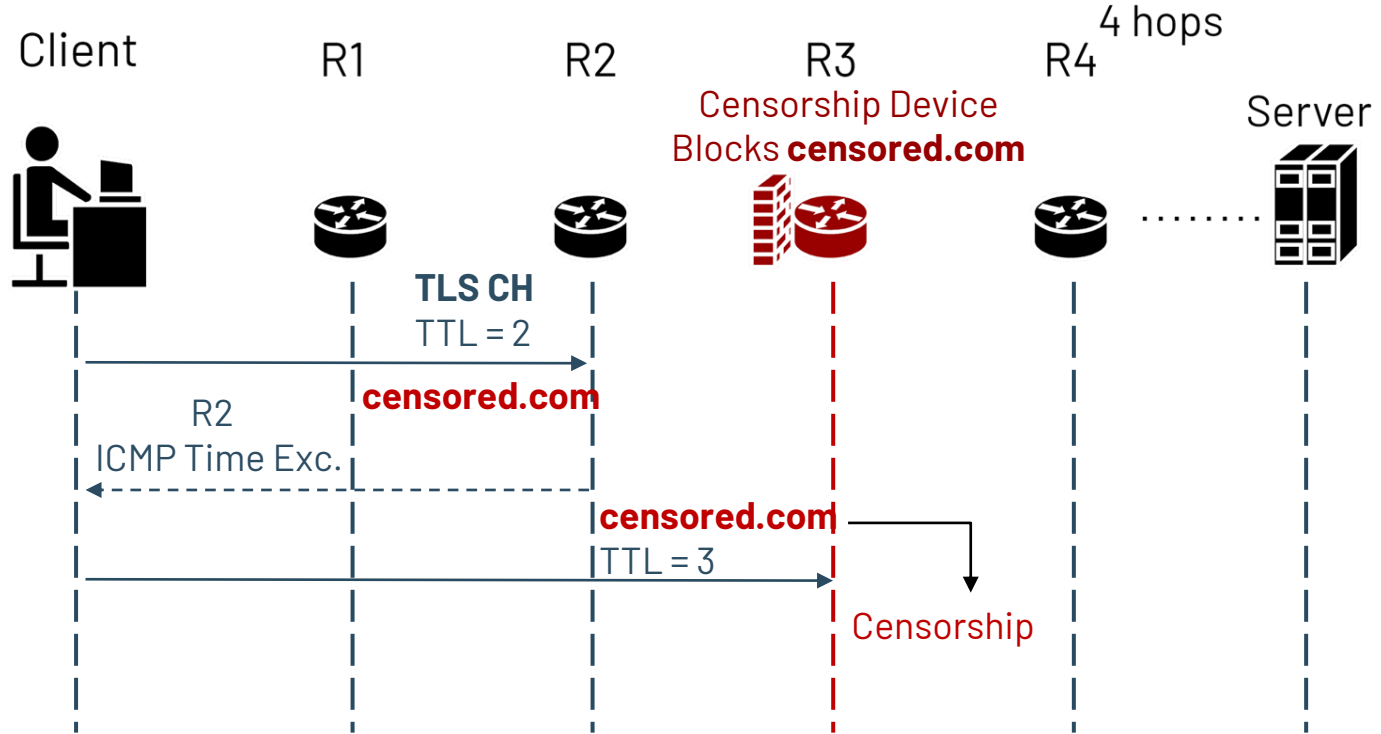
Censorship Traceroute



Censorship Traceroute



Censorship Traceroute



Variety in censorship mechanisms

1

**Censorship methods:
RST injection, packet drops**

2

**Device deployments:
In-path vs On-path**

3

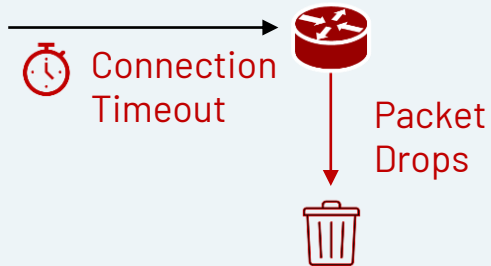
**Specialized censor behavior and
path variance**

Variety in censorship mechanisms

(1) Reset Injection



(2) Packet Drops



1

**Censorship methods:
RST injection, packet drops**

2

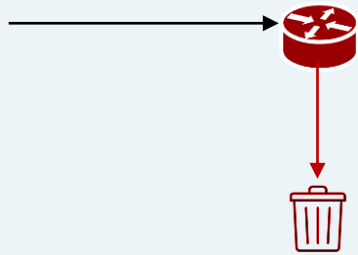
**Device deployments:
In-path vs On-path**

3

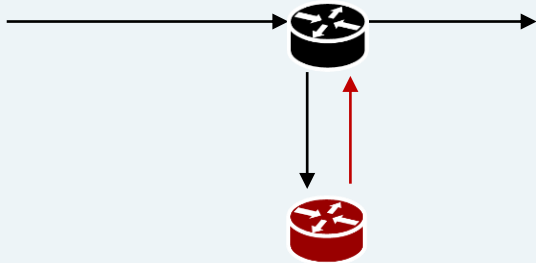
**Specialized censor behavior and
path variance**

Variety in censorship mechanisms

(1) In-Path Devices



(2) On-Path Devices



1

Censorship methods:
RST injection, packet drops

2

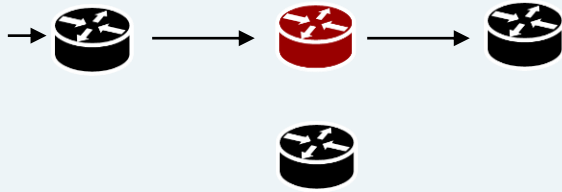
Device deployments:
In-path vs On-path

3

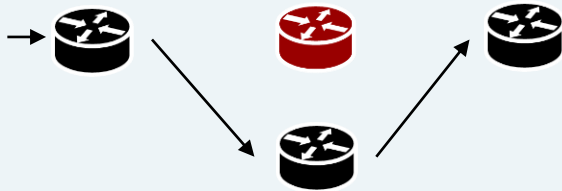
Specialized censor behavior and
path variance

Variety in censorship mechanisms

(1) Path A



(2) Path B



1

Censorship methods – RST injection, packet drops

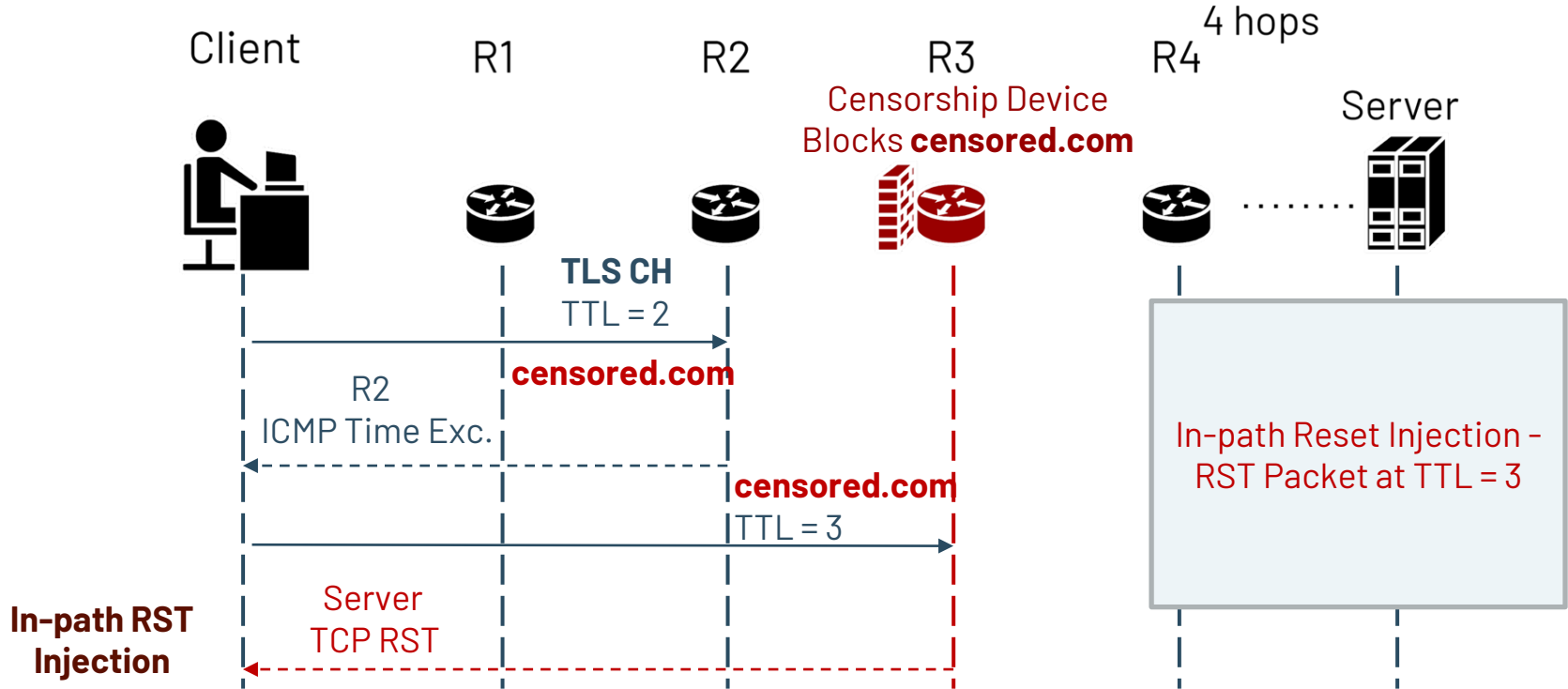
2

Device deployments – in-path vs on-path

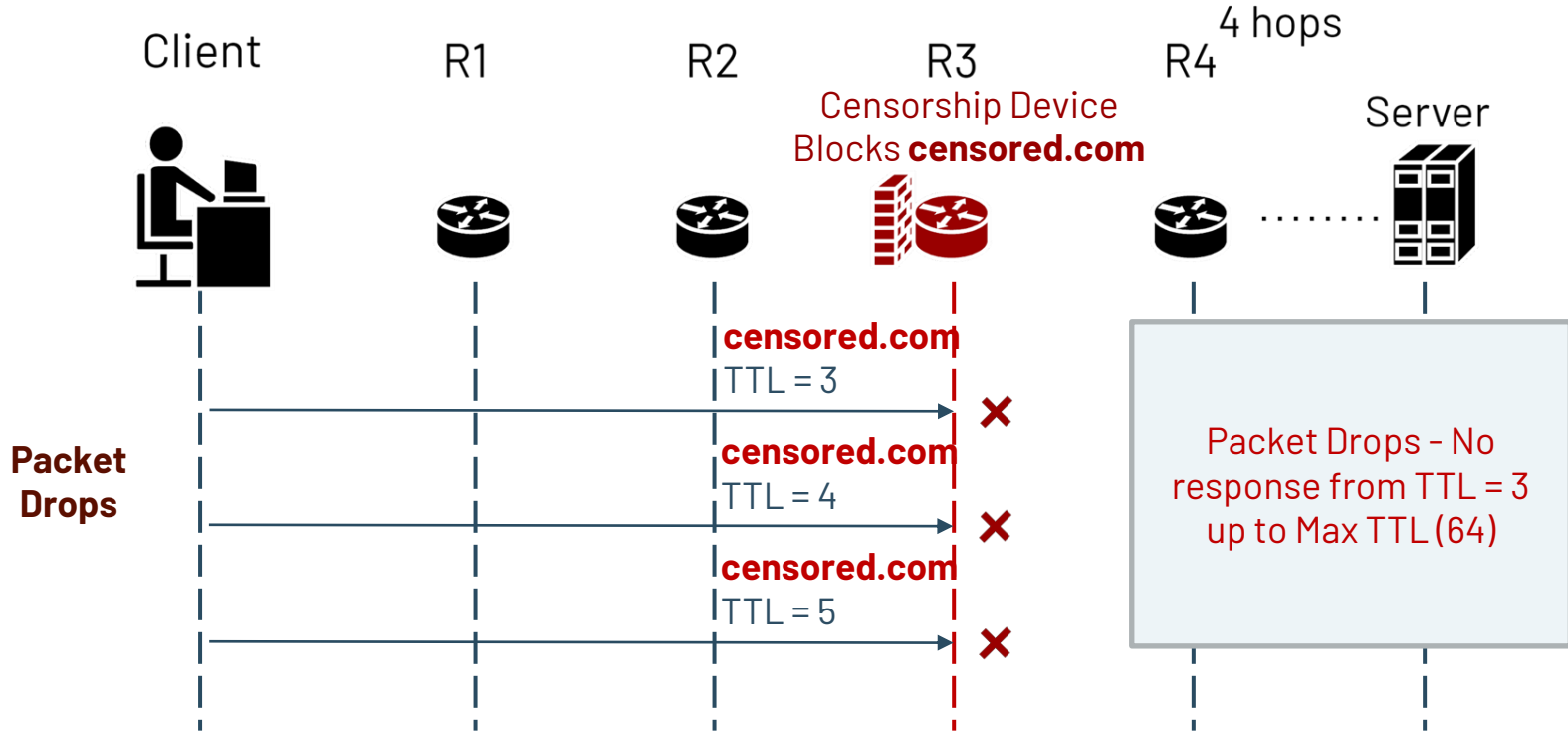
3

Specialized censor behavior and path variance

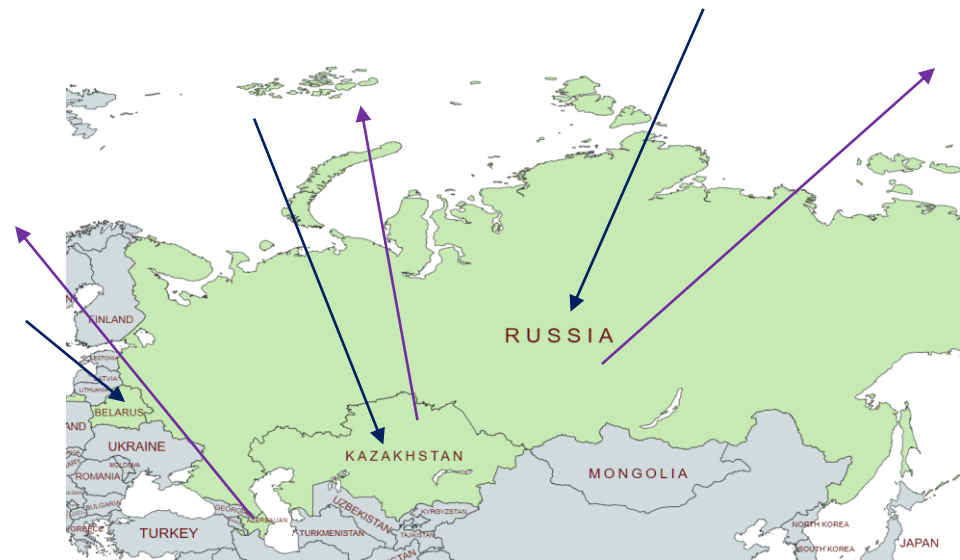
Censorship Traceroute



Censorship Traceroute



CenTrace



→ In-country measurements (Country -> Out)

→ Remote measurements (Out -> Country)

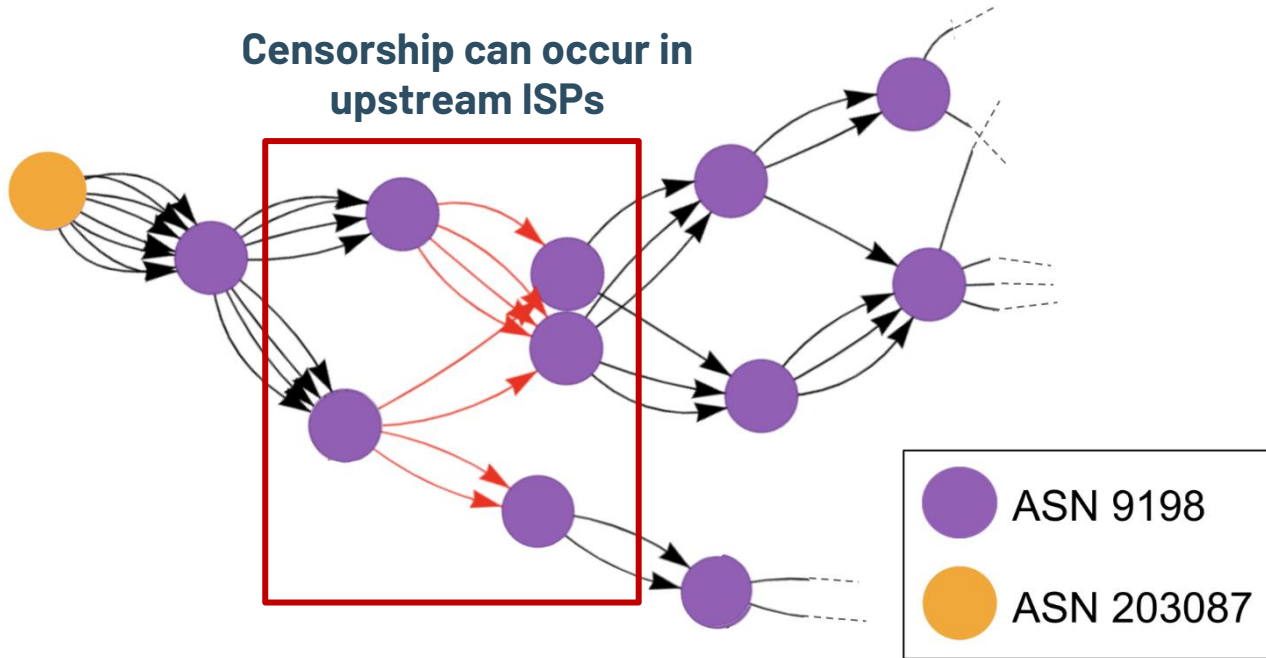
- Conduct in-country and remote measurements in Azerbaijan (AZ), Belarus (BY), Kazakhstan (KZ), Russia (RU)
- HTTP and TLS traceroutes

CenTrace: Finding Blocking Location

	Test CenTrace censored.com	Control CenTrace example.com
1	213.248.87.253	213.248.87.253
2	62.115.137.58	62.115.137.58
3	213.248.75.239	213.248.75.239
4	TIMEOUT	94.20.50.158
5	TIMEOUT	85.132.89.27
	⋮	⋮
15	TIMEOUT	Server - TLS

