CoMo-UPC

TMA evaluation service @ UPC

Pere Barlet-Ros    Josep Sanjuàs-Cuxart

Advanced Broadband Communications Center (CCABA)
Universitat Politècnica de Catalunya (UPC)
{pbarlet,jsanjuas}@ac.upc.edu

3rd COST-TMA meeting
Aachen (Germany), 12 May 2009
The problem

- Several TMA participants working on Anomaly Detection (AD)
- Real packet traces are needed to test novel AD methods
- Several AD algorithms require:
  - Unanonymized IP addresses (or prefix-preserving)
  - Payload inspection (e.g., IDS)
- Traditional solution: Anonymized traffic traces
  - Examples: NLANR, CAIDA, CRAWDAD, ...
The problem

- Several TMA participants working on Anomaly Detection (AD)
- Real packet traces are needed to test novel AD methods
- Several AD algorithms require:
  - Unanonymized IP addresses (or prefix-preserving)
  - Payload inspection (e.g., IDS)
- Traditional solution: Anonymized traffic traces
  - Examples: NLANR, CAIDA, CRAWDAD, . . .

Anonymization is not the right solution!
- Data owners: Privacy concerns
- Researchers: Not enough data
- Lack of recent publicly available traces
Alternative solution: CoMo

- Move the code to the data
  - Instead of publishing anonymized data traces

- Significantly lowers the privacy concerns
  - Traffic data do not leave provider premises
  - Data providers keep the ownership of the data
  - The source code can be inspected by the data owner

- Researchers have (blind) access to unanonymized traffic
  - IP addresses and payloads can be processed . . .
  - . . . but not stored or exported

- The CoMo system is based on this model
CoMo system deployed in the UPC network

Connects 40 departments and 25 faculties (10 campuses)

Continuously monitoring the UPC access link to the Internet
  - 1 Gigabit Ethernet full-duplex
  - Average traffic: ≈ 900 Mb/s (60K flows/s)
UPC traffic

Application Traffic Breakdown: Monitor

Live statistics

- **Appmon**: [http://monitoring.ccaba.upc.edu/appmon/](http://monitoring.ccaba.upc.edu/appmon/)

- **CoMolive!**: [http://monitoring.ccaba.upc.edu/como-live/](http://monitoring.ccaba.upc.edu/como-live/)
CoMo-UPC system

- **AD evaluation service for TMA participants**
  - [http://monitoring.ccaba.upc.edu/como-upc](http://monitoring.ccaba.upc.edu/como-upc)

- **System hardware**
  - Intel Xeon 2.4GHz (dual processor)
  - 2GB RAM

- **Monitoring hardware**
  - Endace DAG 4.3GE

- **Running CoMo v2.0 (development version)**
  - Online collection (unidirectional and bidirectional)
  - Offline with packet traces
## Available packet traces

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
<th>Start (duration)</th>
<th>Contents</th>
<th>Dir.</th>
<th>Mean traffic</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPC-I</td>
<td>11-Dec-2008</td>
<td>10:00 (15 min)</td>
<td>payloads</td>
<td>both</td>
<td>471 Mb/s</td>
<td>53 GB</td>
</tr>
<tr>
<td>UPC-II</td>
<td>11-Dec-2008</td>
<td>12:00 (15 min)</td>
<td>payloads</td>
<td>both</td>
<td>560 Mb/s</td>
<td>63 GB</td>
</tr>
<tr>
<td>UPC-III</td>
<td>12-Dec-2008</td>
<td>16:00 (15 min)</td>
<td>payloads</td>
<td>both</td>
<td>488 Mb/s</td>
<td>55 GB</td>
</tr>
<tr>
<td>UPC-IV</td>
<td>12-Dec-2008</td>
<td>18:30 (15 min)</td>
<td>payloads</td>
<td>both</td>
<td>426 Mb/s</td>
<td>48 GB</td>
</tr>
<tr>
<td>UPC-V</td>
<td>21-Dec-2008</td>
<td>16:00 (1 h)</td>
<td>payloads</td>
<td>both</td>
<td>275 Mb/s</td>
<td>124 GB</td>
</tr>
<tr>
<td>UPC-VI</td>
<td>22-Dec-2008</td>
<td>12:30 (1 h)</td>
<td>payloads</td>
<td>both</td>
<td>573 Mb/s</td>
<td>258 GB</td>
</tr>
<tr>
<td>UPC-VII</td>
<td>10-Mar-2009</td>
<td>03:00 (1 h)</td>
<td>payloads</td>
<td>both</td>
<td>175 Mb/s</td>
<td>79 GB</td>
</tr>
<tr>
<td>CESCA-II</td>
<td>02-Nov-2005</td>
<td>16:30 (30 min)</td>
<td>headers</td>
<td>in</td>
<td>360 Mb/s</td>
<td>8 GB</td>
</tr>
<tr>
<td>CESCA-III</td>
<td>11-Apr-2006</td>
<td>08:00 (30 min)</td>
<td>payloads</td>
<td>in</td>
<td>133 Mb/s</td>
<td>29 GB</td>
</tr>
<tr>
<td>CESCA-IV</td>
<td>24-Oct-2006</td>
<td>09:00 (8 h)</td>
<td>headers</td>
<td>in</td>
<td>750 Mb/s</td>
<td>156 GB</td>
</tr>
<tr>
<td>CESCA-V</td>
<td>25-Oct-2006</td>
<td>09:00 (8 h)</td>
<td>headers</td>
<td>in</td>
<td>719 Mb/s</td>
<td>153 GB</td>
</tr>
<tr>
<td>CESCA-VI</td>
<td>05-Dec-2006</td>
<td>09:00 (8 h)</td>
<td>headers</td>
<td>in</td>
<td>403 Mb/s</td>
<td>139 GB</td>
</tr>
</tbody>
</table>

**Table:** List of available traces in CoMo-UPC (all traces are in ERF format)
Usage procedure

1. Send an email to como-upc@ac.upc.edu with:
   - Your CoMo module source code
   - Traffic to be processed:
     - Online (limited to 30 min)
     - Offline (trace name)
   - Description of your research and use of the data
   - Statement accepting the AUP

2. We run your module on our CoMo system
   - UPC will ensure that your module complies with the AUP

3. We send you back a file with your module results
Acceptable use policy (AUP)

- Private data cannot be stored or exported outside UPC
  - IP addresses, subnets, URLs, payloads (or parts of them), etc.
  - privacy of users and UPC must be respected

- Data obtained from the UPC network:
  - can only be used for scientific or academic research purposes
  - cannot be shared with others without permission from UPC

- Any material (e.g. papers) that contains UPC data must:
  - be provided to UPC and gain explicit permission prior publication
  - properly cite the source of the data

- The institution that receives the data agrees to permanently destroy any data supplied by UPC at any time at UPC request
More information and downloads

- Documentation and instructions:
  - http://monitoring.ccaba.upc.edu/como-upc

- CoMo project and downloads:

- More info also at the TMA portal: