Informing Protocol Design Through Crowdsourcing

Anna Maria Mandalari
amandala@it.uc3m.es

Marcelo Bagnulo
marcelo@it.uc3m.es

Andra Lutu
andra@simula.no

Internet Innovation

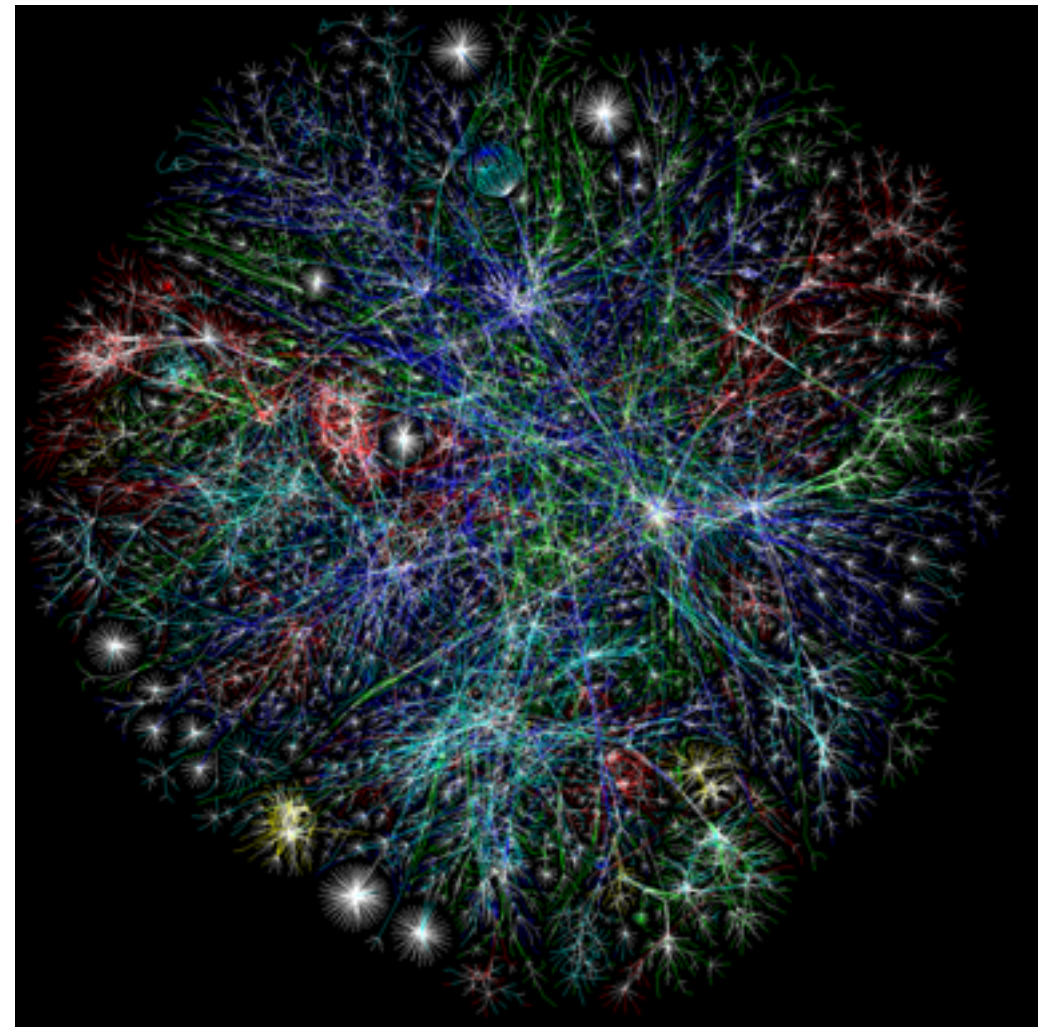The Internet has successfully enabled multiple waves of innovation:

➢ Mobility

➢ Heterogeneity of devices

➢ Video Communication

➢ VoIP

➢ .....

Internet Innovation

The Internet changes dramatically in terms of number and types of its nodes and running applications

Is the Internet Ossified?

Today, many aspects appear to be **"set in stone"**

Criticism: Middleboxes might filter traffic that does not conform to expected behaviors.

Handley, M. (2006). Why the Internet only just works. BT Technology Journal, 24(3), 119-129.

Is the Internet Ossified?

How will Internet react to a new protocol?

Understand the interaction of the new solutions with the middleboxes active along the path.

How to measure a thousand end-users?

- Have a friend in Google(or any other large Internet players)

or

- Get your code to run on a thousand users' machines through another delivery channel

Existing Large-Scale measurements platforms

Existing Large-Scale measurements platforms

- The limited and often special position of testbed nodes

- No possibility to deploy your own test

- Fixed line only

- Access to the results

Crowdsourcing platform

Perform large-scale Internet measurement campaigns

Crowdsourcing platform

Internet Connection Survey

☐ Campaign is finished [restart]   ▤ Submitted tasks   ☒ Results in CSV

Campaign/job ID	3b4ab5ce5e8f	Speed **96** [1-Slow 1000-Fast]	Verify+Rate Verify No Verify/Rate
Work done	**250**/**250** Add positions	You have **2** days to rate tasks	Auto-rating: Verify+Rate Satisfied
Workers will earn	**$0.25**		Folder **DEFAULT** → To ARCHIVE
Takes less than	**9** minutes to finish		
Targeted Countries	[International] -Macedonia -Indonesia -Lithuania -Bangladesh -Egypt -Morocco -Poland -Canada -Australia -Vietnam		

Category: **Surveys** → Up to 10 questions

[?] What is expected from Workers?

1. Go to: http://ametrics2.it.uc3m.es/form.php?campaign={{CAMP_ID}}&worker={{MW_ID}};
2. Answer the questions, selecting a value and then press Submit

3. Once completed, a code will be displayed on your screen, this will be your proof for Microworkers

Note:
DON'T CLOSE the browser until the code is generated.

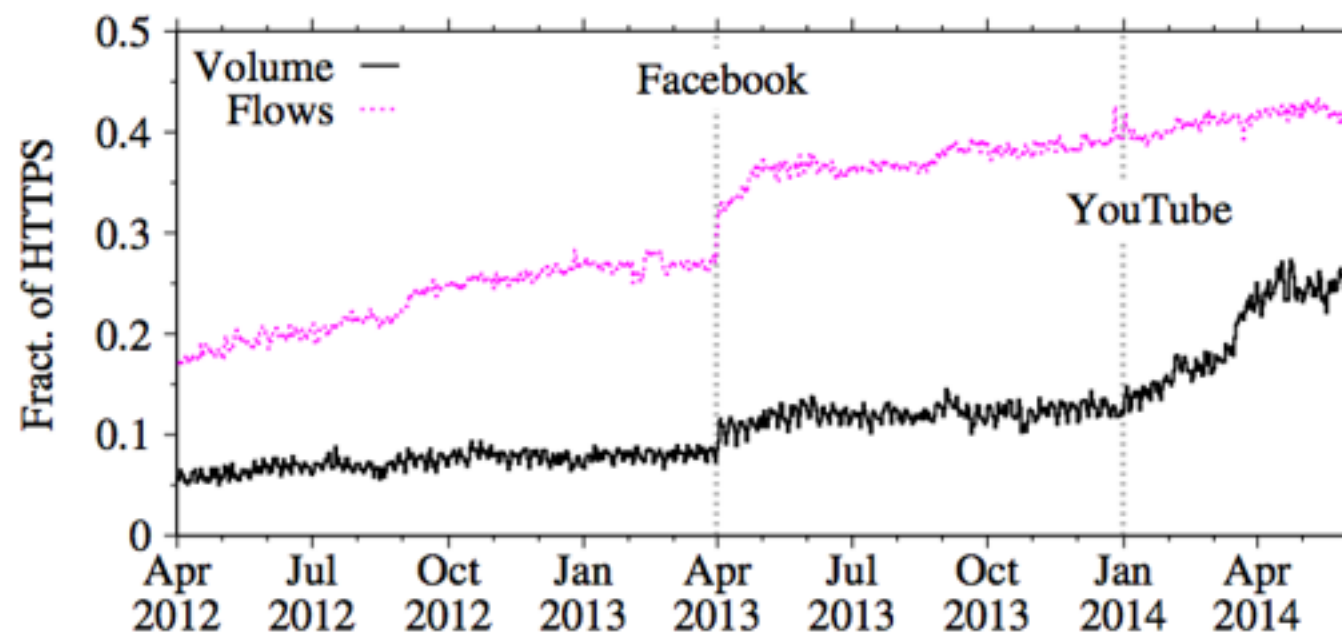**[!] Required proof that task was finished?**

1. The code generated once you completed the survey

Is the Internet Ossified? Case Studies

- The case of the pervasive encryption

- TCP Fast Open

- HTTP/2

The case of pervasive encryption

Many popular applications (e.g., web, Youtube video streaming) have migrated from HTTP to the HTTPS protocol

Challenge: Provide encryption by default for all Internet communications

Naylor, David, et al. "The Cost of the S in HTTPS." Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies. ACM, 2014.

The case of pervasive encryption

Understand the feasibility of pervasive encryption in the Internet.
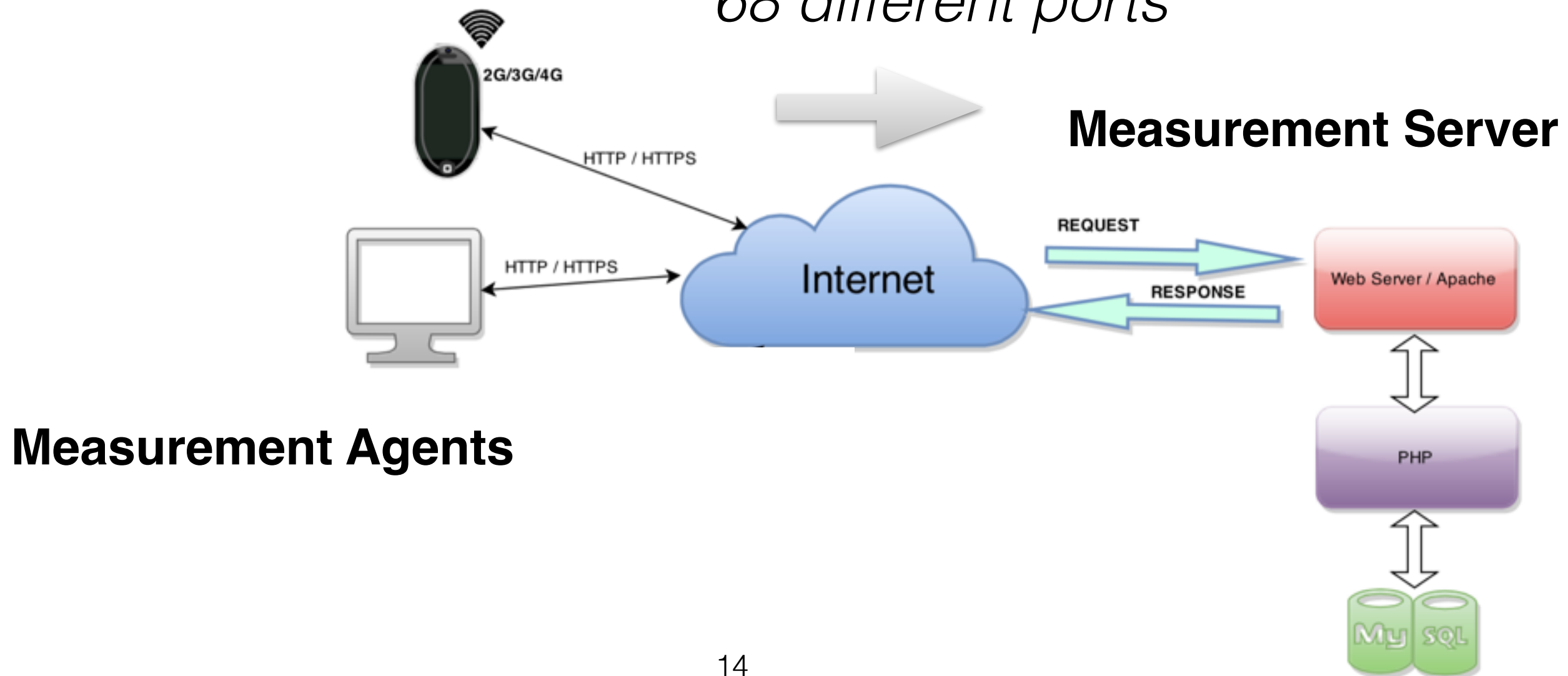
Understand the interaction of middleboxes with the TLS across the different TCP ports that currently use plain text protocols.

Experimental setup

Establish both **HTTP** and **TLS** connections to **68 different ports:**

- 10 well-known ports;
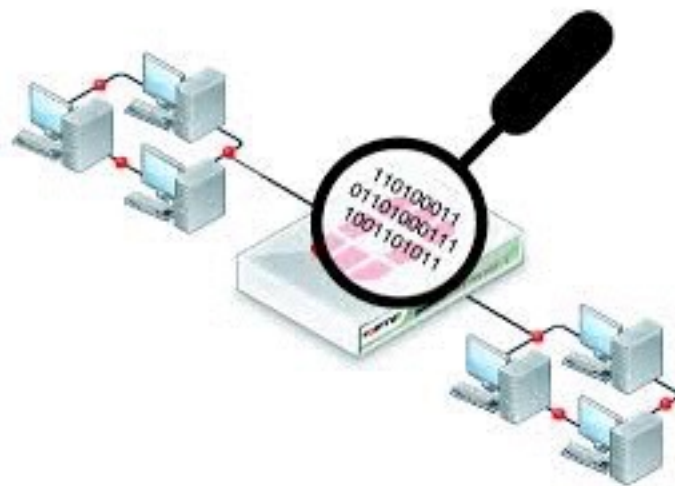- 56 registered ports;
- 2 ephemeral ports.

TLS connections over 68 different ports

Measurement Server

Measurement Agents

2G/3G/4G

HTTP / HTTPS

HTTP / HTTPS

Internet

REQUEST

RESPONSE

Web Server / Apache

PHP

My SQL

Experimental setup: Measurement Server

LAMP STACK

- LAMP model (Linux, Apache Server, MySQL relational database management system, PHP);

- Packets capture.

Experimental setup: Measurement Agent Common Procedure

Limit of crowdsourcing platform: some information may not be available through the platform

- Users connect using a HTTP connection in port 80 to a webpage we provide

Experimental setup: Measurement Agent Common Procedure

- Users connected from Fixed line indicate the place from where they are connecting (Home, Hot Spot, University or or other institution, Company)
- Users connected from Mobile line indicate the technology they are using (2G, 3G, 4G)

Answer to the question, selecting a value and then press Submit.

What kind of Wi-Fi connection are you using?

○ **Public Hot Spot** (if you are connecting from an Internet connection open to the public, such as a coffee bar)

○ **Home** (if you are connecting from home)

○ **Company** (if you are connecting from an office)

○ **University or other institution** (if you are connecting from an University or another institution)

Submit

17

Experimental setup: Measurement Agent Common Procedure

- Metadata collection

```
User-Agent: Mozilla/5.0 (X11; Linux i686 on x86_64; rv:10.0.2) Gecko/20100101 Firefox/10.0.2
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip, deflate
Cookie: .ASPXANONYMOUS=BLAH......;
WRUID=1243657642
DNT: 1
Connection: keep-alive

HTTP/1.1 200 OK
Date: Mon, 23 Apr 2012 20:55:58 GMT
Server: Microsoft-IIS/6.0
X-Powered-By: PleskWin, ASP.NET
X-Powered-By-Plesk: PleskWin
X-AspNet-Version: 2.0.50727
Set-Cookie: ViewMobile=False; path=/; HttpOnly
Set-Cookie: language=en-US; path=/; HttpOnly
Cache-Control: private
Content-Type: text/html; charset=utf-8
Content-Length: 88701
```
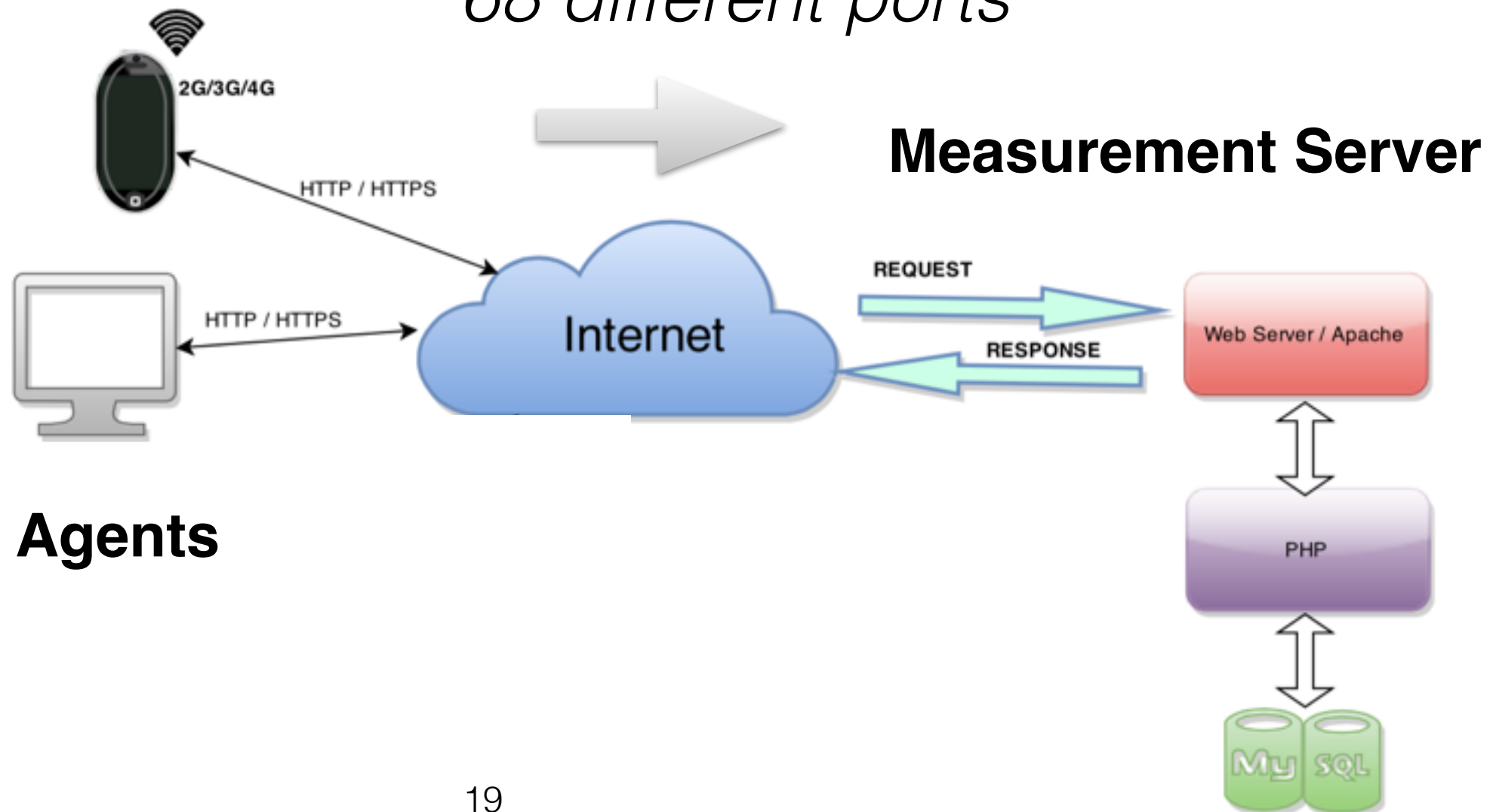
Experimental setup: Measurement Agent Common Procedure

