An Empirical Mixture Model for Large-Scale RTT Measurements

Romain Fontugne[1,2] Johan Mazel[1,2] Kensuke Fukuda[1,3]

[1]National Institute of Informatics [2]Japanese-French Laboratory for Informatics [3]Sokendai

***Abstract*—Monitoring delays in the Internet is essential to understand the network condition and ensure the good functioning of time-sensitive applications. Large-scale measurements of round-trip time (RTT) are promising data sources to gain better insights into Internet-wide delays. However, the lack of efficient methodology to model RTTs prevents researchers from leveraging the value of these datasets. In this work, we propose a log-normal mixture model to identify, characterize, and monitor spatial and temporal dynamics of RTTs. This data-driven approach provides a coarse grained view of numerous RTTs in the form of a graph, thus, it enables efficient and systematic analysis of Internet-wide measurements. Using this model, we analyze more than 13 years of RTTs from about 12 millions unique IP addresses in passively measured backbone traffic traces. We evaluate the proposed method by comparison with external data sets, and present examples where the proposed model highlights interesting delay fluctuations due to route changes or congestion. We also introduce an application based on the proposed model to identify hosts deviating from their typical RTTs fluctuations, and we envision various applications for this empirical model.**

***Index Terms*—RTT, backbone traffic, mixture model**

I. Introduction

From the early years of the Internet, round-trip time (RTT) is a key indicator of network conditions. Consequently, the research community has proposed numerous techniques for accurately estimating RTTs, ranging from Karn's algorithm for TCP congestion control [22], to recent work on detailed hop-by-hop measurements [27]. Algorithms estimating RTT from passive measurements [21], [35] have been particularly valuable to assess the network performance at large-scale. For example, they enabled researchers to measure RTT fluctuations [16], [14] on backbone networks or residential broadband access [26], and to understand the impact of the RTT distribution on TCP flow control [34].

Past work commonly measured RTT from numerous hosts by means of the median RTT, but, several studies [34], [14], [26] controversially reported that the RTT distribution of various flows is characterized by several distinct modes. Figure 1 depicts two RTT distributions of hosts seen at a backbone link and the corresponding median values. One can easily identify four distinct modes (or peaks) in the upper plot and two in the lower one. The values of these modes are of prime importance as they represent the typical RTTs experienced by large populations of hosts, whereas the overall median RTT is here of limited interest. For example, the median value for the lower plot of Figure 1 is significantly varying as the number of hosts for the two peaks is fluctuating, although, the typical RTTs (i.e. the peaks values) for these hosts are constant. Understanding the spatial and temporal dynamics of these typical RTTs is critical for various aspects of networking, including, topology models, geolocation, content delivery and Internet security.

In this work, we emphasize that thorough large-scale RTT studies should shift the focus to multimodal-based analysis in order to understand the typical RTT fluctuations of numerous related hosts at once. Consequently, we propose a model to monitor mixed distributions in RTT measurements. The proposed approach, first, uncovers typical RTTs with a mixture model, then, it correlates the typical RTTs identified at different points in time, finally, it formalizes the time evolution of the typical RTTs as a graph. To the best of our knowledge, this is the first attempt to date that permits systematic and efficient monitoring of RTT distribution mixture. As the proposed model summarizes and characterizes the dynamics of numerous RTT measurements, it allows one to comprehend typical RTT fluctuations experienced on the Internet and identify abnormal delays alterations.

We evaluate the proposed model with RTTs from 12 millions unique IP addresses collected during 13 years of backbone traffic. Using external datasets (i.e. geolocation database and BGP route information) and basic techniques from graph theory, we validate the relevance of resulting graphs in three ways. (1) We demonstrate that the proposed model permits to cluster hosts from the same geographical location, regardless their IP address and corresponding autonomous system (AS). (2) We present an application that identifies RTT fluctuations experienced by a large number of hosts due to infrastructure-wide events (e.g. AS path change or congestion). (3) We also introduce an application to compare raw RTT values of a single IP with the RTT dynamics uncovered by the proposed model, hence, highlight deviating hosts behavior due to local issues (e.g. overloaded host, Internet connectivity issues). The new traffic delay insights uncovered by the proposed method are valuable for various applications based on Internet delays (e.g. server selection algorithm) or inspecting Internet infrastructure (e.g. AS level topology).

II. Background

A. *Traffic Traces*

All results presented in this article are obtained with traffic from the MAWI archive [12], which is a collection of traffic traces captured at a transit link between the WIDE backbone network (AS 2500) and a commercial ISP. The traffic is daily measured between 14:00 and 14:15 JST since January 2001. In this work we analyze 4678 traces accounting for more than 13 years of Internet traffic (Jan. 2001 to Mar. 2014). All traces, with scrambled IP addresses and without packet payload, are freely available on the Internet. In this work, however, we had access to the original traces with unmodified addresses to validate the results of the proposed model.

B. Per-host RTT Estimation

The estimation of the RTTs in the MAWI traces is done with a simple and fast technique based on Karn's algorithm [22]. Namely, we compute samples delay $\delta = \theta_{ACK} - \theta_{SEQ}$ where θ_{SEQ} is the observation time of a TCP packet with a certain sequence number, and θ_{ACK}, the observation time of its corresponding acknowledgment. Retransmitted packets are ignored as the corresponding acknowledgment might be duplicated thus bias the measure. Notice that δ is computed from a pair of packets, and it represents the delay between the MAWI measurement point and the host that sent the acknowledgment. Therefore, a single TCP flow allows us to measure numerous δ for two end-points. Then, the daily RTT of a host A is the median value of all δ corresponding to A. Nonetheless, to compute robust estimations, and because of RTT significant variations [7], we keep only median values that are computed from at least 5 δ values. This operation is repeated for every traces to produce RTT time series for each host. Our dataset consists of RTT estimates from 12 millions unique IP addresses from January 2001 to the end of March 2014.

As we expect end-points within the WIDE network to have significantly lower RTTs than other hosts, we classify hosts into two categories using the routers MAC address. RTT of hosts that are behind the router on the WIDE network are refered as RTT_{IN} and other RTTs are refered as RTT_{OUT}.

