AN INNOVATIVE NETWORK VISIBILITY SERVICE (NOT ONLY) FOR NRENs

Pere Barlet-Ros¹, Josep Sanjuas², Josep Solé Pareta¹, Maria Isabel Gandía³, Chelo Malagón⁴

¹ UPC-BarcelonaTech – {pbarlet,pareta}@ac.upc.edu ² Talaia Networks – jsanjuas@talaianetworks.com ³ CSUC – MariaIsabel.Gandia@csuc.cat ⁴ RedIRIS – chelo.malagon@rediris.es

COMPANIES DEPEND ON NETWORKS
1. Productivity enhancement
2. Entire business models

NETWORK MALFUNCTIONS ⇒ LARGE COSTS
42,000 $/h to 5,600 $/min (Gartner, Ponemon Institute)

SOLUTION: NETWORK VISIBILITY
“When you can measure what you are speaking about, and express it in numbers, you know something about it.” – Lord Kelvin

- understand how the network is used
- identify bandwidth hogs
- detect unwanted applications
- detect anomalies & attacks
- investigate security incidents
- long-term network planning

TRADITIONAL APPROACHES
Deep Packet Inspection (DPI): high visibility at high cost
- instrument network with hardware
- capture & analyze every data packet
- compute any metric of interest

NetFlow / IPFIX / sFlow: low visibility at low cost
- delegate capturing to routers/switches (standardized)
- no access to packet contents - only traffic summaries
- aggregate results and present them to user

NETWORK POLYGRAPH
Leverages a huge body of research works in traffic classification, combines available techniques in a network visibility product

- Capture traffic at one link
- Extract NetFlow & perform DPI
- Train a classification engine
- Extend to all the network

Advantages:
- Main data input: NetFlow – cost-effective, easy to deploy
- Accuracy comparable to DPI, ability to self-assess accuracy

DEPLOYMENT AT CSUC AND REDIRIS
NetFlow as main data input. Auxiliary DPI input to train the application classification engine, which is applied to all the traffic.

MULTI-TENANT: SERVICES ALL INSTITUTIONS
Both an internal tool and a value-added service for institutions connected a NREN.
In production in Anella Cientifica (CSUC), and in deployment at RedIRIS (Red.es).

RESULTS

<table>
<thead>
<tr>
<th>Time</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fri, 04 Feb 2011</td>
<td>94%</td>
</tr>
<tr>
<td>Tue, 08 Feb 2011</td>
<td>96%</td>
</tr>
<tr>
<td>Fri, 11 Feb 2011</td>
<td>98%</td>
</tr>
<tr>
<td>Mon, 14 Feb 2011</td>
<td>100%</td>
</tr>
<tr>
<td>Avg. accuracy</td>
<td>96.76%</td>
</tr>
<tr>
<td>5 retrainings</td>
<td>94% threshold</td>
</tr>
<tr>
<td>Avg. accuracy</td>
<td>97.5%</td>
</tr>
<tr>
<td>15 retrainings</td>
<td>96% threshold</td>
</tr>
<tr>
<td>Avg. accuracy</td>
<td>98.26%</td>
</tr>
<tr>
<td>108 retrainings</td>
<td>98% threshold</td>
</tr>
</tbody>
</table>

CLOUD DEPLOYMENT
Super fast deployment (just a few minutes to configure routers).
Without DPI: runs on default lab-generated training model.
Offered under a subscription model (SaaS) from the cloud.
Currently in use by customers in 3 continents.
Offered by ourselves leveraging the multi-tenancy feature.
Alternatively, on-site deployment (possibly virtualized) is available.

On-line demo: https://polygraph.io