CSE227 – Graduate Computer Security

More Network Censorship!

UC San Diego

Housekeeping

General course things to know

- **Final presentations happen in week 10!**
  - Keep making consistent progress, and feel free to meet with me if you have things to discuss and work on
- Midpoint check-in meetings will happen week 7… thanks for signing up
  - Met with some of you, will meet with all of you by the end of the week!
- Overall, very impressed with the output of work so far, and can’t wait to see more

This week's research goals

General course things to know

- Obviously, you will meet with me this week
- Out of that meeting might come some additional ideas, but remember: now is the time to **work** on your projects and make steady progress towards the end of the quarter
  - Chart out the next four weeks, **execute**
- Goal by end of this week: finalize scope + initial experiments to give you something to aim at for the last N weeks

But first… Canvas

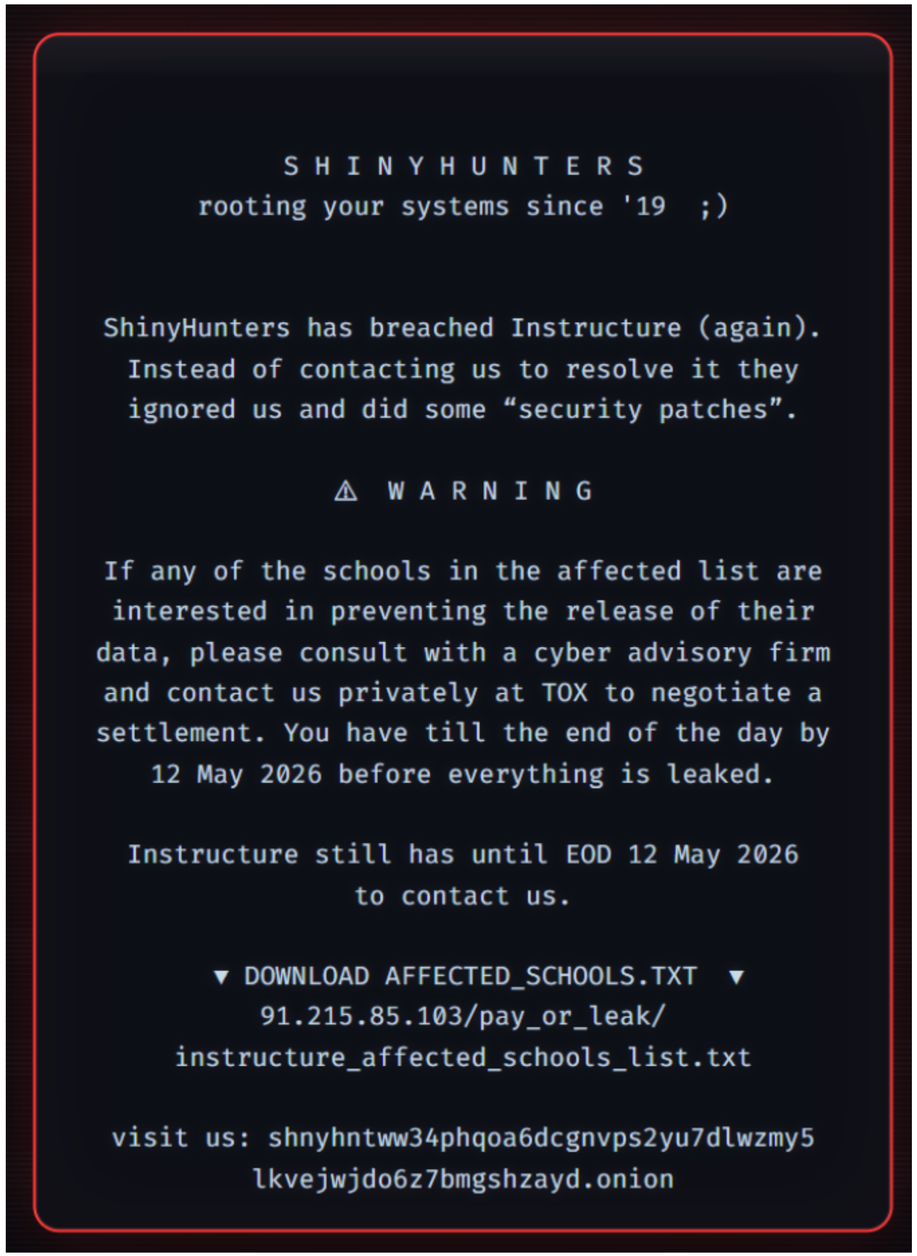
- Canvas was owned by a group of teenage hackers!
- Instructure (the company that runs Canvas) got owned, took down ~10K schools

STATUS UPDATE 5/11/26

We know that concerns about the potential publication of data related to this incident remain top of mind for many customers. We understand how unsettling situations like this can be, and protecting our community remains our top priority.

With that responsibility in mind, Instructure reached an agreement with the unauthorized actor involved in this incident. As part of that agreement:

- The data was returned to us.
- We received digital confirmation of data destruction (shred logs).
- We have been informed that no Instructure customers will be extorted as a result of this incident, publicly or otherwise.
- This agreement covers all impacted Instructure customers, and there is no need for individual customers to attempt to engage with the unauthorized actor.