C. Observations

Figure 1 illustrates the distribution of the RTT_{OUT} and RTT_{IN} for a typical MAWI trace (i.e. 2014/03/07). Both distributions are multimodal; for RTT_{OUT}, we observe four prominent RTT values ranging between 25 and 300 ms, while, RTT_{IN} features two main RTT values between 1 to 10 ms. These observations confirm the expected difference between the range of RTT_{IN} and RTT_{OUT}. Intuitively, prominent values in RTT_{OUT} discriminate hosts from distinct countries, whereas RTT_{IN} highlights different sets of hosts within Japan. Estimating the exact number of typical RTTs from histograms similar to the ones of Figure 1 is hazardous, because the number of local maxima dramatically varies with the resolution of the histogram (i.e. bin size) and the scale of the axes. The following section presents the proposed methodology to systematical uncover the typical RTTs and monitor their time evolution.

III. Methodology

As depicted in Figure 3 the input of the proposed methodology is a set of host RTTs time series, which are analyzed in three key steps:

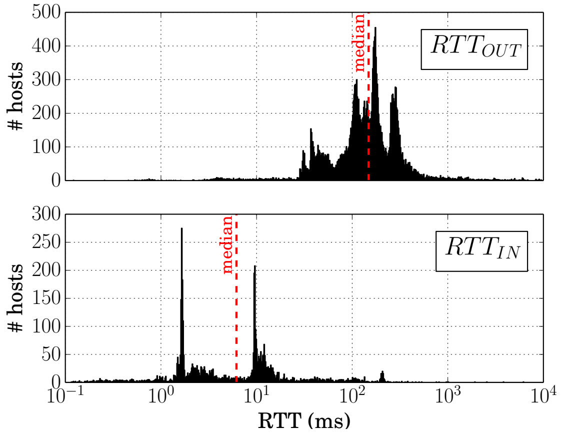


Fig. 1. Distribution of RTT_{IN} and RTT_{OUT} for MAWI traffic collected on the 7^{th} of March 2014.

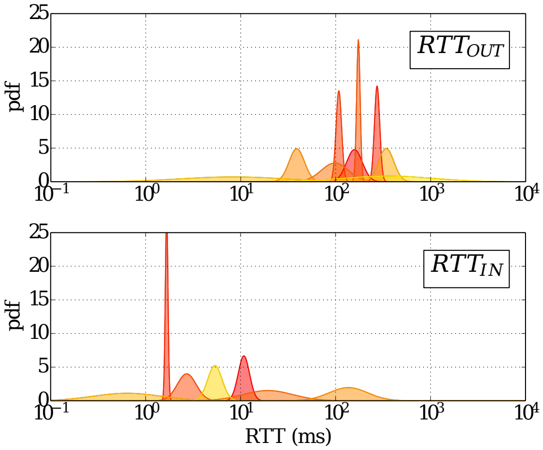


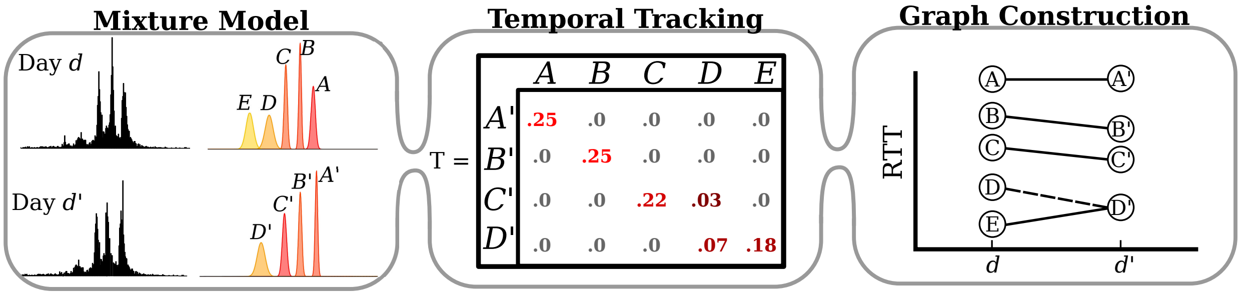
Fig. 2. Results of the log-normal mixture model for RTTs from MAWI traffic collected on the 7^{th} of March 2014.

1) uncover the daily RTT distributions using a mixture model,
2) link RTT distributions from similar sub-population of IPs across time,
3) formalize RTTs time evolution in a graph for further systematical analysis.

Thereby, the result of the proposed methodology is a graph that characterizes the RTT dynamics of hosts with similar Internet delay behavior.

A. Mixture Model

The first step aims to detect the typical RTTs in a trace and estimate their distributions. This is a classical problem in statistics that is addressed by mixture models. For this research, the mixture model should be able to identify an unknown number of mixed component with a low computational cost in order to analyze a large number of RTT measurements. The Dirichlet process mixture model [9] is the natural candidate for this task as it is a generalization of finite



$$T = \begin{array}{c|ccccc} & A & B & C & D & E \\ \hline A' & .25 & .0 & .0 & .0 & .0 \\ B' & .0 & .25 & .0 & .0 & .0 \\ C' & .0 & .0 & .22 & .03 & .0 \\ D' & .0 & .0 & .0 & .07 & .18 \end{array}$$

Fig. 3. Overview of the proposed method using two traces. Left: Typical RTT distributions are uncovered with a mixture model. Center: Uncovered distributions similarities are summarized in the transition matrix, T. Right: The distributions and their dynamics are formalized as a graph.

mixture models. We employ and recommend the recent variational inference algorithm for Dirichlet process mixture model [10] as it is significantly faster that the usual algorithm using Monte-Carlo Markov Chain [20]. The methodology proposed in this article, however, is not bound to a specific mixture model; we only assume that the mixture model uncovers an unknown number of typical RTTs for each trace and estimate the mean and variance (μ, σ^2) of their distributions.

Our experiments revealed that log-normal distribution fits better our RTT measurements than the normal distribution. We confirmed this observation by computing the Bayesian Information Criterion (BIC) for both distributions and found a significant difference in favor of the log-normal distribution. Consequently, we feed the mixture model with the logarithm (base 10) of our RTT measurements and obtain the mean and variance of the RTT distributions in the log-space. The mean RTT of the identified distribution, (μ, σ^2), is hence equal to 10^{μ}.

Figure 2 depicts an example of the mixture model results using the RTTs of Figure 1. Distributions with a small variance emphasize particular RTTs values that are prominent in the trace, whereas, distributions with very large variance give no information about typical RTTs. The mixture model is independently analyzing daily measurements, thus, we obtain for every days a variable numbers of RTT distributions. To track the time evolution of the typical RTTs the next step is to connect the distributions identified at two distinct days.

