Consolidated Review of

Studying Interdomain Routing Over Long Timescales

1. Strengths

This paper proposes a new metric that will allow us to better analyze and understand the nature of interdomain routing changes on the Internet under different circumstances. One of the most significant findings is the technique that can disambiguate intentional routing changes from routing churn. The finding that churn is stationary is also quite interesting

The approach presented by the authors is novel (even though it builds on a metric presented last year at IMC 2012) and looks promising, offering a different angle in the study of inter-domain routing evolution and state (e.g., identification of events).

The paper has some interesting observations.

2. Weaknesses

While RSD and the new defined metric MRSD are definitely interesting, the authors didn't show very many concrete applications of this metric to study Internet routing.

The metric is relatively straightforward and doesn't really reflect any particular insight. It's nice that the paper takes the trouble of defining a metric, and hopefully others will use it, but the definition isn't particularly enlightening. Code is not currently available. Given the definition of the metric, it would be tremendously helpful to have a code reference. The paper is heavy on math/notation and lighter on intuition. Stylistically, given the simplicity of the metric, it would be helpful to see some more intuition.

The paper ends up being a slight variation on the authors' earlier work (two papers on RSD), measuring similarity across time rather than across destinations. From the title and beginning of the paper, I was expecting to learn a lot more about the types of changes that happen on the Internet over time. It seems the authors could go deeper into the topic and write a longer, more interesting paper in the future. The paper doesn't characterize the types of changed observed, which seems like it would be interesting. The paper doesn't start from a strong statement of a problem / motivation, so it is hard to assess. The intro talks about moving away from answering specific questions or from analysis tied to the topology, but I wasn't convinced that these more general questions taught us much about the Internet.

3. Comments

This paper provides a nice addition to the authors' prior long IMC paper [7]. This paper is clear, concise, well written, and makes a small but meaningful contribution over the previous work on RSD. The new metric, TSRD, will prove extremely useful to researchers who want to make more sense of routing changes, and in particular to separate churn from sustained updates. In my view, this is a perfect six-pager

The paper reports: "The first striking aspect of these time series is that they are all are approximately stationary." (Page 4). Shouldn't it be expected that growth of the Internet would have very little local effects? (Since next-hop is, by definition, a local effect, dominated by my choices in providers and peering.) A possible error (or confusing wording): page 4 says: "From this curve we can see that in a time window of 2 years, approximately 50% of the routing decisions persisted (because 50% changed)." By "routing decisions" you mean NEXT HOPS, not complete routing paths, right? In figure 3c, what are standard deviations? (Are these results stable?)

Given that presumably you would like other researchers to use and build on your metric, I think a code release is in order—something that takes routeviews or RIPE RIBs (or updates?) as input and computes TRSD for different subgroups could be tremendously useful. Please release the code!

It is clear that it's a preliminary work, though some more hints and examples of the application of this metric in the conclusion section would make the paper more appealing. Could the peak in late 2012 be related to changes in routing due to the impact of the hurricane sandy?

Many of the time series plots do not seem to show any meaningful trends. Perhaps there are more meaningful ways to present some of these results so that they don't look like noise? Figure 3a and 3b seem particularly devoid of trends.

I found some of the observations interesting, although not that surprising (which is ok!). This paper would be more interested once fleshed out more, showing how your metric can help understand instances of interest, and how interesting changes in your metric (like the last paragraph of 5.2, last paragraph of 5.3) point to something interesting and meaningful.

The intro claims that you show the utility of your metric, but I'm not positive what it was. For example, if, instead of using your metric, I count the number of RouteViews updates in a day, will I see a similar day-of-the-week trend as you find? How much information do you lose by considering a set of next hops from an AS, rather than considering the specific next hop from specific points in the AS? It is perhaps an unlikely case, but two destinations could have distance 0 but completely different routes. Or, perhaps more likely, the distribution in how many routers in an AS use path A vs. use path B could change over time. While there are advantages in moving away from quasi-routers, it is not clear how much you give up. It wasn't clear to me how you used all dumps on a day. Did you do anything to exclude any short-term convergence effects, say? Fig 2: Why is average the right measure? What does the rest of the distribution tell us? And if average is the right way to look at things, why is it in 5.1 instead of part of your metric?

5.1: It would help to contextualize the results if I knew what fraction of ASes even have 2+ hops (or at least 2+ observable hops) to choose from. 5.3 could use more discussion of what is going on. What types of changes do you see? When a change happens for one prefix, does it tend to happen for many prefixes? In standard RSD comparing routes to two prefixes at the same point in time, we can assume the topology is fixed. Over time, the topology is changing. It would be neat if your analysis incorporated that, so we could know how topology changes resulted in path changes. You note that your metric won't change just because the topology grows, but it would be neat to know how many of the changes that you do see are due to changes in underlying topology. Fig 3c: How do the high values on Saturday and Sunday relate to the claim of operators avoiding changes over weekends? You say that 5(a) looks similar for other days, but are the ASes that make the most changes similar across days?

4. Summary from PC Discussion

The PC discussed this paper. The consensus was that the contribution was somewhat incremental, but it was significant enough to merit a short paper. The PC had no recommendations beyond what is already in the reviews.

5. Authors' Response

We would like to thank the anonymous reviewers and our shepherd for all the valuable suggestions. Whenever possible we incorporated them in the final version of the paper. Specifically we: fixed typos; followed most of the presentation suggestions; reinforced that our source code is available and how to obtain it; clarified some parts that the reviewers found confusing; computed standard deviation for Figure 3(c); fixed labels of Figure 3(c); and computed the fraction of ASes that have 2+ next-hop choices in a specific day towards a specific destination. We did prefer however to keep our formal notation in order to avoid ambiguous interpretations. We believe that these modifications improved the quality of our work.

We do not have enough evidence (now) that the peak in the late 2012 is related to Hurricane Sandy, so we did not mention this in the text.

We would like to mention that our goal for this short paper was to present our initial set of results with regard to TRSD and that many of the reviewers suggestions and questions about investigations that would be interesting, are in fact ongoing/future work, which we stated in the concluding remarks.