

Measuring Privacy Threats in China-Wide Mobile Networks

FOCI '18

Mingming Zhang¹, Baojun Liu¹, Chaoyi Lu¹, Jia Zhang¹,
Shuang Hao² and Haixin Duan¹

¹ Tsinghua University, ² University of Texas at Dallas

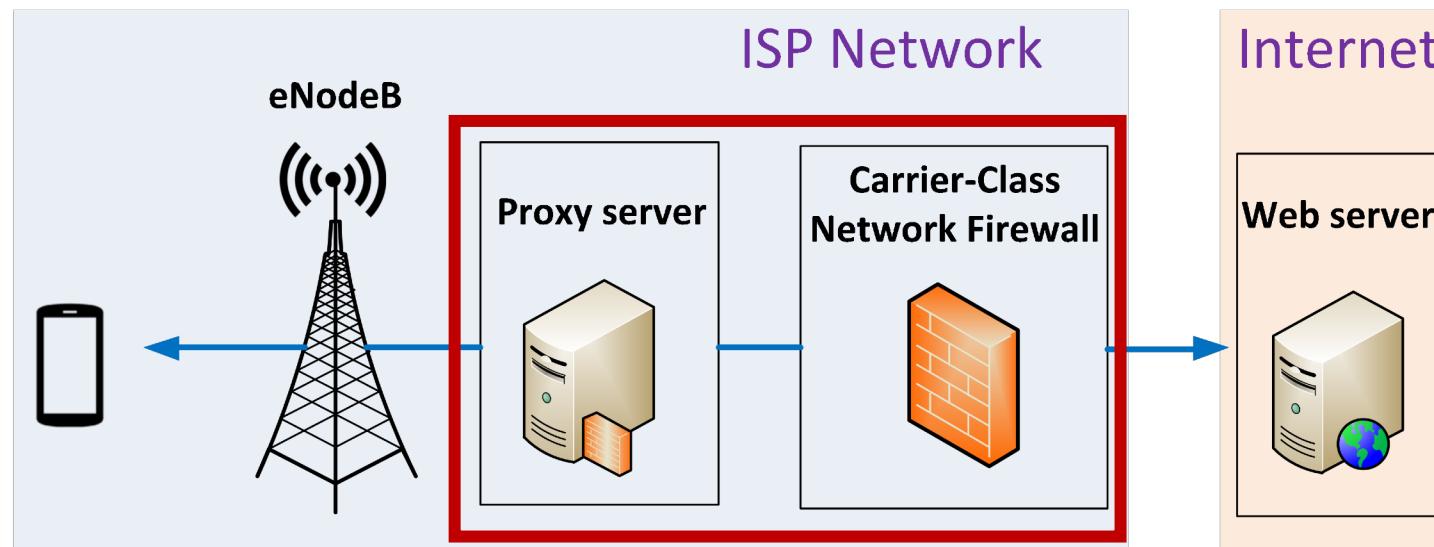


- Background
- Methodology
- Analysis
- Conclusion

Background

HTTP transparent proxies

- Widely deployed by mobile network operators
 - e.g. cache servers, firewalls, NAT devices...
- Enhance network performance and security [Sherry, SIGCOMM'12]
- Violate end-to-end principle



Examples of HTTP traffic manipulations

FOCI '18



Data usage pop-up
Data plan, Recharge service)

HTTP traffic manipulation

- A core area of free communication online in worldwide [Sherry, SIGCOMM'12] [Weaver, PAM'14] [Chung, IMC'16] [Tyson, WWW'17]
- Some transparent proxies leak private data of users and properties of devices [Weaver, SATIN'11]
- Some transparent proxies are vulnerable to known attacks. [Vallina-Rodriguez, MobiSys'15]



HTTP transparent proxies lead to potential security and privacy issues

Questions To Be Answered

- How is that in China-wide?
- How is that in cellular networks?



Our Goal

- Analyze the manipulation of HTTP traffic by transparent proxies
 - From China-wide
 - In cellular networks