Today's lecture

Learning Objectives

- Continue discussing the "Wallbleed" paper
- Discuss the Cuba paper

Where we left off…

- What is network censorship?
- What is DNS blocking?
- What's a middlebox?
- How might you implement DNS blocking if you control a middlebox at a country's border?
  - Why does a middlebox approach here work in a country like China?

DNS Queries and Responses

- What is the Question section?
- What is the Answer section?
- What is a QNAME?
- What is RDATA?
- What do the bytes c00c mean?
- What *device* creates the Answer section?

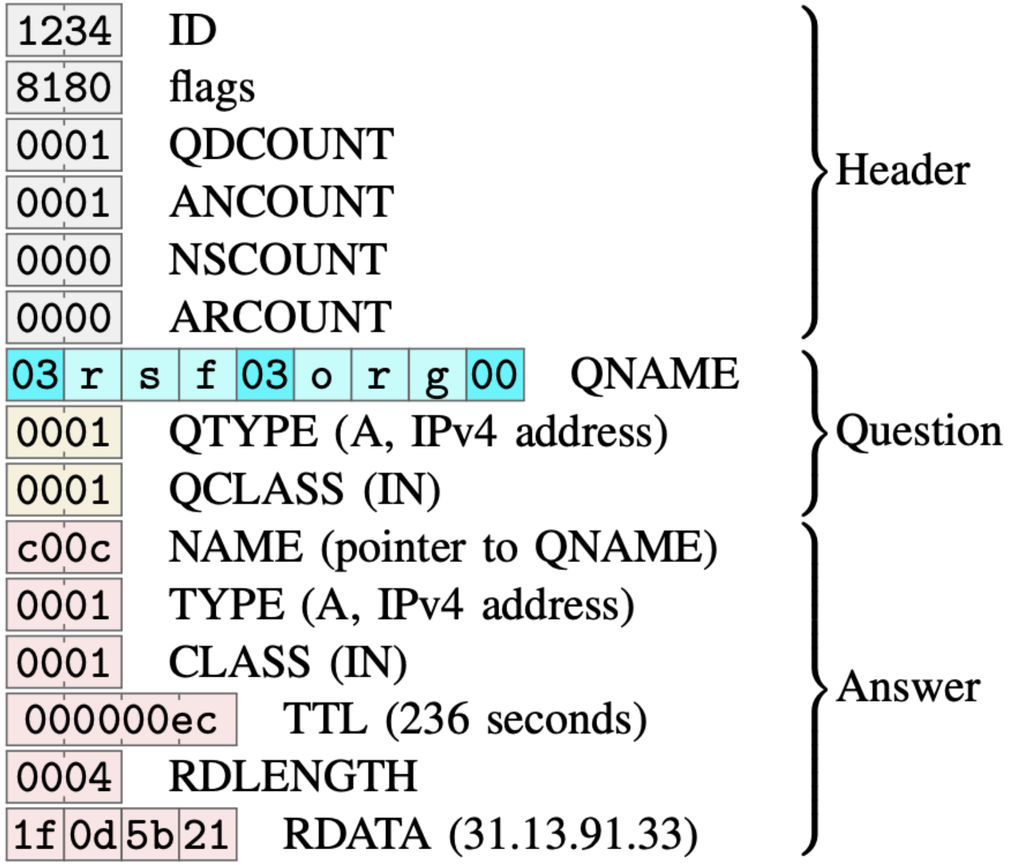


Fig. 1: The structure of an injected DNS response.

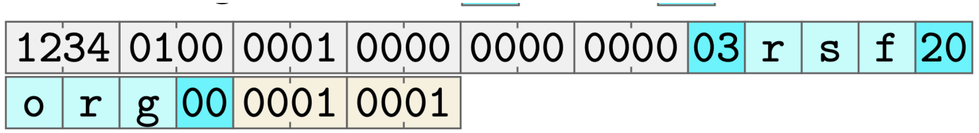
Wallbleed

DNS Name Encoding

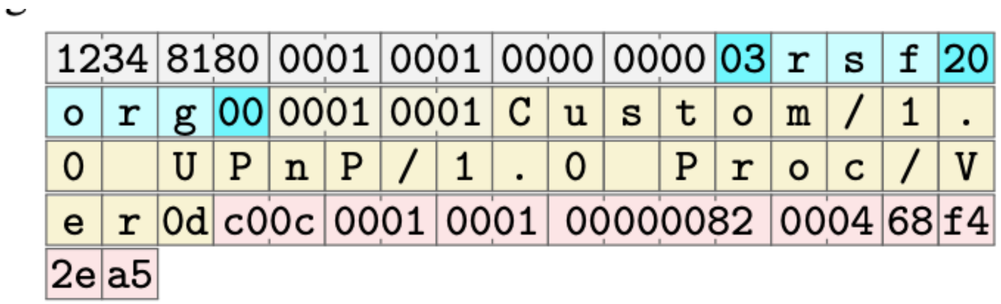
- What is a *label*?
- How are DNS names encoded in DNS packets?
- How does *Wallbleed* leverage this encoding to *leak data?*

Wallbleed

- **Poor implementation of DNS parsing leaks secret memory**
- Authors construct DNS packets where the *label length* does not correspond to the *label itself*
  - Implementation has no bounds checks, so **memory is leaked in response!**
- In this example, *~15 bytes leaked beyond the name in memory*



constructed query



response

Why does this happen?

