An Empirical Characterization Of Internet Round-Trip Times

Daniel S. F. Alves
University of California, Santa Cruz
dalves@ucsc.edu

Katia Obraczka
University of California, Santa Cruz
katia@soe.ucsc.edu

ABSTRACT

End-to-end round-trip times (RTTs), which measure the time when the source transmitted data and when it received confirmation that the data was received, have been used by several Internet applications and protocols as a way to estimate network load and congestion. However, the Internet's ever increasing size and complexity pose many challenges to the study of the RTT. RTT's stochastic nature combined with diverse network topologies, technologies, and workloads are part of the problem, as well as difficulty in acquiring representative RTT samples or testing how RTT measurements are affected by changes in protocols. As part of the answer to these challenges, this paper presents a characterization study of RTT traces collected from both real- as well as simulated networked environments. We verify that temporal- and spatial factors cause RTT behavior to exhibit particular trends. Using rigorous analytical methodology, we also confirm that RTT distributions can be modeled as a power law. We then use RTT power law statistics to validate and fine-tune simulation environments.

1 INTRODUCTION

Latency is a key performance indicator when designing and evaluating computer networks, their protocols, and applications. Of particular interest is the study of the end-to-end round-trip time (RTT), the time between when the source transmits the data and when it receives confirmation that the data was received by the receiver. Several network applications and protocols use the RTT to estimate network load or congestion, and therefore need to measure it frequently. The Transmission Control Protocol, TCP, is the best known example: it bases its error- and congestion-control functions on its ongoing estimation of the RTT instead of relying on feedback from the network.

Understanding Internet RTT behavior has received considerable attention from network researchers and practitioners, and has been the focus of several descriptive- as well as predictive analytics studies. There are, however, many problems surrounding the study of the RTT on the Internet, mostly stemming from the Internet's size and complexity. Firstly, RTTs are inherently stochastic and dependent on a variety of factors, including network topology and traffic load. Compounding the problem is the fact that it is complicated to acquire representative RTT samples or even test how RTT measurements affect different protocols that use those measurements.

Even though there are platforms for collecting RTT measurements under realistic conditions, e.g. PlanetLab [5], access is still constrained and impractical for some studies [21]. Network simulators provide more control over the testing environments as well as higher degree of experimental reproducibility, but much care must be taken to ensure that the observed results correspond to reality; however, some aspects simply cannot be reproduced in simulation [9].

As the Internet evolves and scales up not just in size but also in complexity, it is imperative that the RTT, one of the Internet's main "vital signs", is constantly monitored and studied in order to validate previously observed trends as well as to uncover new ones. Our original motivation for this work is to explore techniques to estimate future RTT behavior based on past and current RTT trends. With that goal in mind, we started by conducting an RTT characterization study using both RTT traces obtained from real Internet traffic as well as traces generated by simulation.

As such, the main contributions of this paper include: (1) verifying that temporal and spatial factors affect RTT behavior in repeatable ways, (2) confirming that the observed RTT distributions can be modeled as a power law and RTT jitter follows a Cauchy distribution, (3) using systematic and robust statistical methodology to verify that the power law adequately fits observed RTTs, and (4) employing RTT power law statistics to validate and fine-tune simulation environments.

The remainder of this paper is organized as follows: Section 2 presents previous research that relates to our work. In Section 3, we describe the datasets used in our study and Section 4 presents our descriptive analysis of these datasets. Section 5 describes our simulation experiments and shows the validation that the resulting RTTs exhibit power law behavior comparable to what was observed from the real RTT datasets. Finally Section 6 concludes the paper with some directions for future work.

2 BACKGROUND AND RELATED WORK

In this section, we provide a brief overview of previous work we find the most related to our paper. We place related work in two main categories, namely descriptive- and predictive studies.

2.1 Descriptive Analysis

The Internet is continuously changing and evolving and as such, its behavior and performance need to be constantly studied and characterized [2].

This work has been partially supported by CAPES-Brazil under the Science without Borders Program, as well as NSF under project CNS 1321151 and a gift from Samsung SDS Research America. The authors would also like to thank Dr. Peter Danzig for proposing the project, providing the commercial dataset, and for the insightful discussions and feedback on our work. We would also like to thank UCSC's Science Internship Project interns Alice Lim, Ishani Karmarkar and Arjun Subramonian for their help with the analysis of the CAIDA dataset.

Q2SWinet'17, November 21–25, 2017, Miami, FL, USA

ACM ISBN ISBN 978-1-4503-5165-2/17/11…$15.00
https://doi.org/http://dx.doi.org/10.1145/3132114.3132123

For example, there exists a large body of work on network traffic's self-similarity properties, including [7, 13, 17]. Self-similarity is strongly related with the presence of power law behavior, differing from the telephony-based models previously adopted. This paper extends this body of work by using more recent developments in the study of power law distributions that allows for a more robust power law identification and goodness of fit testing. We also use this methodology to validate RTT traces generated using network simulation.

The Center for Applied Internet Data Analysis (CAIDA) has been a major player in the quest to understand practical and theoretical aspects of the Internet. Their work on characterizing and analyzing Internet topology, routing, traffic, as well as economics and policy is widely known and recognized. For example, the work presented in [11] compares different Internet distance metrics, namely: number of remaining hops measured in number of IP addresses and in number of Autonomous Systems (ASs) to be traversed, geographical distance to the next hop, and delay.

Shakkottei et al. [20] discuss different methods to collect RTT data without access to client or server. It also presents a study of RTT distribution using different datasets, which found that the peaks in the distribution tend to vary according to the location of the communication end points. It also observed that the RTT distribution has a long tail, which they suggest, is not explainable only by queuing delays.

Other Internet properties have also been studied by the networking research community. A notable example is the use of the Zipf distribution to model different Internet characteristics such as distribution of Web requests and router connectivity. Adamic and Huberman [3] present an overview of the application of the Zipf distribution to model growth, caching, and network connectivity on the Internet.

Garetto and Towsley [10] provide an analysis of Internet traffic's queuing delay. Their work proposes mechanisms to predict queuing delay in a link when transferring files of size modeled by a long-tailed distribution. They validate their results with simulation and real measurements in network, providing a methodology to evaluate performance in networks, and better understanding of influence of long-tailed distributions in this environment.

Rizo-Dominguez et al. [19] focus on RTT jitter, that is, the difference between consecutive measurements, of RTT. Jitter is particularly important for real-time services such as video or voice transmission, since fluctuations in the intervals between arrivals of packets impact user experience as well as buffer requirements. The study models jitter using a Cauchy distribution. In our work, we examine whether this property can be observed in the jitter measurements from the traces and simulations results studied.

2.2 Predictive Analysis

TCP's original RTT estimation mechanism [12] is mainly based on linear filtering, which works well for signals with Gaussian characteristics. However, statistics of RTT show non-Gaussian characteristics and as such other approaches could provide better prediction [8, 14].

Vishwanath and Vahdat [22] proposed a framework to generate realistic traffic patterns. They base their work on the modeling of different levels of behavior in the network, such as user session and application, and extraction of characteristics from observed captured data to recreate distributions for these behaviors.

3 DATASETS

For this work we considered two RTT datasets measured from real Internet traffic. The first one was collected by an Internet traffic management company and covers a 24-hour period in January 19, 2015. The second dataset was collected by CAIDA to study the Internet's topology. While intervals between data collection were longer in the CAIDA trace, data was collected over multiple years, allowing the examination of ongoing trends. Tables 1 and 2 summarize the characteristics of both datasets.

3.1 Commercial Dataset

Table 1: Commercial dataset.

Data collection period	January 19, 2015
Number of entries	637,462,949
Location granularity	Autonomous System (AS) number (client-side only)
Collection method	Requests from clients
Sampling rate	Average 7,378.04 samples per second
Availability	Private
Collection purpose	Commercial traffic optimization

The dataset from the Internet traffic management company covers a period of 24 hours on January 19, 2015 and was measured by Web clients located world-wide. Each entry in this dataset corresponds to one measurement, either RTT or throughput and, among other information, includes source and destination addresses. Data is collected through collaboration with partner companies, who instrument their websites such that when users access their content, browsers collect page usage information. That same mechanism also probes participating systems to measure latency and bandwidth.

Information about the client is anonymized for privacy reasons. Therefore, it is not possible to identify specific connections. But using different levels of available location information, we are able to group measurements by Autonomous System (AS). Information about destinations is even more restricted and consists of an anonymized identication of the content provider that receives object requests.

3.2 CAIDA Dataset

The CAIDA dataset was generated by the Ark project [4] whose purpose was to study the Internet's topology. It employs periodic RTT measurements using "monitors" located around the world. We used a subset of this trace corresponding to the months of January, March, and September in 2008, 2010, 2012, and 2014.