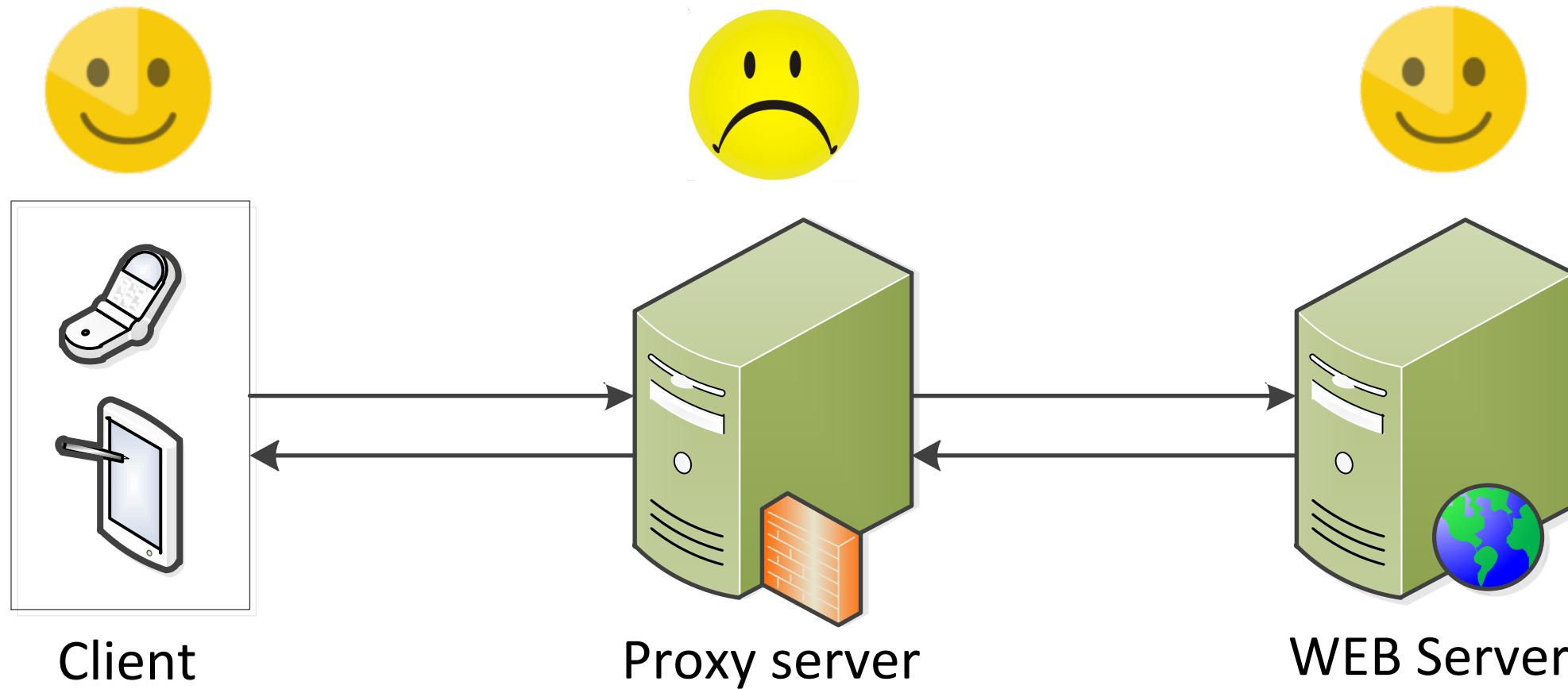
Contents

FOCI '18

- Background
- Methodology
- Analysis
- Conclusion

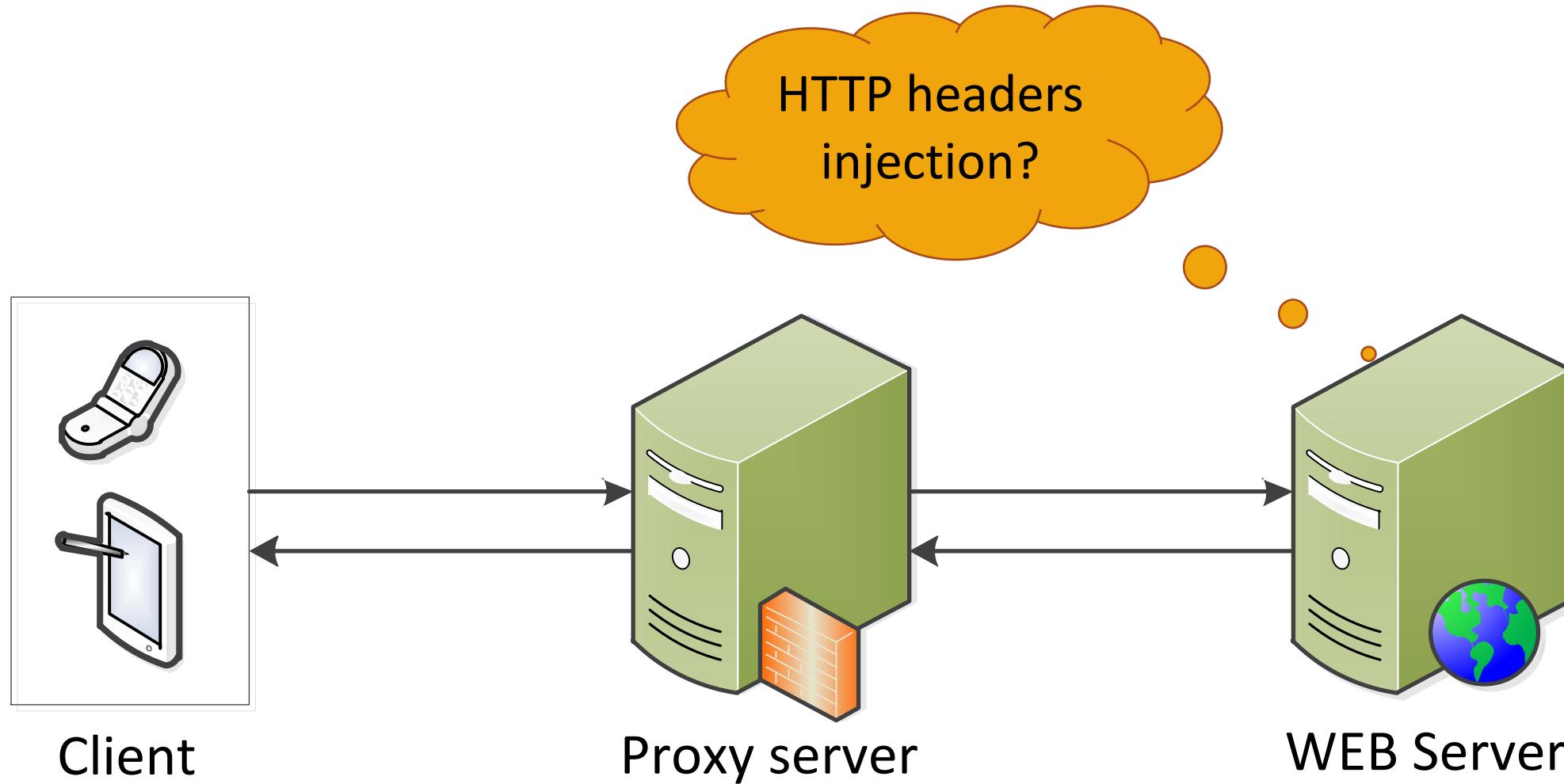
Methodology

- Identify transparent proxies



Methodology

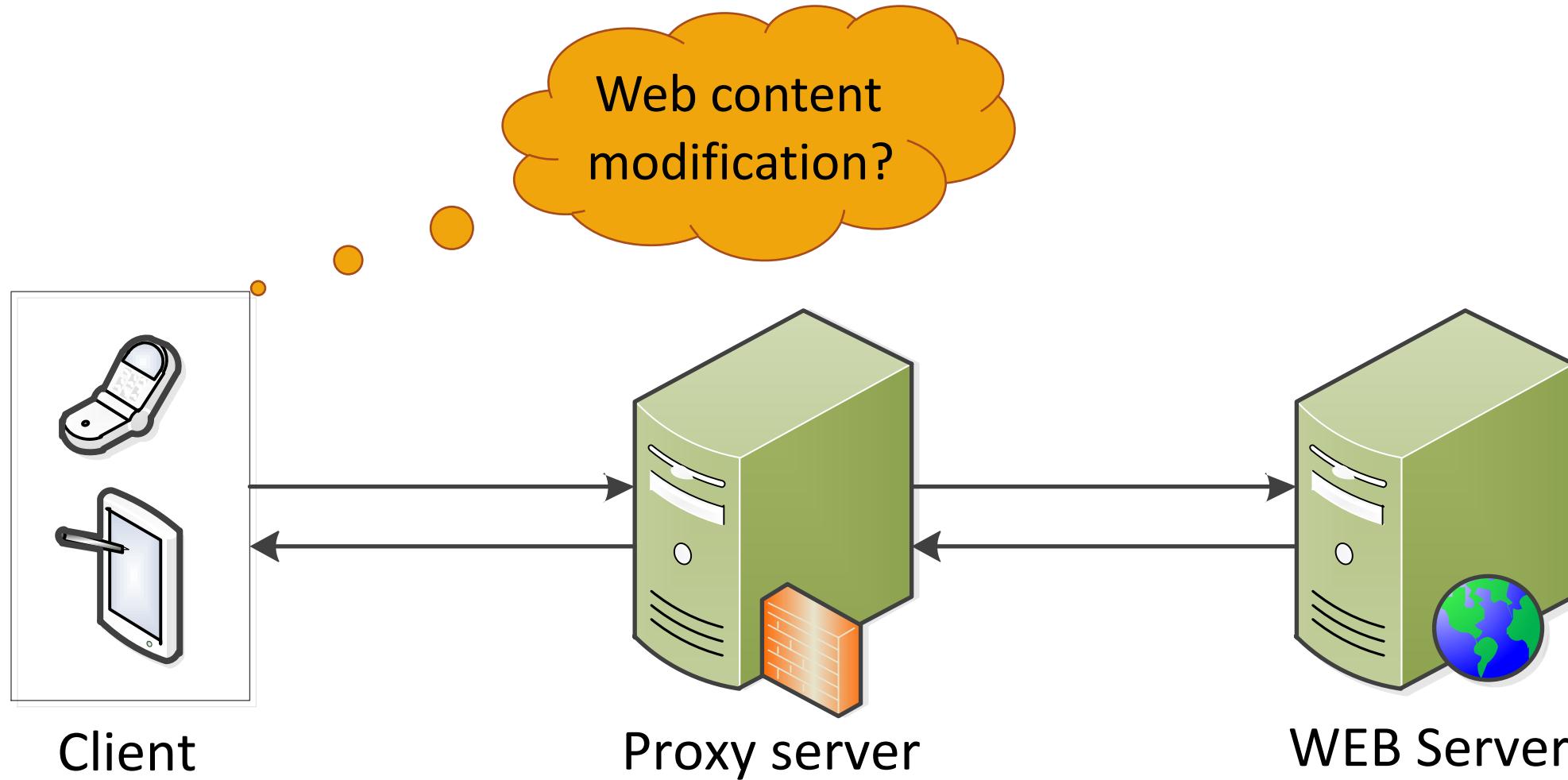
- Identify transparent proxies



Methodology

FOCI '18

- Identify transparent proxies

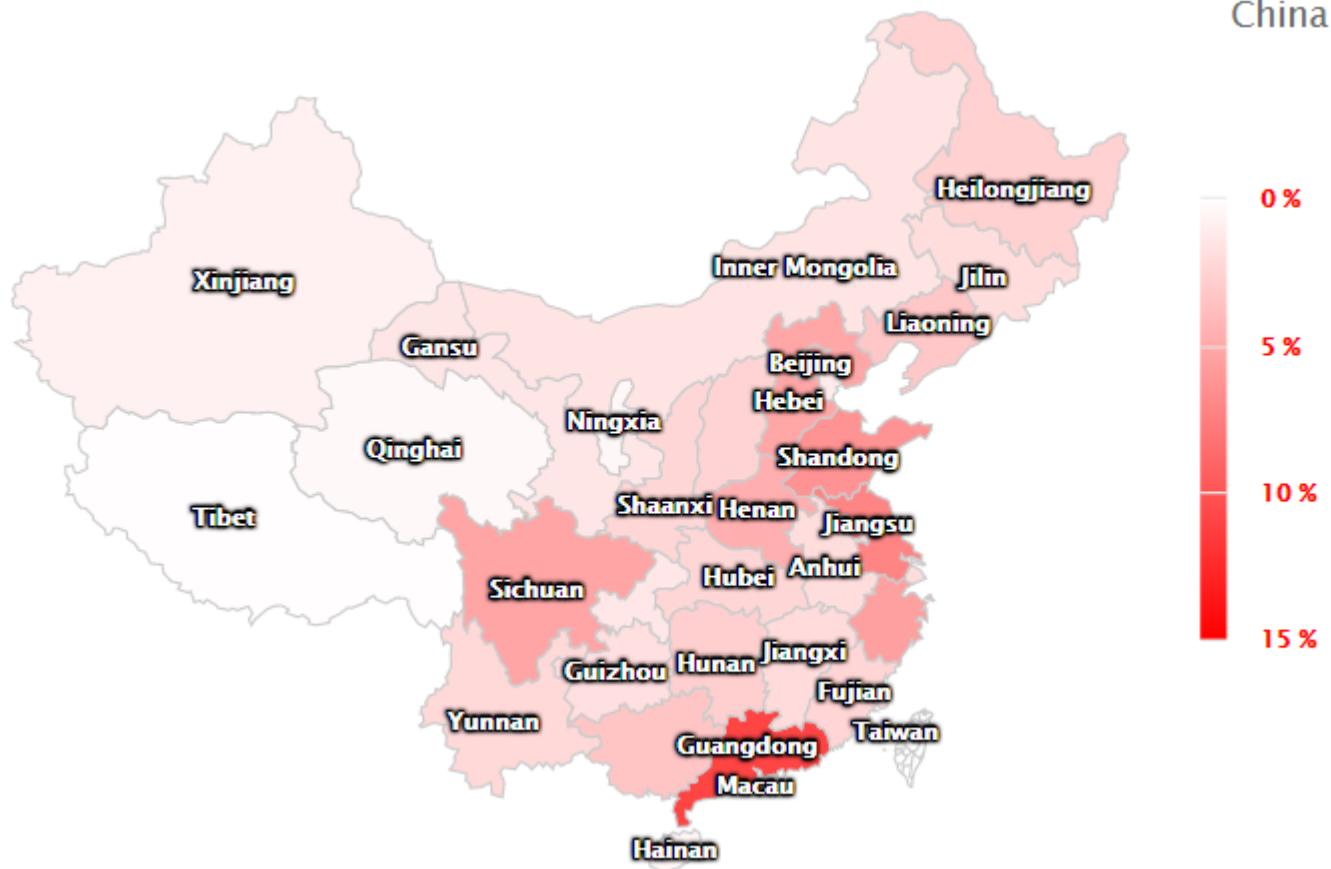


Collected Dataset

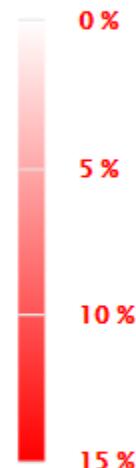
- China-Wide Analysis
 - A mobile network debugging tool of a security software
 - Ethics
 - One-time consent
 - Request our own website
 - Restrict traffic amount
 - Encrypted storage

Collected Dataset

- HTTP traffic originates from China-wide mobile networks
- Filter out invalid traffic



China



| Tests | Count |
|---------------|--------|
| HTTP sessions | 33,439 |
| #IP | 30,810 |
| Provinces | 31 |
| AS | 79 |

Collected Dataset

FOCI '18

- HTTP traffic originates from China-wide mobile networks
- Filter invalid traffic
- Limitation
 - Couldn't partition the data: cellular vs. wi-fi connectivity

Methodology

Identify Manipulation

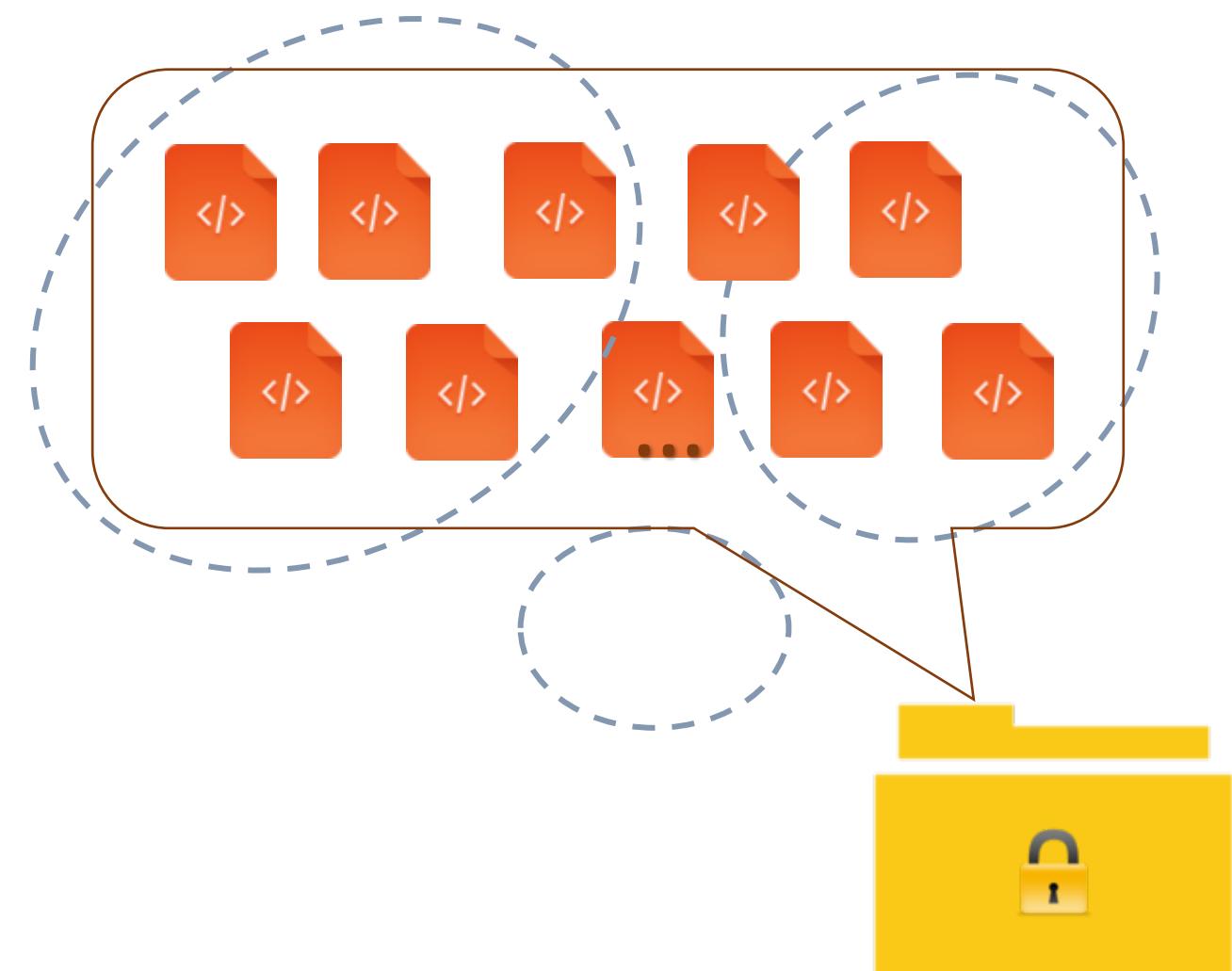
- Webpage modification
- HTTP headers injection

Methodology

FOCI '18

Identify Manipulation

- Webpage modification
 - Hierarchical clustering
 - Classify the similar pages

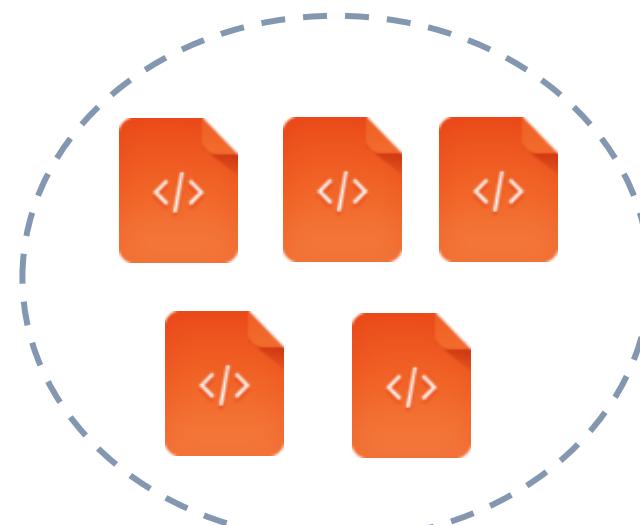


Methodology

FOCI '18

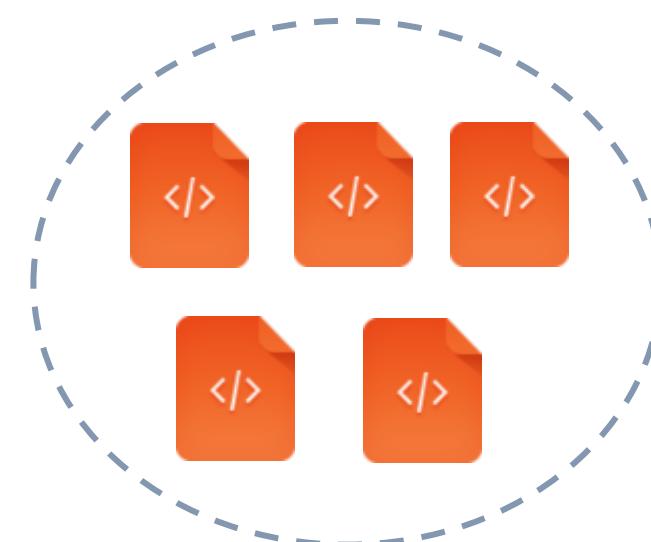
Identify Manipulation

- Webpage modification
 - Hierarchical clustering
 - Classify the similar pages
 - Inspecting sample pages from each cluster manually



Identify Manipulation

- Webpage modification
 - Hierarchical clustering
 - Classify the similar pages
 - Inspecting sample pages from each cluster manually
- HTTP headers injection
 - Jaccard distance between **original headers set** and **captured ones**



$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Contents

FOCI '18

- Background
- Methodology
- **Analysis**
- Conclusion

Scale of Traffic Manipulation

- HTTP session



- 1271 sessions: injecting HTTP headers
- 22 sessions: modifying web contents

- IP address



Scale of Traffic Manipulation

Geo-Distribution (451 IPs in 30 provinces)



Top 8 Provinces

| Province | # Session |
|----------|-----------|
| BJ | 229 |
| HB | 135 |
| JS | 135 |
| JL | 75 |
| HN | 69 |
| SD | 67 |
| GD | 46 |
| SX | 44 |

Scale of Traffic Manipulation

FOCI '18

AS-Distribution

- From three major mobile operators (**China Telecom**, **China Unicom** and **China Mobile**)

TOP 5 ASes

| AS | #Session | ISP |
|-------|-------------|-------------------|
| 4134 | 257 (19.9%) | China Telecom |
| 4837 | 202 (15.6%) | China Unicom |
| 9809 | 128 (9.9%) | China Mobile (GD) |
| 4808 | 114 (8.8%) | China Unicom (BJ) |
| 56046 | 111 (8.6%) | China Mobile (GD) |

