Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 738 – 751
© TÜBİTAK
doi:10.3906/elk-1801-79

On spectral analysis of the Internet delay space and detecting anomalous routing paths

Gonca GÜRSUN*
Department of Computer Science, Faculty of Engineering and Computer Science, Özyeğin University, İstanbul, Turkey

Received: 08.01.2018 • **Accepted/Published Online:** 04.10.2018 • **Final Version:** 22.03.2019

Abstract: Latency is one of the most critical performance metrics for a wide range of applications. Therefore, it is important to understand the underlying mechanisms that give rise to the observed latency values and diagnose the ones that are unexpectedly high. In this paper, we study the Internet delay space via robust principal component analysis (RPCA). Using RPCA, we show that the delay space, i.e. the matrix of measured round trip times between end hosts, can be decomposed into two components: the estimated latency between end hosts with respect to the current state of the Internet and the inflation on the paths between the end hosts. Using this decomposition, first we study the well-known low-dimensionality phenomena of the delay space and ask what properties of the end hosts define the dimensions. Second, using the decomposition, we develop a filtering method to detect the paths that experience unexpected latencies and identify routing anomalies. We show that our filter successfully identifies an anomalous route even when its observed latency is not obviously high in magnitude.

Key words: Internet routing, anomaly detection, content delivery, Internet measurement

1. Introduction

Latency is one of the most important performance metrics. The quality of a wide range of applications, such as server selection in content delivery networks (CDNs), video streaming, and voice over IP, as well as any time-critical applications, require low latency on Internet paths. Therefore there has been great interest in understanding the root causes of high round trip time (RTT) values [1–3].

Besides the physical distance between two end hosts, one key factor that drives latency is routing. Both intradomain and interdomain routing decisions of autonomous systems (ASes) on the paths impact the latencies [1, 4, 5]. In fact, the impact of routing on latency is two-fold. First, phenomena behind routing decisions, together with the physical distance between end hosts, generate patterns in the latency data and result in a low-dimensional delay[1] space [6, 7]. In other words, a matrix of RTT values between end hosts is effectively low-rank.[2] Second, suboptimal routing choices and misconfigurations can increase the latencies on the paths and possibly cause discrepancies in the matrix structure. Leveraging these observations, in this paper, we propose a spectral decomposition of latency matrices to distinguish the regular structure of the latency matrix from the discrepancies. Our goal is to write a given latency matrix as a linear combination of two matrices: a low-rank

*Correspondence: gonca.gursun@ozyegin.edu.tr

[1] We use the terms ‘delay’ and ‘latency’ interchangeably. RTT is the total latency from source to destination and destination to source.

[2] A full-rank matrix is effectively low-rank if the matrix can be approximated well by its first few principal components.

estimated latency matrix that reveals structure in the delay space and an inflation matrix that reveals the noise and the inflation in RTTs that does not fit into the low-rank structure. We decompose latency matrices via a recently developed technique, robust PCA [8], as described in Section 2.

Using this decomposition, we aim to answer the following questions. First, we ask what properties of a given latency matrix contribute to its overall dimensionality. To answer this question, we study the rank values of a bunch of estimated latency matrices and investigate whether there is a correlation between their rank values and some features of the end hosts. We find that the number of unique AS-geolocation of the end hosts on the rows/columns of the matrix determines its rank. Second, we ask whether we can detect anomalous routing paths via our decomposition. We mark any path to be an anomaly candidate if a significant portion of its RTT is estimated as inflation. In other words, for each path, we compute the ratio of the inflation to the estimated latency. If this ratio is higher than a threshold, we investigate the path as a possible anomaly.

Notice that our definition of inflation is different than that of previous work, in which a path is called inflated if the measured RTT is significantly greater than the lower bound RTT computed based on the physical distance. Often, the routing paths are not the physically shortest paths and actual speeds of packets are much slower than the theoretical speed of mediums. Therefore, comparing the observed RTTs with the lower bounds does not pinpoint anomalous routes unless the RTT is obviously much larger than the lower bound. Unlike the previous work, our approach does not set lower bounds. We compute the estimated latency on a path with respect to the current delay state of the Internet via our decomposition. Then we decide whether a path is an anomaly candidate or not by comparing its estimated RTT component with its inflation component.

We summarize our contributions in this paper as follows: a) we show how to decompose a latency matrix via a recent technique, RPCA, into a low-rank and an inflation component; b) we investigate the features of end hosts that result in the low-rank property and we find that both geolocation and the AS of the end hosts define the dimensions of the delay space; c) we propose a method to diagnose the inflated paths by the inflation-to-estimated-latency ratio filters; d) we show that our filter successfully pinpoints the routing anomalies even when the RTTs on the paths are not obviously large; e) we show that our filter successfully pinpoints the routing anomalies even when all measured paths between the end hosts from two regions are inflated; and f) we show how to apply our filter in the case that RTT measurements between some end hosts are missing.

The rest of the paper is organized as follows. We introduce the spectral analysis tool, RPCA, in Section 2 and describe our dataset in Section 3. In Section 4 we study why the delay space is low-dimensional. In Section 5 we present the anomalous routes that we detect in our dataset. We present the related work in Section 6 and conclude in Section 7.

2. Decomposing latency

Latency matrices were shown to be effectively low-rank in previous work [9–17]. This property of latency matrices is used in various applications such as embedding the delay space into low-dimensional coordinate spaces and estimating RTTs between hosts without direct measurements [6, 7, 18–23]. Our goal in this paper is also to leverage the low-rank property in order to distinguish the end hosts that experience estimated latencies with respect to the current state of the delay space from the end hosts that experience unexpected latencies. The first step to do that is to find a low-rank approximation of a given latency matrix.