While this dataset contains less information per entry than the other one, as is evidenced by the count of samples shown in Table 2, it covers a much longer period of time. Data was collected using *scamper*, a tool for probing Internet destinations, which shows the path from a querying computer to the query's destination. The

Table 2: CAIDA dataset.

Data collection period	January, March and September in 2008, 2010, 2012, 2014
Number of entries	157,290,736
Location granularity	City and country (monitoring-side only)
Collection method	*scamper* tool from multiple servers (traceroute based)
Sampling rate	1 sample per second per monitor
Availability	Public
Collection purpose	Topology study of the Internet

dataset is organized according to the querying computers, which are identified through the country and city of their location in a way that it is possible to establish their geographical position. However, the destinations are identified only by IP addresses.

4 RTT SPATIAL AND TEMPORAL CHARACTERIZATION

We examine both datasets with the goals of identifying patterns that can help inform the prediction of future RTTs. More specifically, we will look at how geographical position and the time of the day influence the RTT distribution.

4.1 Commercial Dataset

In the analysis of commercial dataset, we organize the data into six groups: five corresponding to the connection's continent of origin (identified as "market" in the dataset) and one group for connections without this identification. We further group by hour of the day and then describe the distributions of RTT by geographic region over time. We plot time in GMT in order to be able to compare RTT behavior in different time zones.

Figure 1 plots, for each market, the number of RTTs collected every hour over a 24-hour period. A first inspection of Figure 1, which represents the level of network activity generated by each market during one day, already shows clear differences amongst network activity levels of the different markets, as well as daily patterns. In particular we notice that the counts for the North American (NA) and European (EU) markets match with the expected changes in activity throughout the time. The European market shows a drop in counts in the period between 0 and 6 GMT, which corresponds to night in Europe, while the same period presents a peak in the North American market, where it is evening. That period is followed by a rise in activity for the European market, where it is morning, while the American market enters the night with a drop in count.

Figures 2, 3, 4, 5, 6 and 7 plot RTT statistics for each market throughout the day, including the first quartile, the median, the third quartile, as well as the mean. From these graphs, we can observe how daily RTT statistics vary in each market. Notably, based on the similarities of their RTT statistics, we identify three groups, namely: (1) Europe and North America, (2) Asia and Oceania, and (3) South America and Africa. For example, when comparing Figures 2 and 6, we observe that RTT statistics for the North American and European markets are quite similar in terms of quartile values

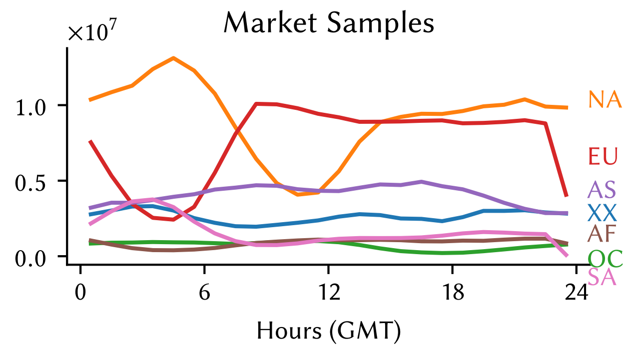

Figure 1: Number of measured RTTs by market plotted over a period of 24 hours. NA represents North America, OC: Oceania, EU: Europe, AS: Asia, AF: Africa, SA: South America, XX: unknown.

(e.g. mean RTT for both markets is around 180 ms), interquartile range, as well as daily trends. The same is true for Asia and Oceania (Figures 7 and 4, respectively). While RTT statistics in North America and Europe do not exhibit significant fluctuation during the day, both Oceania and Asia present visible variations in their 24-hour RTT statistics. Moreover, RTT statistics variation happen around the same times of day for Oceania and Asia.

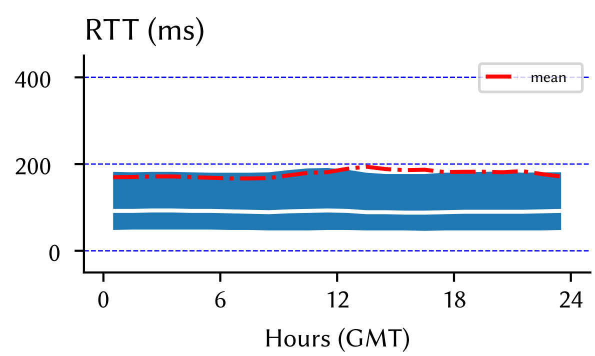

Figure 2: North American market's RTT statistics over 24 hours: mean and interquartile range with median.