B. Temporal Tracking

To find the relationships between two sets of distributions identified at different days, we implement a simple probabilistic model using the parameters uncovered by the mixture model and the IP addresses appeared in both sets. Namely, we derive the transition matrix T describing the similarities between the distributions in both sets.

Let $P^d = (\mu_i, \sigma_i^2)$, $i \in [1, m]$ be the parameters of the m distributions identified on day d, and $P^{d'} = (\mu_j, \sigma_j^2)$, $j \in [1, n]$, be those for the n distributions identified on day d'. Using the probability density function of each distribution,

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

and the RTTs of the hosts monitored both days, X^d and $X^{d'}$, we compute the m-by-n transition matrix

$$T_{i,j} = \frac{f(X^d; P_i^d) \cdot f(X^{d'}; P_j^{d'})}{\sum\limits_k \sum\limits_l f(X^d; P_k^d) \cdot f(X^{d'}; P_l^{d'})}.$$

Consequently, $T_{i,j}$ is the probability of the distributions i and j to stand for the same set of hosts.

C. Graph Construction

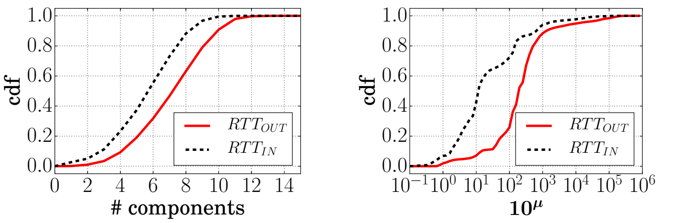
A transition matrix relates distributions identified at two different days, therefore, a sequence of matrices allows us to connect numerous distributions standing for an extended period of time. To ease the manipulation of numerous transition matrices, we merge them in a weighted graph in which the vertices are the identified distributions, and edges are weighted with the probabilities from the transition matrix. We discard edges with null or low weight as they connect distributions that are unlikely related.

Given the distributions $P = P^d, d \in [1, k]$ for k days and all corresponding transition matrices $T^{d,d+\Delta}, d \in [1, k-1], \Delta \in [1, k-1]$. Then, the graph $G = (V, E)$ consists of the set of vertices $V = P$, and the set of edges E connecting vertices that have a transition probability higher than the uniform distribution, $\forall e \in E, e = (P_i^d, P_j^{d'})$ where $i \in [1, m]$, $j \in [1, n]$ and $T_{i,j}^{d,d'} > 1/mn$. Thereby, the proposed model avoids erroneous distributions uncovered by the mixture model (i.e. distributions standing for no typical RTT values), such as the distributions with very large variance of the upper plot of Figure 2.

In summary, the graph G is a network of RTTs distributions that are connected with weighted edges proportional to the hosts they represent. Consequently, a cluster of strongly connected nodes in G represents a set of RTT distributions of correlated IP addresses, and the parameters of the distributions along this path allows one to accurately quantify the RTT time evolution for these IP addresses.

IV. Evaluation

We now present several examples of graph constructed with the proposed model and simple techniques to manipulate them. This evaluation is conducted with the 13 years of traffic presented in Section II. First, we present broad observations



(a) CDF of the number of identified components per day.

(b) CDF of 10^{μ}, the modeled RTT of the identified components (milliseconds).

Fig. 4. Results of the mixture model using all MAWI traces from January 2001 to the end of March 2014.

using the whole dataset. Second, we verify the ability of the model to cluster hosts with similar RTTs by looking at the geographical locations of hosts in sets of strongly connected nodes. Then, we introduce a simple application monitoring the dynamics of the typical RTTs uncovered by the model, and study their relationships with BGP route changes, traffic throughput and network congestion. Finally, we present another application of the proposed model to identify hosts that deviates from their typical RTT fluctuations.

A. Longitudinal study

The number of RTT distributions that are identified for each trace obviously depends on the underlying mixture model (Section III-A) and corresponding parameter values. Using the variational inference algorithm for Dirichlet process mixture model [10], and prior weights equal to 0.1, we obtained a total of 13.5 components per day for both RTT_{OUT} and RTT_{IN}. Figure 4a depicts the cumulative distribution of the number of identified components per day for RTT_{OUT} and RTT_{IN}. We observe that the number of components for RTT_{OUT} is usually larger than the one for RTT_{IN}, meaning that more characteristic RTTs are observed for the hosts that are outside the WIDE network. This is mainly because RTT_{OUT} represents a larger and broader population of hosts that are situated in various networks and geographical locations. The distribution of 10^{μ}, the mean RTT of each component (Fig. 4b), shows that 60% of the components identified in RTT_{IN} feature a mean RTT below 15 ms, whereas for RTT_{OUT} only 10% of the components are below 15 ms. In fact, 50% of the RTT_{OUT} components have a mean RTT between 100 and 400 ms. Because of the propagation delay, we expect most of the hosts represented by RTT_{OUT} to be located overseas and the majority of RTT_{IN} hosts to be in Japan.

B. Geolocation

Figure 5 illustrates the graph obtained with the RTT_{OUT} of the last 10 days of our dataset. For clarity, only edges for consecutive days are shown (i.e. $\Delta = 1$). As explained in Section III-C, strongly connected nodes represent several RTT distributions for similar hosts. Using the Louvain community mining algorithm [11], we found 6 clusters of strongly connected nodes denoted $C0, ..., C5$ in Figure 5. The relevance of these clusters is evaluated using Maxmind's geolocation database, GeoIP City [6]. For each cluster we retrieve a set of representative IP addresses, that is, all IP addresses that appear at least three times in the traces, and verify the corresponding country code with the geolocation database. Table I enumerates each cluster with the corresponding country codes that accounts for more than 3%, and the mean modeled RTT (denoted $\overline{RTT}$) that is the average 10^{μ} value for all nodes in the cluster weighted by the number of IP addresses they represent. For presentation purposes all European countries are grouped together under the *EU* country code.

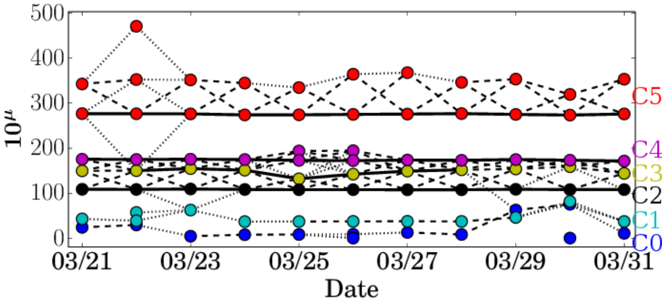


Fig. 5. Graph generated with RTT_{OUT} from 2014/03/21 to 2014/03/31. Only edges connecting consecutive days are displayed. Plain edges mean $T_{i,j}^{d,d'} \geq 0.075$, dashed edges are $0.075 > T_{i,j}^{d,d'} \geq 0.025$, and dotted edges are $0.025 > T_{i,j}^{d,d'}$.