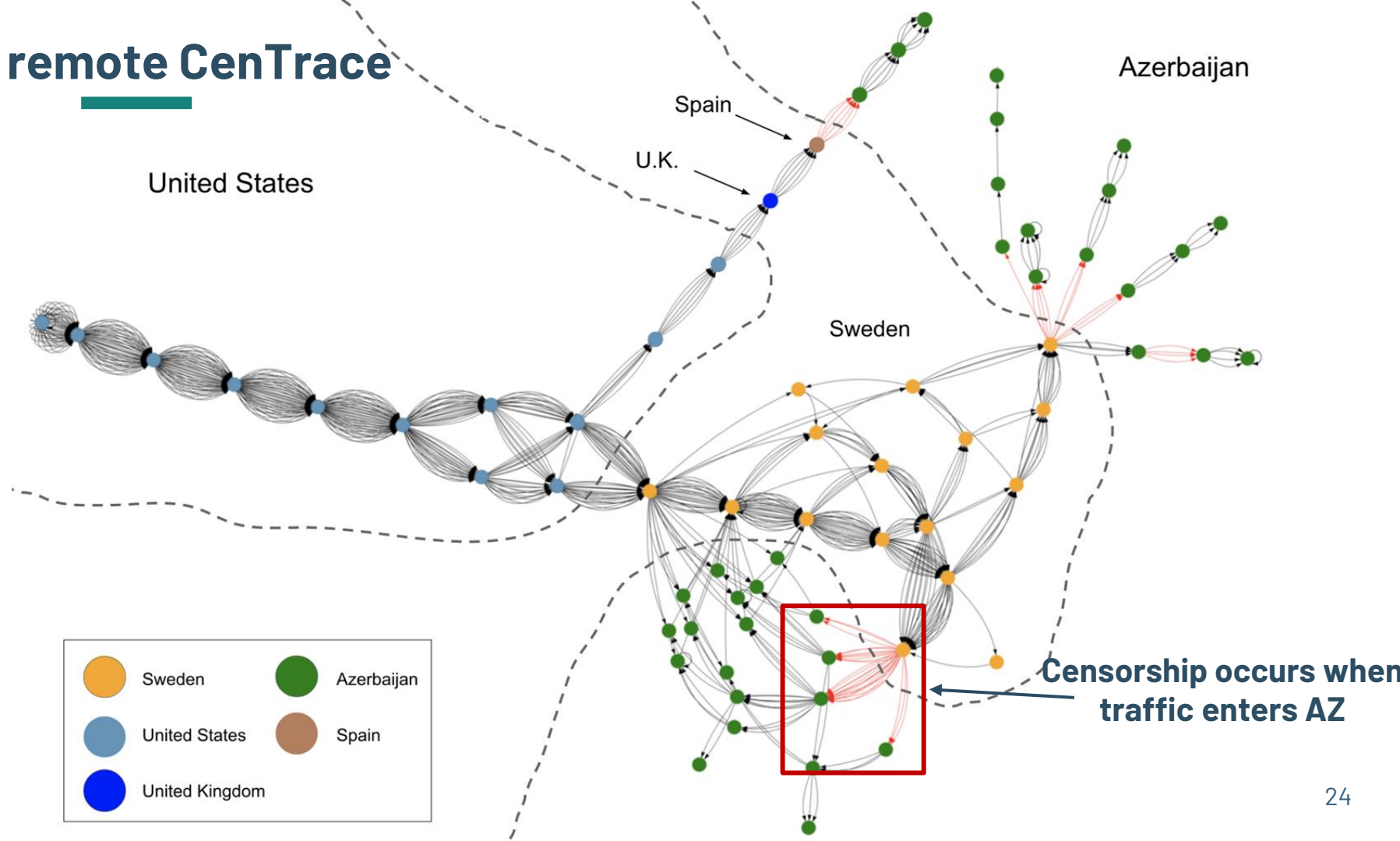
CenTrace: Finding Blocking Location

	Test CenTrace censored.com	Control CenTrace example.com
1	213.248.87.253	213.248.87.253
2	62.115.137.58	62.115.137.58
3	213.248.75.239	213.248.75.239
4	TIMEOUT	94.20.50.158
5	TIMEOUT	85.132.89.27
	⋮	⋮
15	TIMEOUT	Server - TLS

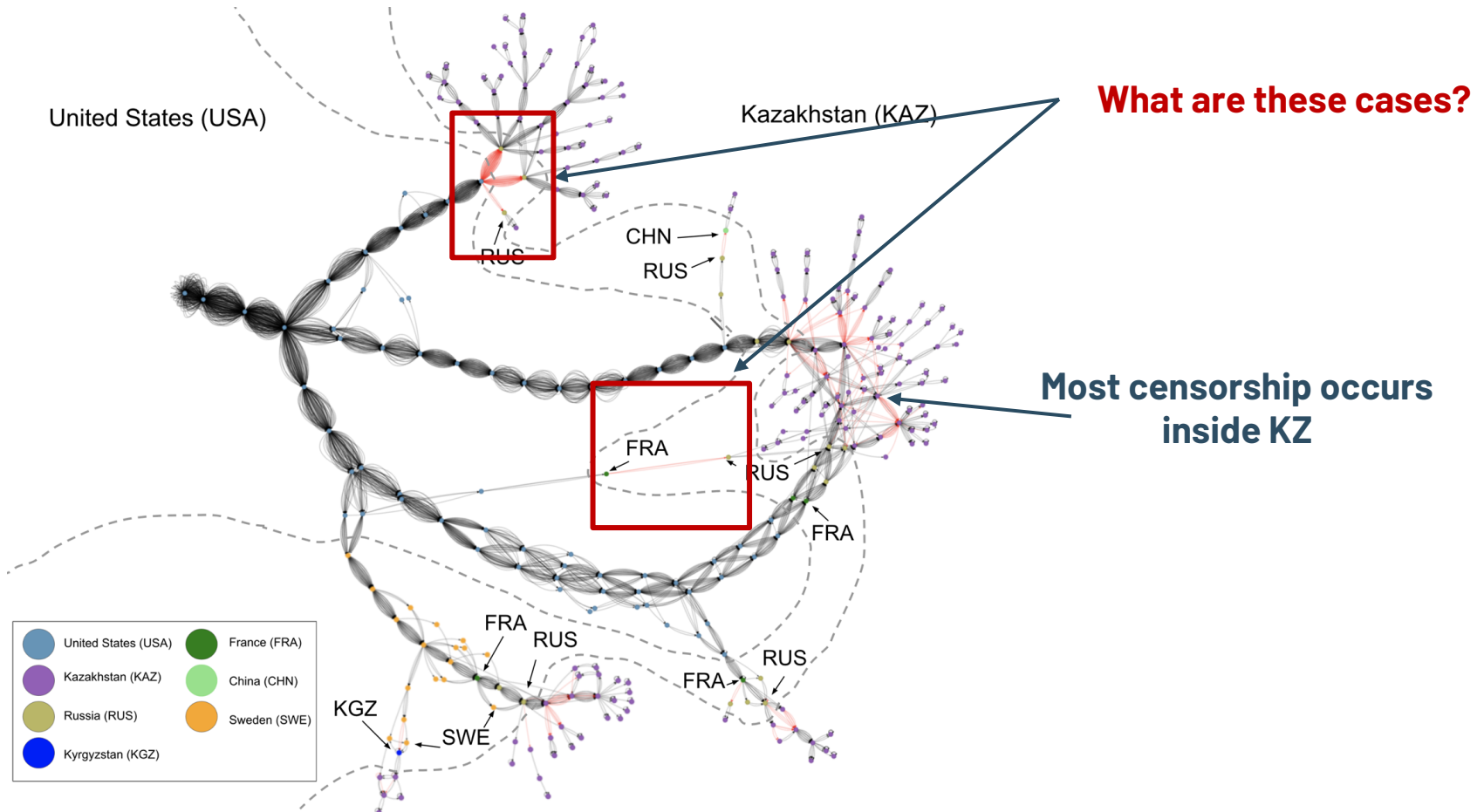


KZ in-country CenTrace

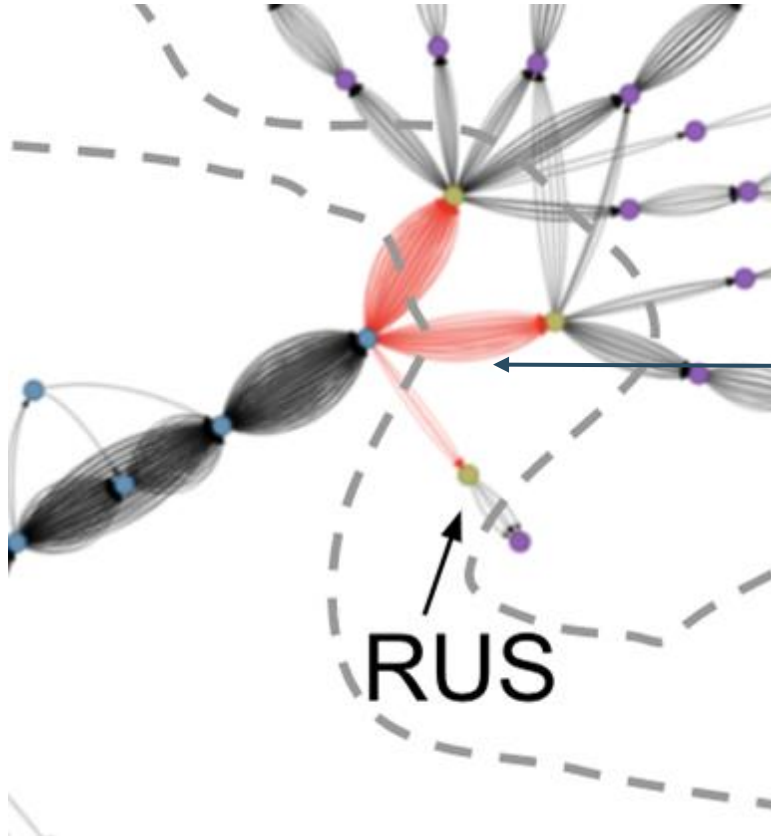
AZ remote CenTrace



KZ remote CenTrace



KZ remote CenTrace



Censorship occurs in Russian AS, even before entering KZ

CenTrace Observations

- Significant portion of remote measurements are **blocked at the endpoint**, indicate local policies
- Some devices exhibit specialized behavior such as **copying TTL values** from offending packet.
- Packet drops in Azerbaijan and Kazakhstan, Resets in Belarus and Russia