- In the background, HTTP and HTTPS connections are performed from the measurement devices to our servers in all the 68 ports
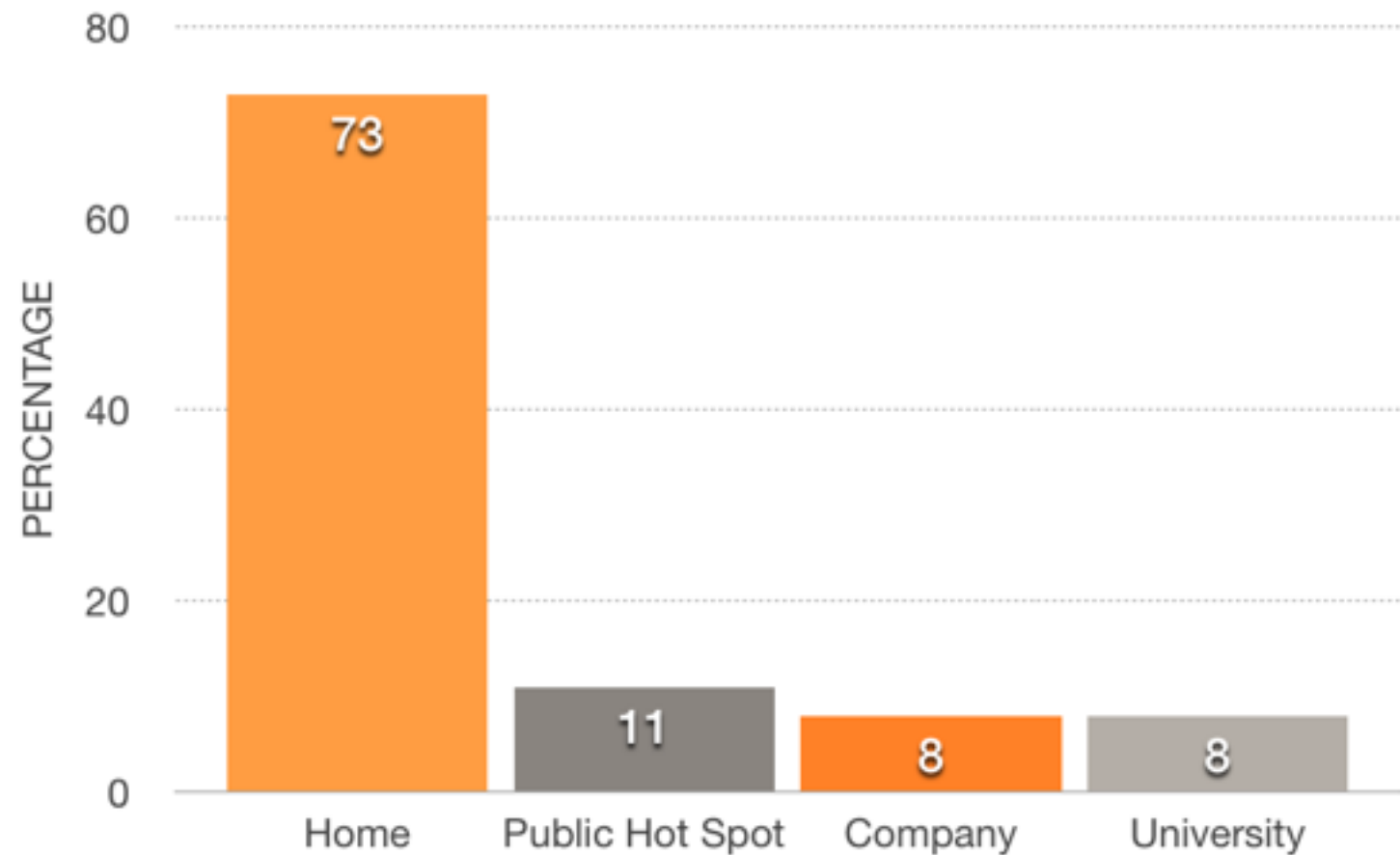
TLS connections over 68 different ports

Measurement Server

Measurement Agents

Data Set

FIXED LINE:

- 1,165 workers;

- 53 different countries;

- 286 ASes.

Data Set

MOBILE:

- 956 workers;

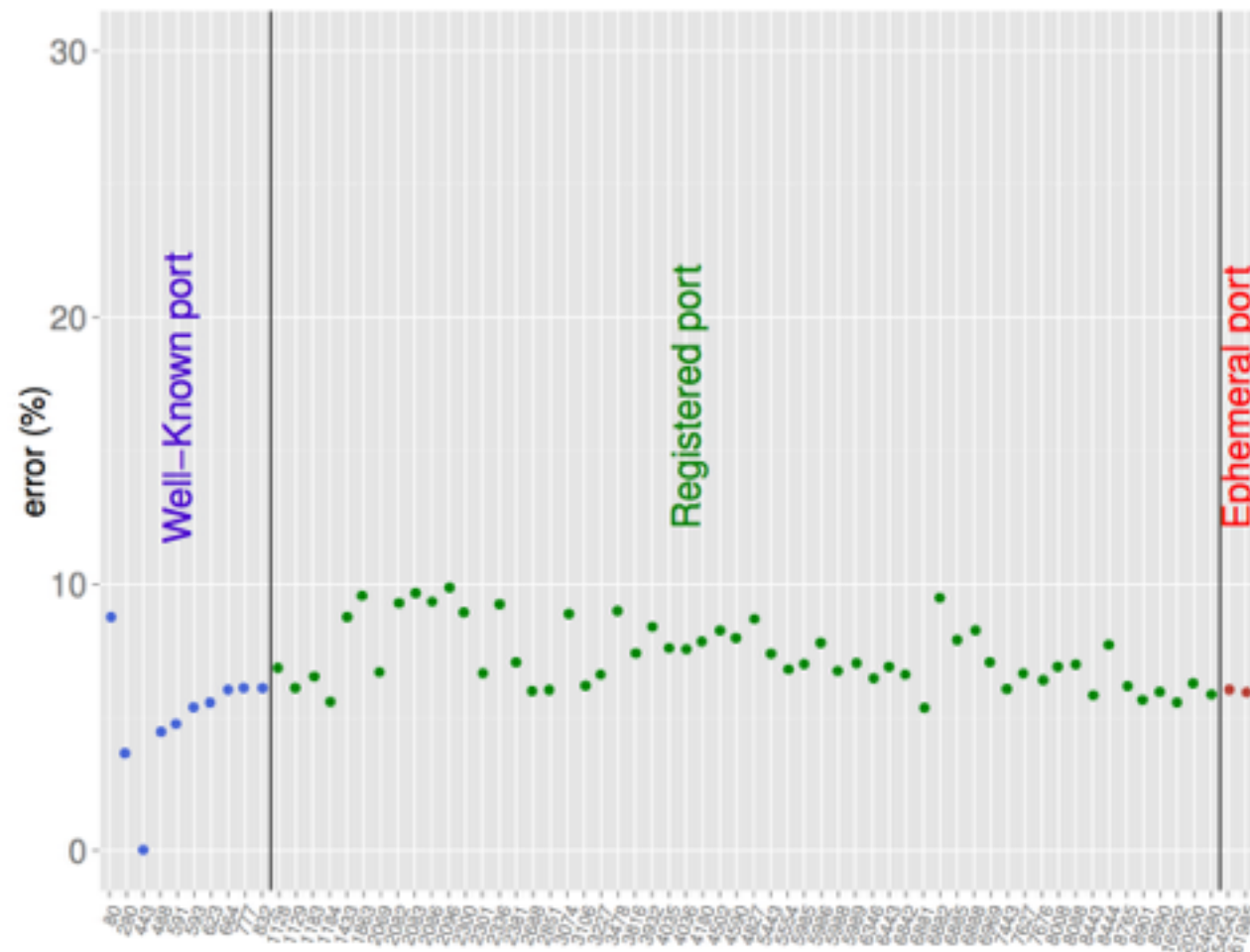- 45 different countries;

- 183 ASes.