	JP	KR	US	CA	EU	CN	$\overline{RTT}$
C0	**98%**						19 ms
C1		**97%**					44 ms
C2			**91%**				108 ms
C3			**73%**			11%	149 ms
C4			**87%**	4%			175 ms
C5			8%		**73%**	3%	289 ms

TABLE I
HOSTS GEOLOCATION BREAKDOWN FOR THE CLUSTERS OF NODES IDENTIFIED IN FIGURE 5.

The two clusters with a $\overline{RTT}$ under 50 ms, $C0$ and $C1$, stand exclusively for end-points respectively in Japan and Korea. A few IP addresses in $C0$ or $C1$ are classified as *US* or *EU* by the geolocation database, however, the RTTs of these IP addresses evoke that they are actually located near the MAWI measurement point in Japan. Further inspections revealed that these IP addresses are assigned to U.S. or European ASes but provide services in Japan. This observation emphasizes the benefits of the proposed empirical approach as opposed to an AS-based approach.

The three clusters $C2$, $C3$ and $C4$ are primarily composed of IP addresses located in U.S. but feature different RTTs. With $\overline{RTT}$ equals to 108 ms, $C2$ stands mainly for hosts located on the West Coast of the U.S., whereas, $C4$ stands principally for the East Coast ($\overline{RTT} = 175$ ms). $C3$ is fluctuating between $C2$ and $C4$, it captures various American IP addresses and several Chinese addresses.

$C5$ is the cluster with highest delays, $\overline{RTT} = 289$ ms, it

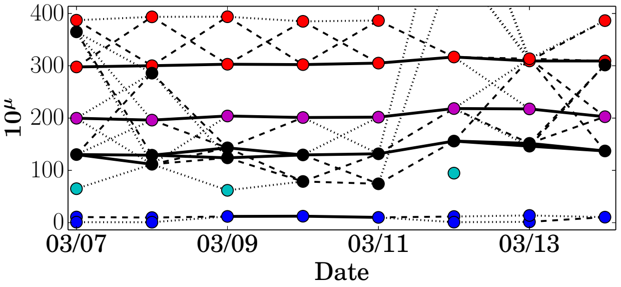


Fig. 7. Graph generated with traffic captured during the week of the Tohoku earthquake (2011/03/07-2011/03/14).

stands mainly for European hosts. Figure 5 shows, however, that the cluster is composed of two typical RTT distributions, one below and one above 300 ms. This bimodal distribution can be systematically identified with the proposed model and a min-cut algorithm, thus, it could be decomposed into smaller clusters to provide more detailed analysis.

We have also inspected the results obtained with RTT_{IN} and the same 10 days, the Louvain algorithm again identified 6 clusters but all are classified entirely in Japan by the geolocation database. Yet the model highlights 5 characteristic clusters with $\overline{RTT}$ ranging from 0.5 to 13 ms and an outlier cluster with $\overline{RTT}$ equals to 123 ms. From $\overline{RTT}$, we infer the location of the corresponding hosts; clusters with $\overline{RTT} < 1$ ms are located in the same building as the MAWI measurement point, $1 \leq \overline{RTT} < 10$ ms in the same urban area as MAWI measurement point, and $\overline{RTT} > 10$ ms represents hosts located in other Japanese cities.

Analysis of RTT_{IN} through the whole dataset highlights clusters whose $\overline{RTT}$ is around 500 ms, these clusters represent Indonesian hosts that were connected to the WIDE network using a satellite link (see Figure 4b, $10^{\mu} = 10^{2.7}$). Therefore, we emphasize the benefits of the proposed model as this type of cluster is easily identifiable in our experiments, but, biases simple median-based analysis.

C. Graph Dynamics

Since RTT changes on the Internet are of prime importance, we have implemented an application based on the proposed model to identify significant RTT variations in MAWI. Namely, we compute a daily score for each cluster, called the $\overline{RTT}$ *fluctuation*, which is the difference between the $\overline{RTT}$ of two consecutive days normalized by the standard deviation of $\overline{RTT}$ from the entire cluster. Consequently, stable RTTs exhibit $\overline{RTT}$ fluctuations close to zero, whereas, significant RTT variations of a cluster are represented by higher values.

Understanding the root causes of these RTT changes is a difficult task because it can be the results of disturbances that appear on distance network where no measurements are available. Since previous studies [33] have reported that BGP routing changes usually affect the RTT values of Internet hosts, we investigate the relationships between BGP updates and the $\overline{RTT}$ fluctuations of each cluster identified with the proposed model.

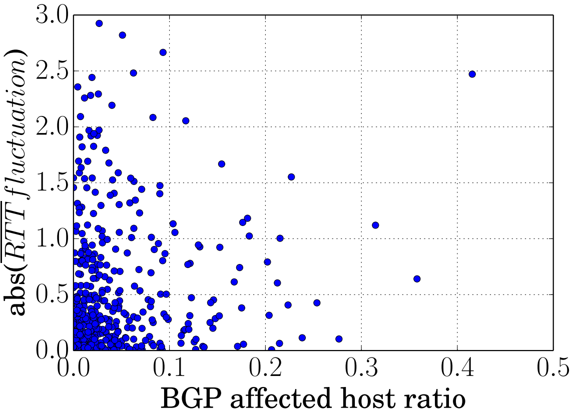


Fig. 8. Comparison of the ratio of IP addresses affected by a BGP route change with the absolute values of the typical RTT fluctuations captured by the proposed model.

We retrieved the BGP Route Information Base (RIB) of the WIDE network from the RouteViews Project [5] in order to seek for every BGP route changes that happened between two consecutive MAWI traces. Using the IP prefix specified by the BGP updates, we can accurately count the number of IPs in our dataset that are affected by BGP route changes. Consequently, for each cluster we compute its daily $\overline{RTT}$ fluctuations and compute the ratio of IP addresses in this cluster that are affected by a BGP route change. Figure 8 depicts for each day and each cluster the absolute $\overline{RTT}$ fluctuation and ratio of IP affected by a BGP update for all RTT measures from the 1^{st} of the January to the 31^{st} of March 2014. Thereby we found that 66% of the BGP updates that affects at least 15% of the cluster IPs exhibit an absolute $\overline{RTT}$ fluctuation higher than 0.15. These results are similar to the ones reported in [33], namely, 72% of the BGP route changes affect Internet hosts RTTs. However, Figure 8 also highlights that a large fraction of the RTT fluctuations captured by the proposed method are not due to BGP updates.