Principal component analysis (PCA) is the most popular tool to find a low-rank approximation for a given matrix and is widely applied on latency matrices. Despite its popularity, PCA has a few drawbacks: it is highly sensitive to arbitrarily large or grossly corrupted observations and missing measurements [24]. Such cases are quite common in RTT measurements, e.g., due to incomplete traces by nonresponsive end hosts, system limitations and failures result in missing and corrupted observations, and anomalies cause arbitrarily large RTT values. PCA is not robust to such cases, i.e. it fails to yield the true underlying structure of the data. A recent technique called robust PCA (RPCA) addresses these drawbacks [8] and suits our decomposition goals well.

2.1. RPCA

Let X be a data matrix. In our context, the rows of X are the sources, the columns are the destinations, and an element stores the RTT value observed between the corresponding source and the destination. RPCA aims to find a decomposition of X such that $X = L + S$. In this decomposition, L is a low-rank matrix generated by the underlying mechanism of the observed data and S is a sparse noise matrix that does not fit into the low-rank property. In our context, L stores the estimated latency values between end hosts and S stores the unexpected inflation (one can also call the values of S as noise).

RPCA decomposes a given X into its L and S under the following conditions:

a) The rank or the column and row spaces of L are unknown,

b) The number of nonempty entries of S is unknown,

c) The locations of nonempty entries of S are unknown,

d) The entries of L and S can be arbitrarily large,

e) The nonempty entries of S are randomly distributed.

In order to find L and S under these conditions, RPCA solves the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & ||L||_* + ||S||_1 \\ \text{subject to} \quad & L + S = X, \end{aligned} \tag{1}$$

where $||L||_*$ is the nuclear norm of matrix L, which is defined as the sum of its singular values, and $||S||_1$ is the l_1 norm of S. In [8] it was shown that solving this optimization problem can exactly recover L and S in polynomial time. Note that writing the optimization over the l_1 norm and the nuclear norm is one of the keys to deal with arbitrarily large, corrupted, and missing data. On the other hand, traditional PCA optimizes over the l_2 norm, which makes it sensitive to large, corrupted, and missing data points. In our analysis, we use the implementation of RPCA provided by the authors of [8] and refer the reader to their paper for further detail on the technique.

Notice that RPCA is not just a variant of PCA. Instead, the difference between two methods is substantial. While both techniques aim at identifying the core low-rank component L from a given measurement matrix X, they do it under completely different assumptions about the additive perturbation. PCA assumes a Gaussian and nonsparse perturbation S and minimizes the l_2 norm, while RPCA assumes the perturbation component S to be sparse regardless of its distribution and therefore minimizes the l_1 norm for S and nuclear norm for L.

3. Datasets

3.1. RTT data

We use a collection of RTT values collected from Akamai CDN [3] via traceroute measurements on 24 January 2016. The measurements are taken from 47 CDN server nodes to 5076 client IPs located in France. The CDN server nodes are spread across 14 unique ASes in three countries: France, the United States, and Japan.

[3] This work was initiated when the author was visiting Akamai Technologies.

Due to the scale of our measurement setup, there are limitations on the number of times a client IP can be tracerouted in a given time period. Such limitations are set by the ISPs in order not to keep client IPs busy. Therefore, we tracerouted each client IP from 20 CDN server nodes on average. That is, we do not have an RTT measurement for all server node and client IP pairs.[4]

In practice, there are several issues with using traceroute data. One issue is that ICMP packets might be deprioritized or simply dropped at the routers. This results in higher RTTs or incomplete traces. To minimize these effects, we take three consecutive measurements from a given server node to a client IP. Then we use the one with minimum latency. We also remove all incomplete traces from our dataset.

Another issue is the asymmetric reverse paths and latency inflation due to long reverse paths. In fact, we find that asymmetry and circuitousness in reverse paths are common. One indication of circuitousness in the reverse path is the significant increase of RTT on a single hop even though the consecutive routers in the forward path are geographically close. We discard these cases from our analysis and focus on the anomalies on the forward paths.

3.2. BGP announcements

We use a collection of BGP tables to analyze the inflated paths. The tables are collected on the same day as traceroute measurements from 297 peer routers in the CDN. They consist of over 790K BGP paths to over 48K prefixes in Europe. Using this dataset, we map our client IPs to their longest matching BGP prefixes. As a result, 5076 client IPs map to 778 prefixes.

3.3. Geolocation mapping

We use the EdgeScape tool of Akamai to map each and every IP address in our dataset to its geographic location.

4. Understanding dimensionality

In this section we ask what makes the latency data low-dimensional. To answer our question we decompose latency matrices, s.t. $X = L + S$. Then we study the rank of the low-rank L components. Our goal is to find which features of the end hosts correlate with the rank values.

We generate X matrices in two levels of granularity. In the first level, each column of X is an individual destination IP. In the second level, we aggregate individual IPs into their prefixes and each column of X represents a BGP prefix.

4.1. IP-level

We generate an X_{IP} matrix for each BGP prefix such that the columns are the individual destination IPs that belong to the prefix and the rows are the server nodes. From each X_{IP}, we discard the rows that have no measurements to any of the destinations on the columns. Then we decompose each matrix X_{IP} into its L_{IP} and S_{IP} and compute the rank of L_{IP}.

Our intuition is that since the routing decisions are made at the level of BGP prefixes, each column vector of a given X_{IP} will be same or very similar to each other. Therefore, the rank of L_{IP} matrices will be correlated with the features of the server nodes in the rows. In order to show that our intuition holds, we plot

[4] In fact, we have measurements from very few server nodes. Therefore, we note that the routing paths in our dataset are not necessarily the ones that the CDN uses to serve content to the end users. In fact, the goal of the CDN is to map end users to a best available server node at a given time and routing paths with large latencies are not used.

the rank values of L_{IP} matrices in Figure 1. Note that, in order not to limit the rank of a matrix by its number of rows or columns, we only consider matrices that have large numbers of individual IPs, i.e. the number of columns in X_{IP} varies between 50 and 222.