- What do the authors posit as the reason this is happening?

Why does this happen?

- What do the authors posit as the reason this is happening?
  - DNS packets are *copied* in memory for downstream processing (e.g., checking name against a blocklist)
  - Implementers built a *custom* name parser, which doesn't perform *bounds checking*
  - So… if you control the bytes, you control how much data is returned!

What did they do…?

- Authors *exploited* Wallbleed for two years…
  - Getting 5.1B *Wallbleed responses*, GBs of data…
- Found **tons** of network protocols, architecture details, *passwords*, cookie values, secret keys….
  - You name it!
- What architecture was the middlebox running on?

Regular Expression	Note	Count	Rate
`ssdp:discover`	SSDP	184M	3.61%
`UPnP/IGD\xml`	UPnP	174M	3.41%
`(?s)[3-4]\xfttt.....-CONTROL`	(§IV-B)	121M	2.37%
`\x45\x00`	(§IV-A)	2.8M	0.05%
`uuid:WAN`	SSDP	34M	0.67%
`Host:␣`	HTTP	21M	0.41%
`(?i)Date:\s* ...`	(§IV-C)	16M	0.31%
`\x7f\x00\x00`	(§IV-D)	2.8M	0.05%
`Cookie:␣`	HTTP	2.0M	0.04%
`RCPT␣TO`	SMTP	72.5k	0.0014%
`&key=`	URL	58.1k	0.0011%
`MAIL␣FROM`	SMTP	42.4k	0.0008%
`&password=`	URL	26.9k	0.0005%

Ethics

- Note…. **huge disclaimer** at the top of the paper:

Statement from NDSS 2025: The NDSS 2025 PC appreciated the technical contributions made in this paper, the confirmation of prior work that is otherwise not (directly) reproducible, and the contributions towards fostering future anti-censorship research, but it also found the paper highly controversial because the experiments that the authors conducted raise ethical concerns. This paper went through scrutiny by various stakeholders beyond the regular PC review, including evaluation by the NDSS'25 Ethics Review Board and consultation of the Steering Committee and ISOC. While the ethical ambiguities were deemed remedied after data aggregation and deletion, the IRB Exempt decision that the authors received from their institution should have been questioned and repudiated by the authors as there are clear human risks involved. Questioning such an IRB decision should be an obligation by researchers in the security community. Additionally, the PC does not consider itself qualified to make a judgment about the legal implications of this work. We acknowledge that there were conflicting opinions during the broader review process on whether the benefits of this research outweigh its risks. We hope that the acceptance of this paper helps the community understand the possible impact of research work, allows better mechanisms to deal with similar cases, and contributes to developing accepted standards on when and how such types of offensive research can be done. The acceptance of this paper does not constitute the PC's endorsement of the used methodology. We advise authors to seek legal advice (from different legislations if applicable) before/while doing security research that may impact critical targets.

Ethics

- Do you think this research is ethical?
  - Why or why not?

Ethics

- Do you think this research is ethical?
  - Why or why not?
- Is it okay to store this data in aggregate over time?
- Is it okay to exploit a known vulnerability?
- Should the authors have disclosed this vulnerability?

Meta-thoughts

- This paper was highly contentious in discussions and remains a very divisive paper. Why?
- What about this paper *surprised* you?
- Other last thoughts?

Break Time + Attendance

https://tinyurl.com/cse227-attend

Codeword:
Cubastic

Digital Discrimination of Users in Sanctioned States: The Case of the Cuba Embargo

Where is Cuba?

Where is Cuba?

Geoblocking

- What is geoblocking?

Geoblocking

- What is geoblocking?
  - Form of censorship where content is equivocated or discriminated based on the *origin of the request*
- Where does geoblocking happen? On the server or the client?

Geoblocking

- What is geoblocking?
  - Form of censorship where content is equivocated or discriminated based on the *origin of the request*
- Where does geoblocking happen? On the server or the client?
  - Server-side — the main focus of this paper
- What are some examples of geoblocking?

Sanctions

- What are sanctions?

Sanctions

- What are sanctions?
  - Either: Things you allow towards an entity or…
  - Penalties you impose on an entity
  - **Contronym —** word that has opposite meanings depending on the context (like *peruse*)
- In this case, we are looking at **penalties imposed on countries** and how *servers* choose to implement them

Sanctions pt. II

- What is the relationship between sanctions and the Internet?

Sanctions pt. II

- What is the relationship between sanctions and the Internet?
  - American Internet providers have a choice: do business (and risk penalty) or don’t do business (and implement measures to geoblock)
  - **Sort of a damned-if-i-do-damned-if-i-don’t situation**
  - Policy is extremely confusing…

Why Cuba?