As discussed in [33], intra-AS route changes (e.g. OSPF route changes) are also a potential source of RTT alterations. Quantifying the impact of these route changes, however, is impractical as it would require access to internal routing information of numerous ASes. Nonetheless, manual inspection of RTT measurements with the proposed model allows us to discover evidences of these intra-AS route changes.

For example, the $\overline{RTT}$ fluctuations for the (red) cluster around 10 ms of Figure 6 are usually close to zero, nonetheless, two abnormally high values drawn our attention to January and March 2010. Indeed $\overline{RTT}$ for this cluster is especially constant over time, but, on the the 27^{th} of January it dropped by 1.5 ms and stayed stable until the 25^{th} of March where it came back to its original value. This behavior has been identified for different clusters throughout the whole

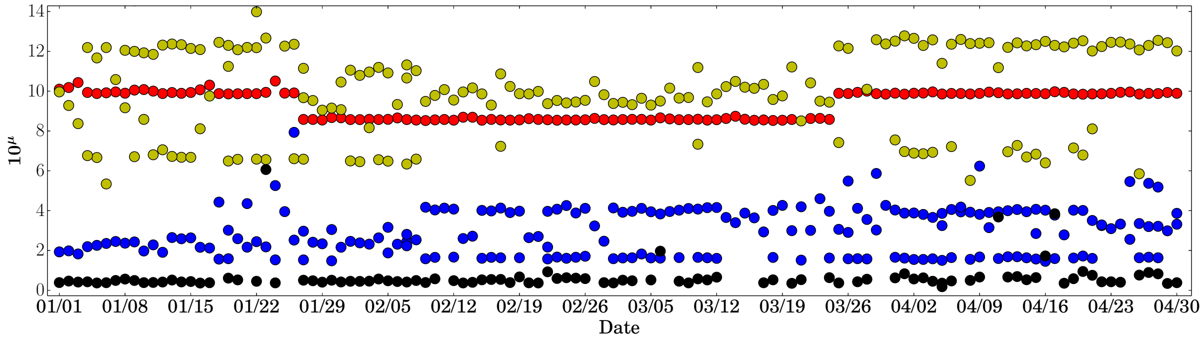


Fig. 6. Typical RTTs, 10^μ, identified in MAWI traces from 2010/01/01 to 2010/04/30 (RTT_{IN}). Example of $\overline{RTT}$ fluctuations due to a route change (see the red cluster around 10ms). The y-axis is in milliseconds.

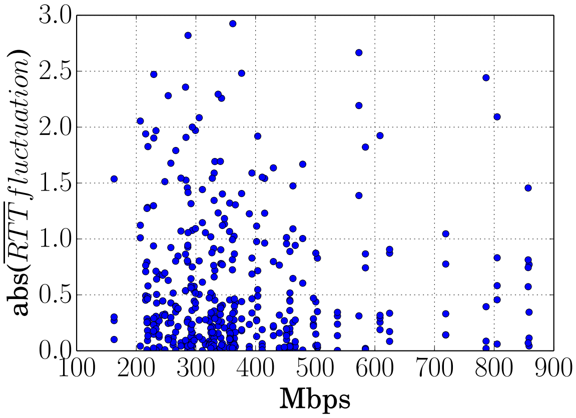


Fig. 9. Comparison of MAWI throughput with the absolute values of the typical RTT fluctuations captured by the proposed model.

dataset, and is very similar to the ones observed during route changes [33]. Finding these route changes without intra-AS routing information is particularly difficult [33], nonetheless, the proposed method enables new alternatives to tackle this task using IP traffic.

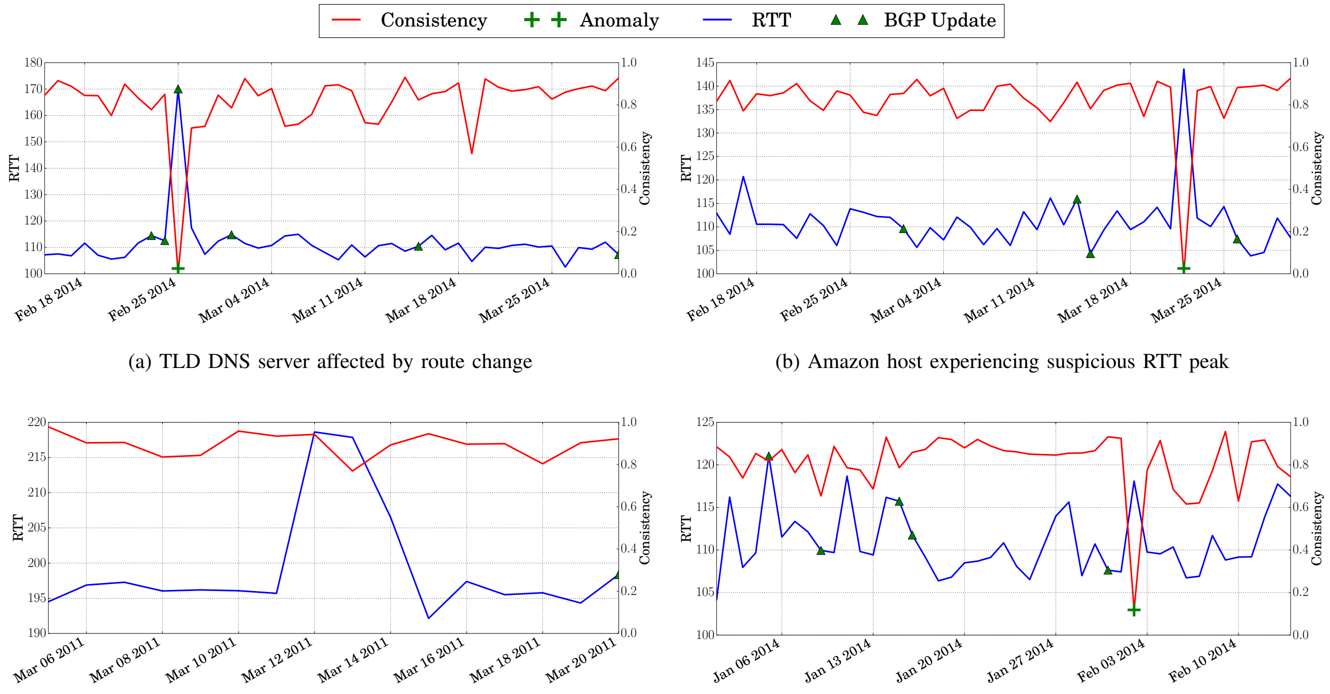
Another obvious source of RTT alterations on the Internet is network congestion. Indeed, RTTs are directly impacted by the routers queuing delay or packet drop. Since systematically evaluating the impact of congestion with router load measurements from numerous ASes is impractical, we propose an indirect study based on the relationships between the average throughput of the MAWI traces and the clusters daily $\overline{RTT}$ fluctuations. Intuitively, voluminous MAWI traces (i.e. high bit-rate) have been captured when congestion happened near the MAWI measurement point or elsewhere. Figure 9 depicts the relationships between the average throughput of the MAWI traces and the clusters daily $\overline{RTT}$ fluctuations for all RTT measurements collected from January to March 2014. Overall, significantly higher $\overline{RTT}$ fluctuations are observed when the average throughput is higher than 500 Mbps, meaning that the proposed model effectively captures the RTT increases due to bottleneck in the network.

