UNIVERSITÉ DE LIÈGE

Faculté des Sciences Appliquées

Tracking Middleboxes in the Mobile World with TraceboxAndroid

Valentin THIRION, Korian EDELINE
& Benoit DONNET
Université de Liège

April 23, 2015

Content

1. Introduction
2. Tracebox
3. TraceboxAndroid
4. Evaluation
5. Shortcomings & Future improvements

Plan

1. Introduction
2. Tracebox
3. TraceboxAndroid
4. Evaluation
5. Shortcomings & Future improvements

Paradigm shift

Middlebox

Any intermediary box performing functions apart from normal, standard functions of an IP router on the data path between a source host and a destination host.

- Implicitly breaks the **end-to-end** paradigm

Paradigm shift

Middlebox

Any intermediary box performing functions apart from normal, standard functions of an IP router on the data path between a source host and a destination host.

- Implicitly breaks the **end-to-end** paradigm
- Normalize network traffic (Network ossification)

Paradigm shift

Middlebox

Any intermediary box performing functions apart from normal, standard functions of an IP router on the data path between a source host and a destination host.

- Implicitly breaks the **end-to-end** paradigm
- Normalize network traffic (Network ossification)
- Incomplete & Non-Collaborative

Paradigm shift

Middlebox

Any intermediary box performing functions apart from normal, standard functions of an IP router on the data path between a source host and a destination host.

- Implicitly breaks the **end-to-end** paradigm
- Normalize network traffic (Network ossification)
- Incomplete & Non-Collaborative
- Do not correctly & completely address challenges of the new paradigm !

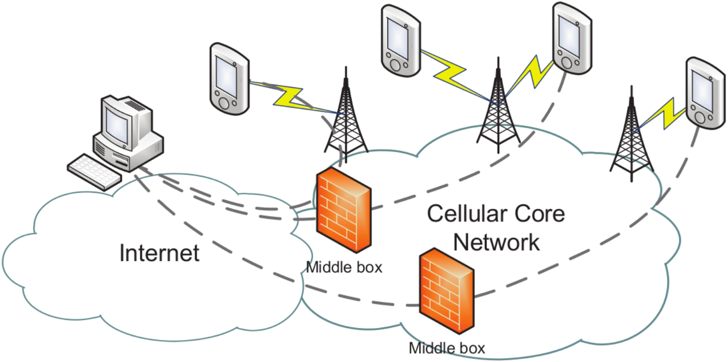
Paradigm shift

Middlebox

Any intermediary box performing functions apart from normal, standard functions of an IP router on the data path between a source host and a destination host.

- Implicitly breaks the **end-to-end** paradigm
- Normalize network traffic (Network ossification)
- Incomplete & Non-Collaborative
- Do not correctly & completely address challenges of the new paradigm !
- Network *disruptions !*

MBs in Cellular Networks



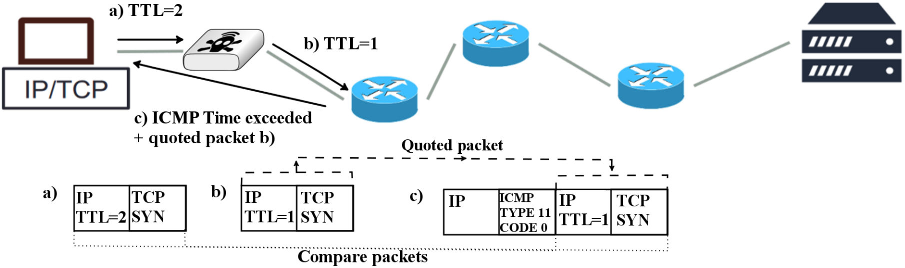
[1]

[1] Zhaoguang Wang et al. “An untold story of middleboxes in cellular networks”. In: *A SIGCOMM Computer Communication Review*. Vol. 41. 4. ACM. 2011, pp. 374–385.