Figure 1 plots the rank of L_{IP} matrices vs. three features of the server nodes. We tag each server node by a) its geolocation, b) the AS that the server node belongs to, and c) both the geolocation and the AS of the server node. First, Figure 1 shows that the rank values are relatively low, i.e. they vary between 1 and 30. Second, we find that the number of unique geolocations or the number of unique ASes of the server nodes is not enough to explain the low dimensionality. Instead, the number of unique geolocation and AS pairs in the rows best correlates with the rank values. Next, we ask whether our findings still hold when we increase the diversity of the columns by bringing various prefixes from different ASes together.

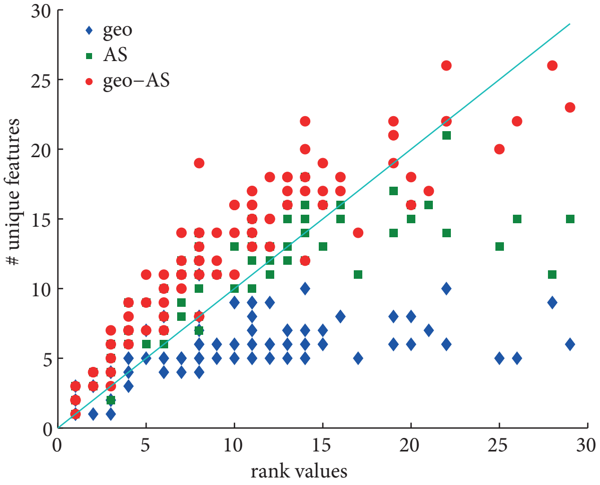


Figure 1. The number of unique features of the server nodes vs. the rank values of the L_{IP} matrices.

4.2. Pfx-level

There are various advantages in aggregating individual IPs into their prefixes, e.g., less noise and smaller data size. Also, such aggregation is natural since routing on the Internet is based on BGP prefixes, i.e. prefixes are atomic with respect to routing. We analyze the dimensionality of the delay space under this aggregation.

We map all IPs into their BGP prefixes in our dataset. Then we generate a matrix X_{FR} s.t. the rows are the server nodes, each column represents one prefix, and an element is the minimum latency measured from the corresponding server node to all IPs within the corresponding prefix on the column. We consider only the large prefixes, i.e. prefixes that have at least 10 IPs mapped to them (details are explained in Section 5). This yields matrix X_{FR} of size 47×80.

We decompose the matrix via RPCA s.t. $X_{FR} = L_{FR} + S_{FR}$. First we find that the rank of L_{FR} is 26. This is exactly the number of unique geolocation and AS pairs of the server nodes in the rows of X_{FR}. To investigate this finding further, we randomly extract 500 submatrices of random sizes from X_{FR}. We decompose each submatrix via RPCA and compute the rank of their L components.

Figure 2 plots the ranks of the L components of the submatrices. We tag each row and column of the submatrices by their features, i.e. their geolocation, AS, and both. Then, for each submatrix, we plot its rank vs. the minimum of the unique features on its rows and columns. Similar to our previous finding, Figure 2 shows that the number of unique geolocation and AS pairs in the rows/columns correlates well with the rank.

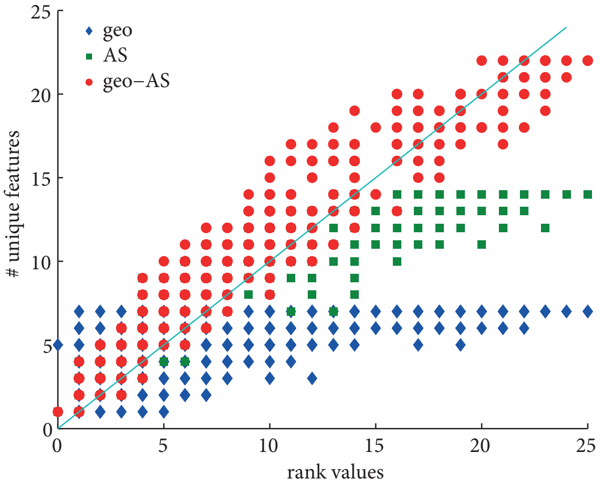


Figure 2. The number of unique features of the server nodes and the destination prefixes vs. the rank values of the L components of the randomly generated submatrices.

This analysis shows that the underlying dimensionality of the delay space is the result of routing choices as the paths that are destined to the same prefixes in the same geolocation and the ASes tend to experience similar latencies.

5. Detecting anomalies

Our goal is to identify anomalous routing paths via spectral analysis of latency values. Note that anomalous routing paths are the ones that have unexpectedly high RTTs due to routing misconfigurations or any suboptimal routing decisions including the ones for load balancing.

5.1. Aggregating RTT data

A given RTT value is composed of three components: transmission delay, propagation delay, and queuing delay. Since each ICMP packet is 32 bytes, the transmission delays in our measurements are very small and can be ignored. Therefore, the delay we see is either propagation delay or queuing delay.

Since our goal is to find routing anomalies, we are interested in high propagation delays rather than high queuing delays. One way to eliminate the measurements with high queuing delays is to aggregate individual client IPs to their BGP prefixes as follows: for each server node we use the minimum RTT measured to any of the client IPs that belong to that prefix. We consider prefixes that have at least 10 IPs mapped to them so that we significantly decrease the likelihood of queuing delays. Then we generate X_{FR} (also described in Section 4), where each column represents a prefix and its entries are the minimum latencies from the server nodes to any of the client IPs in the prefix. We use X_{FR} in anomaly detection as presented below.