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

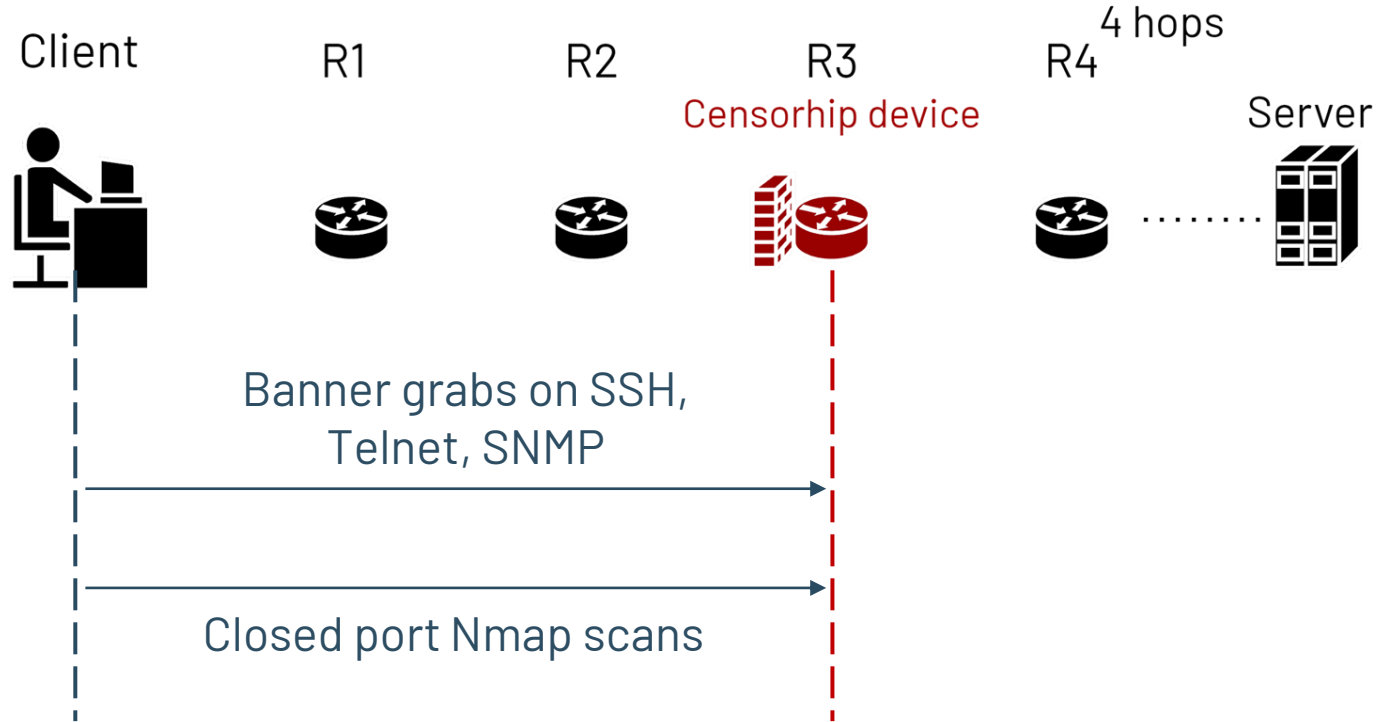
Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

Censorship Device Banners



Censorship Device Banners

- Collect banners on:
 - HTTP
 - TLS
 - SSH
 - Telnet
 - SMTP
 - SNMPv3
- Investigate banners manually and using fingerprint databases (Rapid7 Recog) to identify **commercial** filters

Censorship Device Banners

Device	
Cisco (7)	<p>Do these devices behave the same way?</p>
Fortinet (5)	
Kerio Control (2)	
Palo Alto (2)	
DDoSGuard	
Mikrotik	
Kaspersky	

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

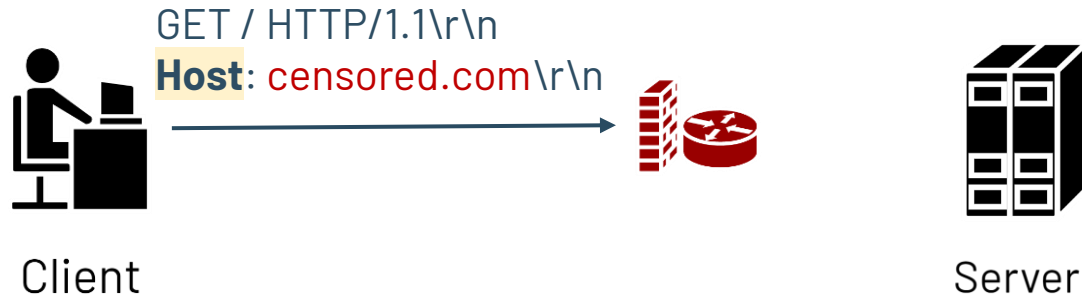
Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

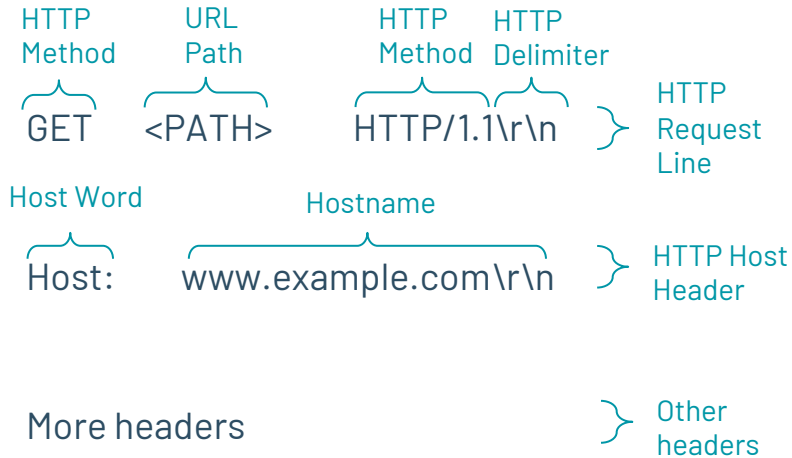
Fuzzing Strategies



Fuzzing Strategies



Fuzzing Strategies: HTTP



~400 fuzzing permutations

Fuzzing Strategies: HTTP

	HTTP Strategy	Examples	Permutations
Alternate	Get Word	POST, PUT	6
	HTTP Word	HTTP/ 1.1, XXXX/1.1	16
	Host Word	HostHeader:	7
	Path	?,z	8
	Hostname	www.example.com www.example.com	5
	Hostname TLD	www.example.net	10
	Hostname Subdomain	m.example.com	10
	Header	Connection: keep-alive	59
Capitalize or Remove	Get Word	GeT, GE	15
	HTTP Word	HtTP/1.1, HTTP/.1	183
	Host Word	HoST:, ost:	79
	HTTP Delimiter	\r	3
Pad	Hostname Padding	**www.example.com*	9

Fuzzing Strategies: HTTP

	HTTP Strategy	Examples	Permutations
Alternate	Get Word (HTTP Method)	POST, PUT, PATCH	6
	Host Word	HostHeader:	7
	Path	?,z	8
	Hostname	www.example.com www.example.com	5
	Hostname TLD	www.example.net	10
	Hostname Subdomain	m.example.com	10
	Header	Connection: keep-alive	59
Capitalize or Remove	Get Word	GeT, GE	15
	HTTP Word	HtTP/1.1, HTTP/.1	183
	Host Word	HoST:, ost:	79
	HTTP Delimiter	\r	3
Pad	Hostname Padding	**www.example.com*	9

