Detecting Unusually-Routed ASes: Methods and Applications

Giovanni Comarela
Boston University
gcom@bu.edu

Evimaria Terzi
Boston University
evimaria@bu.edu

Mark Crovella
Boston University
crovella@bu.edu

ABSTRACT

The routes used in the Internet's interdomain routing system are a rich information source that could be exploited to answer a wide range of questions. However, analyzing routes is difficult, because the fundamental object of study is a set of paths. In this paper we present new analysis tools – metrics and methods – for analyzing AS paths, and apply them to study interdomain routing in the Internet over a recent 13-year period. Our goal is to develop a quantitative understanding of changes in Internet routing at the micro level (of individual ASes) as well as at the macro level (of the set of all ASes). To that end we equip an existing metric (Routing State Distance) with a new set of tools for identifying and characterizing unusually-routed ASes. At the micro level, we use our tools to identify clusters of ASes that have the most unusual routing at each time (interestingly, such clusters often correspond to sets of jointly-owned ASes). We also show that analysis of individual ASes can expose business and engineering strategies of the organizations owning the ASes. These strategies are often related to content delivery or service replication. At the macro level, we show that ASes with the most unusual routing define discernible and interpretable phases of the Internet's evolution. Furthermore, we show that our tools can be used to provide a quantitative measure of the 'flattening' of the Internet.

Keywords

Path-based network, Interdomain routing

IMC 2016 November 14-16, 2016, Santa Monica, CA, USA

ACM ISBN 978-1-4503-4526-2/16/11..

DOI: http://dx.doi.org/10.1145/2987443.2987478

1. INTRODUCTION

Many networks exist in order to provide paths between network nodes, such as highway, air travel and data networks; a prime example is the Internet at the Autonomous System (AS) level. In this case, the complete network structure is not readily available; instead, what is available is a *path-based* network, i.e., a set of nodes and set of paths describing routes taken between nodes.

Path-based networks capture more complex relationships between nodes than do the pair relations in usual networks. Hence, understanding the nature of paths and how paths change over time is more challenging, but is nonetheless potentially valuable in understanding how networks are structured and used.

To treat the Internet as a path-based network, we identify each AS with a node, and associate each prefix with its announcing AS. Each AS is managed by a single organization, which uses the AS for its business purposes; these can include providing Internet access or connectivity, content or cloud service, infrastructure services, or other services. Based on these business goals, organizations make peering and routing decisions that constrain the set of paths over which data flows in the network. Hence, the set of ASes and AS paths can be seen as a path-based network, in which the paths are determined by algorithmic computations realized in BGP (the Border Gateway Protocol) combined with AS policies driven by business strategies. Because path construction in the Internet is a complex process, driven by commercial as well as engineering concerns, it is both important and difficult to fully characterize.

In this paper we develop new tools – metrics and methods – for analyzing path-based networks. Our questions can be broadly grouped into *micro* and *macro* levels. At the micro level, we are interested in identifying and understanding *unusually-routed* ASes – reached through a set of paths that can be considered *unusual*, when compared to the remaining of the network. Furthermore, we are interested in how business and engineering decisions of individual ASes are reflected in their decisions to adopt unusual routing structures.

We develop tools for answering these questions, and using them we show that unusually-routed ASes are

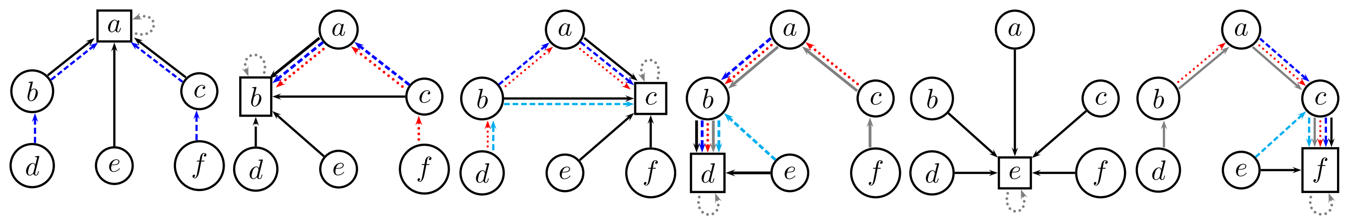

Figure 1: A path-based network over nodes $V = \{a, b, c, d, e, f\}$, broken out to separately show the paths towards each square node. A path is a sequence of directed edges with same line style. Examples of path sets are: $P(f, d) = \{fcabd\}$, $P(b, b) = \{b\}$, and $P(d, c) = \{dbac, dbc\}$. Those path sets entail respectively the following next-hop multisets: $N(f, d) = \{c\}$, $N(b, b) = \{b\}$, and $N(d, c) = \{b, b\}$.

very likely to be economically important. Further, we show that these unusually-routed ASes form *clusters*, and that the clusters often consist of ASes owned by the same organization. When an AS is unusually-routed, it tends to be highly-connected. However, the unusual aspect of such an AS (or group of ASes) is not simply that it is highly-connected (a network property); it is that paths to the AS do not make typical use of the Internet's hierarchy (a path property). Our analysis of individual ASes shows that there are a variety of reasons why an AS may employ unusual routing, and allows us to track how individual ASes adopt unusual routing strategies over time, including infrastructure build-out for content delivery and anycast.

Driven by the results at the micro level we ask questions related to the macro level as well. Here we seek to understand the high-level evolution of the set of all paths in the network over time. We approach this two ways. First, we start from the observation that unusually-routed ASes are often significant Internet businesses. Hence, we identify the set of unusually-routed nodes at each point in time, and use those sets to identify phases of Internet evolution. We show that a segmentation of unusually-routed AS sets yields five phases over 13 years. Digging into the kinds of ASes in each phase, we see that the most unusually routed ASes have shifted over time: from those delivering network operations services, to those involved in content delivery, to those involved in cloud services and domain registry.

The second question at the macro level is driven by the observation that many (but not all) unusually-routed nodes are contributing to the so-called 'flattening' of the Internet [20, 28, 17]. This is the trend of major ASes to move away from hierarchical routing and towards more mesh-like routing. To explore this, we develop metrics to measure quantitatively the process of Internet flattening over time. We show evidence suggesting that Internet flattening has been taking place fairly consistently over the 13-year period of our study, predating the first reports in the literature. We back up our measurements with theoretical and simulation results, and we discuss their robustness due to missing data (i.e., unknown AS-links).

Finally, in addition to the results presented in the paper, in [3] we make available supplementary material, including the proof of two propositions and more extensive documentation of our results, which due to space limitations are not in this manuscript. It also contains instructions for obtaining code and data to reproduce our results.

2. DEFINITIONS

In this section we present the basic definitions that constitute the basis for the tools and methods developed in the next sections. To that end, we borrow RSD (Routing State Distance) from [23], and we equip such concept for cases where next-hops from a source to a destination are not unique. Such modification enables the study of a wider class of path-based networks, at different granularities and over time. Therefore, a contribution of the present work is to show how RSD can be used as part of a larger tool set to take on a new class of problems.

2.1 Path-based networks

The set of paths used in interdomain routing can be abstracted as a *path-based network*. We define a path-based network G as a pair (V, P), where V is a set of nodes and P is a function that, for $a, b \in V$, maps (a, b) to a set of paths $P(a, b)$. Each element $p \in P(a, b)$ is a sequence of nodes from V, forming a directed path starting at a and finishing at b. If $p = a \to v_1 \to \cdots \to v_p \to b$, we say that the *next-hop* of a towards b is v_1; and if $p = a$ we say that a is the next-hop to itself. Hence, the set $P(a, b)$ entails a multiset of next-hops $N(a, b)$, representing the diversity of local decisions made at a in order to reach b. Figure 1 presents an example of path-based network.

Structurally, a path-based network is more complex than a network. A network is a two-way relationship among nodes, and may be encoded in a set of pairs (v_1, v_2), denoting that vertices v_1 and v_2 are connected by an edge. In contrast, path-based network is a three-way relationship among nodes; a set of paths may be encoded in a set of triples (v_1, v_2, v_3) stating that v_3 is a next-hop taken from v_1 on the path to v_2. For instance, in Figure 1, b has a direct connection to reach node c, but b cannot use this connection to reach f.

A path-based network is a particularly suitable tool for modeling the interdomain routing system, because the set of paths used is hard to describe concisely, and because data is constrained to flow only over specific paths. In this context, we consider that the set of nodes is formed by ASes and IP prefixes (as destinations), and the path sets are formed by the AS-paths leading to the AS originating the prefixes. In this case the next-hop representation takes into account the fact that the same AS can choose different next-hops toward the same prefix. At the AS granularity, we model ASes as nodes, and the next-hops as the multiset union over the next-hops toward each prefix originated at the AS. Thus, if a and b are ASes, $N(a, b)$ characterizes the diversity of local decisions of a in order to reach b. In other words, it naturally considers that ASes are not atomic structures, and that by a variety of reasons, there can be different paths from a source to a destination [32, 31].

A limitation of our modeling approach (driven by characteristics of currently available datasets) is that it labels prefixes associated with caches deployed at an access network as originated by the access network, even though the addresses may be related to services offered by other companies. This limitation and its impact on our results are subjects of future work.