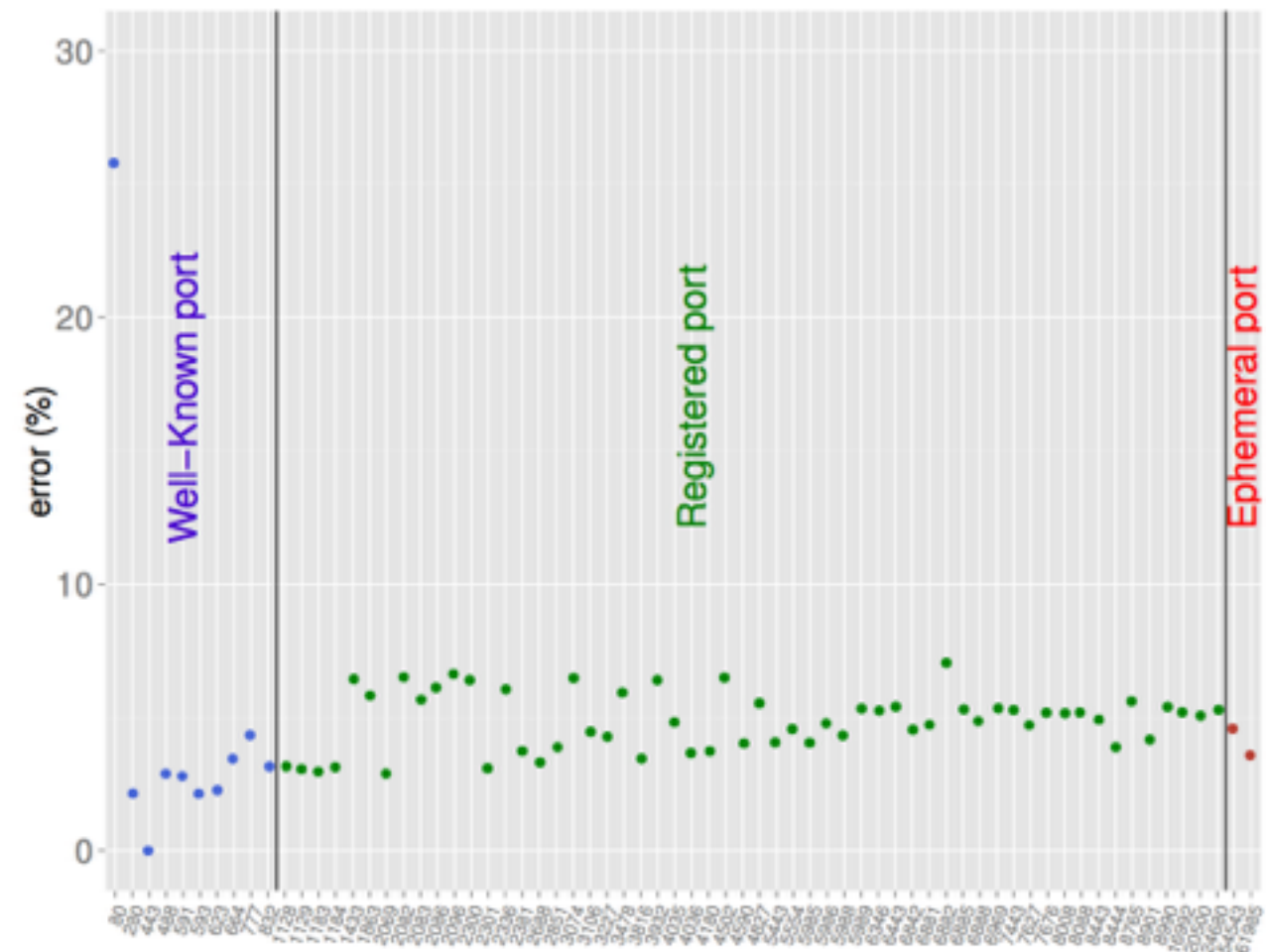Total of 114,228 connections
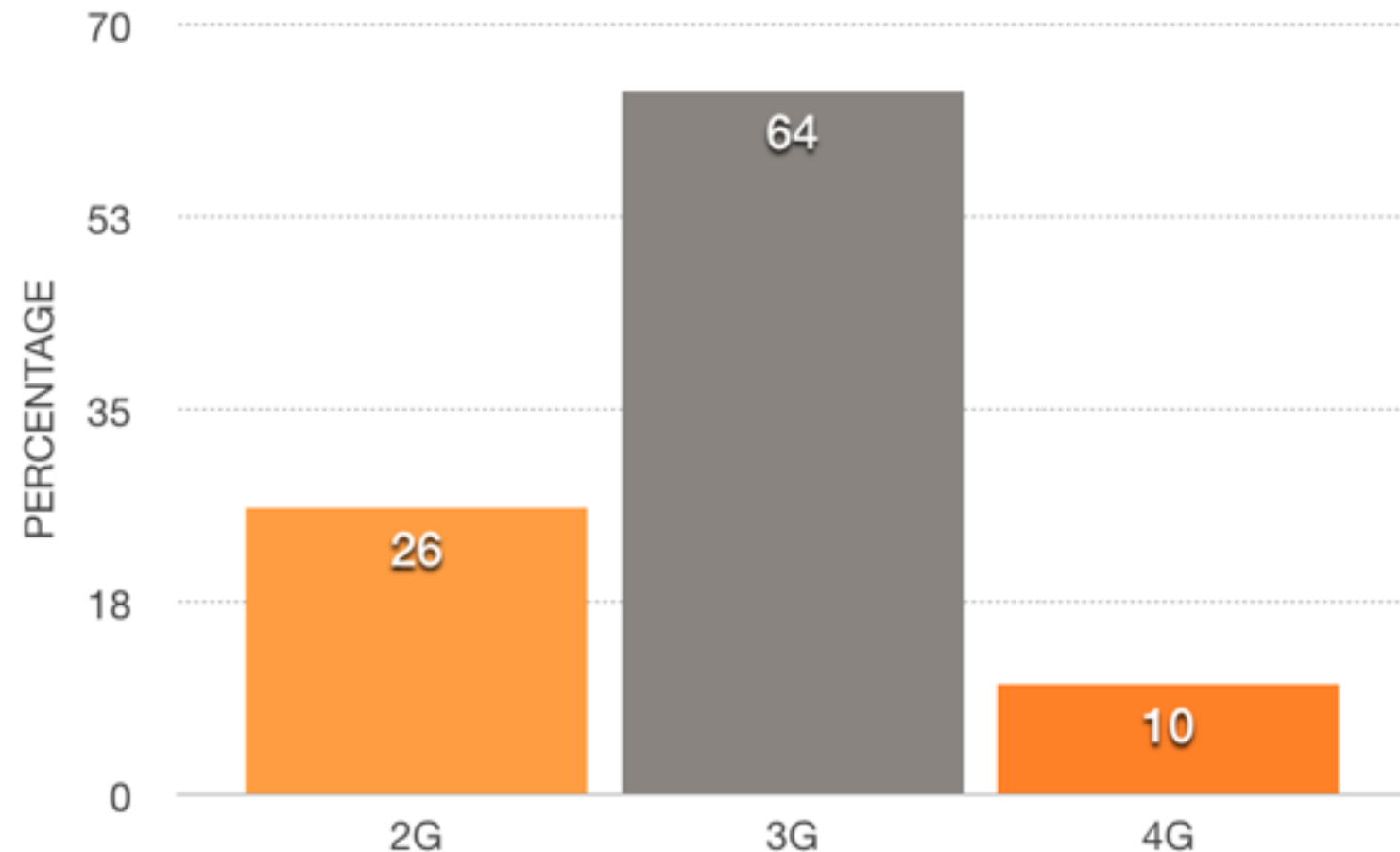
Check workers
User Agent

The data set is freely available on http://it.uc3m.es/amandala/dataset.php

Aggregated results

$$ERROR = (success\ [HTTP] - success\ [TLS])\ /\ success\ [HTTP]$$
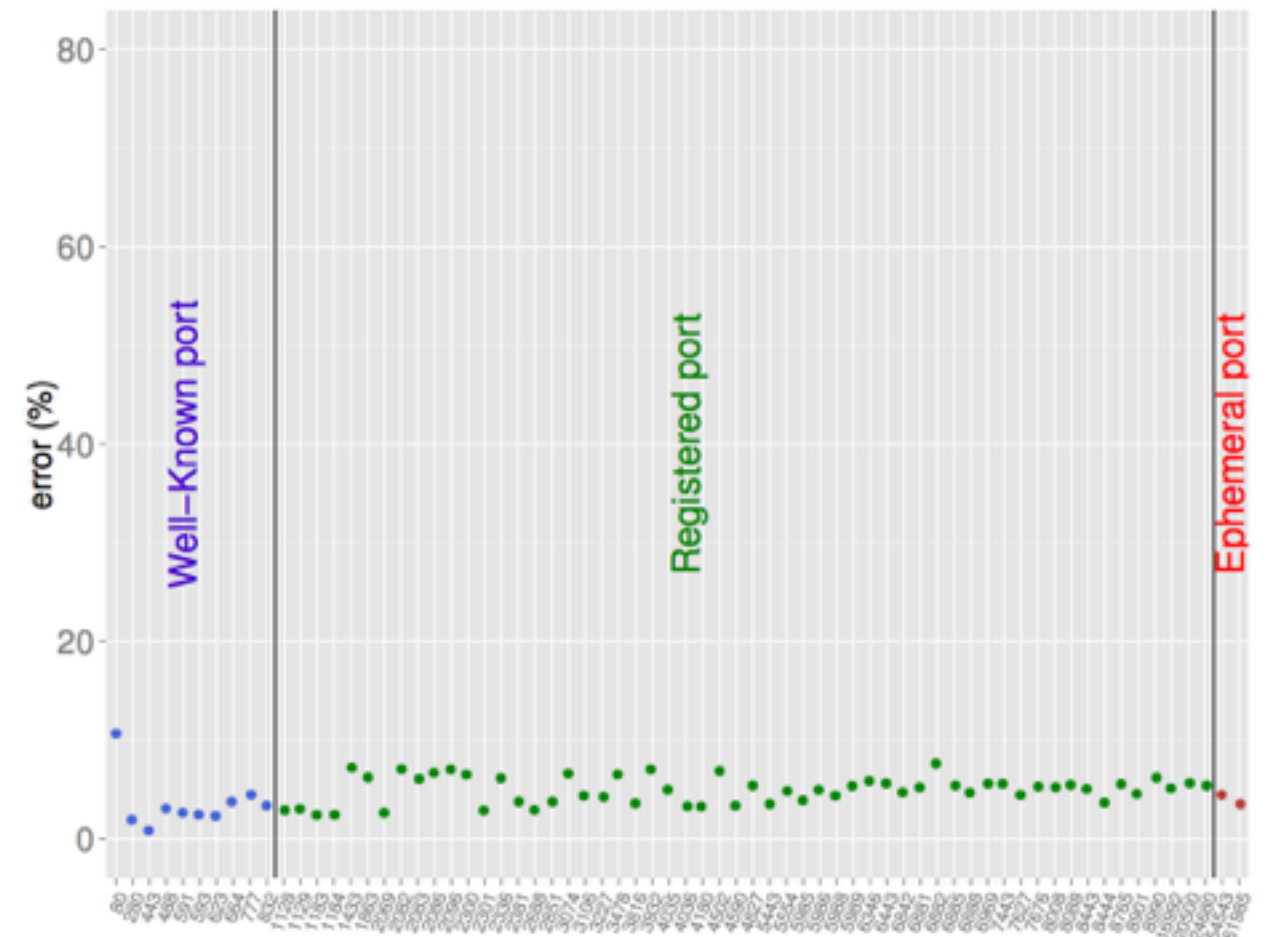