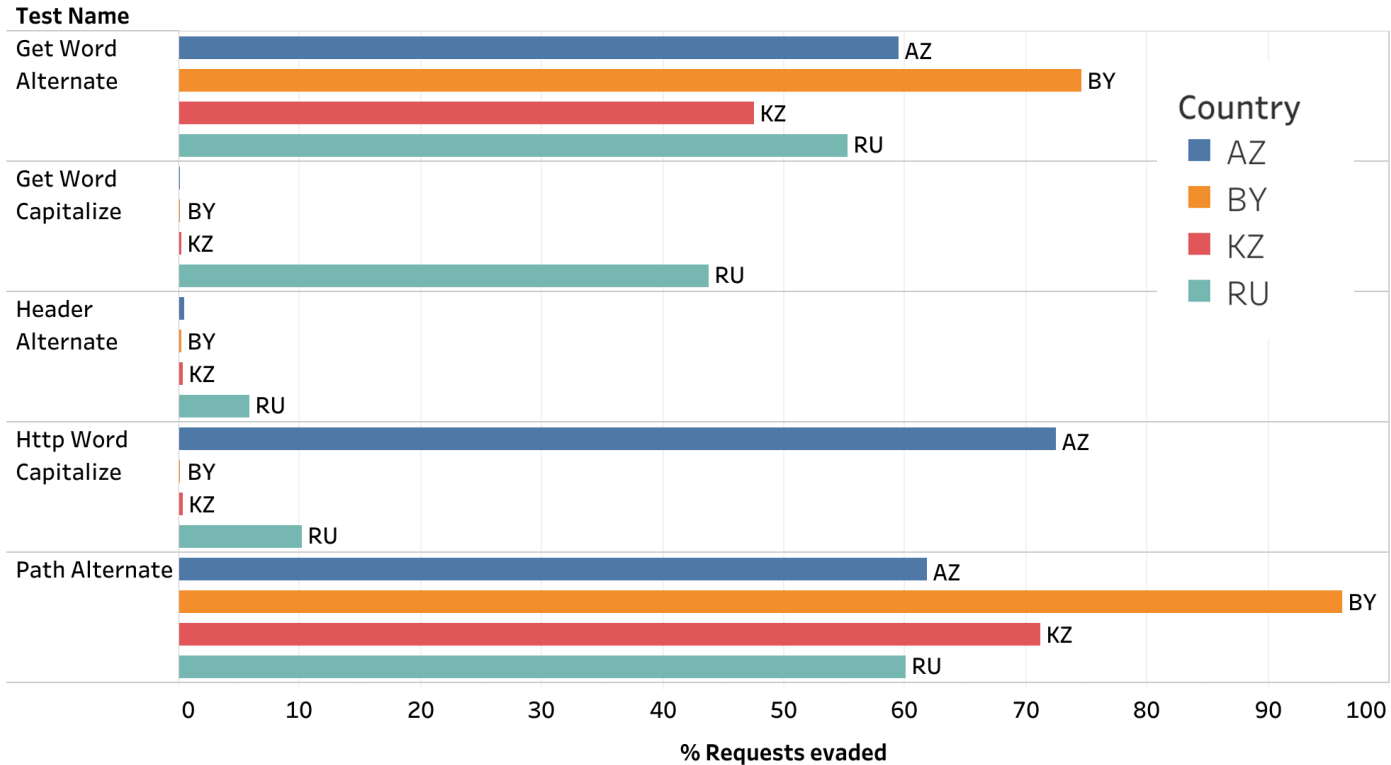
Fuzzing Strategies: HTTP

	HTTP Strategy	Examples	Permutations
Alternate	Get Word	POST, PUT	6
	HTTP Word	HTTP/ 1.1, XXXX/1.1	16
	Host Word	HostHeader:	7
	Path (URL)	?, z	8
	Hostname	www.example.com www.example.com	5
	Hostname TLD	www.example.net	10
	Hostname Subdomain	m.example.com	10
	Header	Connection: keep-alive	59
Capitalize or Remove	Get Word	GeT, GE	15
	HTTP Word	HtTP/1.1, HTTP/.1	183
	Host Word	HoST:, ost:	79
	HTTP Delimiter	\r	3
Pad	Hostname Padding	**www.example.com*	9

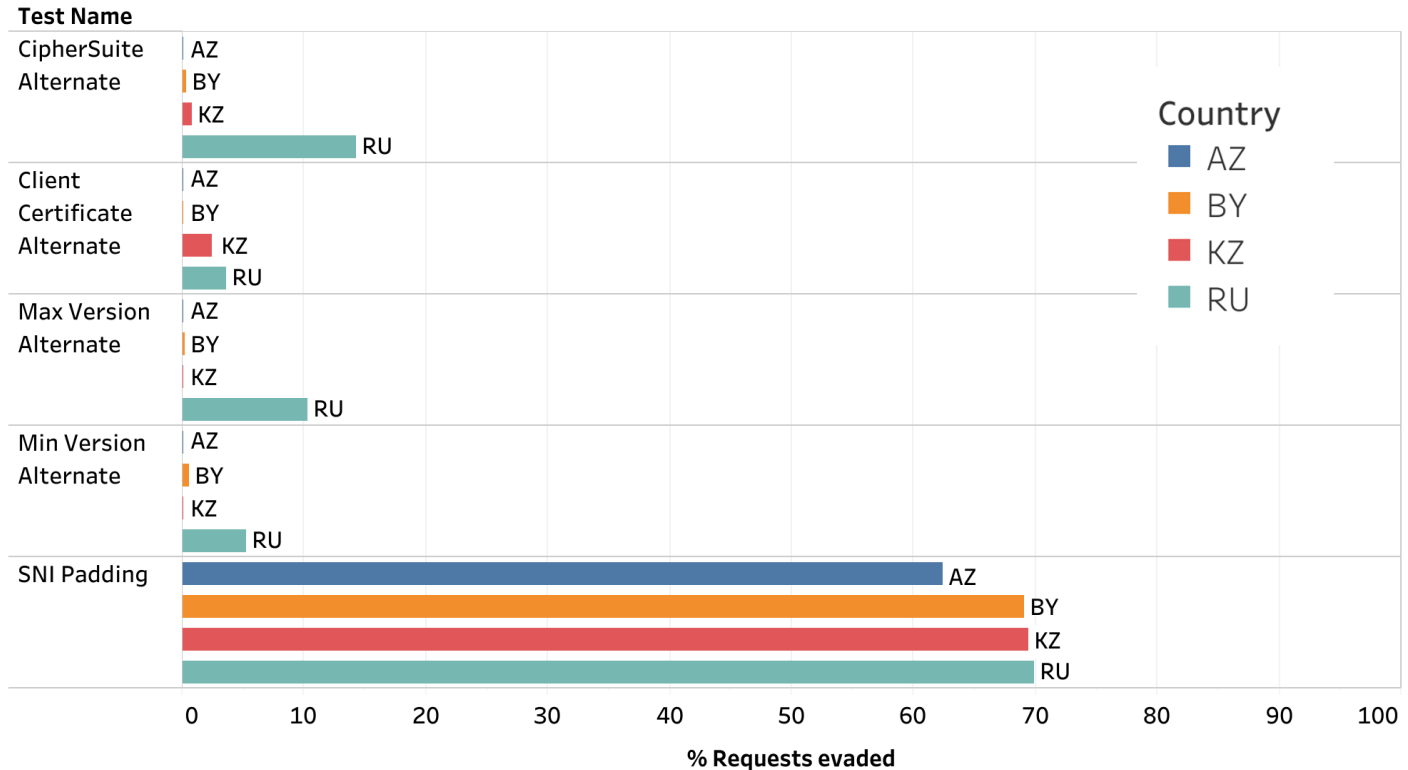
Fuzzing Strategies: HTTP

	HTTP Strategy	Examples	Permutations
Alternate	Get Word	POST, PUT	6
	HTTP Word	HTTP/ 1.1, XXXX/1.1	16
	Host Word	HostHeader:	7
	Path	?,z	8
	Hostname	www.example.com www.example.com	5
	Hostname TLD	www.example.net	10
	Hostname Subdomain	m.example.com	10
	Header	Connection: keep-alive	59
Capitalize or Remove	Get Word (HTTP Method)	GeT, GE	15
	Host Word	HoST:, ost:	79
	HTTP Delimiter	\r	3
Pad	Hostname Padding	**www.example.com*	9

CenFuzz HTTP: Evasion Success Rates



CenFuzz TLS: Evasion Success Rates



We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

Study similarities between censorship devices

We built robust, reusable solutions to:

1

Locate censorship devices

Censorship Traceroute

2

Identify device vendors

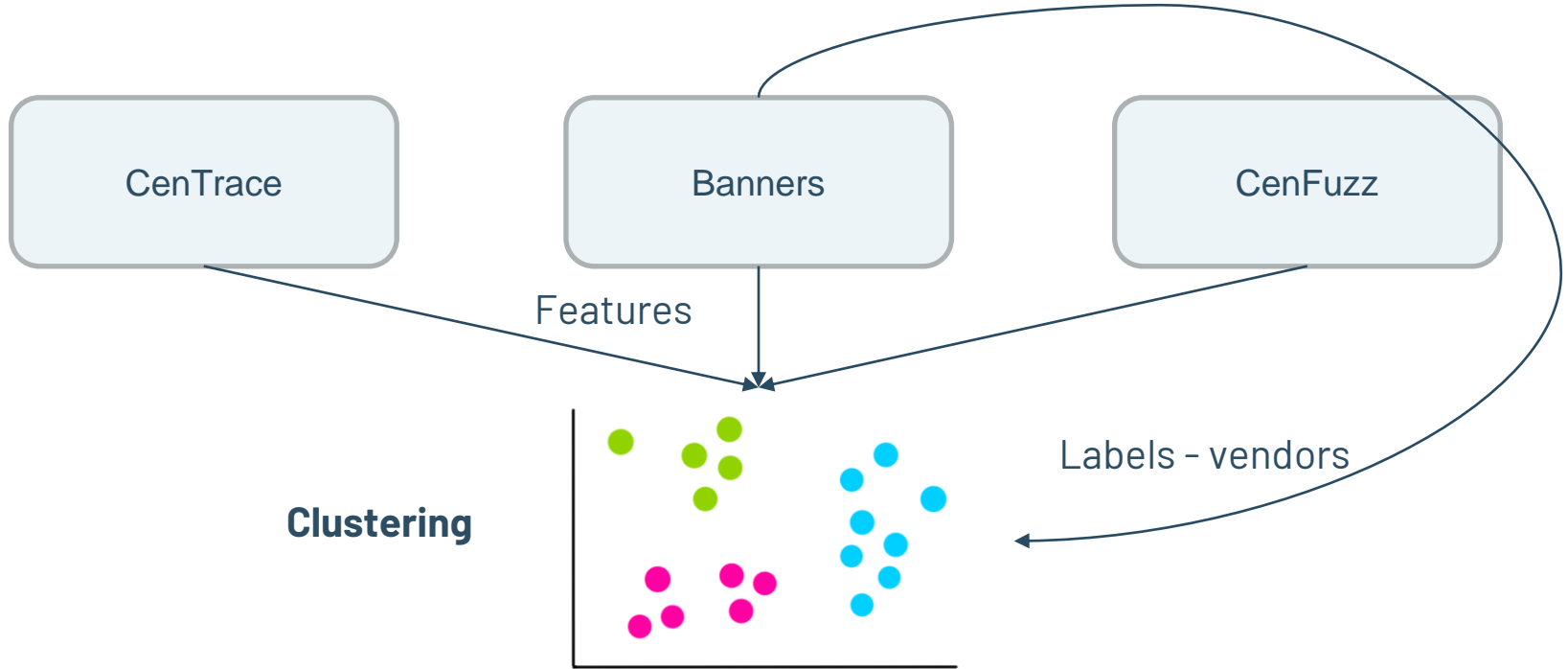
Banner grabs and Clustering

3

Reverse-engineer censorship triggers

Censorship Fuzzer

Clustering Devices

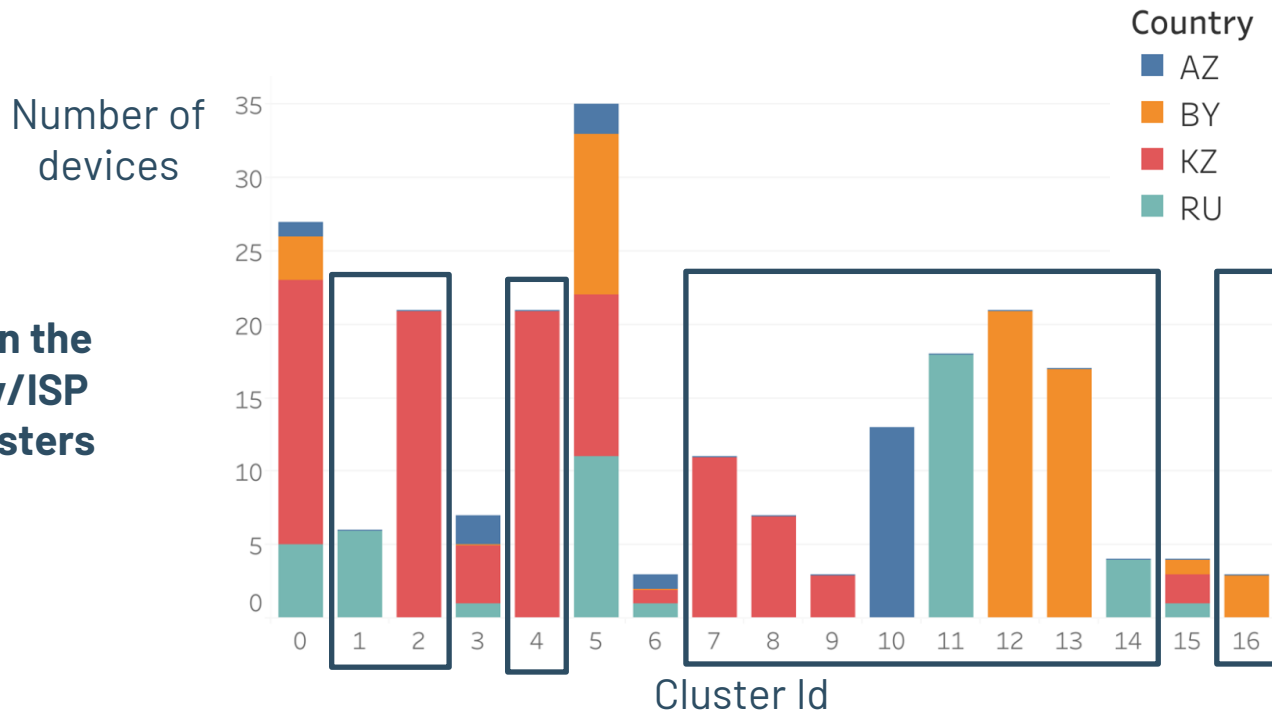


Clustering Devices



Clustering Devices

Devices within the same country/ISP form tight clusters



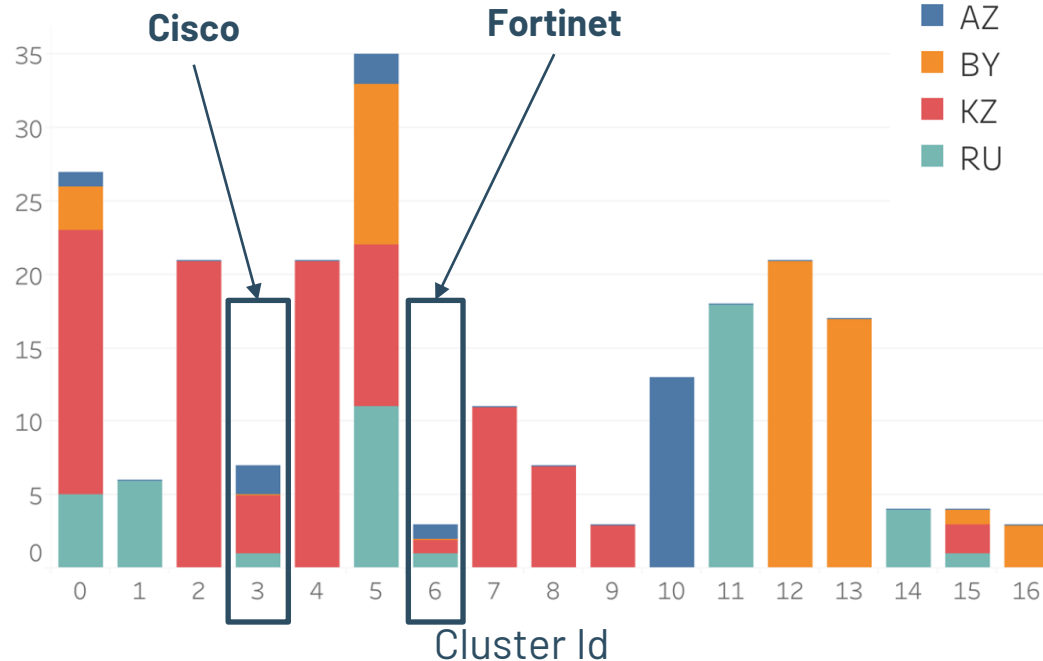
Clustering Devices

Number of devices

Country

- AZ
- BY
- KZ
- RU

Clusters with devices from different countries have same features, indicating cross-country deployment



Our code and data are fully open-source



<https://github.com/censoredplanet/CenTrace>
<https://github.com/censoredplanet/CenFuzz>



Censored Planet report - <https://censoredplanet.org/censorship-devices>
OTF report - <https://www.opentech.fund/news/>



Highlighting policy gaps
Assisting censorship research

What's Next?

- Integrate CenTrace, CenFuzz into Censored Planet, OONI
- Expanded CenTrace: how many end hosts are behind each of the devices?
- Study censorship devices in more countries
- Improve ground truth

Key Takeaways

- Location of censorship is important: **frequently occurs in upstream ISPs or even in other countries**
- Devices can be deployed with different properties: **in-path, on-path, packet drops, copy TTL values**
- **Banners** on popular protocols are useful for identification
- The censorship triggers and other features are **device- or deployment-specific** and can be used to fingerprint them

Key Takeaways

- Location of censorship is important – **frequently occurs in upstream ISPs or even in other countries**
- Devices can be deployed with different properties – **in-path, on-path, packet drops, copy TTL values**
- **Banners** on popular protocols are useful for identification
- The censorship triggers and other features are **device- or deployment-specific** and can be used to fingerprint or identify them

Thank you!

Questions?

Reach out to us at ramaks@umich.edu and monaw@princeton.edu

<https://censoredplanet.org/censorship-devices>

Need more info?

<https://censoredplanet.org/censorship-devices>

Previous Studies

Location



GFW [Marczak et al. (2015), Xu et al. (2011)]
Russia's TSPU [Xue et al. (2021)]
Kazakhstan's HTTPS interception system [Sundara Raman et al. (2020)]
Iran [routeviz]