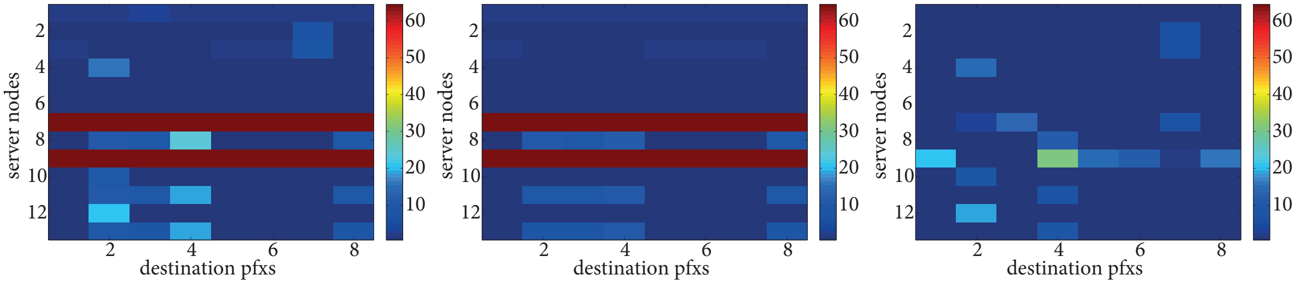
5.2. Detecting inflated paths via ratio filtering

We decompose a latency matrix X into its L and S. As we explain in Section 2, the entries in S represent the inflation that does not fit into the low-rank structure of the latency measurements. In other words, the entries in S are the difference between the measured latency in X and the estimated latency in L.

Having decomposed the matrix, we say that the path from the server node i to prefix j is inflated if the ratio of the inflation to the estimated latency is greater than a threshold τ, i.e. $\frac{S(i,j)}{L(i,j)} > \tau$. In our analysis we set τ to 1. In other words, we investigate a route as an anomaly candidate if half of its measured RTT is estimated as inflation. Then we rank the candidates based on the magnitude of $S(i,j)$ since the larger the inflation is, the more likely it is that the path is anomalous. In practice, we find that paths whose inflation is more than 10 ms are worth investigating.

In addition, we find that on the paths that are cross-continents, the inflation ratio is hardly greater than 1 as the minimum possible RTT is already high. For instance, in our dataset, the minimum latency between Tokyo and Paris is 210 ms and the minimum latency between San Jose, USA, and Paris is 136.4 ms. Therefore, we also investigate all routes on which the magnitude of the inflation $S(i,j)$ is greater than 30 ms.

Next we present three cases of anomalies we detect in our dataset. In the first two cases, we apply our algorithm to the submatrices of X_{FR} s.t. in each submatrix, columns are prefixes from the same AS and rows are the server nodes that have measurements to all prefixes in the columns. This guarantees that a) the rank of the underlying L is very low (ranks vary from 1 to 5), since the prefixes from the same AS and geolocation (in our case, France) are expected to have very similar columns, and b) our results are not biased by the missing measurements. We apply our filter to all such submatrices and below present two of them that have anomalies. In the third case, we apply our method to the entire X_{FR} to test it against missing measurements and relatively higher rank value, i.e. the rank of L_{FR} is 26. We present that even in the case of missing measurements we can detect anomalies.



(a) X (b) L (c) S

Figure 3. Latency values for Case 1 in X (measured latency) and its decompositions, L (estimated latency) and S (inflated latency).

5.3. Case 1: Detecting the anomalous path from a server node when the latency to only one destination prefix is inflated

In this case we study the RTT values to 8 prefixes that belong to Akamai International (AS34164). There are 13 server nodes in our dataset that have measurements to all of these prefixes, i.e. X is 13 × 8 as shown in Figure 3a. The prefixes are listed in Table1 in the order in which they appear in the columns of Figure 3. The heatmaps in Figure 3 visualize the measured latency values in X as well as the latencies in the L and S components. Below are the four anomalies we find in this case.

The first anomalous route is from a server node in Deutsche Telekom (AS3320) in Paris to 2.16.126.0/24. Their path corresponds to row 4 and column 2 of the matrices in Figure 3. The RTT on the path is 16.2 ms and our analysis estimates that the RTT should have been 0.9 ms (see Figure 3b), and therefore it is inflated by 15.3 ms (see Figure 3c).

Looking at the traceroute path from the server node to the prefix, we see that the inflation is due to a detour through Munich and Milan, i.e. the route to the prefix is AS3320 (Paris) → AS3320 (Munich) → AS6762

Table 1. Prefix list for Case 1. Prefixes belong to AS34164. The prefixes appear in the order of column id in Figure 3.

Column id	Prefix	Column id	Prefix
1	184.85.251.0/24	5	2.18.249.0/24
2	2.16.126.0/24	6	23.62.9.0/24
3	2.16.136.0/24	7	92.123.193.0/24
4	2.16.54.0/24	8	96.16.122.0/23

(Milan) → AS6762 (Paris) → AS34164 (Aubervilliers).[5] Looking at the traceroutes of the other seven prefixes, we do not see such a detour. Their routes are AS3320 (Paris) → AS3257 (Paris) → AS34164 (Aubervilliers or Paris). This indicates that there might be a misconfiguration for the prefix 2.16.126.0/24. In addition, looking at our BGP data, we see that AS3320 and AS34164 are peers for many other locations in Europe (Austria, Belgium, Italy, Russia, the Netherlands, etc.) and the direct link between them is used for the prefixes in these locations. Therefore, we infer that shorter routes might be possible for prefix 2.16.126.0/24 if the peering relationship between AS3320 and AS34164 is used.

The second anomalous route is to the same prefix, 2.16.126.0/24, from a server node in Nerim SAS (AS13193) in Paris. Their path corresponds to row 12 and column 2 of the matrices in Figure 3. The RTT on the path is 21.1 ms and our analysis estimates that the RTT should have been 1.1 ms (see Figure 3b), but it is off by 20 ms (see Figure 3c).