2.2 RSD

Additional background to the methods we use in this paper is a metric called *routing state distance* (RSD), which is defined for a path-based network. RSD conceptually measures the dissimilarity of two nodes in terms of *how they are reached* in a path-based network.

Definition 1. *Given $Z \subseteq V$, the Routing State Distance between two network nodes x and y is defined as*

$$\mathrm{RSD}(x, y) = \frac{1}{|Z|} \sum_{z \in Z} d(N(z, x), N(z, y)), \qquad (1)$$

where d is a dissimilarity measure over multisets that assumes values between 0 and 1.

RSD values range between 0 and 1, and if d satisfies the triangle inequality, so does RSD. For d we use the Generalized Jaccard Distance (GJD) between the vectors of frequencies of each next-hop in $N(z, x)$ and $N(z, y)$ normalized by the number of prefixes originated at x and y respectively. GJD has been studied and used in many contexts and is a metric (see e.g., [12]). For two real and non-negative n-dimensional vectors u and w (given that u and w are not simultaneously null vectors), GJD is defined as follows:

$$\mathrm{GJD}(u, w) = 1 - \frac{\sum_{i=1}^{n} \min(u_i, w_i)}{\sum_{i=1}^{n} \max(u_i, w_i)}. \qquad (2)$$

Taking Figure 1 as example, for destinations a and b, the sources d and f completely agree on their next-hops choices toward both destinations; sources a, b and e completely disagree; and source c partially agrees (with $d(N(c, a), N(c, b)) = 0.5$). Therefore, $\mathrm{RSD}(a, b) \approx 0.58$.

Unfortunately, in practice it is not always the case that we can have access to full visibility of the network. Hence it is not possible to obtain $Z = V$ in order to compute RSD. In our work we choose Z such that $N(z, x) \neq \emptyset$ for all $z \in Z$ and $x \in V$; we call Z the set of *sources*. One way of describing $\mathrm{RSD}(x, y)$ is that it approximates the *probability that an arbitrary node in Z uses different next-hops* on the paths to x and y.

Finally, although in this work we use GJD as the distance function d in Equation (1), it is possible to use different distances for different applications. For instance, one can incorporate weights related to other sources of information (e.g., traffic volume and prefix popularity), if it is available.

3. DATA PREPARATION

In this section we describe the data used in our analysis and how we adapt it to measure RSD consistently. We obtain Internet AS paths from two sources: the RIPE Routing Information Service (RIS)[1] and the Route Views[2] projects. We assess the potential impact of missing data (specially AS peering links) in Section 8.

Data acquisition: From each project we obtain all available RIBs (Routing Information Base) for the first three days of each month from January 2003 to December 2015. Each entry in each RIB provides one AS-path record of the form [t; $AS_1, \ldots, AS_s$; p], for day t, AS-path $AS_1, \ldots, AS_s$, and IP prefix p (IPv4 only). The choice of the three first days is arbitrary, and we expect results to hold if one consistently uses any three consecutive days of the month. Considering much longer periods of time can lead to complications when trying to obtain a good approximation of the state of the system (for a given time), thus introducing noise in the results.

Next-hop determination: Each AS-path record provides $(s-1)$ next-hop records of the form [t, AS_i, AS_{i+1}, p, AS_s], $i = 1, \ldots, s-1$. Such record means that at day t, AS_i uses AS_{i+1} as next-hop towards prefix p, which is originated at AS_s. All tuples with $AS_i = AS_{i+1}$ are filtered out. This processing allows us to cover cases where a prefix is originated at more than one AS.

Instability filtering: For each month, we remove next-hop records [t, x, y, p, o] for which (x, y, p, o) appears in only one of the three snapshots of the month. This lowers the influence of short-term routing changes on our data. Each day t is converted to its corresponding month, and resulting duplicate records are removed. We refer to the portion of the resulting dataset for any given day as a single *snapshot*. There are a total of 156 snapshots covering the 13 years of our study.

[1]www.ripe.net/data-tools/stats/ris/
[2]www.routeviews.org

Prefix assignment: Each prefix is assigned to its originating AS. Thus, all next-hop records are converted from $[t,\ x,\ y,\ p,\ o]$ to $[t,\ x,\ y,\ o]$, (i.e., we remove the destination prefix, but keep the destination AS that originates it). Next, all next-hops having the same time, source and destination are aggregated. This yields records of the form $[t,\ x,\ N(x,o),\ o]$, where $N(x,o)$ represents the multiset of next-hops from x towards o. These records define the next-hop function $N(\cdot,\cdot)$ used in Equation (1). Observe that next-hops are still constrained to whole ASes. It is possible to consider finer granularities, as in [23], but as we argue in [16], that approach introduces several complications for temporal analyses.

Selection of sources and destinations: Consistent computation of RSD requires that we know the value of $N(x,o)$ for all x and o over some domain. Accordingly, we heuristically looked for large AS sets Z and V, denoting sources and destinations respectively, such that every $x \in Z$ had at least one next-hop towards all $o \in V$. As a result, each of the 156 snapshots yield sets Z that range in size from 37 (Jan 2003) to 183 (Dec 2015) and sets V that range in size from 14051 (Jan 2003) to 51202 (Dec 2015). For each snapshot the corresponding Z is used as in Equation (1).

Stable source selection: For the static analyses in Section 4 the above data is sufficient. However, for the dynamic analyses in Sections 5-7, variation in the membership of sets Z introduces noise into the results. To reduce the effects of ASes that appear only infrequently in source sets Z, we keep only sources appearing in at least 78 (half) of the 156 snapshots. After this stability filtering the size of Z ranges from 21 (Jan 2003) to roughly 70 in the end of 2015.

4. UNUSUALLY-ROUTED ASes

We start our investigation by exploring the nature of the most unusually-routed Autonomous Systems at any given time. In this section, we develop methods for identifying unusual nodes in a fixed snapshot of the network, and illustrate the value of the methods.

The key idea behind our approach is as follows: we fix the concept of an 'unusually-routed AS' as meaning one for which the set of paths leading to the AS is very different from the set of paths leading to any other AS. A natural similarity measure over paths is defined in terms of the set of next-hops comprising a path. Keeping in mind the view of RSD as 'the probability that an arbitrary node in Z uses different next hops to reach the two destinations' we can formalize an 'unusually-routed AS' as one that *has high RSD to all other ASes.* As an example, one can observe that paths leading to node e in Figure 1 are very different from paths leading to any other node; such fact is captured by RSD, which is 1 (maximum value) between e and any other node. Hence, we say the e is unusually-routed in that path-based network.

Although such reasoning leads to an operational definition of unusually-routed ASes, an important complication arises because an organization may operate multiple ASes. Hence there may be multiple ASes with similarly-unusual patterns of reachability. This necessitates a search for *clusters* of unusually-reached ASes.

4.1 Problem definition

To identify unusually-routed clusters, we seek to find a group of k non-overlapping subsets of V, say $C_1, \ldots, C_k$, meeting three requirements:

1. each C_i is a small set;
2. the elements of a C_i are close to each other, as measured by RSD; and
3. the elements of a C_i are all far from all the elements not in C_i, as measured by RSD.

We formalize these requirements by considering the complete weighted undirected graph H, whose nodes are V and whose edges connect each pair of nodes (x, y) with weight given by $\mathrm{RSD}(x, y)$. Using H, we define the notions of *join* and *disconnect* of $C \subseteq V$ as follows:

DEFINITION 2. *Given $C \subsetneq V$, the* join of C, *denoted by $J(C)$, is:*

$$J(C) = \min_{x \in C, y \notin C} \mathrm{RSD}(x, y).$$

Intuitively, $J(C)$ is the smallest RSD threshold at which another node joins the subset of nodes in C.

DEFINITION 3. *Given $C \subseteq V$, the* disconnect of C, *denoted by $D(C)$, is:*

$$D(C) = \max_{C' \subsetneq C} \min_{\substack{x \in C' \\ y \in C \setminus C'}} \mathrm{RSD}(x, y).$$

Intuitively, $D(C)$ is the largest RSD threshold that internally disconnects the subgraph defined by C in H.

Note that one should expect that for a cluster C of interest $J(C) > D(C)$; this means that each node in C has a closer connection to another node in C than to any node outside C. Using the above definitions we formalize our search for unusual node clusters as Problem 1.

PROBLEM 1. *Given $G(V, P)$ and integers δ and k, find k disjoint sets $C_1, \ldots, C_k$ that*

$$\begin{aligned} \textit{maximize} \quad & \min_{1 \le i \le k} J(C_i) \\ \textit{subject to} \quad & 0 < |C_i| \le \delta, \quad i = 1, \ldots, k, \\ & J(C_i) > D(C_i), \quad i = 1, \ldots, k. \end{aligned}$$

4.2 Algorithm

A solution to Problem 1 yields k clusters $C_1, \ldots, C_k$ that together satisfy the three requirements. The constraint that $|C_i| \le \delta$ ensures that the clusters are small, and the constraint that $J(C_i) > D(C_i)$ means that elements in a cluster are generally close to each other.

Algorithm 1: `FindUnusual`$(G(V,P),\ \delta,\ k)$

$H \leftarrow$ weighted and undirected graph $(V, V \times V, \omega)$, with $\omega((x,y)) = \mathrm{RSD}(x,y)$
$T_H \leftarrow$ a Minimum Spanning Tree of H
$L \leftarrow$ edges of T_H sorted in non-increasing order
foreach e **in** L **do**
  $T_H \leftarrow T_H$ minus edges with weight $\omega(e)$
  $C \leftarrow$ connected components of T_H
  $\mathcal{C} \leftarrow \emptyset$
  for $c \in C$ **do**
    if $|c| \leq \delta$ **then**
      $\mathcal{C} \leftarrow \mathcal{C} \cup \{c\}$
      if $|\mathcal{C}| = k$ **then**
        return $\mathcal{C}$
return $NULL$

The third requirement is met by maximizing the objective function $\min_{1 \leq i \leq k} J(C_i)$. Observe that the C_i's do not necessarily form a partition of V, only the unusual nodes are clustered.

Problem 1 can be solved by a polynomial time algorithm with complexity $O(|V|^2 \log |V|)$ using algorithm with pseudocode shown in Algorithm 1. Note that the running time computation assumes that all pairwise RSD distances $\mathrm{RSD}(x,y)$ are provided as part of the input.

In a high level, the algorithm first builds the complete, weighted graph H. Then, it computes a minimum spanning tree of H and removes edges from the tree in non-increasing order, until it obtains k connected components smaller than δ. When we solve Problem 1 for a particular value of k we refer to the solution as the *top-k* clusters.

Although Problem 1 is different from single-linkage clustering, Algorithm 1 has similarity to the solution strategy for the single-linkage clustering problem [26]. We have the following result about the output of Algorithm 1.

Proposition 1. *For given $G(V,P)$, k and δ, Algorithm 1 finds the optimal solution of Problem 1.*

Proof. [3] □

4.3 Results

Figure 2 is a dendrogram of the unusual AS clusters for the snapshot of December of 2014, obtained by solving Problem 1 for $k = 40$ and $\delta = 50$. (Varying δ up to 200 did not change the results; large values of k generate many unusual clusters). Groups of ASes that were placed in the same cluster in the solution have been given the same color in the figure (all singletons are shown in blue). The figure illustrates a number of points:

Unusual clusters are important. The figure shows that unusual clusters often correspond to important Internet businesses or organizations. Among other significant organizations, Google, Verisign, Motorola, Microsoft, Facebook, and GoDaddy are all represented in the top-40 clusters (note that there are over 47,000 ASes in the dataset). The ASes in the top-40 are quite unusual in terms of how they are reached in the Internet:

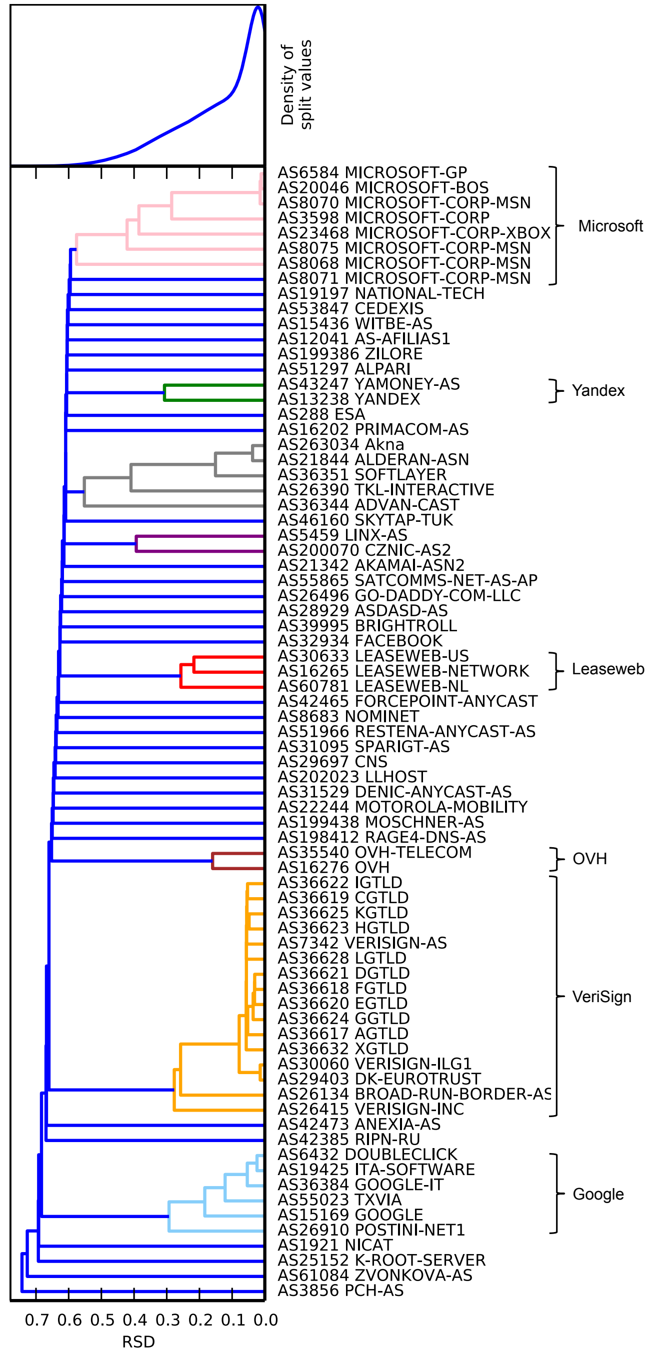

Figure 2: December 2014 dataset. (upper) Density of dendrogram split values across the dataset; (bottom) Dendrogram of the top-40 clusters of unusual ASes. AS descriptions are shortened from [2].

the x axis shows that the minimum RSD between these clusters and any other nodes is in the range 0.6 - 0.74. Comparing to the density plot of split values across the entire dataset, we see that these values are quite rare.

Clusters often reflect organizational boundaries. Figure 2 also shows that clusters of ASes are often owned by the same company; only two clusters consists of ASes that are not completely co-owned. The remaining clusters (Google, Verisign, OVH, Leaseweb, Yandex and Microsoft) all consist of separate ASes that are owned and operated by the same company. For example, the cluster marked 'Google' includes ASes that are nominally registered to Postini, TxVia, and DoubleClick – each of which reflects a prior acquisition by Google. In fact, identifying ASes owned by the same organization is important in analyzing the Internet for business and political reasons [9]. These results suggest that clustering ASes using routing information may be another strategy for inferring co-ownership. Intuitively, when unusual routing strategies arise in groups, it is unlikely to be a coincidence – more likely, the participating ASes are seeking to achieve common business or engineering goals because, for instance, they are owned by the same organization.

Why unusual clusters are unusual. Each AS cluster in Figure 2 has a distance greater than 0.6 from *any other* AS in the dataset. Under our intuitive interpretation of RSD, this means that given an AS x in one of these clusters, at least 60% of the time an arbitrary AS $z \in Z$ chooses a different next-hop as the next step on the path towards x as compared to *any* other AS in the Internet.

To understand how this can occur, one relatively simple case is when the next-hop from z towards x is x itself – i.e., x is 'its own next-hop.' In fact, this can be a good approximation for certain unusual ASes (but not all, as we show later). For instance, in the group of Google ASes, AS15169 (the main Google AS) represents more than 50% of the next-hops towards any AS in the group. In other words, a large fraction of the source ASes in our study, when exchanging data with Google, do so by *directly connecting* to Google. We return to analysis of Google in the next section; here we just note that these results are consistent with a number of other studies [36, 13, 10]. We observed path properties similar to Google's, if not as extreme, for a number of other unusual clusters.

Reviewing the set of organizations in Figure 2, a number of different business goals are evident. Furthermore, we have performed cluster analysis using Algorithm 1 for each month of our 13 year study, and the results (refer to [3] for the extra results, not presented here due to space considerations) show that various organizations have come to adopt unusual routing strategies at different times in the past. This motivates developing tools to analyze how an individual AS's connection and routing strategy evolves over time, which we present in the next section.

5. INDIVIDUAL ASes

When analyzing an individual node, we are concerned with understanding its *path set* – the set of all paths leading to the node. An AS's path set is a useful reflection of its business and engineering strategies. We analyze the strategies employed by individual ASes over time by asking:

1. How does the 'unusualness' of a node vary over time?
2. When are a node's path sets stable, and when are they in flux?
3. What are the key path characteristics that characterize a node's path set?

To answer these questions we make use of three tools, and introduce time-indexing on quantities of interest. First, in order to quantify how unusual a node x is when compared to the network as a whole, we define *average RSD*, denoted $\Delta_t(x)$.

Definition 4. ***(Average RSD)*** *Given a destination x and a time t, define*

$$\Delta_t(x) = \frac{1}{|V_t|-1} \sum_{y \in V_t, x \neq y} \mathrm{RSD}_t(x,y) - \frac{1}{|V_t|(|V_t|-1)} \sum_{y,z \in V_t, y \neq z} \mathrm{RSD}_t(z,y), \tag{3}$$

which measures how far the average RSD between x and the other nodes is from the global average RSD of the network. Next, to identify periods of time when path sets are stable or in flux, we use *temporal RSD* (as defined by the authors of [16]) which we denote by $\tau_{t,t'}(x)$.