Plan

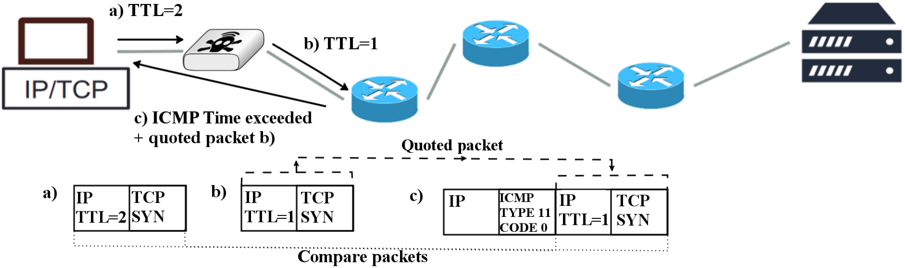
1. Introduction
2. Tracebox
3. TraceboxAndroid
4. Evaluation
5. Shortcomings & Future improvements

Tracebox[2]



[2] Gregory Detal et al. “Revealing middlebox interference with tracebox”. In: *Proceed of the 2013 conference on Internet measurement conference*. ACM. 2013, pp. 1–8.

Tracebox[2]



- Monitoring purposes
- Troubleshooting

[2] Gregory Detal et al. “Revealing middlebox interference with tracebox”. In: *Proceed of the 2013 conference on Internet measurement conference*. ACM. 2013, pp. 1–8.