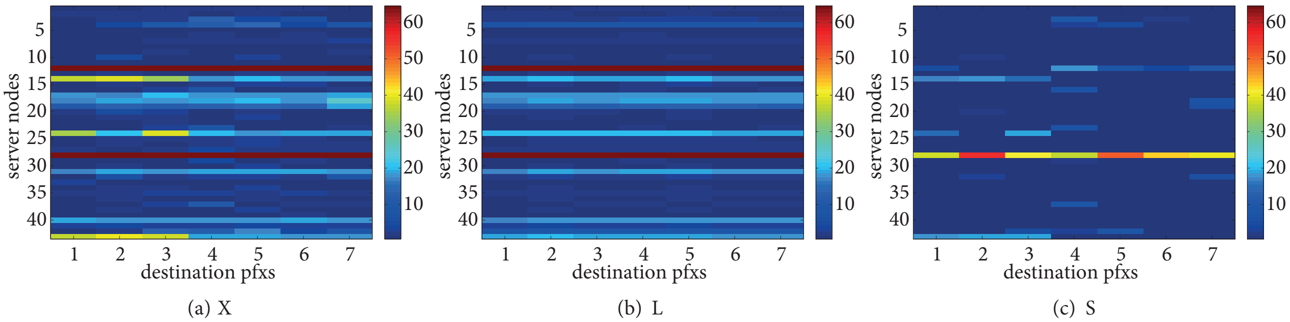
Looking at the traceroute path from the server node to the prefix, we see that the reason for 20 ms of inflation is a detour through Dublin, London, and Amsterdam, i.e. the path is AS13193 (Paris) → AS10310 (Dublin) → AS10310 (London) → AS5580 (London) → AS5580 (Amsterdam) → AS34164 (Aubervilliers). Looking at the traceroutes of the other seven prefixes, we do not see such a detour. Their paths from the same server node go direct, i.e. AS13193 (Paris) → AS1299 (Paris) → AS5580 (Aubervilliers) → AS34164 (Aubervilliers). Therefore, we infer that there is a misconfiguration for prefix 2.16.126.0/24.

The third anomalous route is from a server node in Orange Telecom (AS 5511) in Paris to 2.16.54.0/24. Their path corresponds to row 8 and column 4 of the matrices in Figure 3. The RTT on the path is 24.9 ms and our analysis estimates that the RTT should have been 12.3 ms (see Figure 3b), but it is off by 12.6 ms (see Figure 3c).

Looking at the traceroute path from the server node to the prefix, we see that the inflation is due to a detour via Frankfurt, i.e. the path is AS5511 (Paris) → AS3257 (Frankfurt) → AS3257 (Paris) → AS34164 (Aubervilliers). However, looking at the paths from the same server node to the other prefixes, we see that there is a direct path of AS5511 (Paris) → AS5511 (Aubervilliers) → AS34164 (Aubervilliers).

The fourth anomalous route is from a server node in NTT Comm. (AS2914) in Tokyo to 2.16.54.0/24. This path corresponds to row 9 and column 4 of the matrices in Figure 3. We identify that the path is inflated because of a detour to Hong Kong, Singapore, and Mumbai. We identify a very similar case and discuss it in detail in the next section.

[5] We map all routers on a given traceroute path to their AS and geolocation. Then we represent the consecutive routers that have the same AS-geolocation tag as one hop for readability.



(a) X (b) L (c) S

Figure 4. Latency values for Case 2 in X (measured latency) and its decompositions, L (estimated latency) and S (inflated latency).

Table 2. Prefix list for Case 2. Prefixes belong to AS12670. The prefixes appear in the order of column id in Figure 4.

Column id	Prefix	Column id	Prefix
1	195.167.192.0/20	5	89.225.192.0/18
2	212.99.0.0/17	6	92.103.0.0/16
3	46.218.0.0/16	7	92.103.64.0/18
4	46.218.0.0/18		

5.4. Case 2: Detecting anomalies even when the paths to all prefixes from the same AS and geolocation are inflated

This case demonstrates that our method can catch anomalies even when all the prefixes within the same AS and geolocation are infected. In this case we study the RTT values for 7 prefixes that belong to CompleTel (AS12670) located in France. There are 13 server nodes in our dataset that have measurements to all of these prefixes, i.e. X is 43×7 as shown in Figure 4a. The prefixes are listed in Table 2 in the order in which they appear in the columns of Figure 4.

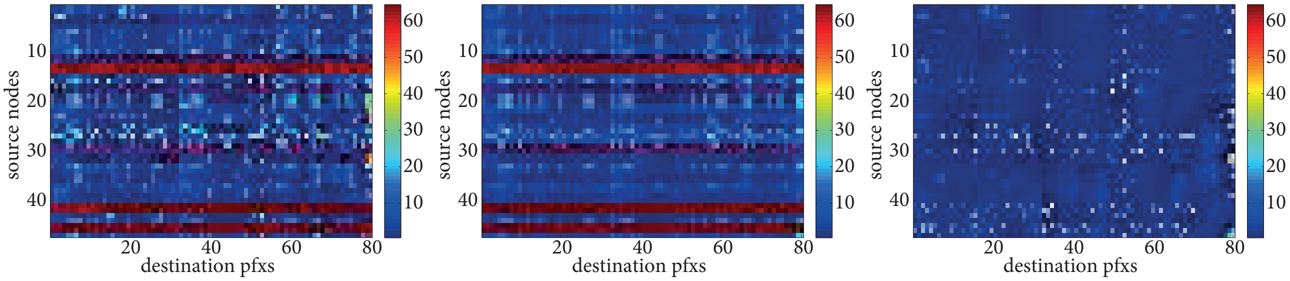
The first anomalous route is between a server node from NTT Communications (AS2914) in Tokyo to all prefixes in the list. This server node corresponds to row 28 in Figure 4. The RTT values from this server node to the prefixes vary from 250 ms to 280 ms. Figure 4b estimates that RTT values should be in the 210–222 ms range; therefore, they are all inflated by 40–56 ms as shown in row 28 of Figure 4c.

Looking at the traceroute paths from the server node to the prefixes, we find that all the paths are longer than estimated. The paths are AS2914 (Tokyo) → AS2914 (Hong Kong) → AS6453 (Hong Kong) → AS6453 (Singapore) → AS6453 (Mumbai) → AS6453(Marseilles) → AS6453(Paris) → AS12670 (Paris).