a) Fixed line

b) Mobile network

25% of the users are not able to perform a TLS connection over port 80 in mobile network.

Proxies

$$ERROR = (success\ [HTTP] - success\ [TLS])\ /\ success\ [HTTP]$$

a) Mobile proxy

b) Mobile non-proxy

70% of the users that use a proxy are not able to perform a TLS connection over port 80 in mobile network.
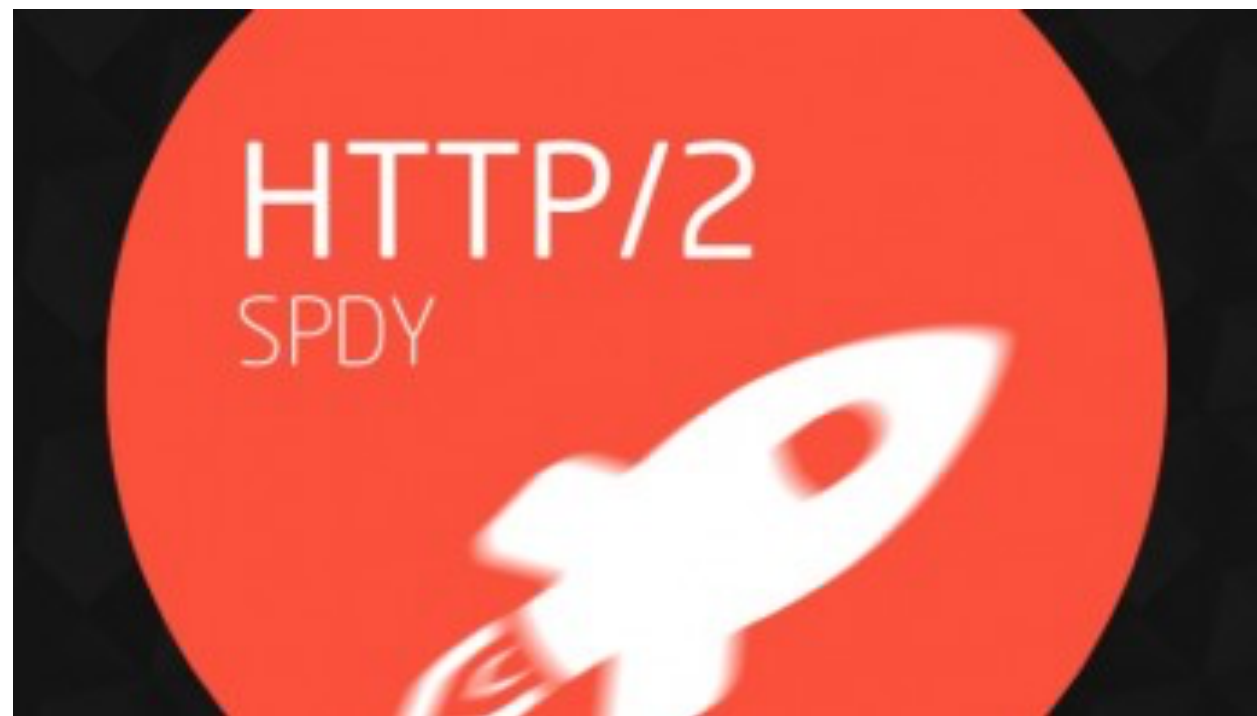
Packets analysis

Analysis	Fixed Line		Mobile	
	SYN(%)	NO SYN(%)	SYN(%)	NO SYN(%)
All	96,8	3,2	36	64
Port 80	88,3	11,7	27,7	72,3
Proxy			22,2	77,8
Non-proxy			12,7	87,3
Proxy (80)			9,6	90,4
Non-proxy (80)			36,4	63,6

When users use a proxy, 90% of the SYN packets are missing

Case study: HTTP/2

From SPDY to HTTP/2

Case study: HTTP/2

HTTP/2:

- binary, instead of textual

- fully multiplexed

- connection for parallelism

- header compression

- pushing and priority

Case study: HTTP/2. Middleboxes compatibility

HTTP/2: To encrypt or not to encrypt?

Some implementations have stated that they will only support HTTP/2 when it is used over an encrypted connection (Google, Twitter…), and currently no browser supports HTTP/2 unencrypted!

Implementation:
- HTTP/2 encrypted using ALPN (H2)
- HTTP/2 in clear using HTTP Upgrade mechanism (H2C)

Again: Middleboxes might filter traffic that does not conform to expected behavior.
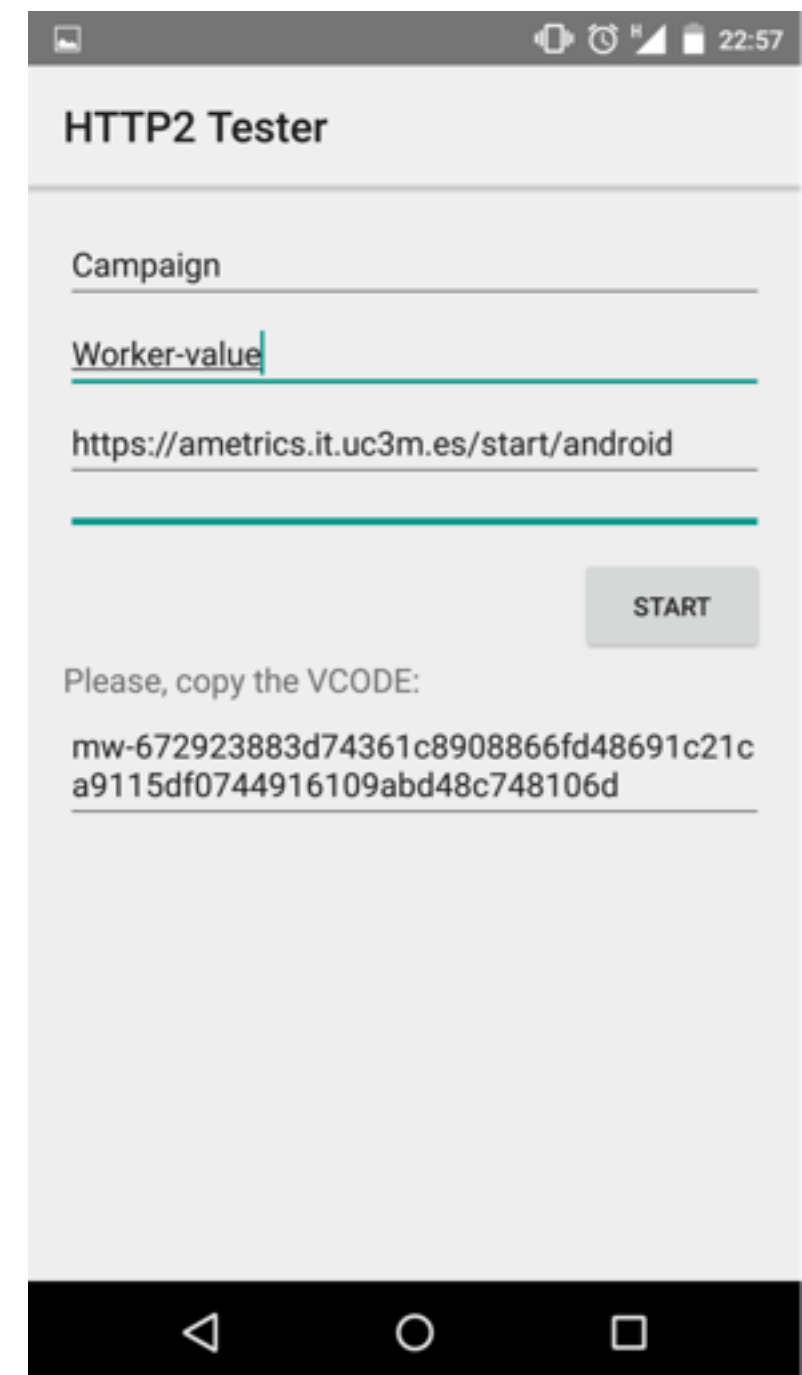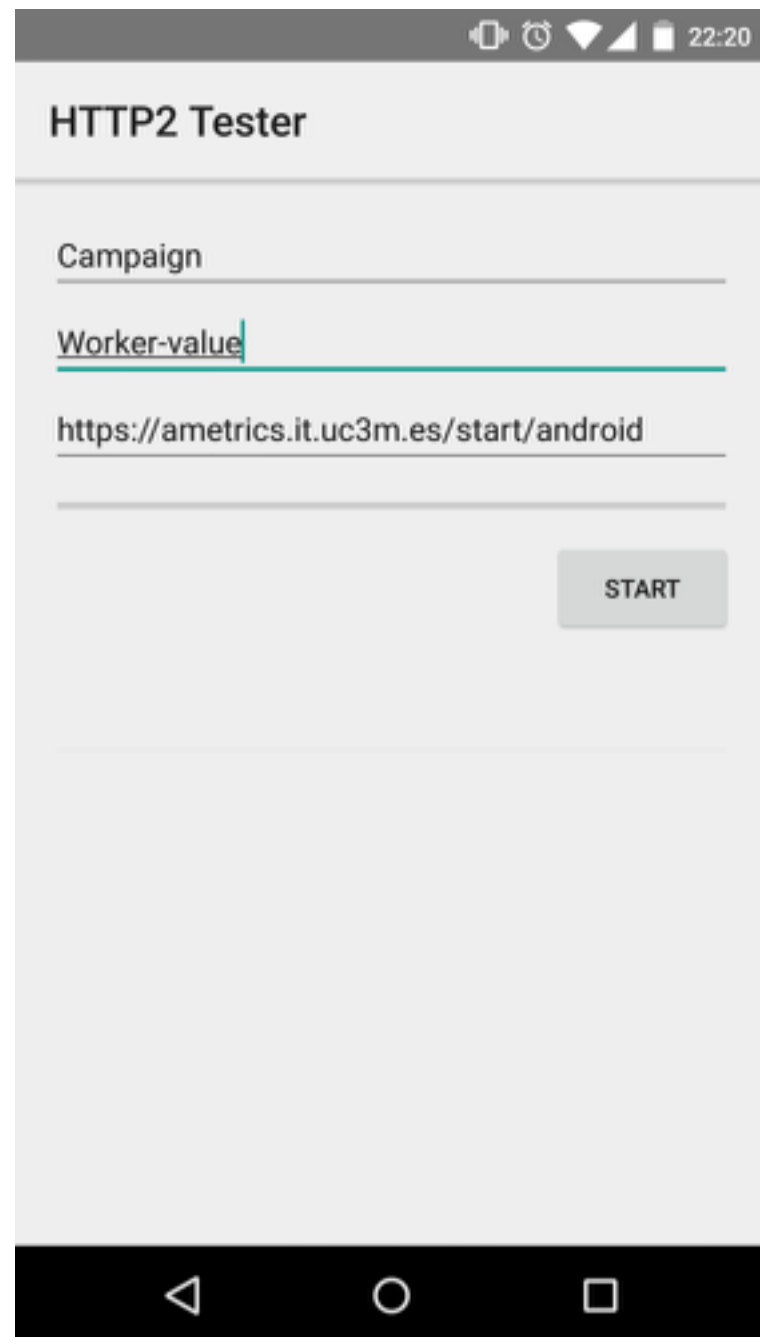
Experimental setup