Manually inspecting abnormal $\overline{RTT}$ fluctuations draws our attention to a different case of network congestion which was observed after the Tohoku earthquake and tsunami disaster from the 11^{th} of March, 2011. Three clusters representing hosts in the U.S. and Europe (similar to $C2$, $C3$, $C4$, and $C5$ in Figure 5 and Table I) synchronously exhibit an $\overline{RTT}$ increase on the 12^{th} of March which is the first MAWI trace following the disaster. Apart from edge networks located at the northern east part of Japan, the Internet infrastructure has been impressively resilient to the earthquake and tsunami. Related studies [13], [17] highlighted that the redundant and over-provisioned backbone network policy in Japan could limit the disaster impact on the total traffic, and the impact on BGP was insignificant. Using the WIDE BGP route information from RouteViews [5], we found that in our dataset between the 11^{th} and the 12^{th} of March no Japanese IP is affected by route changes and less than 4% of the hosts in the U.S. experienced a BGP update. Nevertheless, since the proposed model identified an RTT increase close to 20 ms for all hosts outside Japan, but, no change for Japanese hosts, we infer that the cause of this increase comes from congested routers, or an intradomain route change, of the transit network to the U.S.

D. Anomalous Host Detection

The proposed model identifies Internet hosts with similar RTTs and captures their average RTT fluctuations. The causes of these global fluctuations are mainly due to events happening at the Internet core infrastructure, thus, simultaneously affecting numerous Internet hosts. Characterizing the typical behavior of sets of hosts also permits to identify Internet hosts that deviate from their usual behavior. For example, confronting the RTT values of an IP address with the proposed model allows us to verify if the RTT fluctuations of this address are consistent with the fluctuations of the thousands IP addresses monitored to build the model. In our experiments we verify if the behavior of a monitored IP address is consistent with the behavior of the cluster it belong to with the following *consistency check*.

1) Consistency Check: We verify if a certain host is consistent with a cluster of nodes (identified with the Louvain



(a) TLD DNS server affected by route change

(b) Amazon host experiencing suspicious RTT peak

(c) RTT fluctuation during the Tohoku earthquake

(d) Suspicious RTT fluctuations identified with the consistency check

Fig. 10. Four examples of consistency checks performed with the proposed model.

algorithm) by matching the host RTT values with the RTT distribution represented by each node. Let $P^{d,C} = (\mu_i, \sigma_i^2)$, $i \in [1, m]$ be the parameters of the m distributions of the clusters C, on day d, and X^d the RTT value for the given host on the same day. Using the probability density function of the distributions, the consistency score, $\mathcal{C}^d$, is defined as the maximum probability to match one of the cluster distributions:

$$\mathcal{C}^d = \frac{\max_i(f(X^d; P_i^{d,C}))}{\sum_j f(X^d; P_j^{d,A})}$$

where $P^{d,A} = (\mu_j, \sigma_j^2), j \in [1, n]$ (with $m \leq n$) are all the distributions of all clusters on day d. Therefore, consistency scores close to 1 represent RTT values in accordance with the cluster behavior while low consistency scores highlight contradictory RTT values.

In order to systematically identify deviation of hosts from their cluster behavior, we report as anomalous any consistency score $\mathcal{C}^d$ where $\mathcal{C}^d < \text{mean}(\mathcal{C}) - 3\,\text{std}(\mathcal{C})$.

2) Examples: To illustrate the benefits of the consistency check, Figure 10 depicts the RTT values of four hosts along with their consistency score. For example, Figure 10a shows the RTT measures for a TLD DNS server in February and March 2014. This host features particularly stable RTTs thus its consistency is also stable over time. However, on the 25^{th} of February, a sudden RTT increase led to a drop of the consistency score which is flagged as anomalous. Further inspections revealed that the AS path for this DNS server changed on the same day, and the only other IP affected by the BGP update is also flagged as anomalous on that day. Figure 10b shows another example of anomalous consistency score, but this time for an host from an Amazon network. As there is no BGP update happening on that day for this prefix, and this is the only anomalous IP we found on this network, the low consistency score in this case highlights local issues that affect only this host (e.g. overloaded node, Internet connectivity issues, denial-of-service). These two examples illustrate the efficiency of our consistency check to asses the normal behavior of a host and identify suspicious behaviors. Nonetheless, both cases could be easily identified by analyzing only the RTT raw values as both RTT peaks are rather obvious.

We now present two examples where the proposed model and consistency check surpass a simple raw-RTT-based analysis. The first example comes from data collected during the Tohoku earthquake in March 2011. Figure 10c depicts the typical RTT time evolution observed during the earthquake and the corresponding consistency score. Despite the significant RTT increase measured right after the earthquake, no anomaly is reported by the proposed consistency check. As explained in the previous section, this RTT increase is common to all hosts outside of Japan. Since the model captured this as the typical behavior, and this host behaves accordingly, our consistency check flags no anomaly. Therefore, looking at both the dynamics of the clusters (see Section IV-C) and the consistency of this host, allows us to infer that the earthquake significantly affected numerous hosts RTT but the host of Figure 10c is not behaving differently. In this case, analyzing only the raw-RTT-values is particularly difficult, either this RTT increase would be reported as anomalous, or one would conduct a computationally expensive comparison between these values and the millions other RTT values from other hosts to find this global trend in the dataset. Thereby, the proposed model is an efficient way to carry out numerous comparisons with any given RTT values at lower computational cost.