In order to find whether there is a shorter path between AS2914 (Tokyo) and AS12670 (Paris), we investigate all paths in between them. We find that AS2914 has a presence in many locations in Europe including Paris and therefore shorter paths between AS2914 (Tokyo) and AS12670 (Paris), without redirection via AS6453, are possible. Alternatively, we find that AS2914 makes different next-hop decisions for some other prefixes that are also located in Paris. For instance, for the prefix 83.167.32.0/19 of AS8218 (Neo Telecoms), there is the following path: AS2914 (Tokyo) → AS2914 (Seattle) → AS6461 (Seattle) → AS6461 (Chicago) → AS6461 (New York) → AS6461 (Paris) → AS8218 (Paris). This path is of 210 ms as opposed to the ones via AS6453 that are 250–280 ms. In our BGP data, we find that AS6461 has a peering with AS12670 and could be chosen by AS2914 as an alternative to AS6453. This would lead to a smaller RTT route. We note that the decision of routing via AS6453 might be due to load balancing.

The second anomalous path is from a server node in Orange Telecom (AS5511) in Aubervilliers, France, to the first three prefixes listed in Table 2. This server node corresponds to row 43 in Figure 4. The RTT values from this server node to the prefixes vary from 37.7 to 40.3 ms. Figure 4b estimates that RTT values should be in the 19.3–20.4 ms range; therefore, they are all inflated by around 19 ms as shown in row 28 of Figure 4c.

The traceroute paths from the server node to these three prefixes are AS5511 (Aubervilliers) → AS5511 (Paris) → AS174 (Paris) → AS12670 (Paris). However, the path to the other four prefixes has lower RTT values, i.e. AS5511 (Aubervilliers) → AS5511 (Paris) → AS6453 (Paris) → AS12670 (Paris). Although both paths are 4 hops, the latency on the link between AS5511 and AS6453 is less than the latency on the link between AS5511 and AS6453. Notice that 46.218.0.0/18 is the specific subset of 46.218.0.0/16 and follows a shorter path.



(a) X_{FR} (b) L_{FR} (c) S_{FR}

Figure 5. Latency values for Case 3 in X_{FR} (measured latency) and its decompositions, L_{FR} (estimated latency) and S_{FR} (inflation).

5.5. Case 3: Detecting anomalies in the case of missing measurements

The first two cases test our method when the RTTs between all end hosts are known. Finally, we show how to apply our method in the presence of missing values.

As a preprocessing step, we interpolate the missing measurements as follows. For each missing measurement from a server node i to a prefix j, we replace the missing measurement with the minimum RTT value observed from another server node that is in the same location and AS as the server node i to any prefix that is in the same AS and location as the prefix j. Such interpolation replaces the 0-valued missing entries with the lowest RTT seen between two AS-geolocation regions. This preprocessing step decreases the ratio of missing measurements in X_{FR} from 15% to 1%.

Next we apply RPCA on the interpolated matrix, followed by the same inflation filter. Figure 5 shows the latency values. We find that the server node that has the most anomalous routes is located in Akamai (AS20940) in Aubervilliers. This server node corresponds to row 27 and it has 27 anomaly candidates as shown in Figures 5b and 5c. We find that the paths from this server node make a detour via Amsterdam and Zurich although both the server node and the prefixes are in France. For instance, to prefix 2.20.243.0/24 (column 53), the measured RTT is 21.94 ms. However, the estimated latency in L_{FR} is 9.87 ms and the estimated inflation in S_{FR} is 12.07 ms. The traceroute path to the prefix is AS20940 (Aubervilliers) → AS12322 (Paris) → AS1200 (Amsterdam) → AS13030 (Amsterdam) → AS13030 (Zurich) → AS20940 (Aubervilliers).

This case shows that our method catches anomalous paths even when the magnitude of RTT is relatively low - that is, when the inflation is not obvious. The average latency from this server node to all prefixes is 18.3 ms. Comparing the RTT value of the anomalous path, 21.94 ms, with the rest of the paths from the same server node to the same region would not detect the anomaly. However, our method successfully pinpoints the anomalous paths by the inflation-to-estimated-latency ratio filter. In conclusion, we show that with the help of a simple interpolation step, our method successfully detects anomalies even when a significant portion of the RTTs are missing.

5.6. Discussion

The challenge with this kind of anomaly detection work is that there is no available ground truth of latency values for the paths. For the paths between a source, a destination pair is selected based on the relationship between ASes and the current state of the paths at the time. Therefore, we do not necessarily know which paths are the best in terms of latency or even what lower bound of the latency is possible between a source and destination pair.

One way to assess the goodness of a path could be comparing it with a rough estimate of the geographic latency between the source and the destination. We call the geographic latency between the source and the destination the ratio of their geographic distance in miles (or kilometers) over the speed of the underlying medium. The assumption is that the packets follow the shortest geopath from the source to the destination. However, it is well known that this is rarely the case for Internet routing, i.e. the paths are picked based upon various policies of the ASes and therefore delays are almost always higher than the theoretical geolatencies. In addition, since the routing decisions and the state of the paths change, getting a sense of the feasible latency lower bounds at a given time is difficult. In fact, this observation is our motivation behind using spectral analysis techniques to detect inflated paths. Our approach decides whether a path is possibly inflated with respect to the delays on other similar paths.

We use our RPCA method as an anomaly candidate suggestion tool. After decomposing the latency matrix into S and L, we rank the paths based on their S/L ratios and then we filter the suspicious ones by the τ threshold (as discussed earlier in Section 5). We expect that all anomalous paths are above the threshold so that we do not miss any anomalies. In other words, we want high true positives and low false negatives. Therefore, the choice of τ is important. For our dataset, we set τ to 1 ms and manually investigate all paths above 1 ms by using the traceroute and the geomapping data to verify whether they are likely to be anomalous. Note that choice of τ may change for some other delay matrix. In fact, τ defines the acceptable latency threshold and may vary.