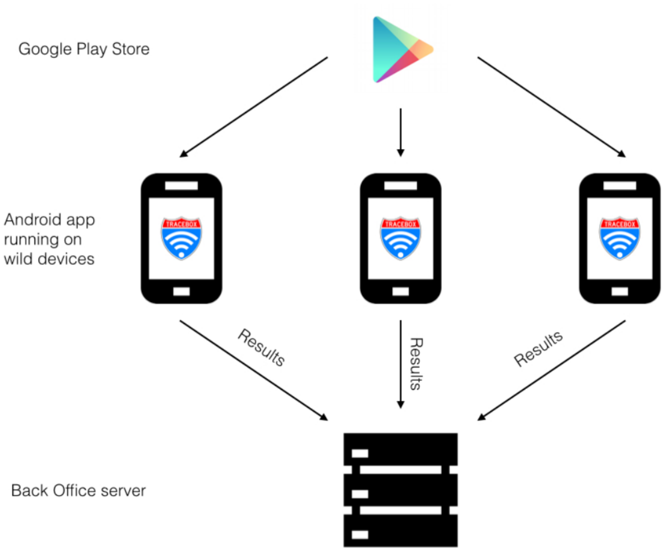
Advantages

- Server-independant
- Detect multiple modifications
- Lightweight

Plan

1. Introduction
2. Tracebox
3. TraceboxAndroid
4. Evaluation
5. Shortcomings & Future improvements

Crowd-Sourcing

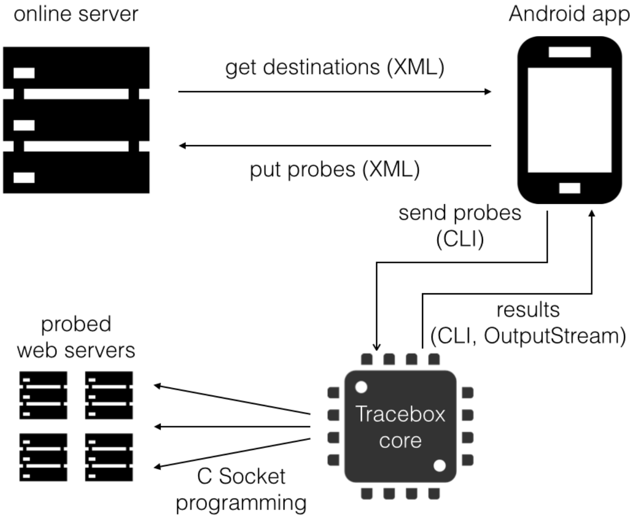


Google Play Store

Android app
running on
wild devices

Back Office server

System Overview

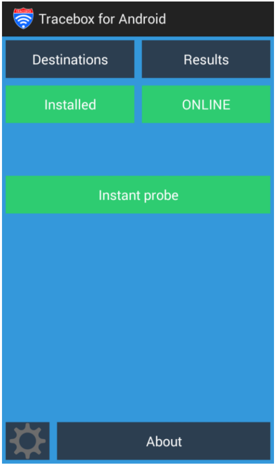


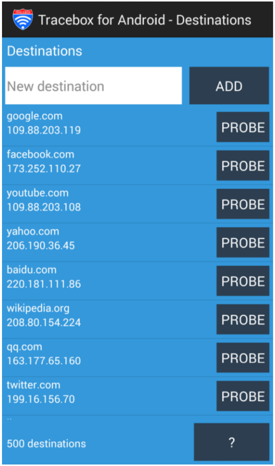
Probing Mechanisms

Three probing mechanisms are available in the app menu:

- Instant probing
- Background probing
- Custom probing

The App





google.com
109.88.203.119

facebook.com
173.252.110.27

youtube.com
109.88.203.108

yahoo.com
206.190.36.45

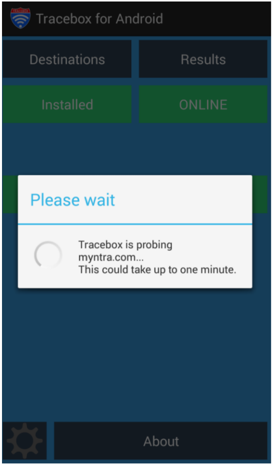
baidu.com
220.181.111.86

wikipedia.org
208.80.154.224

qq.com
163.177.65.160

twitter.com
199.16.156.70

500 destinations



Please wait

Tracebox is probing
myntra.com...
This could take up to one minute.

Plan

1. Introduction
2. Tracebox
3. TraceboxAndroid
4. Evaluation
5. Shortcomings & Future improvements

Data Overview

The principal information are:

- Network Paths & MB locations
- MB modifications

Data Overview

The principal information are:

- Network Paths & MB locations
- MB modifications

But also:

Data Overview

The principal information are:

- Network Paths & MB locations
- MB modifications

But also:

- Network type
- Carrier
- Cellular network characteristics
- Battery consumption
- User location
- Time

Dataset

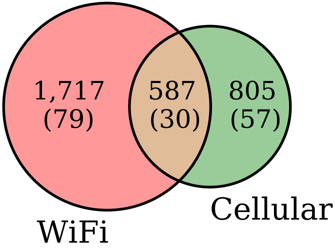
- **What ?** A dataset composed of 1,756 probes.
- **When ?**

  Between May, 2014 and September, 2014.

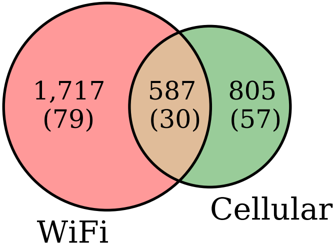
Dataset

- **What ?** A dataset composed of 1,756 probes.
- **When ?**
  Between May, 2014 and September, 2014.
- **From where ?**
  From a few users in Belgium, Italy, USA, China, and Nigeria.
- **To where ?** Alexa top-500.
- **Which carriers ?** O2, Mobile Vikings, E-Plus, BASE, T-Mobile, Movistar, KPN and more.

Paths & MBs locations

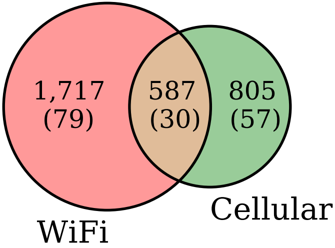


Paths & MBs locations



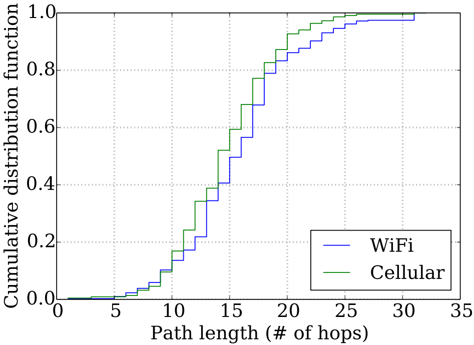
- 576 among 606 (95.05%) paths are crossing at least one rewritting middlebox.