- Why study Cuba?

Why Cuba?

- Why study Cuba?
  - Doing this work is challenging, risky, and ethically dubious… wanted to study in a country with minimal risks to participants and researchers
  - Cuba has relatively higher civil-liberty posture than other countries with sanctions
  - It’s close by!

A word on methods…

- This is a mixed-methods study. What does this mean?

A word on methods…

- This is a mixed-methods study. What does this mean?
  - Contains both qualitative + quantitative methods — both of which help to tell the fuller story of the paper
- What are some benefits of qualitative research?
- What are some benefits of quantitative research?

A word on methods…

- This is a mixed-methods study. What does this mean?
  - Contains both qualitative + quantitative methods — both of which help to tell the fuller story of the paper
- What are some benefits of qualitative research?
- What are some benefits of quantitative research?
  - **In sum, you choose the right tool for the job, and sometimes you want both tools!**

Qualitative Study

- This paper first seeks to understand the everyday experiences of Cuban users
- Users had a hard time disambiguating censorship from geoblocking. What's the difference?

Qualitative Study

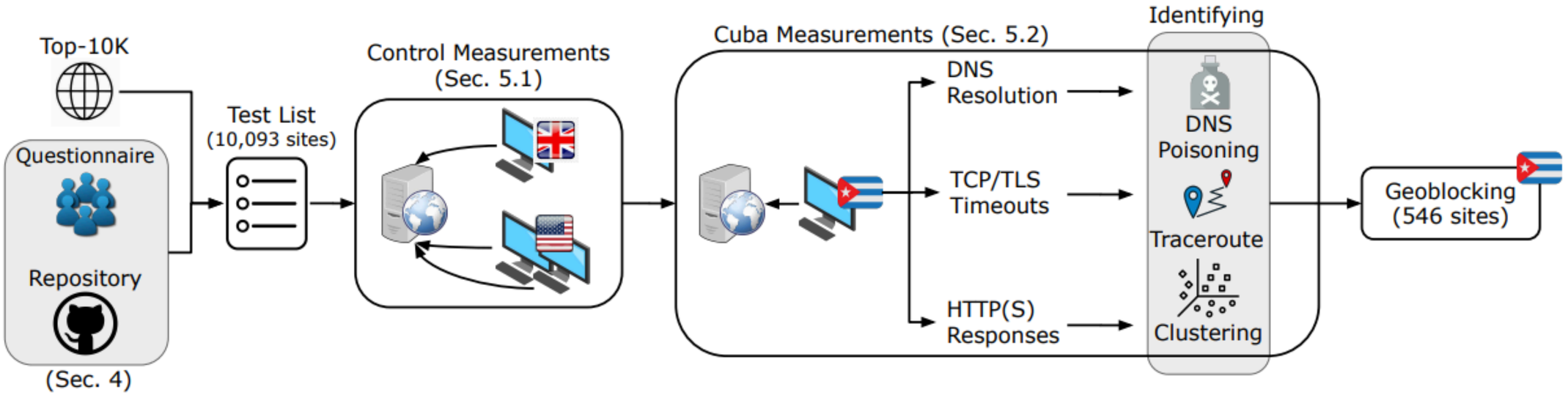
- This paper first seeks to understand the everyday experiences of Cuban users
- Users had a hard time disambiguating censorship from geoblocking. What's the difference?
  - Not all geoblocking is censorship; not all censorship is geoblocking!
- Interviews, in general, highlighted the risks and damages of geoblocking:

"In theory, these sanctions should only affect the government, but they affect all citizens directly, greatly limiting our chances of prospering and being able to develop like any other citizen in the world." (R8)

Measurements

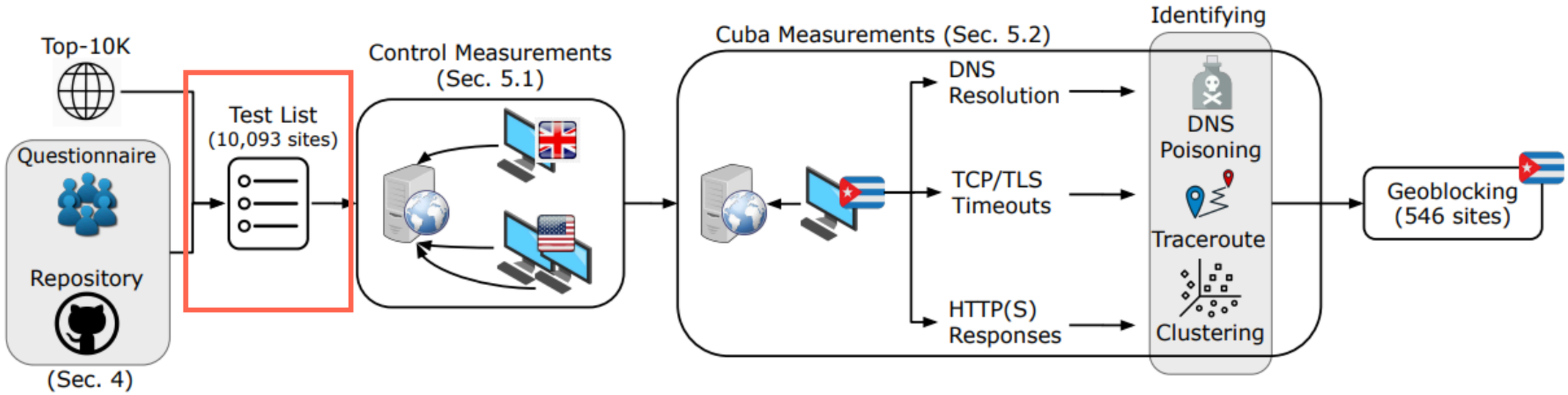
- Authors conducted what's called a *differential study* between a **control measurements** and **test measurements.**
  - What were the control measurements?
  - What is the test measurement?
- What issues did the authors run into with conducting test measurements?
- What *network services* were the authors investigating in this paper?