The last example is the RTT values of an host in the U.S. where the third highest RTT value has been solely reported as anomalous. As shown in Figure 10d the highest RTT value for this host is caused by a BGP update. This path update has affected numerous hosts, hence, this is captured by the model as a typical behavior and is not reported as anomalous. However, the RTT increase of the 2^{nd} of February is reported as anomalous because most of the other hosts behave differently on that day. Without the proposed model, the analysis of this complex situation would be laborious and a potential source of false positive alarms.

V. Related Work

Internet delays and RTTs have received a lot of attention from the networking research community. Various RTT estimation algorithms have been proposed to accurately measure RTTs, including, end-hosts measurements techniques [22], RTT estimation from network link passive measurements [21], [35], and hop-by-hop estimations using IP timestamp option [27] or ICMP timestamp [8].

Researchers have also carried out numerous efforts to collect Internet delay measurements; For example, the CAIDA's Skitter and Archipelago (Ark) projects [4], [2], the AMP project [1], or PingER [3]. These projects rely mainly on large probe deployments to collect delay measurements from diverse geographical locations.

These various RTT measurement techniques have enabled the design of diverse applications. For example various applications identify the geolocation of an IP address using RTT measurements [15], [23], [18], [25]. In cyber security, metrics measuring the impact of DoS attacks are usually taking into account RTT values [24] or similar delay measurements [28].

Similar to the evaluation conducted in our work, several works inspect the relationships between route changes and RTT fluctuations. These works usually rely on RTT active measurements and compare RTT variations with intra and inter domain routing changes [30] or BGP updates [36], [33].

To the best of our knowledge, RTT measurements models have been rarely proposed in the past. Zhang et al. proposed, DS^2 [37], a tool analyzing numerous end-to-end delay measurements obtained with the King tool [19], and summarizing these measurements to estimate delay between arbitrary end hosts. DS^2 is a valuable help for designers of large-scale distributed systems that are relying on overlay networks. This approach is orthogonal to the model proposed in this paper, as DS^2 disregards the time evolution of Internet delays.

The model proposed in this paper supplements the vast RTT literature, by simultaneously analyzing delay measurements from numerous hosts and characterizing their time evolution.

VI. Discussion

Our results support the benefits of the proposed model to monitor and investigate spatial and temporal RTT dynamics. In fact, by dissecting the RTT distributions, the proposed model gives great insights into the typical delays experienced by a large population of IP addresses. The identification of hosts deviating from typical behaviors uncovered by the model further emphasizes the advantages of the proposed model and its benefits over raw RTT analysis.

Since the proposed model is designed for low memory usage and computational complexity, we successfully conducted all our experiments on a commodity computer. Moreover, as the model follows a statistical approach, it is inherently suitable to sampled traffic.

In this article we presented results using fixed parameters values for the mixture model, however, further experiments revealed that the resolution of the model can increase with more sensitive settings. Overall, we found that the quality of RTT, hence, the model accuracy, decreases with respect to the distance between the measurement point and the monitored hosts. The noise added at each hop indeed makes RTT measures inaccurate, therefore, in our experiments hosts from France and U.K. are not distinguishable, whereas we observed clusters of hosts located in different Japanese cities.

The empirical approach presented in this article has the advantage of monitoring RTTs as it is experienced by Internet users, unlike active measurements that could be mislead by load balancing [29]. Nevertheless, the proposed model is not limited to passive measurements, it is also suitable for active measurements analysis.

As emphasized in Section IV detecting route change is an intuitive application of the proposed model. Nonetheless, this model was originally designed to detect hosts with a sudden RTT increase due to DoS attacks [24]. As shown with the consistency check results of Fig.10, modeling common RTT variations allows us to build reference data, and reports hosts deviating from this common behavior. Interestingly, this approach to detect DoS (still under development) allows one to detect hosts under attack from a distant monitored network where no malicious traffic may be observed. Using the proposed model, we can accurately identify the side effects of the attack by taking into account the various events happening on the network (e.g. route changes, or congestion) that can affect the RTT measurements.

The application domain of the proposed model is, however, not limited to these examples. In principle, any work involving numerous measurements of end-points delays could take advantage of this model. The model leverages large-scale measurements analysis by computing a coarse grained view of the RTTs dynamics, thus, avoiding the burden of enormous datasets. Therefore, the proposed model has potential benefits in various research domains including overlay networks [37], server selection [32], routing [30], [36], [33], geolocation [18], security [24] and Internet outage detection [31].

From our experience we found that results from applications based on the proposed model are, however, particularly difficult to evaluate. Since RTTs are the sum of various networks delays, applications may identify events appearing outside of the monitored network. Due to the lack of data source for certain networks (e.g. OSPF messages, and routers load), carrying out comprehensive evaluations for these applications presents real challenges.

VII. Conclusions

Delays on the Internet are of prime importance for various time-sensitive applications. This article proposed an empirical model to uncover and monitor the typical RTTs of a large population of connected hosts. The result of the model consists in a graph summarizing the time evolution of the uncovered typical RTTs. Consequently, we analyzed RTTs to millions of IP addresses using the proposed model and basic graph theory techniques. The model is evaluated with external datasets, such as geolocation database and BGP route information. We also presented results where the model captured important RTT fluctuations caused by route changes or congestion, and introduced an application to detect hosts exhibiting anomalous RTT values.

Acknowledgments

We thank Patrice Abry and Kenjiro Cho for initial discussions that encouraged us to develop this work. We also thank Jean Lorchat and Cristel Pelsser for their valuable feedback and help with BGP data. This research has been supported by the Strategic International Collaborative R&D Promotion Project of the Ministry of Internal Affairs and Communication in Japan (MIC) and by the European Union Seventh Framework Programme (FP7/2007- 2013) under grant agreement No. 608533 (NECOMA). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the MIC or of the European Commission.

References

[1] Active Measurement Project. *http://research.wand.net.nz/software/amp.php*.

[2] Ark. *http://www.caida.org/projects/ark/*.

[3] PingER. *http://www-iepm.slac.stanford.edu/pinger/*.

[4] Skitter. *http://www.caida.org/tools/measurement/skitter/*.

[5] The RouteViews project. *http://www.routeviews.org/*.