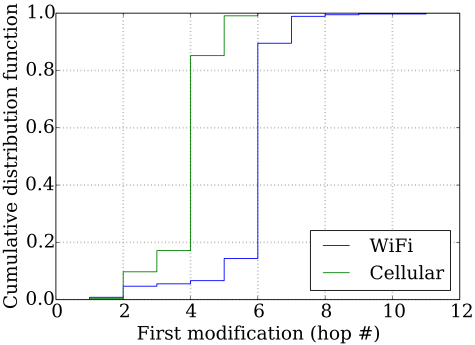
Paths & MBs locations



- 576 among 606 (95.05%) paths are crossing at least one rewritting middlebox.
- 361 among 389 (93.04%) WiFi paths
- 215 among 218 (98.62%) cellular network paths

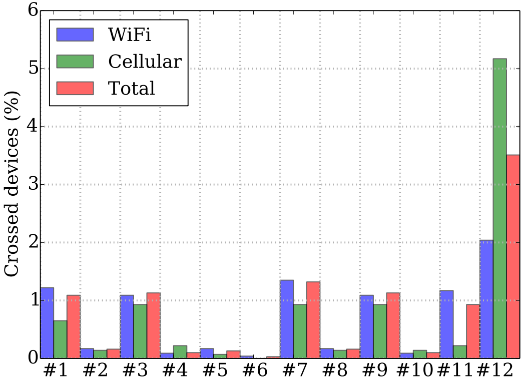
Paths & MBs locations





MB Modifications

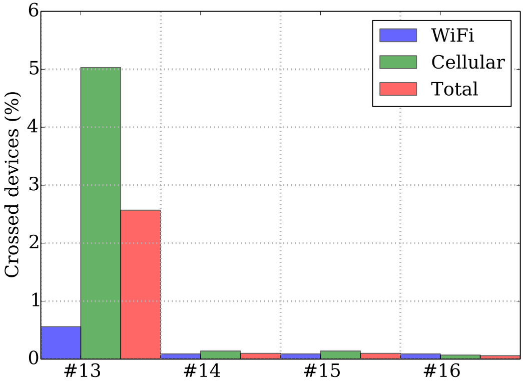
TCP & IP modifications



Label	*Field*
# 1	IP::ToS
# 2	IP::TotalLength
# 3	IP::ID
# 4	IP::Flags
# 5	IP::Protocol
# 6	IP::Checksum
# 7	TCP::SourcePort
# 8	TCP::DestPort
# 9	TCP::SeqNumber
# 10	TCP::Offset
# 11	TCP::WindowSize
# 12	TCP::Checksum

MB Modifications

TCP Options modifications



Label	*TCP option*
# 13	TCP::Option_MSS
# 14	TCP::Option_WS
# 15	TCP::Option_SACK
# 16	TCP::Option_MPTCP

Plan

1. Introduction
2. Tracebox
3. TraceboxAndroid
4. Evaluation
5. Shortcomings & Future improvements

Raw sockets

But we need to forge network & transport headers and to read ICMP messages:

- Raw sockets.
- `CAP_NET_RAW` POSIX capability.
- Root the device for a *proof-of-concept*.
- Find a solution for longterm use.

Raw sockets

But we need to forge network & transport headers and to read ICMP messages:

- Raw sockets.
- `CAP_NET_RAW` POSIX capability.
- Root the device for a *proof-of-concept*.
- Find a solution for longterm use.

In the next version:

- Unprivileged UDP
- Unprivileged TCP with regular options
- Require rooting for more options.