Definition 5. ***(Temporal RSD)*** *Given a destination x and two points in time t and t', define*

$$\tau_{t,t'}(x) = \frac{1}{|Z_{t,t'}|} \sum_{z \in Z_{t,t'}} d(N_t(z,x), N_{t'}(z,x))), \tag{4}$$

where the distance function d is the same as used in Equation (1) and $Z_{t,t'} = Z_t \cap Z_{t'}$.

Finally, in order to examine a node's path characteristics in detail, we define *next-hop distribution,* denoted $\eta_t(x,y)$.

Definition 6. ***(Next-hop Distribution)*** *Given network nodes x and y, and time t, we define $\eta_t(x,y)$ as the fraction of times that x appears as a next-hop towards y at time t:*

$$\eta_t(x,y) = \frac{\sum_{z \in Z_t} \textit{frequency of } x \textit{ in } N_t(z,y)}{\sum_{z \in Z_t} |N_t(z,y)|}. \tag{5}$$

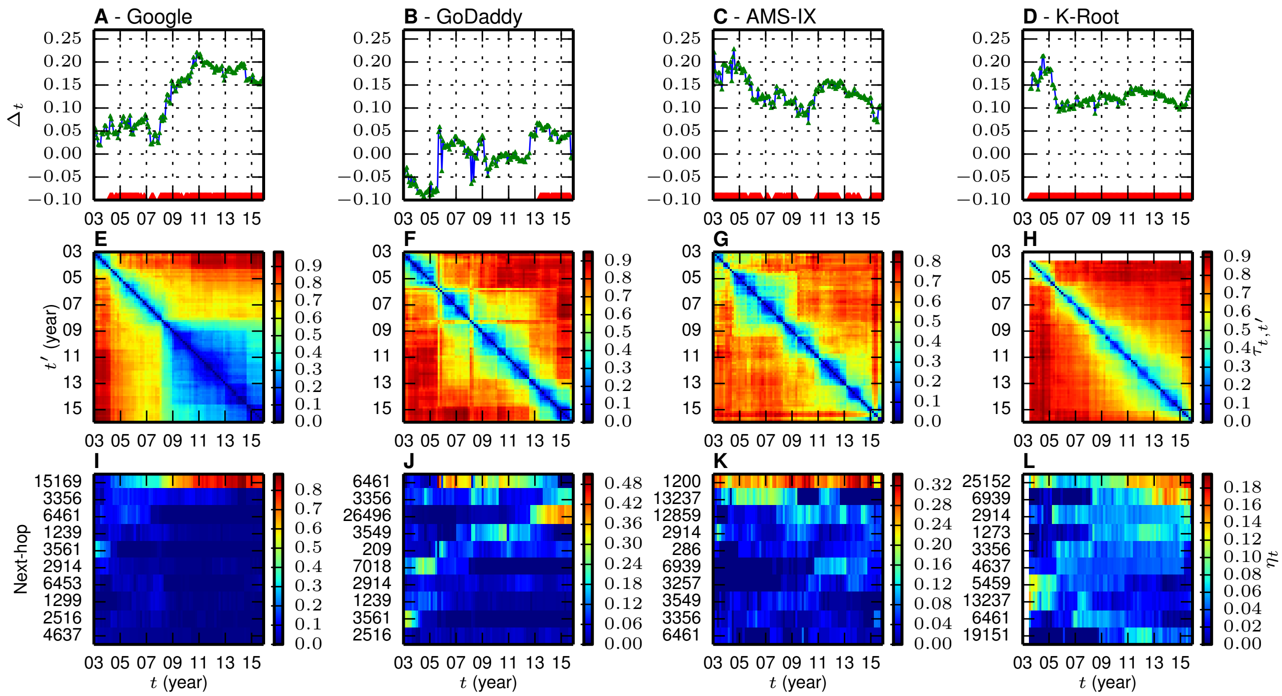

Figure 3: Columns correspond to ASes operated by Google (AS15169), GoDaddy (AS26496), Ams-IX (AS1200) and K-Root Server (AS25152). The top row shows Average RSD (Δ_t) time series; the red triangles at the bottom indicate when the AS was in one of the top-40 groups of unusual ASes. The middle row shows Temporal RSD ($\tau_{t,t'}$) comparing all possible pairs t and t' over the 13 years of study. The bottom row shows next-hop distribution (η_t) of the top 10 next hops used to reach the corresponding AS.

We use four ASes as case studies, operated by Google (AS15169), GoDaddy (AS26496), Ams-IX-Amsterdam (Ams-IX, AS1200) and the K-Root Server (AS25152). We choose these four ASes to illustrate key features seen across the entire dataset. (A more extensive analysis of 1000 ASes is presented in [3]). Together the organizations operating these four ASes conduct a very wide range of activities, but each AS is in the top-40 unusual set at certain times. Results for these ASes are presented in Figure 3, and are discussed individually below.

Google. Founded in 1998 as a search engine, Google became a major content provider on acquisition of YouTube in October 2006. Figure 3A shows that this acquisition coincided with the beginning of a dramatic increase in its average RSD. This increase is understood via Figure 3I, which shows that the fraction of ASes whose next-hop on the path to Google is Google itself began to increase starting in late 2006, and has increased steadily to the present. In the most recent dataset, over 80% of the sources we studied connect directly to Google and use it as the next-hop towards Google prefixes.

The analysis shows that during the post-2006 period Google built out a worldwide network of locations connecting directly with many ASes, bypassing traditional hierarchical Internet routing. Since content delivery (e.g., YouTube video) involves the transfer of much higher volumes of traffic than does search engine service, this build-out was presumably motivated by a desire to avoid the costs of paying transit providers, and to avoid network delay and congestion by being closer to customers. From mid-2008 onward Google appears continuously in the top-40 unusual ASes. Corroborating this view, the Temporal RSD plot in Figure 3E shows two distinct phases, before and after 2008, corresponding to Google's transformation into business that has a major content delivery component.

Besides Google, other companies that are known to be expanding their network infrastructure are also present in Figure 2. These include ASes owned by Microsoft and SoftLayer, agreeing with results in [13]. Amazon, which is also discussed in [13], is occasionally identified as unusual in our results. Using our tools to analyze AS16509 (the main Amazon AS), we observe that the popularity of AS16509 as a next-hop to itself is consistently increasing over time, but that at present Amazon still heavily relies on large transit providers.

GoDaddy. GoDaddy is the world's largest domain registrar, managing the assignment of many of the names held in the DNS. GoDaddy started in 1997 and began a period of rapid growth in early 2005, which included adding content-hosting services for small businesses.

Figures 3B, F, and J illuminate GoDaddy's expansion strategy. On commencing growth in 2005, GoDaddy began expanding its set of connections, leading to a jump in Δ_t and a diversification of the next-hops used to reach AS26496 (visible also in Figure 3F). However, in early 2013, GoDaddy abruptly shifted to an unusual connection pattern, connecting directly to many networks and cutting out the intermediaries that had been dominant in 2005-2013. This is visible in the sharp rise in Δ_t in early 2013, and the sudden dominance of AS26496 as its own next-hop in Figure 3J. This shift suggests an intensive effort by GoDaddy to build new network infrastructure in order to reach its customers over shorter paths. The pre-2013 period is also useful to illustrate how the output of our tools for ASes that do not behave unusually. In such case, one can observe lower RSD and higher dependency on transit providers, when compared to the post-2013 period.

Ams-IX. Ams-IX operates Internet eXchange Points (IXPs), locations established to facilitate connections between ASes. Over 700 ASes currently make use of the Amsterdam IXP as a location to connect to other ASes. Figures 3C, G, and K show that, unlike the previous two examples, routing to the Ams-IX AS (AS1200) has been unusual over almost the entire period of study. Figure 3K shows that this is because, as with the previous two cases, AS1200 is the most common next-hop on paths to itself. The difference in the temporal patterns of Δ_t compared to Google and GoDaddy arise because direct-connection occurs not due to a large infrastructure build-out, but rather to the hundreds of ASes that connect directly to Ams-IX at their IXP. These connections enable Ams-IX to provide coordination and support for the participating ASes.[3]

K-Root Server. As a final example we discuss AS25152, which operates the K-Root Server. This server is an important element of the Domain Name System (DNS), so ensuring the constant availability of the services provided by K-Root is important for smooth operation of the Internet and the Web.

To provide highly available service K-Root uses anycast, which consists of creating multiple hosts with the same IP address and connecting those hosts to the Internet in different locations. The normal action of the routing system then serves to direct any request for service sent to the K-Root IP address to a nearby K-Root host; if a K-Root host fails, the routing system naturally adapts to direct data to other K-Root hosts instead.

