On NDN and (“lack of”) Measurement

Thomas Silverston
National Institute of Information and Communications Technology (NICT)
ICT Testbed Research, Development and Operation Lab
P2P-TV Measurement Experiments and Traffic Analysis

- Experiment Testbed
- Traffic Analysis
- Novel Mechanisms

[Measuring P2P IPTV Systems],
ACM NOSSDAV 2007

[Traffic Analysis of P2P IPTV Communities],
Elsevier Computer Networks 2009
with A. Dainotti (Caida)

Data set (Anonymized)
Motivation

• Internet is mostly used to access content
  – Video: 90% of global consumer traffic by 2018
    • [Cisco VNI 2015]

• Users are interested with content, not location
  – TCP/IP (host-to-host communication)

• Information Centric Networking
  – Named-Data Networking [CoNext 2009]
  – Host-to-content communication
    • Packet address refers to content and not location
  – In-Network Caching

• New “network layer” for Future Internet
  – Data at the core of the communication
Open Issues
1. Caching at each Nodes
2. Routing/Forwarding toward Content
3. Security issues (information-leakage)
Caching Strategies in NDN

- **Popularity-based strategies**
  - MPC: Most-Popular Content Caching Strategy
    [IEEE ICC 2013]
    - Cache only popular Content
  - SACS: Socially-Aware Caching Strategy
    [IFIP Networking 2014]
    - Cache Content from popular users (Planet Lab experiments)
    - Infer User Traffic from Social Network dump (IEEE ICC 2014)
NDN Performances Evaluation

• Architecture evaluation
  – How many Cache Nodes in NDN to be efficient?’
  – Comparison with Client/Server, CDN architecture

• Trade-off 50% of cache nodes for higher performances
  – Deployment at reduced infrastructure cost for ISPs
Routing in ICN/NDN

• Routing scheme for NDN
  – Flooding (i.e.: wasting resources)
  – NLSR: in-path caching

• SRSC: SDN-based Routing Scheme for NDN
  [IEEE Netsoft 2015] Controller-based (anycast routing)
Routing in ICN/NDN

- Implementation on NDNx (NFD)
- Deployment on virtual Testbed with Docker
- Request: Zipf, etc.

Abilene

Geant
Security in NDN
Information-leakage

• One of the main security threat in Internet

• Cyber Espionage
  – Targeted Attacks (phishing, malware, website, external memory device)

• Examples: Sony, Target
  – $100 M upgrading systems
  – 46% drop in benefits
    *[Understanding Targeted Attacks: The Impact of Targeted Attacks]*
Targeted Attacks

Understand a full picture of the targeted email attack to implement the effective countermeasures!

- Infects PC via emails
- Probes network
- Steals Information

**Countermeasures**

Train employees?
Human errors

**Source:**
IT Security Center
IPA: IT Promotion Agency

- Fraud emails are just an initial phase to seek entry
  - They establish communication channels to enable remote control from the outside
  - True attack: steal and/or destroy targeted information through remote control

- It’s a whole system-wide design issue
  - Change the system design to one that expects and prepares for deep infiltration of the system

**Core of Attack:** NOT the spread of infection
BUT spread of infiltration

**Inside Operation Prevention** (incl. Exit Control)

**Entry Control**
Information-leakage through NDN packets

- **Interest/Data** packets are “Request/Reply”
  - Content name, etc.
- **Data** can be **filtered out** out by network admin.
  - White/Black lists of (un)authorized content names
    - *CustomerList, BankingInfo*, etc.
- **Interest** packets are sent out the network to external publishers as requests (“free” names)
  - Malwares can use *Interest* to leak Information through Targeted Attacks
Information-leakage Countermeasure with Data

Gatekeeper can prevent information leakage through Data packet (reply messages)

Rules to Publish Content

1) Gatekeeper has white list of public contents
2) Every new content is checked by gatekeeper to register it into white list
3) Any content cannot be accessed unless it is listed in white list

