Day 19: Middleboxes
The End-User Perspective

Courtesy: Z. Wang et al., Vyas Sekar, Naseo Rodriguez
Middle Boxes Allow Network Evolution

Performance, Security, Compliance, Accountability

Percentage of Methods Used to Exfiltrate Data

- Exposed Private Web Application Interface (1.5%)
- HTTP File Upload Site (1.5%)
- Malware Capability: IRC (2%)
- Malware Capability: SMTP (4%)
- SQL Injection (8%)
- Native FTP Client (10%)
- Native Remote Access Application (27%)
- Malware Capability: FTP (17%)
Classes On: MPTCP/SPDY

• Middleboxes are bad, because they
  – Strip out TCP Options
    • Makes evolving protocols hard
  – Have build in assumptions about behavior
    • Drop packets that violate this behavior
Last Class: Managing Middleboxes

• Is Challenging:
  – Dynamic Network Policies
  – Hard to Ensure Middlebox Traversal
  – Inefficient resource utilization

• Existing Solutions:
  – Virtualize the Middleboxes
    • Improves Resource utilization
    • Can make configuration someone else’s policy
  – Introduce Centralized Control
    • Ensures Middlebox Traversal
How Do Middleboxes Affect You?
Background on cellular network
Why carriers deploy middleboxes?
They Protect You from Attacks. Is that all?
HTTP Header Enrichment

- Technique that allows ISP-enforced proxies to extend/inject HTTP headers for:
  - Performance Enhancement
  - Content Customization
  - Advertising and user-tracking
How does HTTP Header Enrichment work?

GET /index.html HTTP/1.1
Host: www.example.com

GET /index.html HTTP/1.1
Host: www.example.com

x-acr: 486E03F2A285E07F5A981152DB80BB4932022388EC34B2243
928;ncc=310410;type=Dyna
How does HTTP Header Enrichment work?

- The Middlebox intercepts all packets
- Analyzes the packet’s payload
- Re-writes the payload to include the user information
User Implications

- HTTP Header Enrichment may become a privacy threat for mobile users:
  - ISPs may **leak** sensitive user and device data
  - ISPs may enable **user-tracking** (unique IDs)
Privacy-compromising headers

**Definition:** HTTP headers leaking sensitive information that identify uniquely:

- the device (e.g., IMEI)
- the user (e.g., IMSI/MSISDN)

Identified in 5 mobile operators
## Privacy-compromising headers

<table>
<thead>
<tr>
<th>Vodacom (ZA)</th>
<th>Phone #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orange (JO)</td>
<td>MSISDN</td>
</tr>
<tr>
<td>Smart (PH)</td>
<td></td>
</tr>
<tr>
<td>Vodacom (ZA)</td>
<td>IMEI</td>
</tr>
</tbody>
</table>
Two Classes of Middleboxes

• Needs header Information (TCP/IP headers)
  – Firewall, NAT, Load-balancers

• Needs access to the Payload
  – HTTP Header is considered payload
  – Proxy, Caches, DPI, Transcoders
  – Advertising, user tracking, parental controls
SSL

- The end-points agree on a key and encrypt all payload data.
Which of these are affected by SSL

- Needs header Information (TCP/IP headers)
  - Firewall, NAT, Load-balancers

- Needs access to the Payload
  - HTTP Header is considered payload
  - Proxy, Caches, DPI, Transcoders
  - Advertising, user tracking, parental controls
One Implication of SSL

- Some middleboxes are unable to examine or manipulate the packets.
  - WHY?
- Good: no more ads, or user tracking
- Bad: no more protection against worms or malware
State of the Art Solution

- Middlebox has special keys.
  - Keys allow the Middlebox to pretend to be “example.com”
- These keys are installed in the client
- When packets are sent:
  - Middlebox intercepts uses its keys to decrypt and manipulate
  - Then opens a new connection to example.com
Issues?
Alternate Approaches

• Approach 1
  – Encrypt different parts of the content with different keys
  – Give MB keys for the appropriate portions
Alternate Approaches

• Approach 2: Only for DPI
  – Encrypt the “attack” signature, e.g. worm signature
  – Middlebox compares encrypted packets with the encrypted signature
Discussions