K-Root has been operated using anycast since near the beginning of our study period, which explains why it has consistently been an unusual AS. In contrast to the three previous cases, the unusual routing towards K-Root is not because its AS is its own most common next-hop (as can be seen in Figure 3L); K-Root is only connected in less than 20 locations at present. However the connection of the K-Root AS at multiple locations means that it is not "near" any other *single* AS from the perspective of a majority of other ASes.

This shows that anycast generates routing patterns that meet our definition of unusual. Hence it suggests that these methods might be useful for detecting anycasted ASes in general, with the advantage of being lightweight and requiring only passively collected information (routing tables). As future work we intend to evaluate such possibility and compare the results with other works, e.g., [15] and [14]. For instance, initial investigation showed a considerable overlap between ASes listed on Figure 2 and the results of [14].

Taken together, these case studies paint a varied picture of why and how an organization decides to make use of an unusual routing strategy for its ASes. In some cases, such as Google and GoDaddy, the goal is to build out a worldwide infrastructure to provide high performance, high reliability, or both. In other cases, the need to provide coordination services for many other ASes (at an IXP) leads to unique paths to the coordinating AS. And yet another reason for adopting unusual routing is when using anycast, which does not lead to the announcing AS using paths similar to any other particular AS in the view of most other ASes.

6. EVOLUTION OF UNUSUAL ASes

Section 4 showed that unusually-routed ASes are often significant Internet organizations, and Section 5 showed that organizations make distinct decisions to adopt unusual routing strategies at particular times. This suggests that macro-level insights into the overall business and engineering goals driving Internet routing can be obtained by examining the sets of unusually-routed ASes over time.

We characterize the evolution of unusually-routed ASes as follows. We consider the set of unusual nodes at a given time to be the union of the top-k clusters returned as the solution to Problem 1. We collect all such sets over time into an $m \times T$ binary matrix $\mathbf{X}$, where m is the total number of unusual nodes over all timesteps, $T = 156$ is the number of timesteps, and $x_{it} = 1$ if and only if node i is unusual at time t.

Our initial results suggest that the unusually-routed ASes in the Internet have formed a set of distinct phases over time. This evidence is shown in Figure 4A. For each snapshot we construct its set of unusual ASes, based on the top-k clusters for $k = 45$, leading to the dataset $\mathbf{X}$ (details on how k is selected are described later in this Section). Figure 4A then plots for each pair of unusual sets A and B, the overlap distance ($O_d(A, B)$), representing the fraction of the smaller set that is not contained in the larger set. It is possible to identify a noisy diagonal block structure in the heat map, suggesting that the sets of unusual ASes have gone through

[3]More details at https://ams-ix.net/technical/specifications-descriptions/as1200-peering.

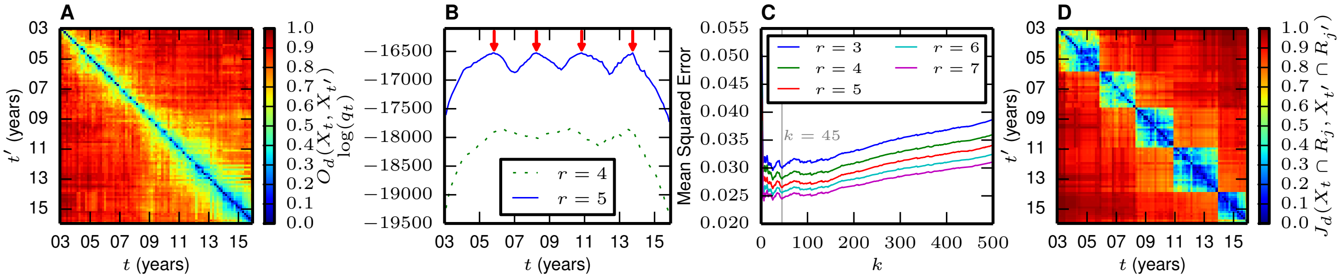

Figure 4: Phase-based analysis of unusual ASes for the past 13 years. Each month t is characterized by its set of top-45 unusual AS clusters, denoted by X_t. (A) O_d, overlap distance between sets of all unusual ASes at different times. (B) Segmentation scores $\log(q_{r,t})$, showing the relative quality of t as a segment boundary. Red arrows show actual segment boundaries determined by Problem 2. (C) Mean Squared Error of the segmentation model, given the data for different r and k values. (D) J_d, Jaccard distance between sets of unusual ASes after using segment representatives to filter noise.

periods of stability that have alternated with short periods of more rapid change.

This leads to the following goals, which drive the analyses in this section: we seek to find a good segmentation of the sets of unusual ASes over time, and we seek to use that segmentation to understand evolution of routing strategies among important Internet organizations.

6.1 Problem definition

We say that $\mathbf{X}$ has a good r-segmentation if we can partition the columns of $\mathbf{X}$ into r segments, in which any two columns of the same segment of $\mathbf{X}$ show strong similarity.

To formalize this problem, let S be a set of $r+1$ segment boundaries $1 = s_0 < s_1 < \cdots < s_r = T+1$, and $\mathbf{P} \in [0,1]^{m \times r}$, where p_{ij} represents the probability that $x_{it} = 1$ when $s_{j-1} \leq t < s_j$. If $\mathbf{X}$ has a good r-segmentation, much of the information in $\mathbf{X}$ can be captured by a model $\mathcal{M} = (S, \mathbf{P})$. Hence we segment $\mathbf{X}$ by finding a model $\mathcal{M}$ that has high likelihood. Under the assumption of independence of the x_{it}'s the task of finding a good $\mathcal{M}$ is formalized in the following problem definition:

Problem 2. *Given an integer $r > 0$ and an $m \times T$ binary matrix $\mathbf{X}$, find a model $\mathcal{M} = (S, \mathbf{P})$ that maximizes the likelihood function of $\mathcal{M}$ given by:*

$$\mathcal{L}(\mathcal{M}|\mathbf{X}) = \prod_{j=1}^{r} \prod_{t=s_{j-1}}^{s_j - 1} \prod_{i=1}^{m} (p_{ij})^{x_{it}} (1 - p_{ij})^{1 - x_{it}}. \quad (6)$$

Hence we cast the problem of finding a good r-segmentation to the problem of maximizing $\mathcal{L}(\mathcal{M}|\mathbf{X})$, in which the latent variables are the $r-1$ segment boundaries and the probability matrix $\mathbf{P}$.

6.2 Algorithm

Our algorithm for solving Problem 2 is adapting the standard dynamic-programming recursion [7] for segmenting the sequence of observations stored in $\mathbf{X}$ to the likelihood function we describe in Equation (6). The algorithm has running time $O(T^2(r+m))$, and since T is small (156), execution takes only a few minutes.

An important step in identifying a good r-segmentation for $\mathbf{X}$ is the choice of r. To assess the quality of a solution to Problem 2 for a given r, we ask whether the computed segment boundaries are sharp. We compute a score $q_{r,t}$ that indicates the relative quality of all r-segmentations that contain t as a segment boundary. More specifically,

$$q_{r,t} = \frac{\sum_{S \in \mathcal{S}_{r,t}} \max_{\mathbf{p}} \mathcal{L}((S, \mathbf{P})|\mathbf{X})}{\sum_{S \in \mathcal{S}_r} \max_{\mathbf{p}} \mathcal{L}((S, \mathbf{P})|\mathbf{X})}, \quad (7)$$

where $\mathcal{S}_r$ is the set of all possible segmentation boundaries of $1, \ldots, T$ in r segments, and $\mathcal{S}_{r,t}$ is the set of all possible segmentation boundaries in which t is present as a boundary. Intuitively, the presence of distinct peaks in $q_{r,t}$ that correspond to the boundaries obtained by solving Problem 2 suggests that the obtained segment boundaries are well-localized and distinct.

Although computing $q_{r,t}$ requires going over *all* possible segmentations of the data into r segments, we adopt techniques from [27, 37] to compute the values of the above scores in polynomial time using dynamic programming. In fact, the running time of the algorithm for this task is the same as the time required to compute the optimal segmentation ($O(T^2(r+m))$).

6.3 Results

Computing $q_{r,t}$ for different values of r suggests that $r = 5$ is the smallest r with sharp segment boundaries. Figure 4B shows the $q_{r,t}$ for $r = 4$ and 5, showing as an example that segment boundaries for $r = 4$ are not as distinct as they are for $r = 5$. To choose k we examined the MSE (Mean Squared Error) of the model $\mathcal{M}$ given the data $\mathbf{X}$. As showed in Figure 4C, increasing r always decreases the MSE, but across all curves there is a clear pattern change after $k = 45$, with the MSE being minimum (or close to) at that point.

For $k = 45$ and $r = 5$ the segment boundaries obtained by solving Problem 2 are December 2005, May 2008, December 2010 and November 2013. Inspection

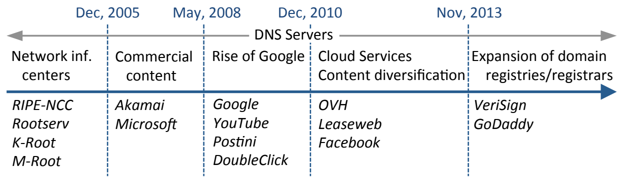

Figure 5: Segment representatives reflecting evolution of unusually-routed ASes.

of the resulting segments shows that the dominant unusual ASes differ across segments. To demonstrate this, we remove all ASes that do not appear in more than half of the snapshots of at least one segment. This has the effect of filtering noise due to ASes that only infrequently appear as unusual. While $m = 1192$ distinct ASes appear as unusual at some point over the 13 years of the study period, only 104 ASes[4] appeared in at least half of the snapshots of a segment. We define the *representative* of segment j, denoted by R_j, as the set of ASes that appear in the unusual sets of more than half of the snapshots in segment j.