Pipeline


Top-10K
Questionnaire
Repository
(Sec. 4)
Test List
(10,093 sites)
Control Measurements
(Sec. 5.1)
Cuba Measurements (Sec. 5.2)
DNS
Resolution
TCP/TLS
Timeouts
HTTP(S)
Responses
Identifying
DNS
Poisoning
Traceroute
Clustering
Geoblocking
(546 sites)

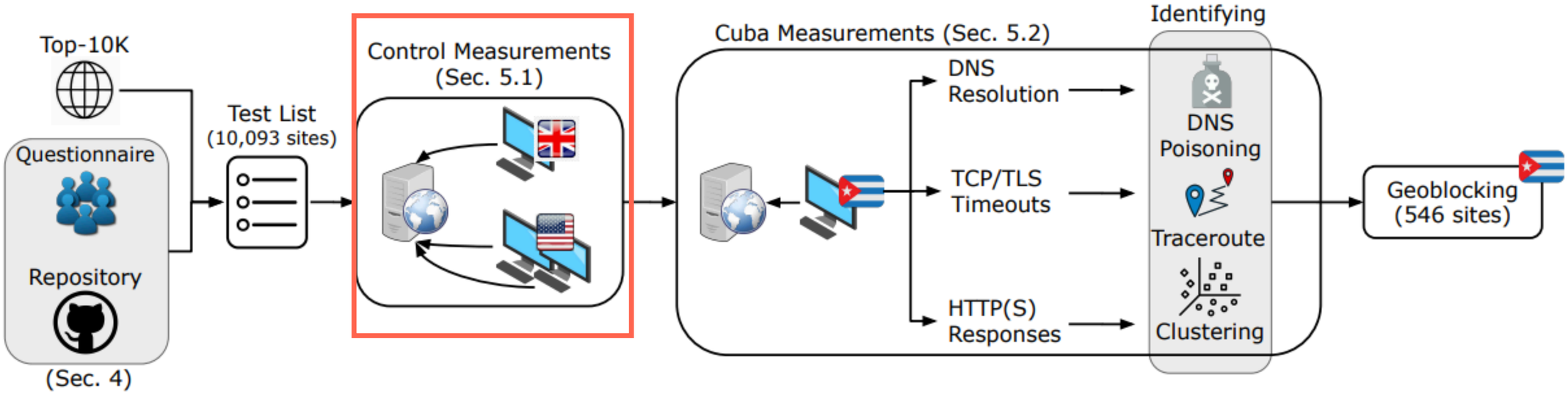
Pipeline

Where did the authors get the domain list from?



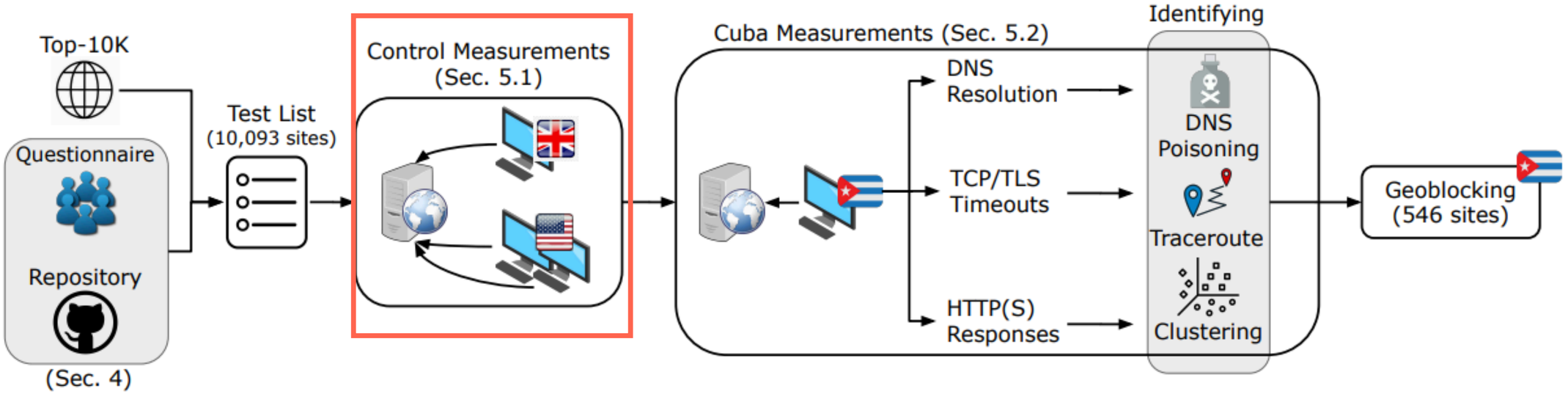
Pipeline

Control measurements performed *three* DNS lookups per domain. Why?