Scale of Traffic Manipulation

FOCI '18

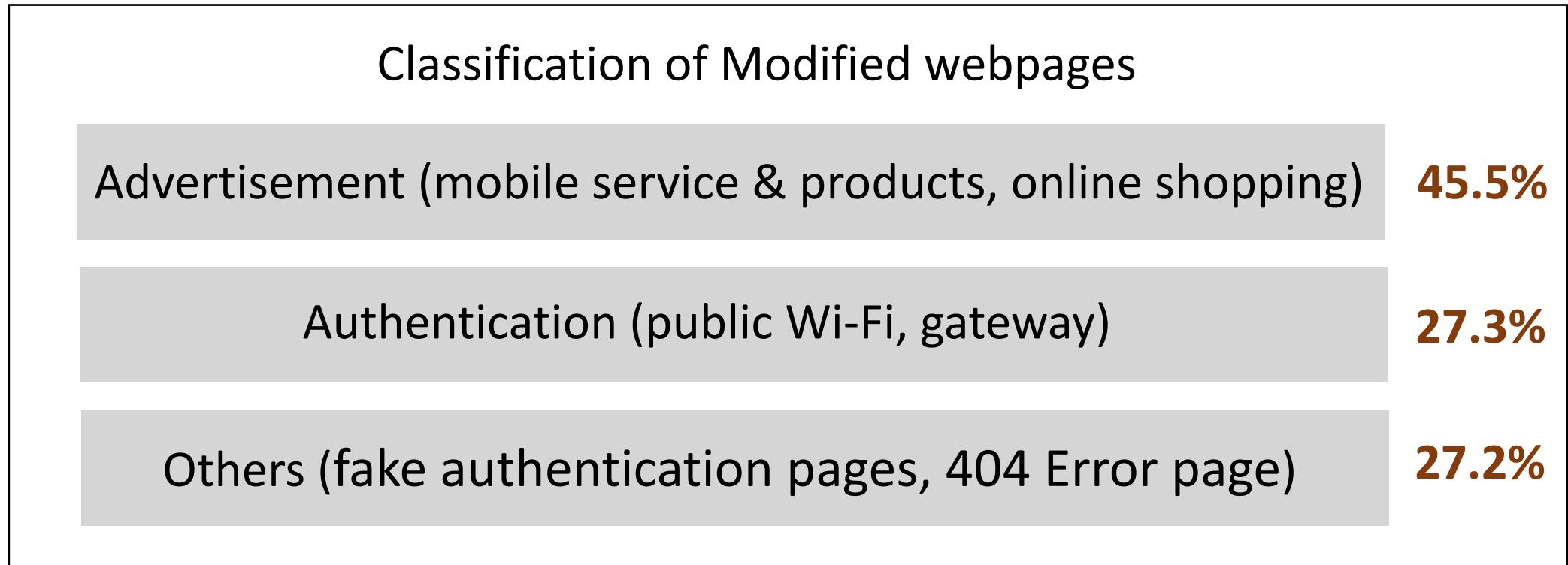
Network Operator

- Top 12 operators related to HTTP traffic manipulation.
- 90% manipulated traffic are found in networks of the top 3 ISPs.

| Organization | QTY | Organization | QTY |
|---|-----|--|-----|
| China Mobile | 524 | Beijing Founder Broadband Network | 3 |
| China Unicom | 325 | Shanghai Anchang Network Security | 2 |
| China Telecom | 317 | ZhengZhou GIANT Computer Network | 2 |
| CNISP-Union Technology (Beijing) | 15 | Beijing flash newsletter cas telecommunication | 1 |
| Zhejiang Taobao Network | 8 | BeiJing New-Billion Telecom Technology | 1 |
| BeiJing Guoxin bilin Telecom Technology | 4 | Beijing yiantianxia Network Science&Technology | 1 |

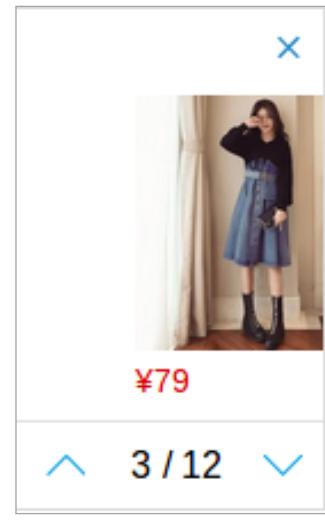
Modification of HTML Contents

- 22 modified web pages from 30K samples

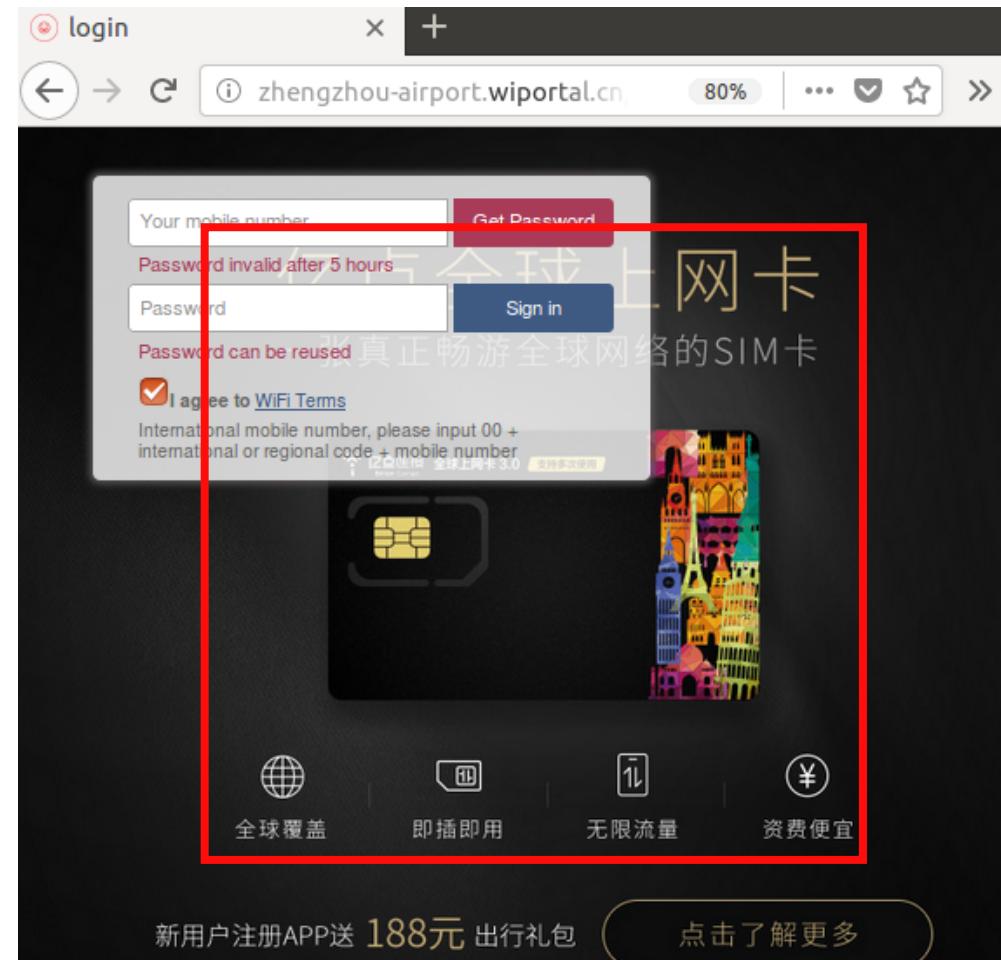


1. Advertisement (10 of 22)

- Services of mobile operators
- Finance & stock service
- Products of online shops