Our RTT statistics also reveal distinct interquatile ranges for these three groups of markets. We notice that the North American and European markets exhibit the lowest RTT statistics, which indicate connections with generally low response times, while the South American and African markets show the opposite. The Oceanic and Asian markets are in-between, with the Asian market showing slightly higher values. The most likely explanations for this are: (1) disparity in the markets' underlying networking infrastructure between these markets, and (2) difference in how many connections originating in a given market stay "local", i.e. how many connections have both the client and server in the same market. We

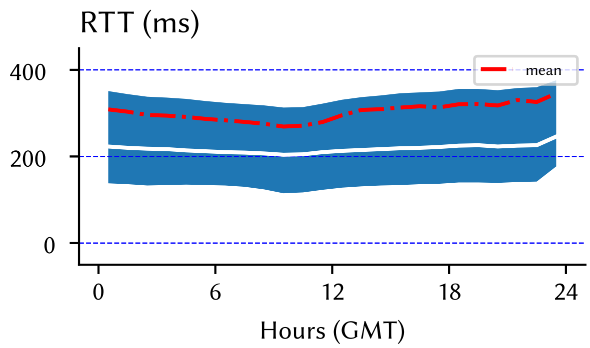

Figure 3: South American market's RTT statistics over 24 hours: mean and interquartile range with median.

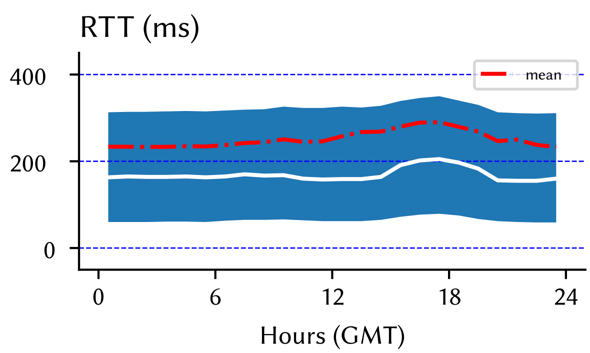

Figure 4: Oceanic market's RTT statistics over 24 hours: mean and interquartile range with median.

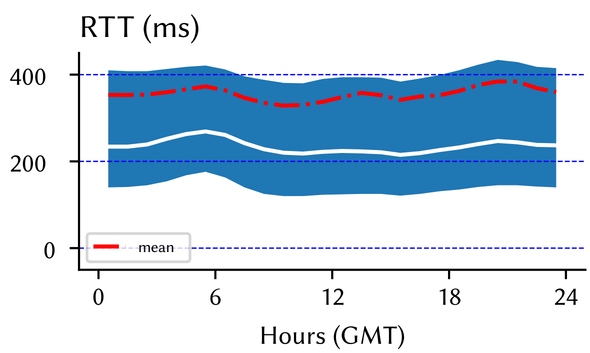

Figure 5: African market's RTT statistics over 24 hours: mean and interquartile range with median.

contend that in the South American and African markets there are a large number of non-local connections. In order to address this particular issue more accurately, we would need to consider the

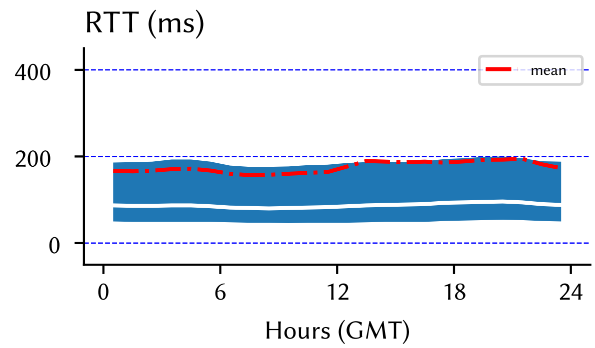

Figure 6: European market's RTT statistics over 24 hours: mean and interquartile range with median.

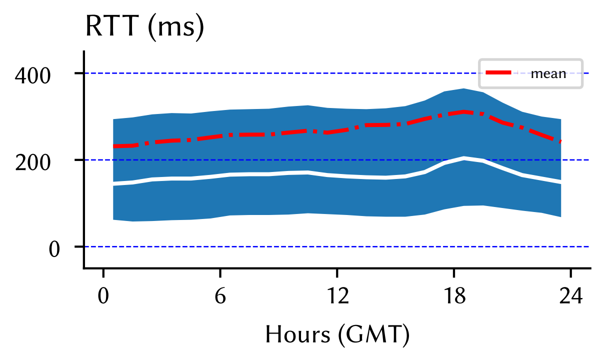

Figure 7: Asian market's RTT statistics over 24 hours: mean and interquartile range with median.

effect of Content Distribution Networks (CDNs). However, from the information available in the commercial dataset, we are not able to determine whether content was served by the origin server or by a CDN.

