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

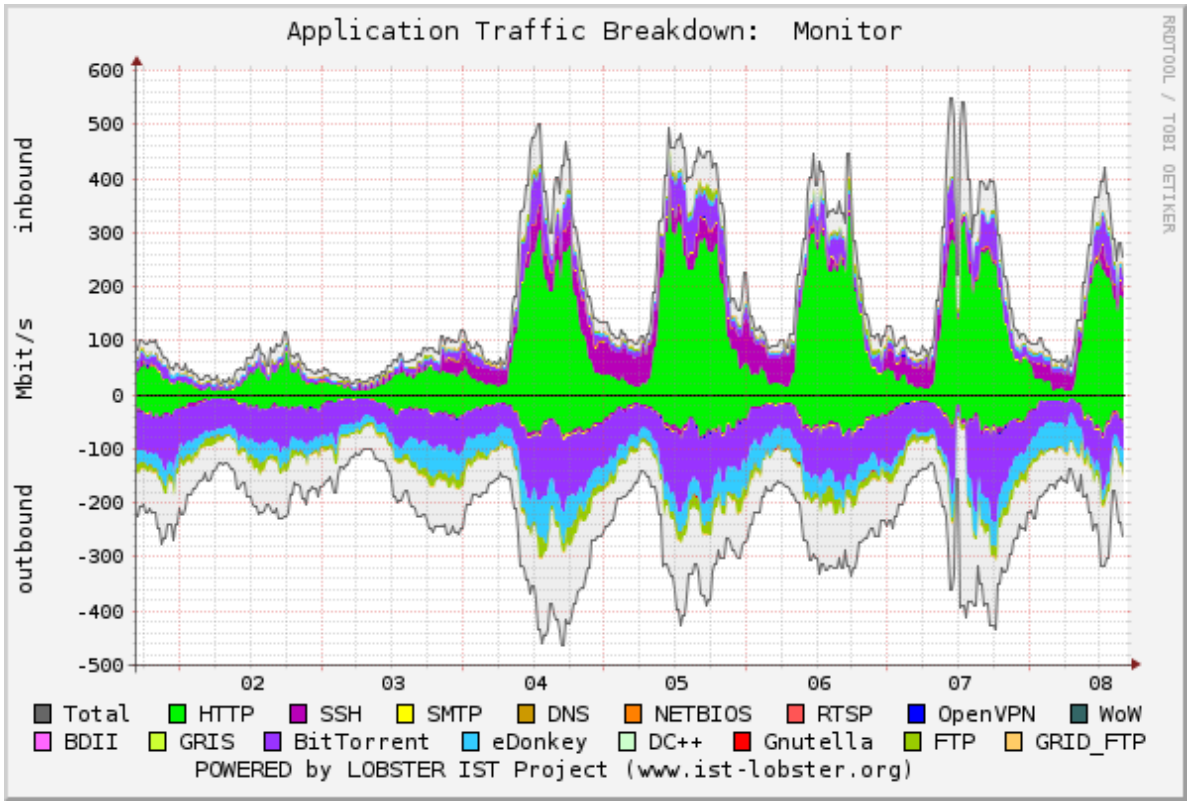
Alternative solution: CoMo

- 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

UPC network

- 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: $\approx$ 900 Mb/s (60K flows/s)

UPC traffic



Live statistics

- Appmon: `http://monitoring.ccaba.upc.edu/appmon/`
- CoMolive!: `http://monitoring.ccaba.upc.edu/como-live/`

CoMo-UPC system

- AD evaluation service for TMA participants
  - `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

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

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:
  - `http://como.sourceforge.net`
- More info also at the TMA portal:
  - `http://www.tma-portal.eu/resources/tools/como`