Figure 4D shows Jaccard distance among the sets of unusual ASes over time, after denoising by intersecting the unusual ASes in each snapshot with the respective segment representative R_j. The figure shows that representative nodes are present consistently throughout each segment, and that there is little overlap in the representative nodes across segments.

To interpret the nature of the phases shown in Figure 4, we manually select some well-known ASes that are characteristic of a segment's representative set and appear in that set for the first time. These are shown in Figure 5, and a reflect a number of aspects of Internet organization over the past 13 years.

First, spanning all segments are DNS root servers that use anycast to ensure high availability and performance. This demonstrates the crucial role that DNS plays in the Internet and the importance of reliable DNS service.

In the period before December 2005, dominant unusual nodes tend to be associated with network operations, as exemplified by DNS servers and network information centers. The period from December 2005 to May 2008 exhibits a shift towards the building of infrastructure for delivery of commercial content. Microsoft and Akamai are notable for their emergence as unusual ASes in this period. The third segment marks a turning point in which Google and its associated businesses – YouTube, Postini, and Doubleclick, among others – launch a very significant buildout of infrastructure. As shown in the previous section, the goal of this buildout was to minimize the use of the Internet as a hierarchy, instead constructing short, direct paths from many of its customers to its network. Although Google has appeared as unusual in earlier segments, one can note that the emergence of Google as representative in this segment is in accordance with the transformations observed through Figure 3E. By December 2010, a number of other large content and cloud providers followed Google's lead, and expanding their own infrastructures, including Facebook and OVH. Finally, after November 2013 a new set of organizations emerge with unusual routing structures, mostly related to the commercial exploration of DNS services, exemplified by Verisign and GoDaddy.

[4]Those ASes are listed on [3].

7. GLOBAL PATH CHARACTERISTICS

Our last set of macro-level investigations concerns the time evolution of global changes to the entire path set of the network. At the highest level, our goal is to measure the extent to which the set of all AS paths in the Internet reflects a shift away from hierarchical and toward mesh-like (flat) routing.

One of the results of Section 5 is that when an AS deviates from hierarchical routing, the RSD between that AS and other ASes increases. In fact, we can formalize this effect using the following proposition:

PROPOSITION 2. *Let $G_1(V, P_1)$ and $G_2(V, P_2)$ be path-based networks, where P_1 and P_2 are composed by shortest paths overlaid on a complete graph and on a tree respectively. Then*

$$\mathrm{RSD}(x, y) = 1, \text{ for } G_1, \text{ and} \tag{8}$$

$$\mathrm{RSD}(x, y) = \frac{1 + \mathrm{treeDist}(x, y)}{|V|}, \text{ for } G_2, \tag{9}$$

where $\mathrm{treeDist}(x, y)$ *is the length of the shortest path (counting edges) between x and y in a tree.*

PROOF. [3] □

Proposition 2 shows that RSD is capable of distinguishing the extremes of hierarchical and flat routing schemes. In particular, the RSD between two nodes assumes its highest value for flat routing: for G_1 the pairwise distances assume value 1. In contrast, RSD values in hierarchical routing are relatively low – roughly $O(\log(|V|)/|V|)$ for G_2 – assuming that the underlying tree of G_2 is balanced.

These observations suggest an approach for characterizing the entire path set of the network in order to assess Internet flattening over time. To that end, we introduce the following metric.

DEFINITION 7. ***(Global RSD)*** *Given a path-based network $G(V, P)$, define*

$$\Delta(G) = \frac{1}{|V|(|V|-1)} \sum_{x,y \in V,\ x \neq y} \mathrm{RSD}(x, y). \tag{10}$$

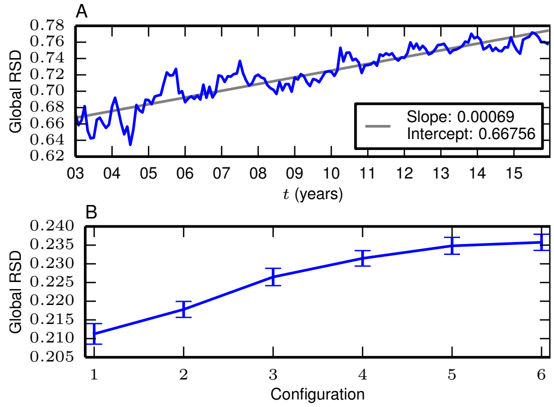

Figure 6: (A) Global RSD over the 13-year period of study. Line is fit using weighted least squares (using the variance of pairwise RSD at each time). (B) Global RSD for the six ITER configurations of our simulation study. Configurations 1 and 6 represent hierarchical and mesh-like structures respectively. Error bars are 95% confidence intervals.

The motivation for Global RSD is that if Internet flattening is taking place, then average RSD values across the entire network should be increasing.

Figure 6A shows Global RSD for each of the snapshots in our dataset. The figure shows a fairly steady growth of Global RSD over the 13 years of our study. According to Proposition 2, such growth is consistent with a shift from tree-like to mesh-like routing. Interestingly, we observe that Global RSD growth had been taking place well before Internet flattening was first reported by [20].

One concern in interpreting Figure 6A is that the idealized routes considered in Proposition 2 differ considerably from the more complex set of routes used in the Internet. To explore whether RSD can reflect flattening in routing patterns more representative of the Internet, we turn to simulation. For this purpose we use ITER [17], an agent-based simulator specifically developed to shed light on the transition of the Internet from a hierarchical scheme to one closer to a peering mesh. ITER uses an agent-based formulation in which ASes individually adjust their connections to other ASes to meet business strategies such as profitability.

The authors in [17] vary three parameters in ITER to study the transition from hierarchical to mesh-like routing. These are: the number of regions that a content provider can span (R), the fraction of traffic generated by content providers (C), and the traffic threshold for peering (α; we refer the reader to [17] for details).[5] The authors of [17] use settings of (R, C, α) equal to (1, 0.1, 1) and (6, 0.6, 10) to study hierarchical and flat routing structures respectively. To study varieties of networks ranging from hierarchical routing to flat routing, we use six parameter settings that interpolate between these two extremes. For each of the six configurations we execute 100 runs of the simulator. Each run results in an AS topology with annotated link relationships (peer or customer-provider). Using these, we employ standard algorithms to infer non-valley-prefer-customer paths [19], and then compute global RSD as in Definition 7.

Figure 6B presents Global RSD averaged over each of the 100 runs. The figure shows that RSD responds smoothly to the transformation of the network from a hierarchical to a flat structure. Thus we conclude that the effect predicted by Proposition 2 extends to the situation in which the set of paths is more representative of actual Internet routing.

Taken as a whole, the results in this section suggest that Global RSD has been growing over the 13 year period of our study. Furthermore, we conclude that Global RSD changes in a predictable way as the set of paths in a network transitions from hierarchical to mesh-like, and that the changes in Global RSD we observe are consistent with reports of such changes taking place in the Internet.

8. DATA CONSIDERATIONS

The data sources for our study (Route Views and RIPE RIS) determine the visibility we have on the Internet's routing structure. A natural question therefore concerns the impact of that visibility on our results. In that regard, we address two questions:

1. The monitoring points for Route Views and RIPE RIS yield full visibility mainly for Internet transit providers. Are such points good locations for observing global Internet evolution?
2. The nature of the Route Views and RIPE RIS monitoring points is such that many unobserved links are of the peering type, particularly those formed in IXPs (Internet eXchange Points) [24, 6, 4, 33]. How does this affect our results?

8.1 Monitor locations

To understand whether the transit providers fully visible through Route Views and RIPE RIS are good locations for observing global routing changes, we first observe that the degrees of freedom in making routing decisions are much greater for highly-connected transit providers than for the majority of ASes that are at the edge of the network. For example, consider a stub AS that is single-homed to its provider, through which it reaches every destination. In that case, the next-hop from that AS to *every other AS* is constant. In other words, its routing table is constrained to carry the same entry for every destination.

Another way of stating this is that a subset of paths carries most of the information used in RSD analysis.

[5] We use here the notation from [17] for consistency. ITER parameters in this section should not be confused with the same symbols in other sections.