In order to provide a more microscopic RTT analysis, we zoom in to the Autonomous System (AS) level and study connections originating in a given AS. We zoom in even further and extract all measurements between an AS and the provider most frequently accessed by connections originating within the AS.

Going beyond these statistics, we also selected measurements of RTT between AS 7922 and certain providers in order to try to study behavior of a connection. We selected AS 7922 because it was the AS with the most of measurements with distinct providers. We further isolated all measurements between this AS and its most accessed provider in order to study the distribution of RTT jitter. We want to confirm the Cauchy behavior as mentioned in [19], which is relevant for real-time applications in particular.

In Figure 8 we see the results of this analysis with the original distribution and its fit. We fit the data to the Cauchy distribution using the median and the interquartile range of the measurements as location and scale parameters. Furthermore, we repeated this analysis with shuffled measurements before calculating differences and with discarded measurements, and in both cases we observed similar fit, indicating that measurement order or loss of measurements does not have a strong impact on the identification of the underlying distribution.

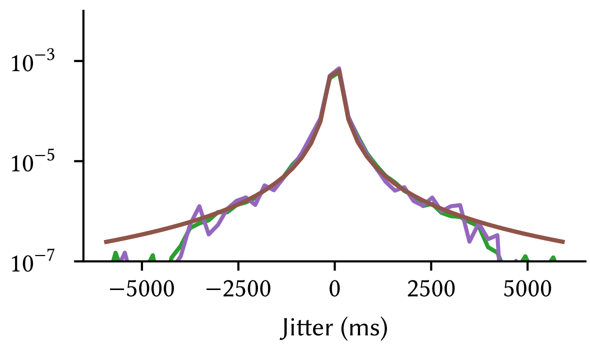

Figure 8: Density estimation of the distribution of differences between RTTs and its best fit.

4.2 CAIDA Dataset

The CAIDA dataset was collected by the Archipelago (Ark) Measurement Infrastructure as part of the Internet Topology Discovery project [4] with the use of *scamper*, a utility that probes destinations in the Internet, showing the path from the probe's origin to its destination. As such, it generates "traceroute-like" measurements which are used to compose Internet topology maps. The dataset contains measurements collected continually since 2007, although the last two years of data are not available to the general public. As described in Section 3, we use data measured during the months of January, March and September of 2008, 2010, 2012, and 2014. We selected March and September to have two months spaced half a year apart, and January because it marks the beginning of the calendar year.

Each entry in the CAIDA dataset contains: the URL of the connection's source, the destination IP address, the RTT between the source and the destination, as well as the IP addresses of intermediate nodes and the corresponding RTTs. In our analysis, we do not account for the intermediate RTTs as they would bias the analysis by including multiple occurrences of nodes closer to the source of the request.

Since we can infer geographical information from the each RTT's source URL, we can cluster measured RTTs by continent, similarly to what we did with the commercial dataset. Even though this dataset has a much lower density of measurements per time unit, its longer duration allows us to look for longer-term trends (i.e., trends transcending a day) that the commercial dataset cannot reveal. The data as illustrated in the following graphics show some examples of similarity trends between consecutive dates. Figure 9 shows measurements of the third quartile of the RTT measured for five consecutive days by a computer in Canada. Figure 10 shows measurements for the third quartile of the RTT measured for five different computers in five different days and continents.

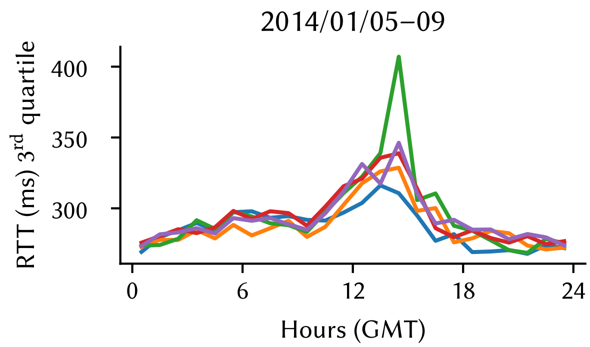

Figure 9: Measurements for the third quartile from a computer in Canada for 5 consecutive days.