Triggers



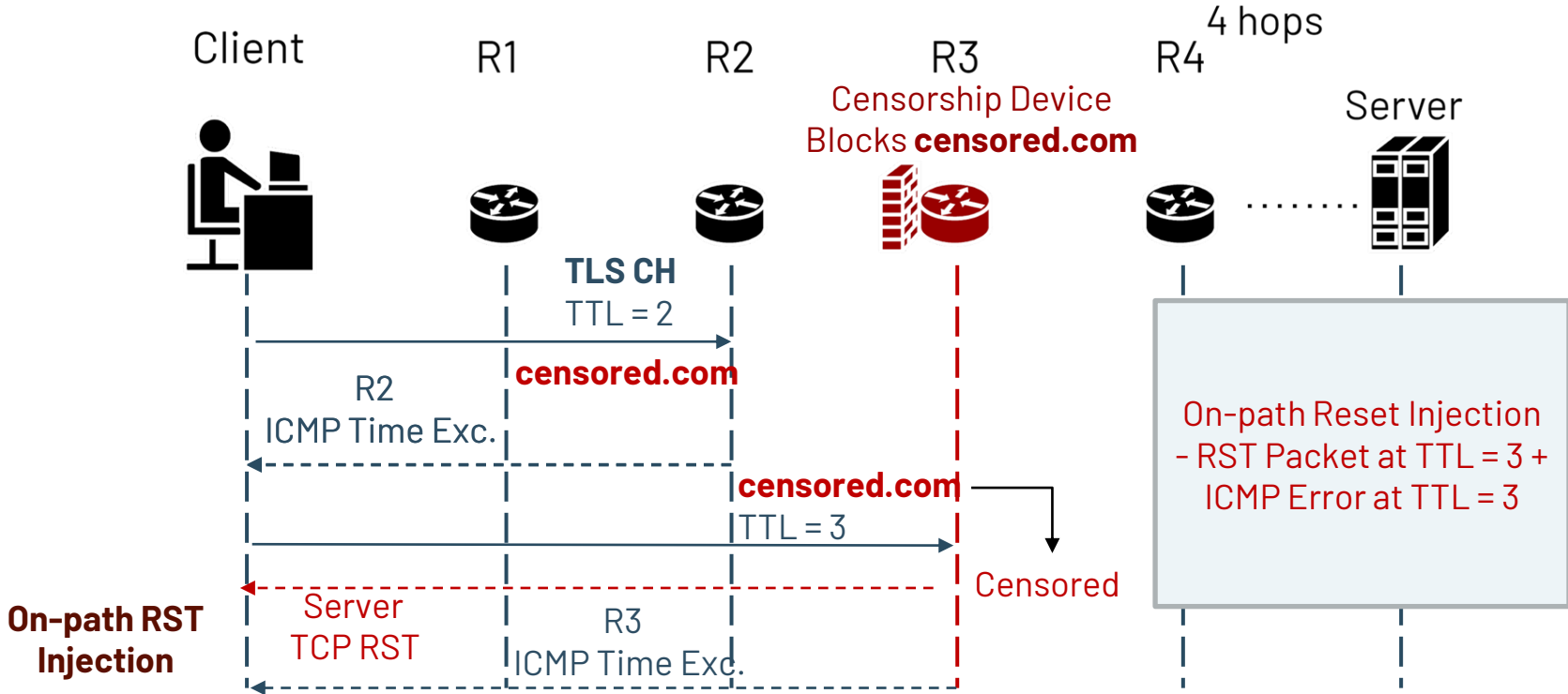
Circumvention [Bock et al. (2019), Li et al. (2017)]
Fuzzing [Jermyn et al. (2017)]

Identity

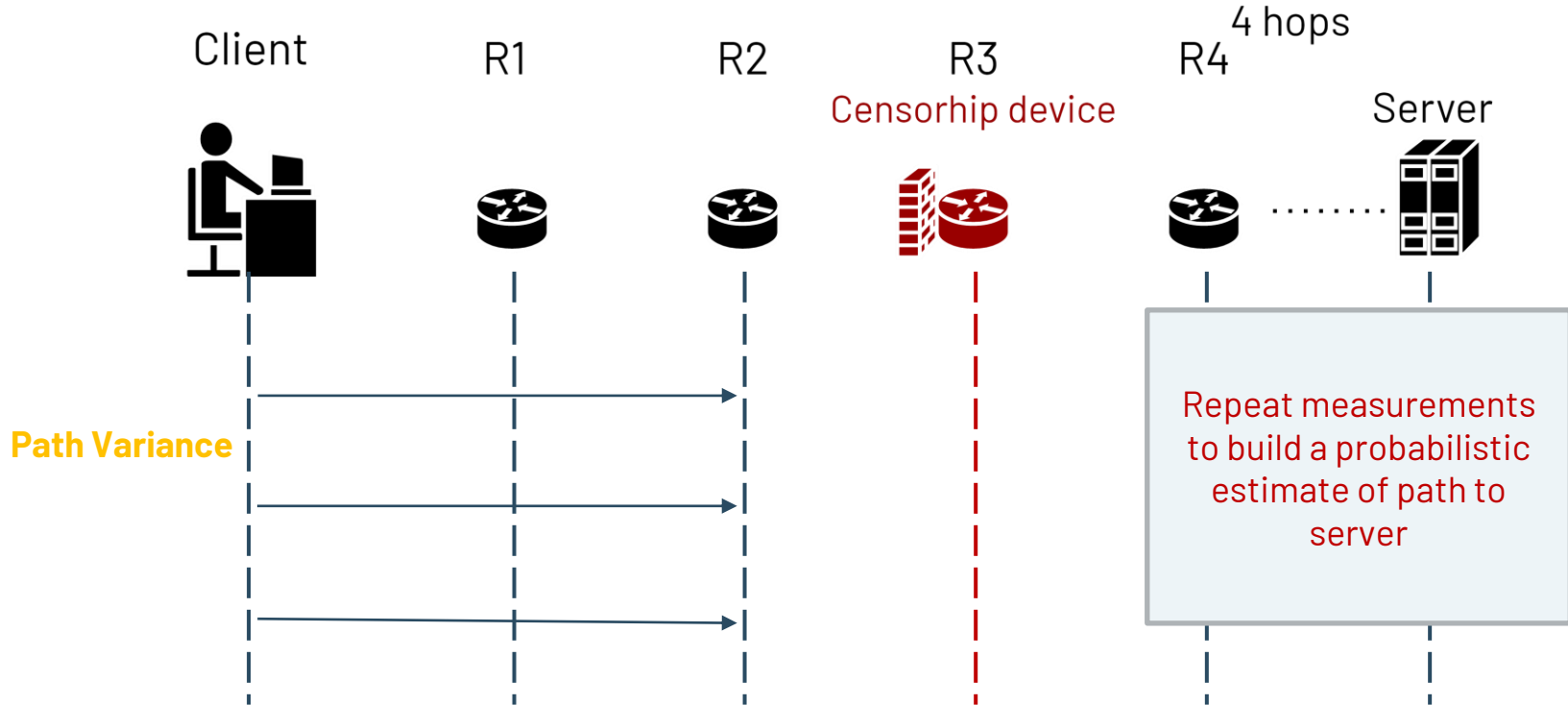


Network Signatures [Planet Netsweeper (2018), Planet Blue Coat (2013), Bad Traffic (2018), Dalek et al. (2013)]
Blockpages [Sundara Raman et al. (2020)]

Censorship Traceroute



Censorship Traceroute

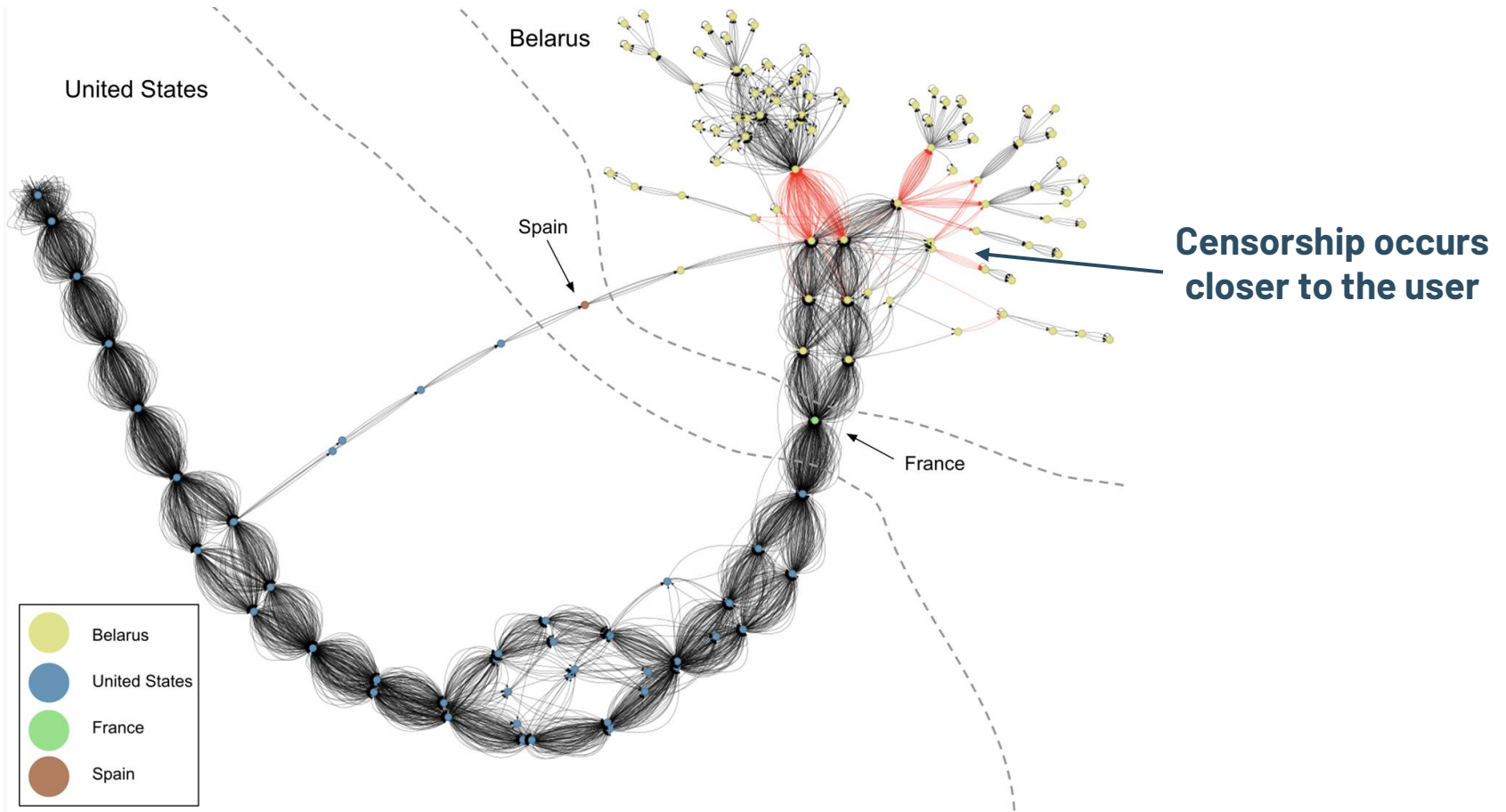


CenTrace Measurements

Co.	In-Country Measurements			Remote Measurements			
	Clients	CenTraces	Blocked CenTraces	Endpoints	Endpoint ASNs	CenTraces	Blocked CenTraces
AZ	1	18	6	29	10	227	96
BY	-	-	-	123	19	1,040	287
KZ	1	14	8	95	29	868	748
RU	1	14	0	1,291	498	10,488	418