- **H2**: Browser campaign

- **H2 without ALPN**: Android app campaign

- **H2C**: Android app campaign

- **H2C without Upgrade**: Android app campaign

Check workers are not using WiFi

Experimental setup

Establish both **HTTP** and **HTTP/2** connections to **68 different ports** using an Android app

Experimental setup

Android App Testing: Download + Install

☐ Campaign is running [pause] [stop] [clone and edit]   ☰ Submitted tasks   ☒ Results in CSV

Campaign/job ID	cc42bd7f0832	Speed **100** [1-Slow 1000-Fast]	**Verify+Rate** **Verify** No Verify/Rate
Work done	**53/⁶⁰** Add positions	You have **7** days to rate tasks	Auto-rating: Verify+Rate Satisfied
Workers will earn	**$0.40**		Folder **DEFAULT** → To ARCHIVE
Takes less than	**12** minutes to finish		
Targeted Countries	[International] -Bangladesh -Indonesia -India -Nepal		
	retarget		

Category: **Mobile Applications (iPhone & Android)** → Download + Install

? What is expected from Workers?

1. Go to https://ametrics.it.uc3m.es/help/android/{{CAMP_ID}}/{{MW_ID}}
2. Follow the instruction to download and execute the app

3. Once completed, a code will be displayed on your screen, this will be your proof for Microworkers. Copy and paste this code in the Microworkers proof box

! Required proof that task was finished?
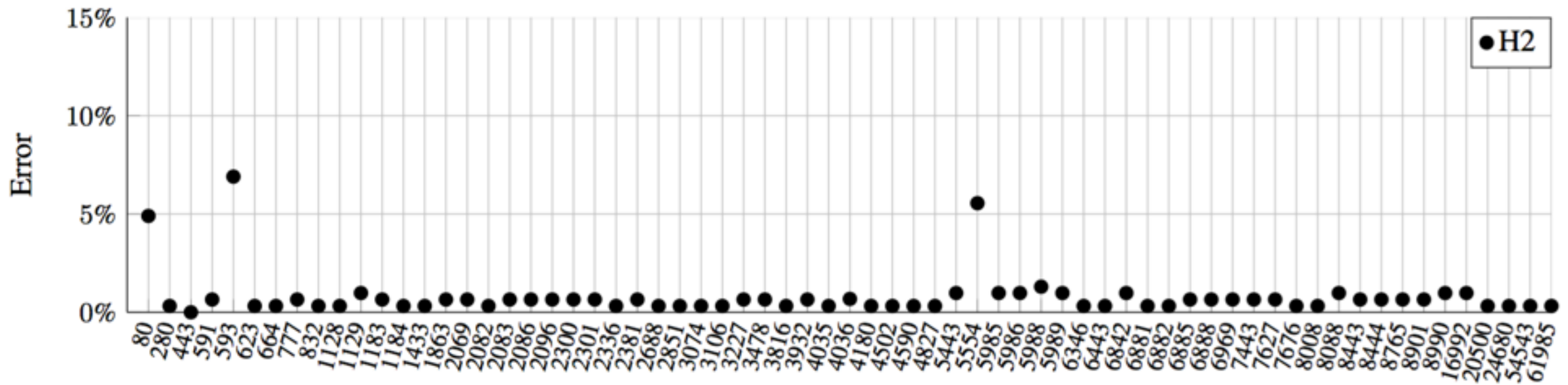
1. VCode generated once you completed the test