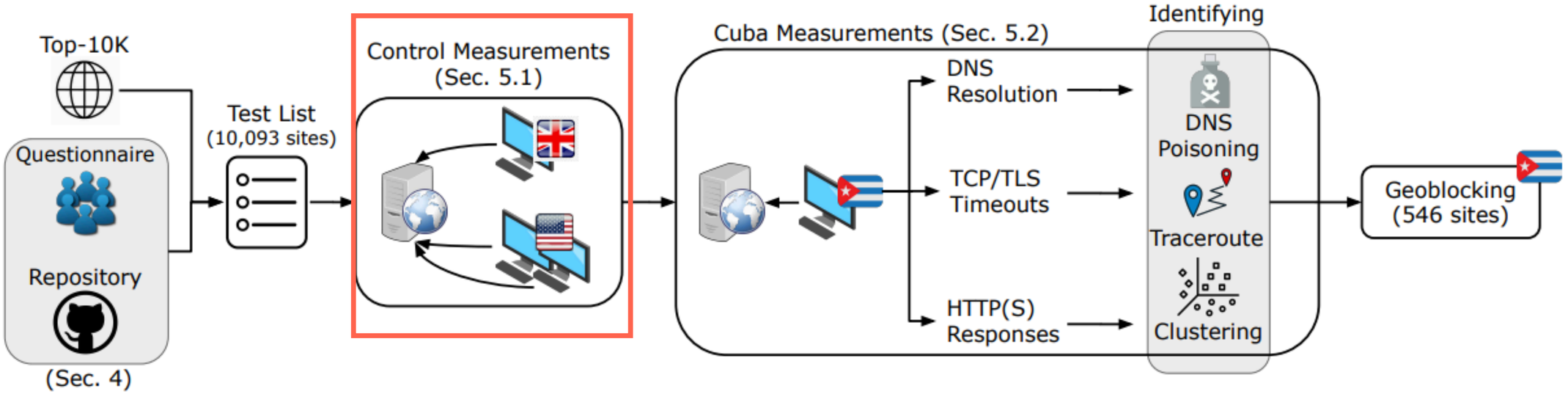
Pipeline

Control measurements attempted *three* TCP connections per domain. Why?



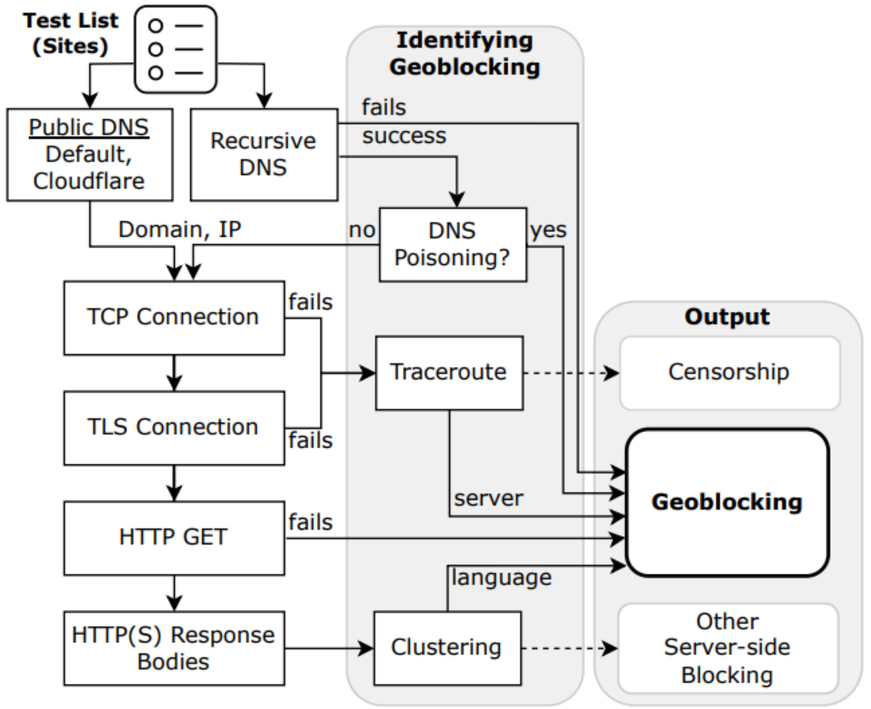
Pipeline

Authors found only ~6400 domains passed control TCP. Why?



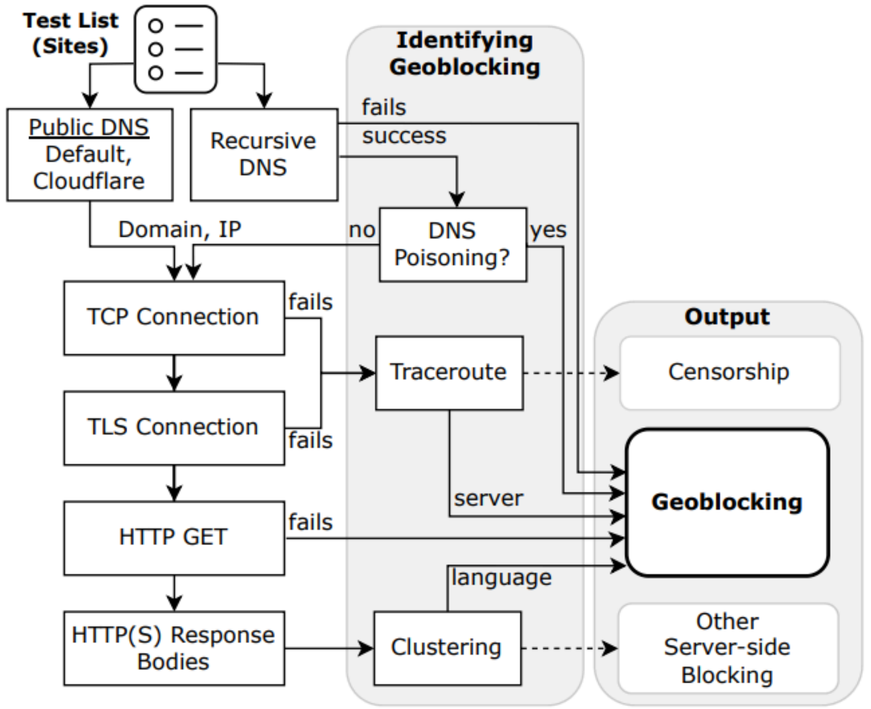
Geoblocking in Cuba

- First step in analysis is DNS resolution — looking for an *IP match* to the control domains.
  - What is DNS poisoning, and how did the authors analyze for it?
- Authors use traceroutes to identify TCP manipulation. What is traceroute, and why is it useful here?
-

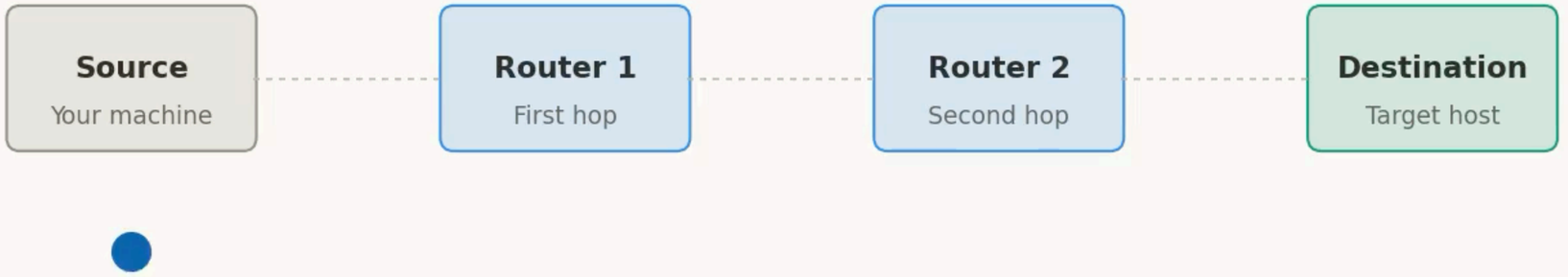


Geoblocking in Cuba

- First step in analysis is DNS resolution — looking for an *IP match* to the control domains.
  - What is DNS poisoning, and how did the authors analyze for it?
- Authors use traceroutes to identify TCP manipulation. What is traceroute, and why is it useful here?
  - Shows you the *path* your packet takes through the network, identifies any *drops*
  - How does traceroute work?