Block Types: TCP RST injection, Blockpage injection, Packet Drops

BY remote CenTrace

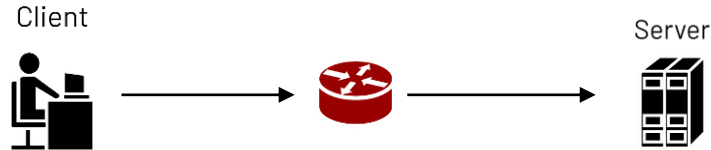


Censorship Type and Overall Location

Result Code	Block Type	Country			
		AZ	BY	RU	KZ
Path (Client -> Endpoint)	Timeout	93	32	128	618
	RST/FIN	0	251	33	0
	HTTP	0	0	0	1
At Endpoint	Timeout	8	0	8	129
	RST/FIN	1	4	97	2
	HTTP	0	0	0	4
Past Endpoint (Endpoint ->)	Timeout	0	0	1	2
	RST/FIN	0	0	150	0
No ICMP	RST/FIN	0	0	1	0

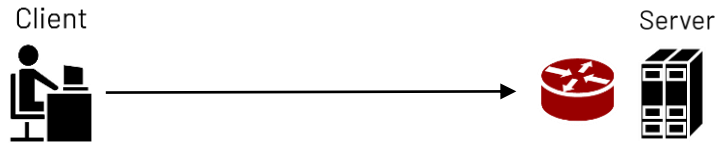
Censorship Type and Overall Location

Result Code	Block Type	Country			
		AZ	BY	RU	KZ
Path (Client -> Endpoint)	Timeout	93	32	128	618
	RST/FIN	0	251	33	0
	HTTP	0	0	0	1



Censorship Type and Overall Location

Result Code	Block Type	Country			
		AZ	BY	RU	KZ
Path (Client -> Endpoint)	Timeout	93	32	128	618
	RST/FIN	0	251	33	0
	HTTP	0	0	0	1
At Endpoint	Timeout	8	0	8	129
	RST/FIN	1	4	97	2
	HTTP	0	0	0	4



Censorship Type and Overall Location

Result Code	Block Type	Country			
		AZ	BY	RU	KZ
Path (Client -> Endpoint)	Timeout	93	32	128	618
	RST/FIN	0	251	33	0
	HTTP	0	0	0	1
At Endpoint	Timeout	8	0	8	129
	RST/FIN	1	4	97	2
	HTTP	0	0	0	4
Past Endpoint (Endpoint ->)	Timeout	0	0	1	2
	RST/FIN	0	0	150	0
No ICMP	RST/FIN	0	0	1	0

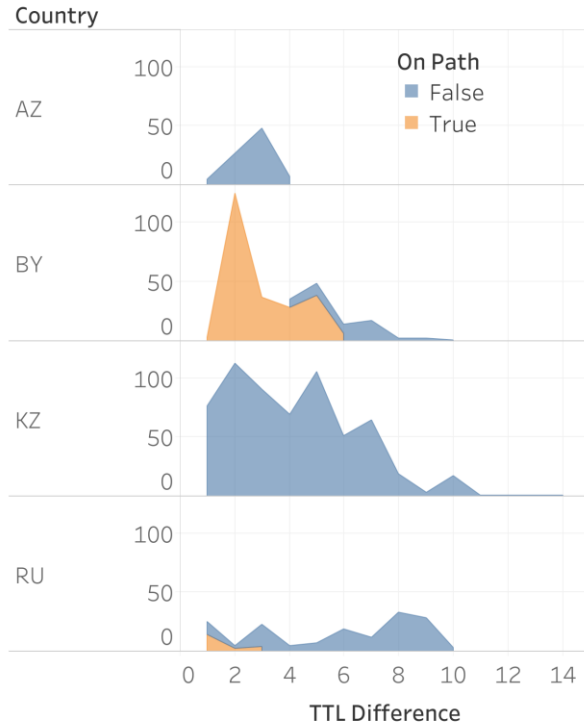
Censorship Type and Overall Location

Result Code	Block Type	Country			
		AZ	BY	RU	KZ

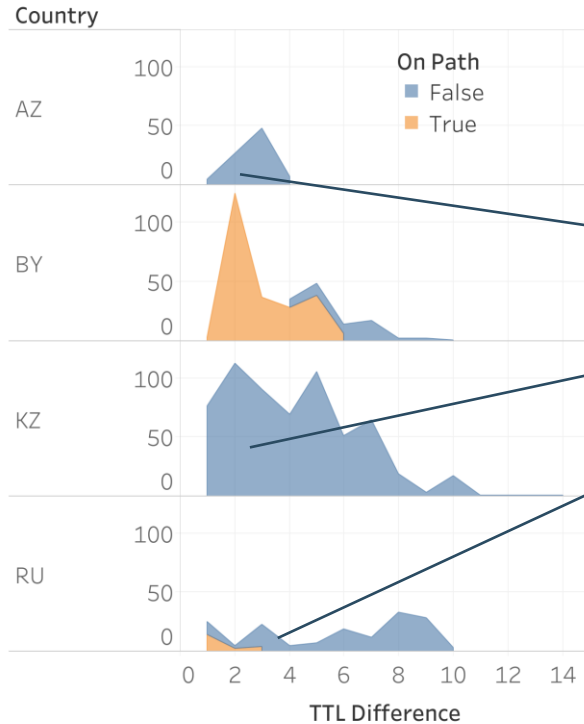


Past Endpoint (Endpoint ->)	Timeout	0	0	1	2
	RST/FIN	0	0	150	0
No ICMP	RST/FIN	0	0	1	0

In-path vs On-path and distance from Client

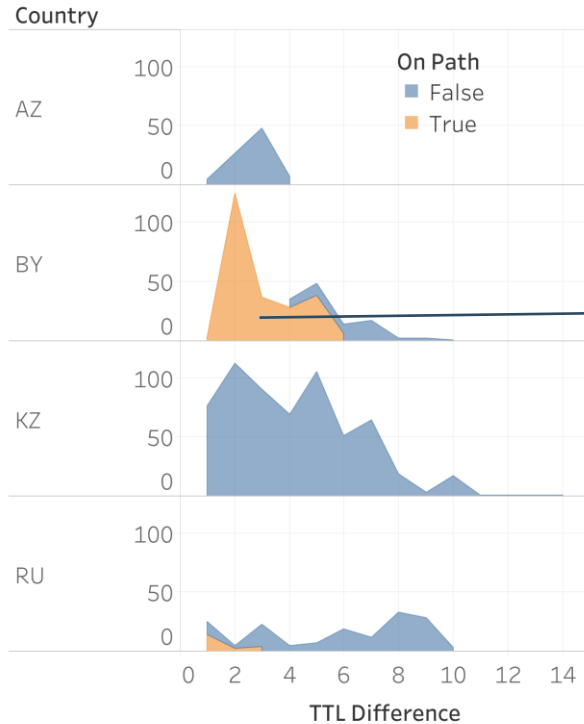


In-path vs On-path and distance from Client



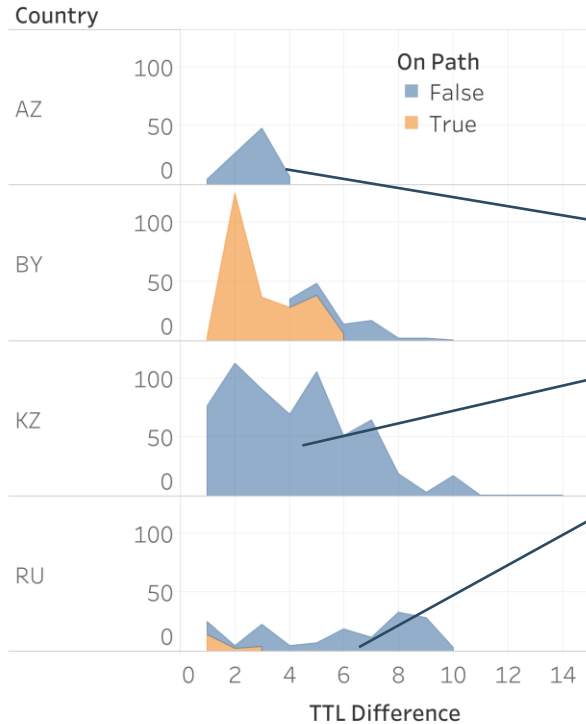
Most censorship happens close to the endpoint, but there is a tail

In-path vs On-path and distance from Client



Censorship in Belarus is mostly on-path

In-path vs On-path and distance from Client



Most other censorship is in-path
→ We can possibly extract IPs
of middleboxes

Fuzzing Strategies: TLS

Handshake Header

Handshake Type (Client Type)

Length

Version

} TLS Version

Client Random

Session ID

Cipher Suites Length

} List of Cipher Suites

Cipher Suites

Compression Methods

Extensions Length

Extension

Type: server_name

Length

Server Name Indication Extension

Server Name list length

Server Name Type: host_name

Server Name Length

Server Name: www.example.com

} Server Name

Traceroute

IPv4 Header