30

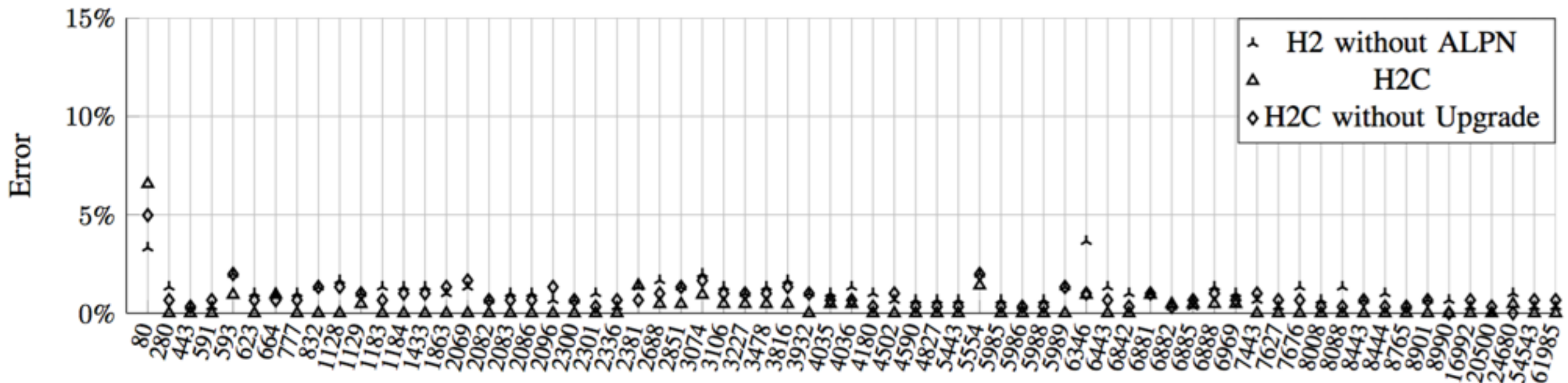# Data Set

- **628 workers;**

- **38 different countries;**

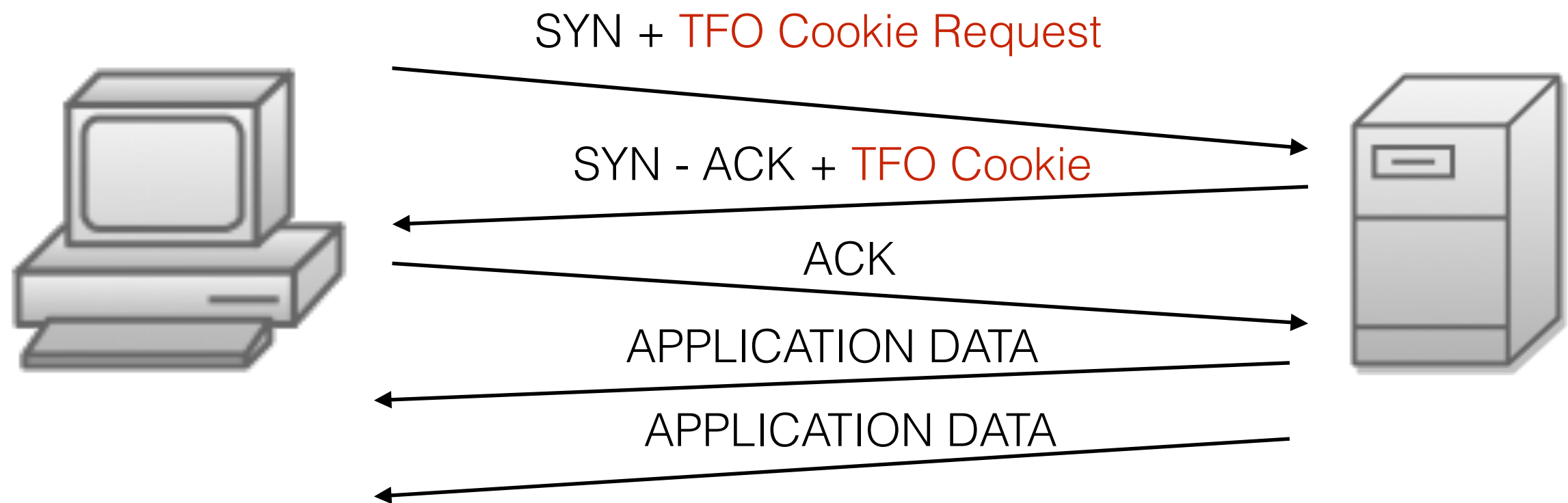- **40 ISPs.**

The data set is freely available on http://www.it.uc3m.es/amandala/http2.html

Results

- Browser (Fixed line)

- Android app (Mobile network)

Case study: TCP Fast Open (TFO)

SYN + TFO Cookie Request

SYN - ACK + TFO Cookie

ACK

APPLICATION DATA

APPLICATION DATA

S. Radhakrishnan, Y. Cheng, J. Chu, A. Jain, and B. Raghavan. TCP Fast Open. In Proceedings of the Seventh Conference on emerging Networking EXperiments and Technologies, page 21. ACM, 2011.

TCP Fast Open (TFO)

SYN + TFO Cookie + APPLICATION DATA

SYN - ACK (SYN & DATA)

APPLICATION DATA

ACK

S. Radhakrishnan, Y. Cheng, J. Chu, A. Jain, and B. Raghavan. TCP Fast Open. In Proceedings of the Seventh Conference on emerging Networking EXperiments and Technologies, page 21. ACM, 2011.

Experimental setup

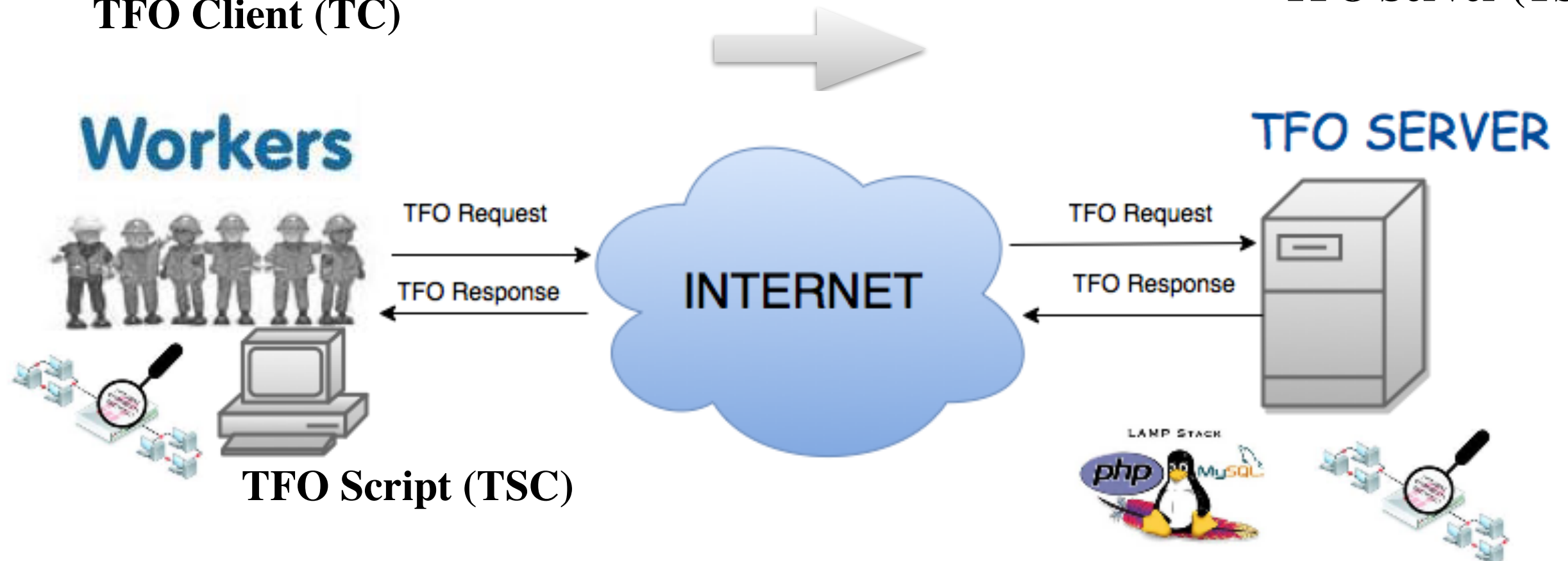Establish **TFO** connections to **68 different ports:**

- 10 well-known ports;
- 56 registered ports;
- 2 ephemeral ports.

TFOExplorer

*TFO connections over
68 different ports*

TFO Client (TC)

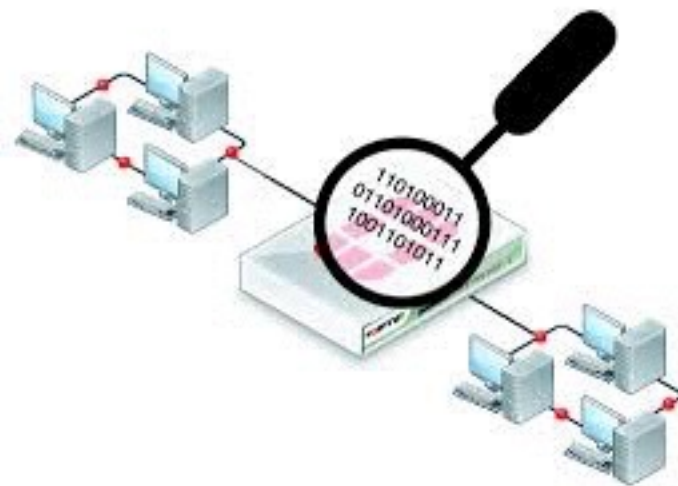**TFO Server (TS)**

TFO Script (TSC)

TFOExplorer: TFO Server

Needed Administrator Privileges

- NGINX

- LEMP model (Linux, NGinx Server, MySQL relational database management system, PHP);

- Packets capture.

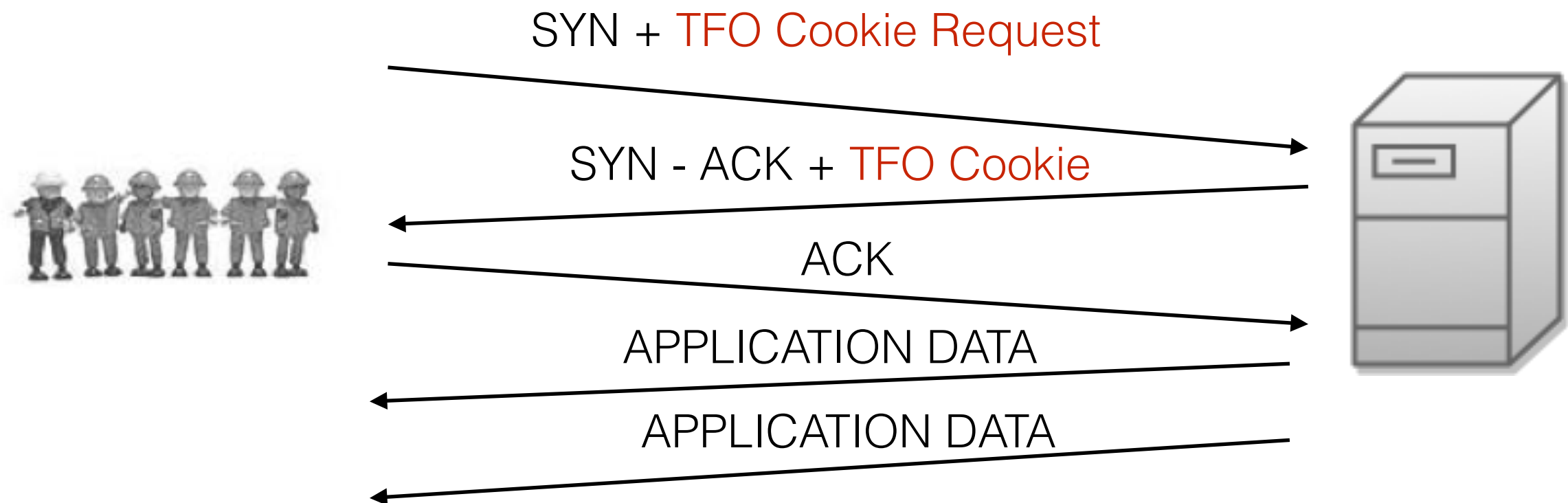*Check workers are not using a virtual machine*