6. Related work

Understanding the delay space and detecting RTT inflations are of great interest in the literature. The works in [5, 25] showed that half of the paths are inflated due to routing policies, while [4] studied the possible root causes of path inflation and showed that intradomain routing decisions and peering policies are the major reasons. The authors of [2] studied the impact of routing changes on network delay and jitter using network topology. Similar to our findings, the work in [2] showed that intradomain routing decisions can cause as severe latency inflation as interdomain routing decisions. The authors of [26] studied how routing parameters impact path optimality. Our work is complementary to all these studies as our method provides a list of anomaly candidate paths to investigate.

In addition, our work relates to the studies on Border Gateway Protocol (BGP)-based anomaly detection techniques. BGP is the default interdomain routing protocol that manages connectivity among ASes. Accidental and malicious activities such as misconfigurations, failures, worm attacks, and prefix hijackings can induce severe connectivity loss and delay. For instance, the work in [27] analyzed one large-scale routing anomaly incident: Chinese ISP prefix hijacking. The analysis highlighted the challenge of understanding the root causes of such incidents and the importance of robust detection techniques. In [28], BGP anomaly detection techniques were surveyed. Unlike our work, most techniques are based on either AS path information and BGP messages or the topology of the AS graph: [29, 30] used BGP update messages to pinpoint the root cause of anomalies, while [31] enhanced the use of path information with Internet Routing Registry (IRR) data. The authors of [32] used path-based analysis of BGP paths. They applied supervised learning techniques to identify worm events and black holes. The authors of [22, 33] provided visualization tools to present routing anomalies. Similar to our work, these studies used traceroute data to generate AS paths. Unlike our work, they did not consider latency information on the paths. In [34] graph mining techniques were used to detect control plane anomalies from AS topology graph. In [35] a crowd-based anomaly detection framework was proposed where clients share latency information with each other to detect anomaly route candidates. Similar to our work, the authors of [35] used RTT information on the routing paths. Unlike our work, the proposed approach is distributed and relies on clients truthfully sharing information.

Finally, the low-rank property of latency matrices is used in various studies. The authors of [18, 20] proposed methods for predicting RTT based on this observation, while [7] uses dimensionality reduction as a way of estimating RTTs between hosts without direct measurements. In [19] the low-rank property was used as a way to infer proximity between hosts, while [21] presented a latency estimation system that uses matrix factorization, which also uses the low-rank structure of latency data. In addition, [6, 22, 23, 36] showed that latency space can be embedded into low-dimensional coordinate spaces. All of these work are similar to ours as they leverage the highly structured nature of the latency data. Unlike these works, however, we use the low-rank property to detect routing anomalies on the Internet.

7. Conclusion

In this paper, we study the delay space of the Internet via robust principal component analysis. We find that the dimensionality of the delay space is well correlated with the geolocation and ASes of the end hosts. Then we show how to leverage the low-rank property of the delay space to identify anomalous routing paths. We show that our method successfully identifies the anomalies even when all prefixes from the same region are infected and in the presence of missing latency measurements.

Acknowledgment

We thank KC Ng from Akamai Technologies for feedback and insightful discussions.

References

[1] Krishnan R, Madhyastha HV, Srinivasan S, Jain S, Krishnamurthy A, Anderson T, Gao J. Moving beyond end-to-end path information to optimize CDN performance. In: ACM 2009 Internet Measurement Conference; 4–6 November 2009; Chicago, IL, USA. New York, NY, USA: ACM. pp. 190–201.

[2] Pucha H, Zhang Y, Mao ZM, Hu YC. Understanding network delay changes caused by routing events. SIGMETRICS Perform Eval Rev 2007; 35: 73–84.

[3] Zarifis K, Flach T, Nori S, Choffnes D, Govindan R, Katz-Bassett E, Mao ZM, Welsh M. Diagnosing path inflation of mobile client traffic. In: Springer 2014 Passive Active Measurement; 10–11 March 2014; Los Angeles, CA, USA. Berlin, Germany: Springer. pp. 23–33.

[4] Spring N, Mahajan R, Anderson T. The causes of path inflation. In: ACM 2003 SIGCOMM; 25–29 August 2003; Karlsruhe, Germany. New York, NY, USA: ACM. pp. 113–124.

[5] Tangmunarunkit H, Govindan R, Shenker S. Internet path inflation due to policy routing. SPIE 2001; 4526: 188-195.

[6] Abrahao B, Kleinberg R. On the Internet delay space dimensionality. In: ACM 2008 Internet Measurement Conference; 20–22 October 2008; Vouliagmeni, Greece. New York, NY, USA: ACM. pp. 157-168.

[7] Tang L, Crovella M. Virtual landmarks for the Internet. In: ACM 2003 Internet Measurement Conference; 27–29 October 2003. Miami Beach, FL, USA. New York, NY, USA: ACM. pp. 143–152.

[8] Candés EJ, Li X, Ma Y, Wright J. Robust principal component analysis? J ACM 2011; 58: 11-37.

[9] Lakhina A, Crovella M, Diot C. Diagnosing network-wide traffic anomalies. SIGCOMM Comput Commun Rev 2014; 34: 219–230.

[10] Roughan M. Simplifying the synthesis of Internet traffic matrices. SIGCOMM Comput Commun Rev 2005; 35: 93-96.

[11] Zhang Y, Roughan M, Lund C, Donoho DL. Estimating point-to-point and point-to-multipoint traffic matrices: An information-theoretic approach. IEEE ACM T Network 2005; 13: 947–960.

[12] Chang H, Jamin S, Mao ZM, Willinger W. An empirical approach to modeling inter-as traffic matrices. In: ACM 2005 Internet Measurement Conference; 19–21 October 2005; Berkeley, CA, USA. Berkeley, CA, USA: ACM USENIX Association. pp. 12-25.

[13] Chang H, Jamin S, Willinger W. To peer or not to peer: modeling the evolution of the Internet's as-level topology. In: IEEE 2006 INFOCOM; 23–29 April 2006; Barcelona, Spain. Piscataway, NJ, USA: IEEE. pp. 1–12.

[14] Erramill V, Crovella M, Taft N. An independent-connection model for traffic matrices. In: ACM 2006 Internet Measurement Conference; 25–27 October 2006; Rio de Janeiro, Brazil. New York, NY, USA: ACM. pp. 251-256.

[15] Zhang Y, Roughan M, Willinger W, Qiu L. Spatio-temporal compressive sensing and Internet traffic matrices. SIGCOMM Comput Commun Rev 2009; 39: 267–278.

[16] Bharti V, Kankar P, Setia L, Gürsun G, Lakhina A, Crovella M. Inferring invisible traffic. In: ACM 2010 Conference on Emerging Networking Experiments and Technologies (Co-NEXT); 30 November–3 December 2010; Philadelphia, PA, USA. New York, NY, USA: ACM. pp. 22:1–22:12.

[17] Gürsun G, Crovella M. On traffic matrix completion in the Internet. In: ACM 2012 Internet Measurement Conference; 14–16 November 2012; Boston, MA, USA. New York, NY, USA: ACM pp. 399-412.

[18] Dabek F, Cox R, Kaashoek F, Morris R. Vivaldi: A decentralized network coordinate system. SIGCOMM Comput Commun Rev 2004; 34: 15-26.

[19] Liao Y, Du W, Geurts P, Leduc G. Dmfsgd: A decentralized matrix factorization algorithm for network distance prediction. IEEE ACM T Network 2013; 21: 1511-1524.

[20] Madhyastha HV, Anderson T, Krishnamurthy A, Spring N, Venkataramani A. A structural approach to latency prediction. In: ACM 2006 Internet Measurement Conference; 25–27 October 2006; Rio de Janeiro, Brazil. New York, NY, USA: ACM. pp. 99-104.

[21] Mao Y, Saul LK. Modeling distances in large-scale networks by matrix factorization. In: ACM 2004 Internet Measurement Conference; 5–7 November 2004; Vancouver, Canada. New York, NY, USA: ACM. pp. 278–287.

[22] Wong T, Jacobson V, Alaettinoglu C. Internet routing anomaly detection and visualization. In: 2005 International Conference on Dependable Systems and Networks; 28 June–1 July 2005; Yokohama, Japan. Piscataway, NJ, USA: IEEE. pp. 172–181.

[23] Zhang B, Ng TSE, Nandi A, Riedi R, Druschel P, Wang G. Measurement based analysis, modeling, and synthesis of the internet delay space. In: ACM 2006 Internet Measurement Conference; 25–27 October 2006; Rio de Janeiro, Brazil. New York, NY, USA: ACM. pp. 85–98.

[24] Ringberg H, Soule A, Rexford J, Diot C. Sensitivity of PCA for traffic anomaly detection. SIGMETRICS Perform Eval Rev 2007; 35: 109–120.

[25] Savage S, Collins A, Hoffman E, Snell J, Anderson T. The end-to-end effects of Internet path selection. In: ACM 1999 SIGCOMM; 31 August–3 September 1999; Cambridge, MA, USA. New York, NY, USA: ACM. pp. 289–299.

[26] Mühlbauer W, Uhlig S, Feldmann A, Maennel O, Quoitin B, Fu B. Impact of routing parameters on route diversity and path inflation. Computer Networks 2010; 54: 2506–2518.

[27] Hiran R, Carlsson N, Gill P. Characterizing large-scale routing anomalies: a case study of the China Telecom incident. In: Springer-Verlag 2013 Passive Active Measurement; 18–19 March 2013; Hong Kong. Berlin, Germany: Springer. pp. 229-–238.

[28] Al-Musawi B, Branch P, Armitage G. BGP anomaly detection techniques: a survey. IEEE Commun Surv Tut 2017; 19: 377–396.

[29] Deshpande S, Thottan M, Ho TK, Sikdar B. An online mechanism for BGP instability detection and analysis. IEEE T Comput 2009; 58: 1470–1484.

[30] Zhang K, Yen A, Zhao X, Massey D, Wu SF, Zhang L. On detection of anomalous routing dynamics in BGP. In: Springer 2004 International Conference on Research in Networking; 9–14 May 2004; Athens, Greece. Berlin, Germany: Springer. pp. 259–270.

[31] Sriram K, Borchert O, Kim O, Gleichmann P, Montgomery D. A comparative analysis of BGP anomaly detection and robustness algorithms. In: IEEE 2009 Cybersecurity Applications Technology Conference for Homeland Security; 3–4 March 2009; Washington, DC, USA. Piscataway, NJ, USA: IEEE. pp. 25-38.

[32] Ganiz MC, Kanitkar S, Chuah MC, Pottenger WM. Detection of interdomain routing anomalies based on higher-order path analysis. In: IEEE 2006 International Conference on Data Mining; 18–22 December 2006; Hong Kong. Piscataway, NJ, USA: IEEE. pp. 874–879.

[33] Fischer F, Fuchs J, Vervier PA, Mansmann F, Thonnard O. Vistracer: A visual analytics tool to investigate routing anomalies in traceroutes. In: ACM 2012 International Symposium on Visualization for Cyber Security (VizSec); 15 October 2012; Seattle, WA, USA. New York, NY, USA: ACM. pp. 80–87.

[34] Moriano P, Iyer S, Camp LJ. Characterization of Internet Routing Anomalies through Graph Mining. Technical Report. Bloomington, IN, USA: Indiana University School of Informatics, Computing, and Engineering, 2017.

[35] Hiran R, Carlsson N, Shahmehri N. Crowd-based detection of routing anomalies on the Internet. In: IEEE 2015 Conference on Communications and Network Security; 28–30 September 2015; Florence, Italy. Piscataway, NJ, USA: IEEE. pp. 388-396.

[36] Lim H, Hou JC, Choi CH. Constructing Internet coordinate system based on delay measurement. In: ACM 2003 Internet Measurement Conference; 27–29 October 2003; Miami Beach, FL, USA. New York, NY, USA: ACM. pp. 129-142.