Future Improvements

- More app flexibility and in-app information
- Drop policies
- HTTP-level behavior
- Public dataset

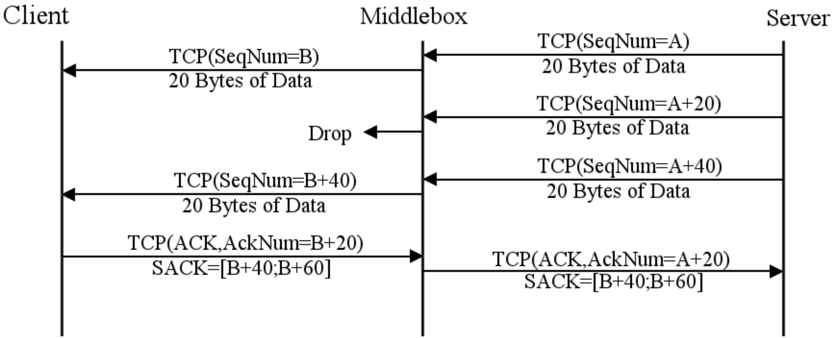
Interested ?
Send me an email at **korian.edeline@ulg.ac.be** to be notified when the new version is released.

Thank you !

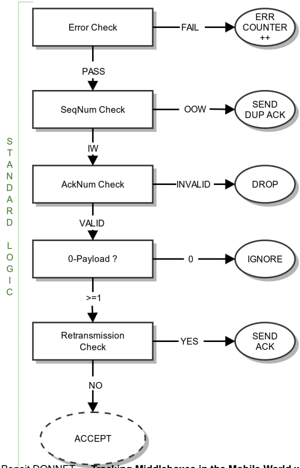
System Evaluation

Case	*Samsung Galaxy SII*	Arnova 10d G3
Memory	10.8Mb	6.45Mb
CPU (app)	< 1 %	< 1 %
CPU (instant probe)	12.5 %	12.5%

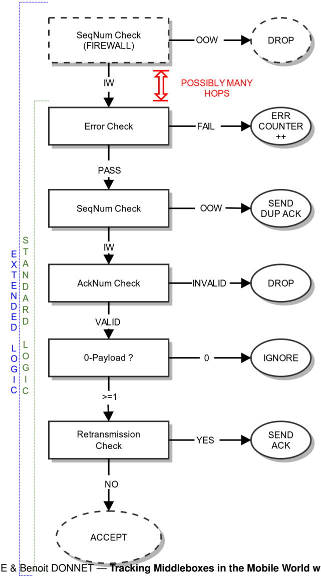
ISN Randomizer



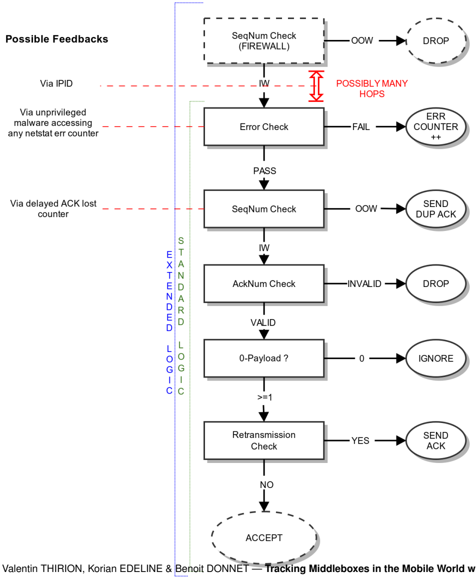
Security: TCP Validation Logic



Security: TCP Validation Logic



Security: TCP Validation Logic


Possible Feedbacks
Via IPID
Via unprivileged malware accessing any netstat err counter
Via delayed ACK lost counter
EXTENDED LOGIC
STANDARD LOGIC
SeqNum Check (FIREWALL)
OOW
DROP
IW
POSSIBLY MANY HOPS
Error Check
FAIL
ERR COUNTER ++
PASS
SeqNum Check
OOW
SEND DUP ACK
IW
AckNum Check
INVALID
DROP
VALID
0-Payload ?
0
IGNORE
>=1
Retransmission Check
YES
SEND ACK
NO
ACCEPT