[6] MaxMind GeoIP City. *http://www.maxmind.com/en/city*, 2012.

[7] J. Aikat, J. Kaur, F. D. Smith, and K. Jeffay. Variability in TCP round-trip times. In *Proceedings of IMC'03*, pages 279–284. ACM, 2003.

[8] K. G. Anagnostakis, M. Greenwald, and R. S. Ryger. cing: Measuring network-internal delays using only existing infrastructure. In *INFOCOM 2003*, volume 3, pages 2112–2121. IEEE, 2003.

[9] C. Antoniak. Mixtures of dirichlet processes with applications to bayesian nonparametric problems. *The annals of statistics*, 2(6):1152–1174, 1974.

[10] D. M. Blei, M. I. Jordan, et al. Variational inference for dirichlet process mixtures. *Bayesian analysis*, 1(1):121–143, 2006.

[11] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, (10):P10008, 2008.

[12] K. Cho, K. Mitsuya, and A. Kato. Traffic data repository at the WIDE project. In *USENIX 2000 Annual Technical Conference: FREENIX Track*, pages 263–270. USENIX Association, 2000.

[13] K. Cho, C. Pelsser, R. Bush, and Y. Won. The Japan earthquake: the impact on traffic and routing observed by a local ISP. In *Proceedings of CoNEXT Special Workshop on Internet and Disasters*. ACM, 2011.

[14] B.-Y. Choi, S. Moon, Z.-L. Zhang, K. Papagiannaki, and C. Diot. Analysis of point-to-point packet delay in an operational network. *Computer Networks*, 51(13):3812–3827, 2007.

[15] F. Dabek, R. Cox, F. Kaashoek, and R. Morris. Vivaldi: A decentralized network coordinate system. *ACM SIGCOMM Computer Communication Review*, 34(4):15–26, 2004.

[16] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and S. Diot. Packet-level traffic measurements from the Sprint IP backbone. *Network, IEEE*, 17(6):6–16, 2003.

[17] K. Fukuda, M. Aoki, S. Abe, Y. Ji, M. Koibuchi, M. Nakamura, S. Yamada, and S. Urushidani. Impact of Tohoku earthquake on R&E network in Japan. In *Proceedings of CoNEXT Special Workshop on Internet and Disasters*. ACM, 2011.

[18] P. Gill, Y. Ganjali, B. Wong, and D. Lie. Dude, wheres that IP?: circumventing measurement-based IP geolocation. In *Proceedings of the 19th USENIX conference on Security*, pages 241–256. USENIX Association, 2010.

[19] K. P. Gummadi, S. Saroiu, and S. D. Gribble. King: Estimating latency between arbitrary internet end hosts. In *Proceedings of IMW'02*, pages 5–18. ACM, 2002.

[20] H. Ishwaran and L. F. James. Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 96(453), 2001.

[21] H. Jiang and C. Dovrolis. Passive estimation of TCP round-trip times. *ACM SIGCOMM Computer Communication Review*, 32(3):75–88, 2002.

[22] P. Karn and C. Partridge. Improving round-trip time estimates in reliable transport protocols. *ACM SIGCOMM Computer Communication Review*, 17(5):2–7, 1987.

[23] E. Katz-Bassett, J. P. John, A. Krishnamurthy, D. Wetherall, T. Anderson, and Y. Chawathe. Towards IP geolocation using delay and topology measurements. In *Proceedings of IMC'06*, pages 71–84. ACM, 2006.

[24] K.-c. Lan, A. Hussain, and D. Dutta. Effect of malicious traffic on the network. In *Proceedings of PAM'03*. Springer, 2003.

[25] R. Landa, R. Clegg, J. Araujo, E. Mykoniati, D. Griffin, and M. Rio. Measuring the relationships between internet geography and rtt. In *Proceedings of Computer Communications and Networks (ICCCN) 2013*, pages 1–7. IEEE, July 2013.

[26] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On dominant characteristics of residential broadband internet traffic. In *Proceedings of IMC'09*, pages 90–102. ACM, 2009.

[27] P. Marchetta, A. Botta, E. Katz-Bassett, and A. Pescape. Dissecting round trip time on the slow path with a single packet. In *Proceedings of PAM'14*, pages 88–97. Springer, 2014.

[28] J. Mirkovic, A. Hussain, B. Wilson, S. Fahmy, P. Reiher, R. Thomas, W.-M. Yao, and S. Schwab. Towards user-centric metrics for denial-of-service measurement. In *Proceedings of the 2007 workshop on Experimental computer science*, page 8. ACM, 2007.

[29] C. Pelsser, L. Cittadini, S. Vissicchio, and R. Bush. From Paris to Tokyo: on the suitability of ping to measure latency. In *Proceedings of IMC'13*, pages 427–432. ACM, 2013.

[30] H. Pucha, Y. Zhang, Z. M. Mao, and Y. C. Hu. Understanding network delay changes caused by routing events. *ACM SIGMETRICS Performance Evaluation Review*, 35(1):73–84, 2007.

[31] L. Quan, J. Heidemann, and Y. Pradkin. Trinocular: Understanding internet reliability through adaptive probing. *ACM SIGCOMM Computer Communication Review*, 43(4):255–266, Aug. 2013.

[32] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker. Topologically-aware overlay construction and server selection. In *INFOCOM 2002*, volume 3, pages 1190–1199. IEEE, 2002.

[33] M. Rimondini, C. Squarcella, and G. Battista. Towards an automated investigation of the impact of bgp routing changes on network delay variations. In *Proceedings of PAM'14*, pages 193–203. Springer, 2014.

[34] S. Shakkottai, N. Brownlee, A. Broido, and k. claffy. The RTT distribution of TCP flows on the Internet and its impact on TCP based flow control. Technical report, Cooperative Association for Internet Data Analysis (CAIDA), Mar 2004.

[35] B. Veal, K. Li, and D. Lowenthal. New methods for passive estimation of TCP round-trip times. In *Proceedings of PAM'05*, pages 121–134. Springer, 2005.

[36] F. Wang, Z. M. Mao, J. Wang, L. Gao, and R. Bush. A measurement study on the impact of routing events on end-to-end internet path performance. *ACM SIGCOMM Computer Communication Review*, 36(4):375–386, 2006.

[37] B. Zhang, T. Ng, A. Nandi, R. H. Riedi, P. Druschel, and G. Wang. Measurement-based analysis, modeling, and synthesis of the Internet delay space. *IEEE/ACM Transactions on Networking*, 18(1):229–242, Feb 2010.