Data Set

- **46 workers;**

- **18 different countries;**

- **22 ISPs.**

*The data set is freely available on http://www.it.uc3m.es/amandala/tfocampaign.html*

Results

- TC is able to perform a TFO connection (label Successful)

SYN + TFO Cookie Request

SYN - ACK + TFO Cookie

ACK

APPLICATION DATA

APPLICATION DATA

Results

- Middleboxes drop packets with unknown TCP options and we receive the SYN without option (label No option SYN);

SYN + TFO Cookie Request
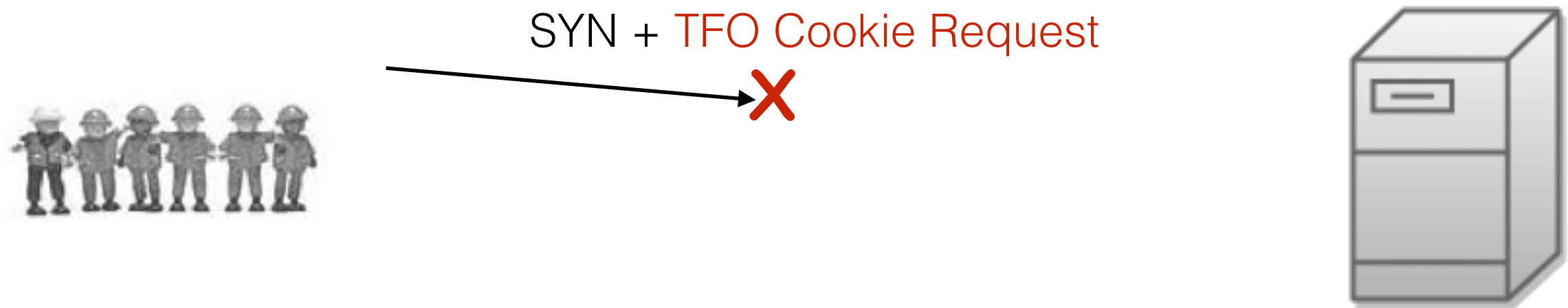
SYN

SYN - ACK

ACK

APPLICATION DATA

APPLICATION DATA
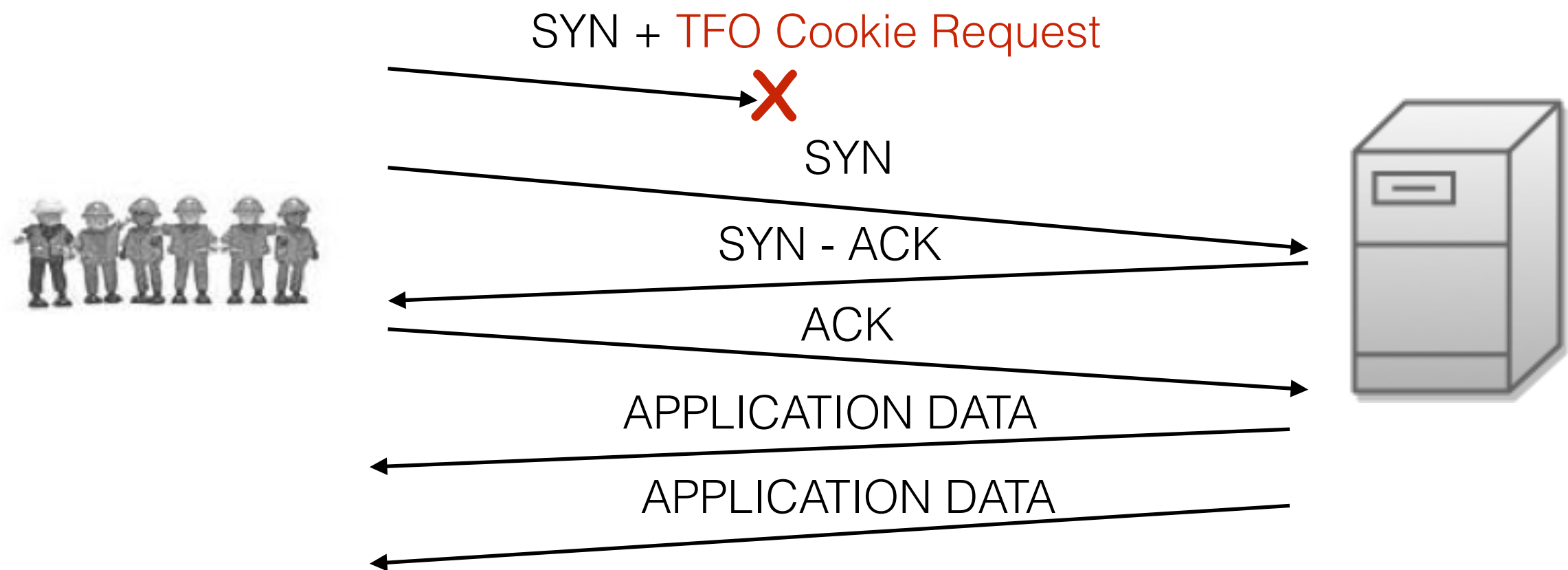
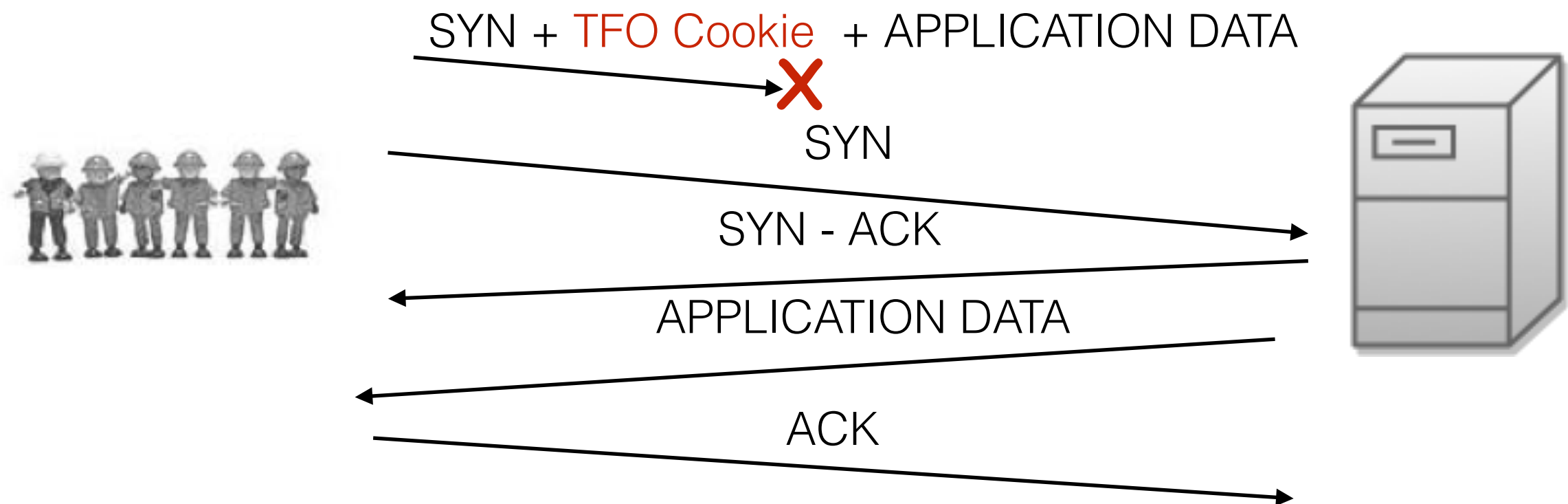# Results

- Middleboxes drop packets with unknown TCP options and we do not receive the SYN without option (label No option no SYN);

SYN + TFO Cookie Request

Results

- Middleboxes drop packets with data in the SYN packet and we receive the SYN without data (label No data SYN);

SYN + TFO Cookie + APPLICATION DATA

✗

SYN

SYN - ACK

APPLICATION DATA

ACK

Results: TFO CONNECTION

TFO behavior	Number of workers	Number of workers (%)
Successful	19	41,3
No option SYN	18	39,13
No option no SYN	0	0
No data SYN	8	17,39
No data no SYN	1	2,18

39% of the users are not able to perform a TFO connection over port 80 in mobile network.

Results: DATA IN THE SYN

TFO behavior	Number of workers	Number of workers (%)
Successful	23	50
No data SYN	23	50
No data no SYN	0	0

50% of the users are not able to perform a TFO connection over port 80 in mobile network.

Crowdsourcing Platform Guidelines

Type of test	Users	Time	Cost	Task Complexity
Browser	Thousands	Days	0,05-0,25 $	2/5
Android app	Hundreds	Days	0,40-1,00 $	3.5/5 (no rooted devices)
Linux app	Dozens	Week	0,80-1,00 $	5/5

Crowdsourcing Platform Guidelines

- Download code: Google play ✓
  Windows ✗   Linux ✗ ✗

- Best moment to setup a campaign: Weekend!

- Surveys: 60% of the workers lie

- Short times are better

Conclusion

- Overcome several of the limitations of the crowdsourcing platforms to perform network measurements

- It is probably feasible to roll out TLS protection for most ports except for port 80, assuming a low failure rate (5%)

- New protocols at application layer need to be encrypted, in particular in mobile networks!

- New studies and more vantage points for TCP/UDP protocols are needed

References

http://metrics-itn.eu/

http://www.it.uc3m.es/amandala/

ANY QUESTIONS?

FAQ

www.it.uc3m.es/amandala/faq.pdf