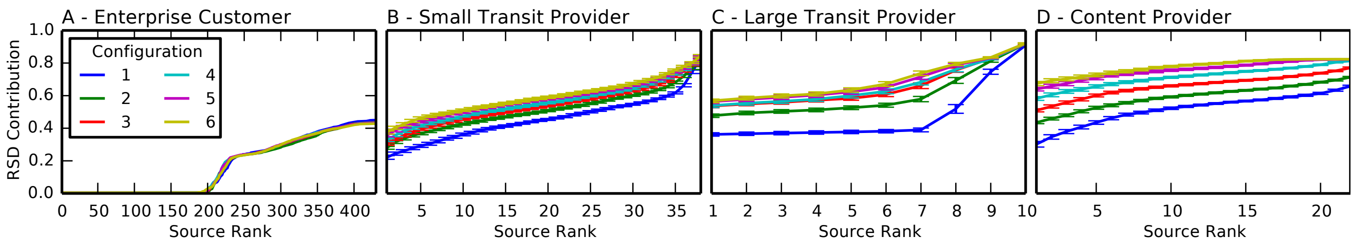

Figure 7: RSD contribution for different AS categories and parameter configurations. Error bars are 95% confidence intervals – Omitted for the leftmost figure to improve visibility.

For example, the next-hops from a stub AS provides no *information* to measurements of RSD, since $N(x, .)$ is a constant function for a single-homed stub AS x.

To demonstrate this effect in realistic Internet routing, we provide further analysis of our ITER simulation runs. ITER places ASes into four different categories: Enterprise Customers (EC), Small Transit Providers (STP), Large Transit Providers (LTP) and Content Providers (CP). Depending on their category, ASes have different functions to optimize and interconnection strategies. We use the default configuration of 500 ASes, consisting of 430 enterprise customers, 22 content providers, 38 small transit providers, and 10 large transit providers. In the default configuration, about half of the enterprise customers will be singly-homed, and half will be multi-homed.

To investigate the contribution that an AS as a source makes in measuring Global RSD, we have Definition 8.

DEFINITION 8. ***(RSD contribution)*** *Given an AS s, its contribution as a source to global RSD is defined as*

$$\nu(s) = \frac{1}{|V|(|V|-1)} \sum_{x,y \in V,\ x \neq y} d(N(s,x), N(s,y)), \tag{11}$$

where d and N are defined as for Equation (1).

For each of the 100 runs within each simulator configuration, we ranked ASes according to their ν values within categories, and then averaged the 100 values for each rank.

Figure 7 presents $\nu(s)$ for each of the four categories of ASes, considered as potential monitoring points. The figure shows the results for each simulator configuration, ranging from hierarchical routing (Configuration 1) to flat routing (Configuration 6). The figure shows clearly how the "role" of an AS within the Internet affects the amount of RSD information contributed by the AS. Figure 7A shows that, across all configurations, the first 200 or so EC ASes are single-homed enterprise customers. These ASes have *no effect* on any RSD values. The next 230 or so ASes are multi-homed enterprise customers, which make modest contributions to RSD measurements. However, there is no impact of these measurements on Global RSD across routing configurations. That is, the significant differences in RSD contribution across the ITER configurations occur essentially *only* among the transit providers and content providers. Hence all of the information necessary to construct Figure 6 as a measure of Internet flattening can be obtained solely by measuring a *small subset* of the most informative AS paths.

Figures 7B and C show that transit providers make much larger contributions to overall RSD measures. Furthermore, these monitoring locations allow the observation of distinct differences across the various routing configurations. Finally, we note that the greatest variation in RSD contribution across routing configurations is made by content providers, which makes sense since content providers are mainly driving the evolution of routing from hierarchical to mesh-like structure.

We conclude that transit providers and content providers that have many neighbors represent particularly informative monitoring points for measuring RSD. However, we acknowledge that this conclusion is subject to limitations of ITER, which in some cases do not capture real-world complexities. For example, ITER does not include peering of ASes in the EC category.

8.2 Peering links

Fully understanding the impact of missing peering links on our results is difficult in the absence of historical data that includes peering links. However, we are able to assess the impact of missing link data at two specific time points, ten years apart. We make use of two datasets from different sources and times; each contains many AS links from IXPs not present in our original data. The first is from May 2005 [24], and the second from September 2015 [1]. Each of these datasets is the result of a measurement campaign designed to capture missing AS links, typically within IXPs.

When comparing the new and old datasets we identified 10504 and 27412 new undirected links for 2005 and 2015 respectively. The main challenge in integrating these datasets is that they contain links rather than paths. We could simply add the corresponding link as a single next-hop, corresponding to a one-hop path, but that is too conservative; the presence of the link can be used to infer additional next-hops as well. Accordingly, for each new AS link (x, y) (we take same actions from y to x), we chose to add y as the next-hop from x for all of the destinations in the *customer cone* of y, as defined in [29].[6] From the customer cone of y we did not consider ASes that were in the customer cone of x and peers of x. These resulting inferred next-hops are consistent with traditional Internet routing models, which despite their limitations [5], we use as an approximation of what one would expect to be visible if we had access to the routing table from x. In many cases such a strategy may not yield the exact set of next-hops that actually involve the link (x, y) – for example, when a customer of y is multi-homed the path to it from x may not pass through y. However we emphasize that the goal of this exercise is not to obtain precise AS topologies.

By following the steps presented in Section 3 (except the last one) we obtained new next-hop matrices containing entries related to data from RIPE, Route Views and the new AS-links. When comparing these new matrices with the original ones (with only data from RIPE an Route Views), we observed only minor changes in our results. More specifically, global RSD changed by less than 1.1% and the sets of unusual ASes (for $k = 45$ and 100, and for $\delta = 50$) changed by less than 15% (for both the 2005 and 2015 datasets). Therefore, the impact of missing links in the next-hop matrices analyzed in this work is not significant, and minor variation should be expected.

However, these results do not prove that unknown AS-links are unimportant. The methods that we use rely on vantage points (sources) from which paths to most Internet locations are visible (such as are provided by Route Views and RIPE). Adding the new data unfortunately does not provide additional vantage points of this type. The reason is that many of the new AS-links are related to IXPs, and paths crossing IXPs typically reach certain customer cones, but not the majority of the Internet. As a result, it is difficult to practically assess the impact of the missing paths that cross IXPs. We note that these initial studies did reveal small areas with high RSD values in the new next-hop matrices. We conclude that a full assessment of the impact of missing IXP links on RSD values is an open question and a valuable direction for future work.

9. RELATED WORK

Little previous work has focused on analysis tools for path-based networks. However the conclusions of our study relate to a number of previous efforts.

RSD. The starting point for our work is the notion of RSD as introduced in [22] and developed more fully in [23]. While those papers showed that RSD is useful for inferring missing data and for identifying similarly-routed sets of ASes ("local atoms"), in this paper we build an additional set of tools (metrics and algorithms) on top of RSD to explore the nature of ASes that are anomalous with respect to the set of routes used to reach them.

[6]http://www.caida.org/data/active/as-relationships/

AS path analysis. A number of studies have carefully analyzed AS paths, to characterize their behavior [8, 34] or to understand routing policies [32, 31]. In contrast to those papers, our work provides metrics and methods that are useful in analyzing path-based networks in general, and focuses on different questions.

Internet flattening. The so-called 'flattening' of the Internet has been the subject of a number of previous studies. The authors of [20] were the first to document the phenomenon, using `traceroute` measurements from 50 nodes. The authors of [28] document flattening using a larger set of networks, and focus on changes in traffic patterns (we do not address traffic volumes in our work). The authors of [29] show that the peering density of customer cones is increasing, which translates into a shift away from a tree-like structure. In the case of one specific AS that is contributing to flattening, [13] showed that Google had more than 5000 peers by March of 2015, and that the number is increasing. In contrast to these efforts, we focus on developing metrics to directly analyze the paths in a path-based network, and using those metrics we quantify flattening effects at the macro level (across all paths) as well as the micro level (with respect to individual ASes)

Graph analysis of interdomain topology. A number of studies have analyzed the AS-level Internet by using graph metrics in static or dynamic fashion (e.g., [18, 11, 30, 21] – among many others). Graph analysis of the Internet can provide insight about the underlying structure over which data flows, but it does not take into account the economic relationships between ASes, nor the paths that data is actually allowed to traverse. Our work avoids the ambiguities of using a graph to capture routing, by analyzing the set of paths used rather than the underlying graph.

Incompleteness of the known Internet topology. A considerable amount of work has exposed the problem of missing information (links and nodes) in common sources of data related to the Internet topology [35, 24, 6, 4, 33, 25]. While our work is subject to the same limitations as previous studies, we show evidence in Section 8 that our results do not appear to be strongly affected by missing data.

10. CONCLUSIONS

The set of paths in a path-based network has rich information content; analyzing them can expose nodes with unusual path sets, identify how a node's path set has changed over time, and uncover network-wide transitions over time. We present metrics and efficient algorithms for exploring each of these aspects of path-based networks, and we illustrate their use in studying routing in the Internet over the 13 year period ending December 2015.

We show that, at the level of individual ASes, characterizing paths can expose relationships among ASes in terms of managing organizations, and can reflect how an organization's business goals and the associated engineering strategies shift over time. In particular, we show that organizations deviate from hierarchically-routed connections for a number of reasons, from content delivery (exemplified by Google), to service reliability (exemplified by DNS servers), to inter-node coordination (exemplified by Ams-IX).

At the level of the whole network, characterizing paths can expose phases in the sets of ASes that are most unusually routed over time. Applying this analysis to the Internet details how content delivery began to drive Internet flattening in late 2005, expanding in 2008 with Google's move into video delivery, and maturing by mid 2010.