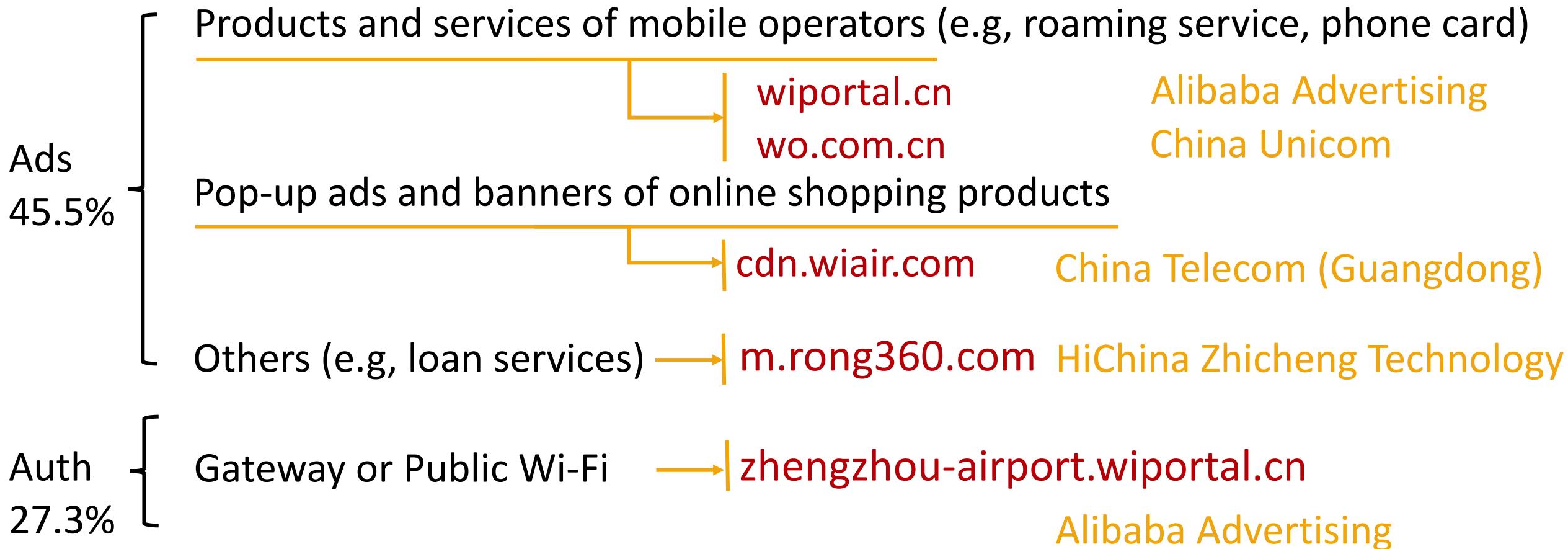


2. Authentication (6 of 22)



Modification of HTML Contents

- Who is behind the modification? (22 modified web pages)



Modification of HTTP Headers

- 1,271 HTTP sessions are injected with 43 types of headers
- These headers embed privacy data of users or devices
 - Location
 - IP address
 - Device serial number (e.g., IMEI)
- 3 categories
 - Identify users
 - Track users
 - Specials types

Dataset Analysis

Modification of HTTP Headers

- Headers for identifying mobile users (11 kinds in total)

| Header | Type | Organization | Count |
|-------------------------|--------------------|---|-------|
| x-IMEI* | IMEI | ChinaMobile (GD) | 12 |
| x-IMSI* | IMSI | ChinaTelecom, ChinaUnicom | 6 |
| x-up-calling-line-id | Phone # | ChinaTelecom (SH, SN, QH, SC, XJ, GS, BJ, SD, LN, YN, NM, ZJ, AH), ChinaMobile (GD), ChinaUnicom (BJ, JL, LN) | 50 |
| X-Nokia-CONNECTION_MODE | Connecting mode | ChinaMobile (GD) | 11 |
| x-up-bear-type | Communicating Type | ChinaMobile (GD), ChinaTelecom(BJ, SH, SX, QH, SC, XJ, GS, YN), ChinaUnicom (BJ, NM) | 122 |
| x-huawei-NetworkType* | Communicating Type | ChinaUnicom, ChinaTelecom | 6 |

Modification of HTTP Headers

- Headers for tracking mobile users (9 kinds in total)

| Header | Type | Organization | Count |
|-----------------|-----------------------|---|-------|
| X-Forwarded-For | Client IP | Farahoosh Dena, ChinaMobile (GD, SD), ChinaTelecom(SH, SX, SC, QH, XJ, GS), PT Telkom, ChinaUnicom (JL, LN, XJ) | 139 |
| X-Nx_remoteip* | Client IP | ChinaTelecom (QH, SC) | 3 |
| x-huawei-NASIP* | Gateway configuration | ChinaUnicom | 5 |
| x-source-id | Gateway configuration | ChinaUnicom (JL, LN), ChinaMobile (GD), ChinaTelecom (SH, SN, QH, SC, XJ, YN, NM, JS) | 62 |
| Cdn-Src-Ip* | Client IP | CNISP-Union, ChinaUnicom (LN) | 24 |

Modification of HTTP Headers

- Special header
 - Compromised *Content-Type*
 - The value of 2 sessions have been modified to probes of a vulnerability (Struts2, CVE-2017-5638).
 - OGNL codes

