Demo Abstract: Towards extracting semantics by visualizing large traceroute datasets

Emden R. Gansner(1) · **Balachander Krishnamurthy**(1) · **Walter Willinger**(1) · **Fabián E. Bustamante**(2) · **Mario A. Sánchez**(2)

Received: 30 August 2012 / Accepted: 8 November 2012

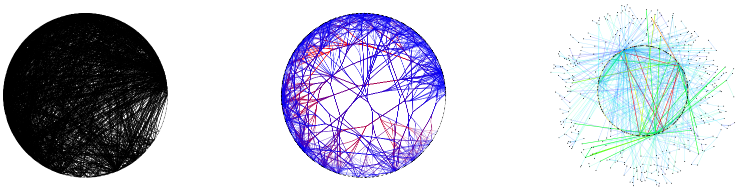
Measurement projects that have relied on traceroute as a primitive have generated hundreds of millions of traceroute samples and used them for network debugging, troubleshooting and Internet mapping. We advocate the use of visualization as a means to extract semantics from large sets of such traceroute data. We have implemented a prototype environment, VITASCOPE, for analyzing and visualizing traceroutes as AS-level paths. The basic graph functionality of VITASCOPE is built on Graphviz, an open-source platform for viewing and manipulating graphs. VITASCOPE takes a raw collection of AS-level traceroute data and converts it into an edge-weighted undirected graph, which serves as the principal representation of the data. A graph is created by constructing a node to represent each unique AS found in the traceroutes. Edges are added to the graph for each unique pair of consecutive ASes appearing in a traceroute; each edge is weighted by the number of traceroutes traversing the edge. To help reduce clutter and identify high-level trends among the edges, various edge bundling algorithmsare used.

In this demo, we illustrate some basic aspects of VITASCOPE by using traceroute measurements that resulted from four large-scale traceroute campaigns performed on different measurement platforms: Caida's Archipelago (Ark) [2], Looking Glasses (LG) [1], and Ono [3] and Dasu [4] from Northwestern University. Given the increasing importance of Internet eXchange Points (IXP) for Internet connectivity, we focus our analysis on IXPs, selecting subsets of these traceroute data that traversed one of the world's largest IXP (i.e., DE-CIX in Frankfurt, Germany).

Figure 1a shows the peering fabric as seen exclusively by Dasu. For this figure, we first placed the (known) member ASes of this IXP at fixed positions around a circle. We then identified the unique IXP peering links in the

---

(1) AT&T Labs-Research, Florham Park. E-mail: {erg,bala,walter}@research.att.com.
(2) Northwestern University, Evanston. E-mail: {fabianb,msanchez}@eecs.northwestern.edu.
First IMC Workshop on Internet Visualization (WIV 2012), November 13, 2012, Boston, Massachusetts, USA.



(a) Looking in | (b) Looking in with bundling | (c) Looking out

Fig. 1 DE-CIX peering fabric as seen exclusively by Dasu traceroutes: (a) IXP peering fabric; (b) Same as (a) but with bundling; (c) The AS-level Internet as seen by traceroutes that traversed the "important" IXP peering links.

traceroute samples included in the Dasu dataset, and plotted a straight line between the pairs of ASes involved in each peering.

A visually more appealing view of the same data that still maintains much of the peering-specific semantics is shown in Fig. 1b where some form of bundling is applied to the individual peering links in the first figure. The bundling keeps together peerings that involve member ASes that are nearby on the circle and indicates by color (from very light blue to darker shades of blue to red) whether only a few, many, or a large number of peering links are bundled together. While the use of simple bundling methods already enables fine-grained comparions of different inferred peerring fabrics, applying more sophisticated bundling techniques can shed light on the peering policies of the member ASes at this IXP (e.g., open vs. restrictive/selective policy).

Alternatively, instead of focusing on the peering links between member ASes at this IXP, we can also consider the entire AS-level path from source to destination AS along which each of our traceroutes traveled. If a given AS-link is traversed by multiple traceroute samples, we weight the link by the number of traceroute samples that traveled along it.

Using our current platform, Figure 1c shows the contributions made by the Dasu dataset alone. Comparing such simple visuals across the different datasets allows for a qualitative assessment of the number of non-IXP ASes involved and the degree of diversity of AS-level traceroute paths produced by the different traceroute campaigns. Moreover, applying appropriate bundling techniques in this context has the potential of revealing why certain ASes or AS links see many more traceroute probes than others. We view comparing such simple visuals (or their bundled counterparts) as a first step towards identifying particular features that beg further explanations and could be explored in more detail within the context of our platform.

References

1. Brice Augustin, Balachander Krishnamurthy, Walter Willinger. IXP Mapping Project. http://www-rp.lip6.fr/ augustin/ixp/lg-list.txt.
2. CAIDA. Macroscopic topology measurements. http://www.caida.org/, June 2010.
3. Choffnes, D. R., and Bustamante, F. E. Taming the torrent: A practical approach to reducing cross-ISP traffic in peer-to-peer systems. In *Proc. of ACM SIGCOMM* (2008).
4. Dasu. ISP characterization from the network edge. http://www.aqualab.cs.northwestern.edu/projects/dasu.