Our results also suggest some directions for future work. First, we note the need of further investigation with regard to missing data, since from the RIPE RIS and Route Views projects we cannot obtain all Internet paths. Second, by considering only ASes and the prefixes they originate, we are not able to distinguish addresses associated with caches deployed inside access networks (by content providers, content delivery networks, etc.). One possible approach to circumvent this limitation is to consider destinations as services, instead of ASes (e.g., google.com instead of AS15169). To that end, it is necessary to obtain paths towards a large variety of services, requiring active measurements – for instance, from looking glass servers and RIPE Atlas.[7]

Third, we showed that groups of unusual ASes often correlate with ASes that are owned by an organization. Hence, routing information can be used to support other techniques of co-owned ASes detection (e.g., [9]). In this context, it may be worth investigating an AS that joins/leaves a unusual group in order to decide whether it was a simple routing strategy change or reflecting acquisition or sale of another company. Similarly, unusual ASes may be a good starting point to detect anycast adoption. There are other approaches to that end (e.g., [15] and [14]), but addresses announced by these unusual ASes can be obtained without the need of active measurements.

Fourth, the RSD framework can also benefit organizations. At the operational level some *what if* questions can be asked. For example, what if RSD for a set of ASes/prefixes changes more than some pre-established threshold? It can be a first sign of prefix hijacking or malicious re-routing, thus triggering the need of further investigation. The framework can also provide valuable information to an organization about its competition. More specifically, observing the movement of a groups of ASes in RSD space can expose how their business and engineering strategies behave and change over time. Naturally, these questions demand the ability to produce results in real (or near to real) time, which could be achieved by implementing our tools on top of BGPStream.[8]

Finally, we note that the tools presented here can be applied whenever a node set is equipped with a set of paths; for example, they can be applied to any path-selection strategy used in a network, such as shortest-path or maximum-flow. Based on the results in this study, we believe that these tools show promise for the analysis of other path-based networks, including the movement of goods in transportation networks and the movement of information in social networks.

[7] https://atlas.ripe.net

[8] https://bgpstream.caida.org

11. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under grant numbers CAREER-1253393, CNS-1012910, CNS-1618207, IIP-1430145, IIS-1218437, IIS-1320542 and IIS-1421759. We are grateful to the anonymous reviewers and to our shepherd, Alex Snoeren, for their valuable comments. We also thank Emile Aben, Amogh Dhamdhere, Constantine Dovrolis, Ethan Katz-Bassett and CAIDA for providing tools, data and comments.

12. REFERENCES

[1] IXP-Country-Jedi dataset http://sg-pub.ripe.net/emile/ixp-country-jedi/.
[2] http://bgp.potaroo.net/cidr/autnums.html.
[3] http://cs-people.bu.edu/gcom/dsr/.
[4] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger. Anatomy of a Large European IXP. In *Proceedings of SIGCOMM*, pages 163–174, 2012.
[5] R. Anwar, H. Niaz, D. Choffnes, I. Cunha, P. Gill, and E. Katz-Bassett. Investigating Interdomain Routing Policies in the Wild. In *Proceedings of IMC*, pages 71–77. ACM, 2015.
[6] B. Augustin, B. Krishnamurthy, and W. Willinger. IXPs: Mapped? In *Proceedings of IMC*, pages 336–349. ACM, 2009.
[7] R. Bellman. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences, PNAS*, 1952.
[8] A. Broido and k. claffy. Analysis of RouteViews BGP data: policy atoms. In *NRDM*, 2001.
[9] X. Cai, J. Heidemann, B. Krishnamurthy, and W. Willinger. An organization-level view of the Internet and its implications (extended). Technical report, ISI-TR-2009-679, USC/ISI, 2012.
[10] M. Calder, X. Fan, Z. Hu, E. Katz-Bassett, J. Heidemann, and R. Govindan. Mapping the expansion of Google's serving infrastructure. In *Proceedings of IMC*, pages 313–326. ACM, 2013.
[11] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir. A model of Internet topology using

k-shell decomposition. *PNAS*, 104(27):11150–11154, 2007.

[12] F. Chierichetti, R. Kumar, S. Pandey, and S. Vassilvitskii. Finding the Jaccard Median. In *Proceedings of SODA*, pages 293–311. SIAM, 2010.

[13] Y.-C. Chiu, B. Schlinker, A. B. Radhakrishnan, E. Katz-Bassett, and R. Govindan. Are We One Hop Away from a Better Internet? In *Proceedings of IMC*, pages 523–529. ACM, 2015.

[14] D. Cicalese, J. Auge, D. Joumblatt, T. u. Friedman, and D. Rossi. Characterizing IPv4 Anycast Adoption and Deployment. In *Proceedings of Co-NEXT*. ACM, 2015.

[15] D. Cicalese, D. Joumblatt, D. Rossi, M.-O. Buob, J. Augé, and T. Friedman. A fistful of pings: Accurate and lightweight anycast enumeration and geolocation. In *Proceedings of INFOCOM*. IEEE, 2015.

[16] G. Comarela, G. Gürsun, and M. Crovella. Studying interdomain routing over long timescales. In *Proceedings of IMC*, pages 227–234. ACM, 2013.

[17] A. Dhamdhere and C. Dovrolis. The Internet is Flat: Modeling the Transition from a Transit Hierarchy to a Peering Mesh. In *Proceedings of Co-NEXT*, pages 21:1–21:12. ACM, 2010.

[18] B. Edwards, S. A. Hofmeyr, G. Stelle, and S. Forrest. Internet topology over time. *CoRR*, abs/1202.3993, 2012.

[19] L. Gao and F. Wang. The extent of AS path inflation by routing policies. In *Proceedings of GLOBECOM*, pages 2180–2184. IEEE, 2002.

[20] P. Gill, M. Arlitt, Z. Li, and A. Mahanti. The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse? In *Proceedings of PAM*. Springer-Verlag, 2008.

[21] E. Gregori, L. Lenzini, and C. Orsini. k-Dense communities in the Internet AS-level topology graph. *Computer Networks*, 57(1):213 – 227, 2013.

[22] G. Gürsun, N. Ruchansky, E. Terzi, and M. Crovella. Inferring Visibility: Who's (Not) Talking to Whom? In *Proceedings of SIGCOMM*, pages 151–162. ACM, 2012.

[23] G. Gürsun, N. Ruchansky, E. Terzi, and M. Crovella. Routing State Distance: A Path-based Metric for Network Analysis. In *Proceedings of IMC*, pages 239–252. ACM, 2012.

[24] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy. Lord of the Links: A Framework for Discovering Missing Links in the Internet Topology. *IEEE/ACM Trans. Netw.*, 17(2):391–404, Apr. 2009.

[25] A. Khan, T. Kwon, H.-c. Kim, and Y. Choi. AS-level Topology Collection Through Looking Glass Servers. In *Proceedings of IMC*, pages 235–242. ACM, 2013.

[26] J. Kleinberg and E. Tardos. *Algorithm Design*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.

[27] M. Koivisto, M. Perola, T. Varilo, W. Hennah, J. Ekelund, M. Lukk, L. Peltonen, E. Ukkonen, and H. Mannila. An MDL method for finding haplotype blocks and for estimating the strength of haplotype block boundaries. In *Proceedings of the 8th Pacific Symposium on Biocomputing, PSB*, pages 502–513, 2003.

[28] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet inter-domain traffic. *SIGCOMM Comput. Commun. Rev.*, 41(4), Aug. 2010.

[29] M. Luckie, B. Huffaker, A. Dhamdhere, V. Giotsas, and k. claffy. AS Relationships, Customer Cones, and Validation. In *Proceedings of IMC*, pages 243–256. ACM, 2013.

[30] P. Mahadevan, D. Krioukov, M. Fomenkov, X. Dimitropoulos, k. c. claffy, and A. Vahdat. The Internet AS-level Topology: Three Data Sources and One Definitive Metric. *SIGCOMM Comput. Commun. Rev.*, 36(1):17–26, Jan. 2006.

[31] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig. Building an AS-topology model that captures route diversity. In *SIGCOMM*, 2006.

[32] W. Mühlbauer, S. Uhlig, B. Fu, M. Meulle, and O. Maennel. In search of an appropriate granularity to model routing policies. In *SIGCOMM*, 2007.

[33] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. The (in)Completeness of the Observed Internet AS-level Structure. *IEEE/ACM Trans. Netw.*, 18(1):109–122, Feb. 2010.

[34] J. Rexford, J. Wang, Z. Xiao, and Y. Zhang. BGP routing stability of popular destinations. In *SIGCOMM IMW*, 2002.

[35] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush. 10 lessons from 10 years of measuring and modeling the Internet's autonomous systems. *IEEE JSAC*, 2011.

[36] Y. Shavitt and U. Weinsberg. Topological trends of Internet content providers. In *Proceedings SIMPLEX*, pages 13–18. ACM, 2012.

[37] E. Terzi. *Problems and Algorithms for Sequence Segmentation*. PhD thesis, University of Helsinki, 2006.