Network Administrator

Gatekeeper

Normal Agent

The Internet

Attacker

Comp1/Pub/Info1

Enterprise Network

Comp1/Priv/Info1

Firewall

Employee A
Targeted Attacks in NDN

1. C&C server (Control malware via bots)
2. Bot
3. Malware

Interest Name can be used to leak information through Targeted Attacks (request messages)
URLs Dataset

- Web Crawling of 7 main organizations
  - Amazon, Ask, Stackoverflow, BBC, CNN, Google, Yahoo
  - Common Crawl Data Set repository
- 1.73B URLs -> 7M for each organization

![URLs Parameters Diagram](image)

<table>
<thead>
<tr>
<th>Parameters (RFC 1808)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of &lt;PATH&gt;</td>
<td>Number of ‘/’ in &lt;path&gt;</td>
</tr>
<tr>
<td>Length of &lt;QUERY&gt;</td>
<td>Similarity of characters in &lt;PATH&gt;</td>
</tr>
<tr>
<td>Length of &lt;FRAGMENT&gt;</td>
<td>Similarity of characters in &lt;QUERY&gt;</td>
</tr>
<tr>
<td>Length of Directory</td>
<td>Similarity of characters in &lt;QUERY&gt;</td>
</tr>
<tr>
<td>Length of File</td>
<td>Similarity of characters in &lt;FRAGMENT&gt;</td>
</tr>
</tbody>
</table>
Average Frequencies in Path, Query, and Fragment

- Calculated average frequencies of characters in path, query and fragment of the URLs in all the organizations
URLs Similarity

Legitimate names exceed average similarity

<table>
<thead>
<tr>
<th>Organization</th>
<th>Average $C_{Path}$</th>
<th>Average $C_{Query}$</th>
<th>Average $C_{Fragment}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>0.76</td>
<td>0.73</td>
<td>0.5</td>
</tr>
<tr>
<td>Ask</td>
<td>0.76</td>
<td>0.86</td>
<td>0.57</td>
</tr>
<tr>
<td>Stackoverflow</td>
<td>0.77</td>
<td>0.76</td>
<td>0.4</td>
</tr>
<tr>
<td>BBC</td>
<td>0.74</td>
<td>0.56</td>
<td>0.6</td>
</tr>
<tr>
<td>CNN</td>
<td>0.81</td>
<td>0.54</td>
<td>0.63</td>
</tr>
<tr>
<td>Yahoo</td>
<td>0.72</td>
<td>0.64</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Average 0.75  0.68  0.55
Anomaly Detection in NDN

- Prevent Information-leakage
  - Internet security threat through Targeted Attacks
- Web Organizations Crawling (Google, CNN, etc.)
  - Statistics on URLs (names) and HTTP traffic
- Malicious Names filtering in NDN (15% misdetection names)
  - [IEEE Lanman 2016] with D. Kondo (UL), Prof. Asami (U. Tokyo), Prof. Tode (U. Pref. Osaka) and Prof. O. Perrin (UL)
  - [NOM WS – Infocom 2017] D. Kondo (UL), Prof. Asami (U. Tokyo), Prof. Tode (U. Pref. Osaka) and Prof. O. Perrin (UL)
  - One-Class SVM
Project ANR Doctor (2014-2017)
http://www.doctor-project.org/

- Deployment of new network functions and protocols (e.g.: NDN) in a virtualized networking environment (e.g.: NFV)
  - Monitoring, managing and securing (using SDN for reconfiguration)
- Partners: Orange, Thlaes, Montimage, UTT, LORIA/CNRS (900k€)
- NDN/HTTP proxy designed in the project
Conclusion

- NDN Architecture
  - Caching: popularity-based
  - Routing: Controller-based
  - Security: Name-Anomaly Detection in NDN

ありがとうございました

thomas@nict.go.jp