Traceroute: discovering the path hop by hop



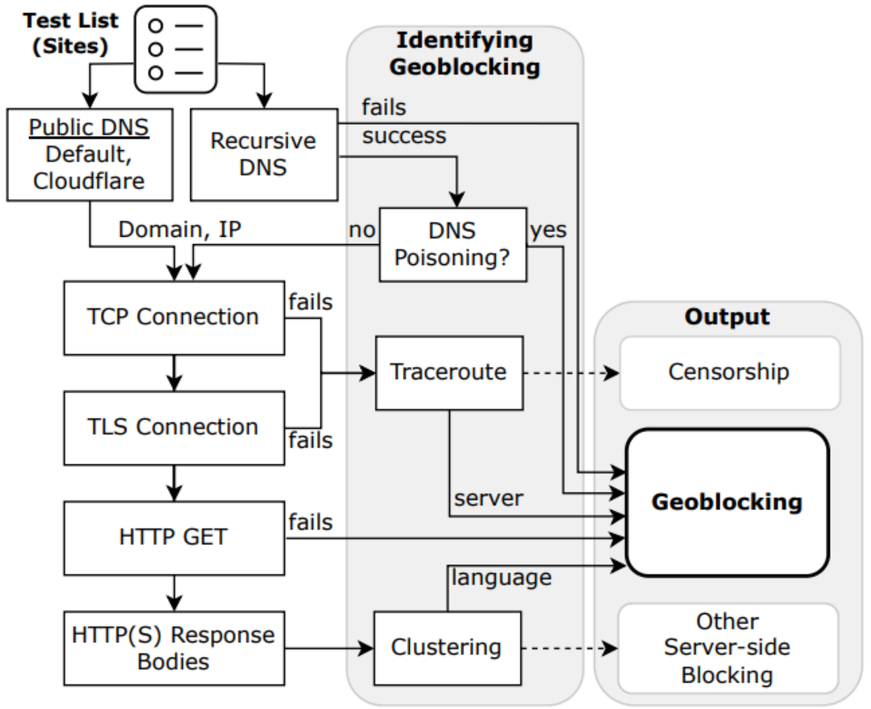
Probe 1 — TTL=1

Probe traveling, TTL counting down…

Outbound probe ICMP Time Exceeded Destination reply

Geoblocking in Cuba

- The authors next try to identify if the HTTP response received is a **blockpage.** What is a blockpage?
- How do the authors try to identify blockpages?
  - How do the authors know the blockpage is coming from the server (and not some other device on path?)



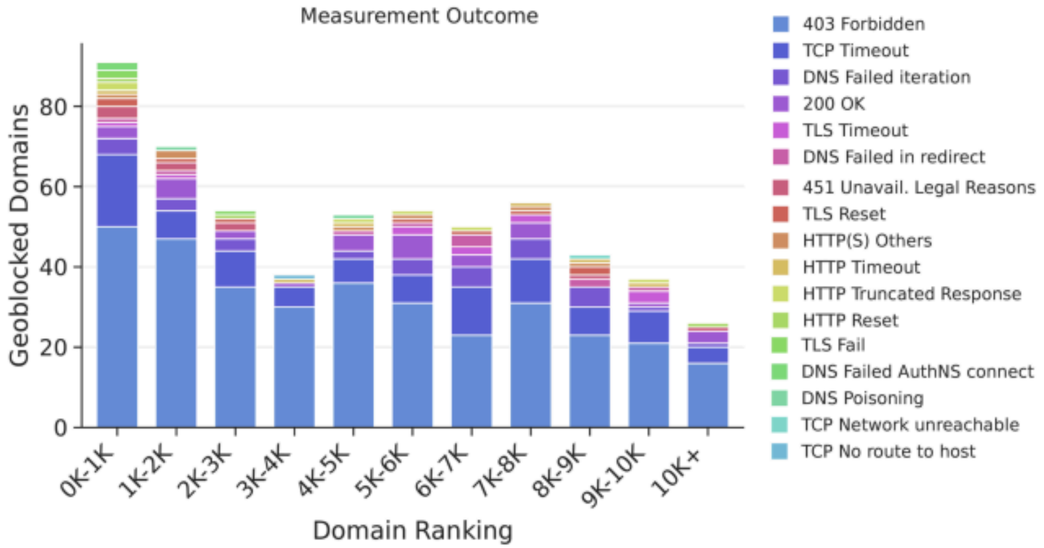
What they found

- Geoblocking was happening at almost every layer
  - Most commonly, at the HTTP response layer
- What was the most common mechanism for geoblocking a TCP connection?

Network Stages (5)	Measurement Outcomes (17)	# Domains
DNS Fails 37/546 (**6.8%**)	Failed iteration	33
	Failed AuthNS connect	2
	Manipulated IP	2
TCP Fails 97/546 (**17.7%**)	Timeout	95
	Network unreachable	1
	No route to host	1
TLS Fails 23/546 (**4.2%**)	Timeout	11
	Reset	9
	Fail	3
HTTP Fails 24/546 (**4.4%**)	Timeout*	6
	Reset*	3
	Truncated Response	6
	DNS fail in redirect*	10
HTTP Responses 395/546 (**72.3%**)	403 Forbidden	347
	451 Unavailable for Legal Reasons	9
	200 OK	32
	Others *e.g.* 404, 503	7
Geoblocked Domains		**546**

What they found

- Geoblocking is *most prevalent* in the Tranco 2K — but diverse mechanisms of geoblocking across all ranks
  - att.com, ebay.com, exacttarget.com, trello.com… famous and popular American services blocked in Cuba
- **88% of geoblocked domains do not give notice of *why* blocking is occurring**



Discussion

- Is geoblocking a harm worth paying attention to? Why or why not?
  - What can be done about geoblocking?
- What did this paper remind you about *trust* on the Internet?
  - Who are the entities you have to trust when browsing the Internet?
- What did you think about this paper? Yay? Nay?

Next time…

- **CRIME!**
  - Namely, measuring crime… hooray
- Keep working on your projects!