Modification of HTTP Headers

- Special headers
 - Compromised *Content-Type*

```
8 def exploit(url, cmd):  
9     payload = "%{#_=multipart/form-data'}."  
10    payload += "(#dm=@ognl.OgnlContext@DEFAULT_MEMBER_ACCESS)."  
11    payload += "(#_memberAccess?#"  
12    payload += "(#_memberAccess=#dm):"  
13    payload += "((#container=#context['com.opensymphony.xwork2.ActionContext.container'])."  
14    payload += "(#ognlUtil=#container.getInstance(@com.opensymphony.xwork2.ognl.OgnlUtil@class))."  
15    payload += "(#ognlUtil.getExcludedPackageNames().clear())."  
16    payload += "(#ognlUtil.getExcludedClasses().clear())."  
17    payload += "(#context.setMemberAccess(#dm)))."  
18    payload += "(#cmd='%s')." % cmd  
19    payload += "(#iswin=(@java.lang.System@getProperty('os_name')))."  
20    payload += "(#cmds=(#iswin?{'cmd.exe','/c',#cmd}:{'/bin/bash','-c',#cmd}))."  
21    payload += "(#p=new java.lang.ProcessBuilder(#cmds))."  
22    payload += "(#p.redirectErrorStream(true)).(#process=#p.start())."  
23    payload += "(#ros=(@org.apache.struts2.ServletActionContext@getResponse().getOutputStream()))."  
24    payload += "(@org.apache.commons.io.IOUtils@copy(#process.getInputStream(),#ros))."  
25    payload += "(#ros.flush())}"  
26  
27 try:  
28     headers = {'User-Agent': 'Mozilla/5.0', 'Content-Type': payload}  
29     request = urllib2.Request(url, headers=headers)  
30     page = urllib2.urlopen(request).read()
```

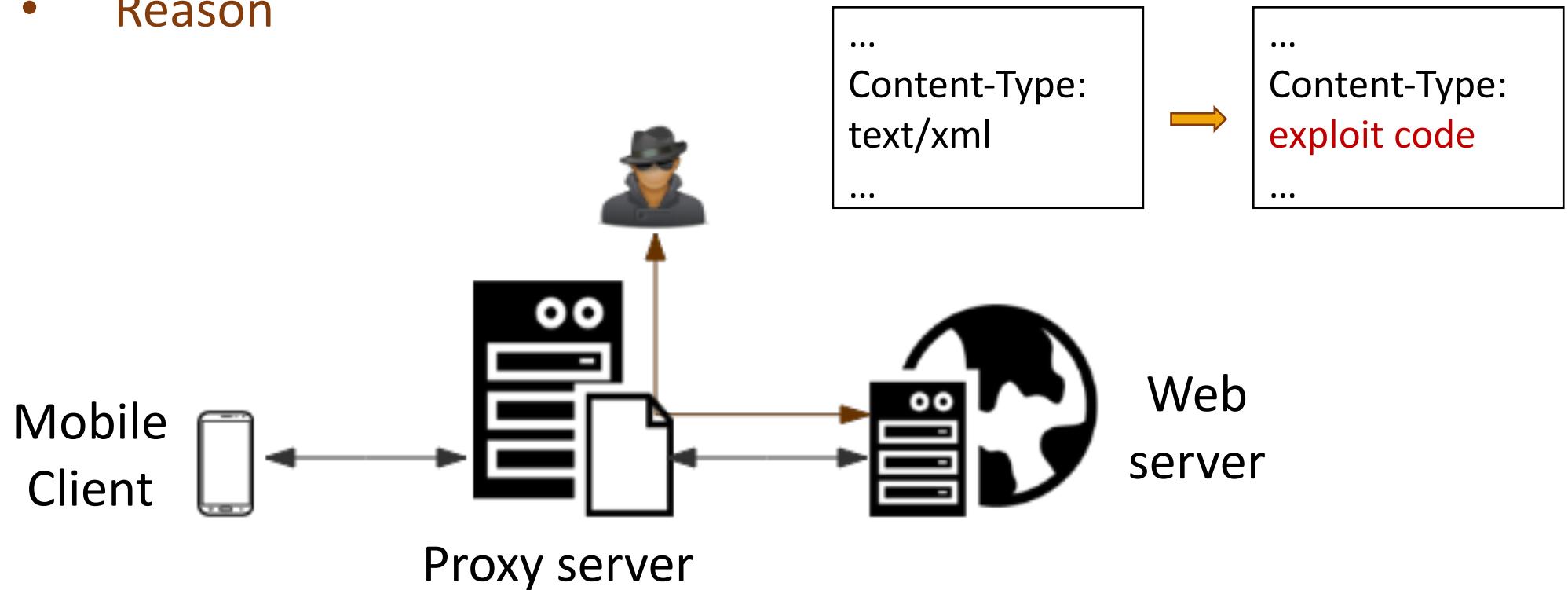


- whoami
- nMaskCustomMuttMoloz
- ...

Payload += “(#cmd=%s).” % cmd

Modification of HTTP Headers

- Special headers
 - Compromised *Content-Type*
 - Reason



Contents

FOCI '18

- Background
- Methodology
- Analysis
- Conclusion

Contribution

- A measurement study on
 - manipulation of HTTP traffic by transparent proxies
 - in cellular network from China-wide

HTTP traffic manipulation

- 3.86% of collected HTTP traffic are modified
- Two ways
 - web contents modification
 - HTTP headers injection

Motivations of manipulating HTTP traffic

- Advertising
 - E.g., ads injected to web pages
- Malicious behaviors
 - E.g., exploit code
- User tracking or identifying
 - E.g., user-related and device-related headers

Future work

- Exact location of traffic manipulation
 - TTL limited requests
 - In-path vs. on-path injections

Measuring Privacy Threats in China-Wide Mobile Networks

Mingming Zhang, Baojun Liu¹, Chaoyi Lu¹, Jia Zhang¹,
Shuang Hao² and Haixin Duan¹

zmm18@mails.tsinghua.edu.cn