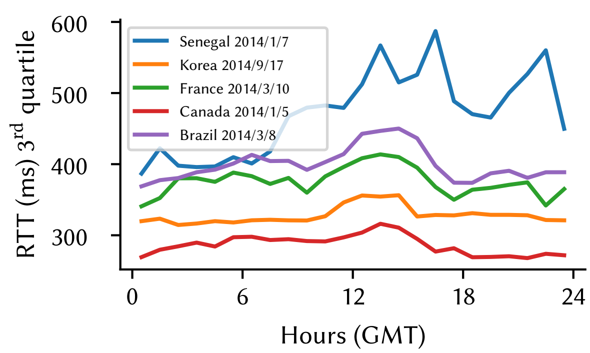

Figure 10: RTT measurements for the third quartile from different countries on different days of the year.

Even though the data is less dense than in the other dataset, the prevalence of this pattern in different regions and days of the year confirms the influence of daily patterns on the distribution of RTT. Those graphics also show how different places have different patterns of RTT, in agreement with the previous data that we analyzed. Furthermore it can be noted that that the variations in RTT throughout the day show similarities despite being measured on different times of the year and regions, which suggest a common cause for those variations. The aggregate measurements of each month also show little variation. Figure 11 also shows a comparison of empirical CDF and CDF of the expected Cauchy, showing high similarity.

4.3 Power Law

An important property of the statistical distribution of RTT measurements is its power law characteristics as identified in [20]. However, as described in [6], showing that a dataset exhibits power

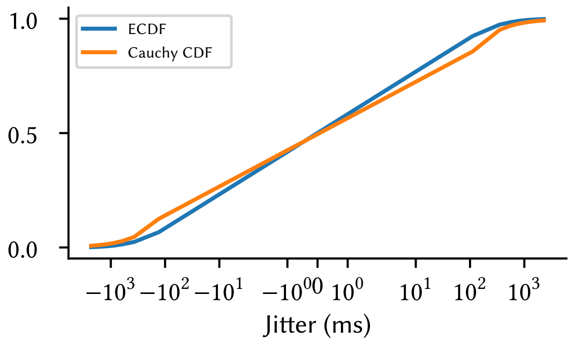

Figure 11: Comparison between empirical CDF for measured Jitter and expected Cauchy fit of one hour of measurements of a computer in the USA.

law behavior is not trivial. In fact, similar to natural phenomena that exhibit power law behavior, Internet connections' RTT distributions exhibit power law behavior only for RTT values larger than a certain threshold.

In order to confirm whether RTT distributions collected from more recent Internet activity still exhibits power law characteristics, we employ the approach proposed in [6], which provides a mode systematic and robust statistical methodology to identify the "power law threshold". It also proposes a "goodness-of-fit" test to verify that the power law is an appropriate model.

The basic idea behind the approach proposed in [6] is to use the Kolmogorov-Smirnov (KS) statistic to identify the RTT distribution's "power law threshold" as follows. The KS statistic is defined as the supremum of the distance between the empirical cumulative distribution $F_n(x)$ and the presumed empirical distribution $F(x)$ for all values x in a set X, that is:

$$\mathrm{KS}\,(X) = \sup |F_n(x) - F(x)|$$

We compute, for each measurement r, the KS statistic as the supremum of the difference between the empirical RTT CDF and the Pareto distribution with estimated minimum value $x_{min} = r$ and scale parameter

$$\hat{\alpha} = 1 + n\left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}\right]^{-1}$$

Where n is the number of samples in X greater or equal to x_{min} and x_i is one of those samples. We select then the r that resulted in the minimum KS value, that is:

$$\hat{x}_{min} = \operatorname*{argmin}_r \mathrm{KS}(X_r), X_r = \{x \in X : x \geq r\}$$

After that we need to validate whether the resulting power law model is a good fit for this portion of the dataset. In order to do so we generate synthetic datasets of the same size as the original collection of samples using as parameters the size of the dataset partitions, the limit between partitions, the estimated shape factor of the power law, and the measurements that do not follow the power law distribution. As a result we have a dataset S with values generated according to a random variable R such that:

$$|S| = |X|$$

$$s \in S \rightarrow f_R\,(s) > 0$$

$$f_R\,(s) = \begin{cases} \frac{1}{|X|} & \text{if } s < \hat{x}_{min} \text{ and } s \in X \\ \frac{|i\{s \in X : s \geq \hat{x}_{min}\}|}{|X|} f_P(s) & \text{otherwise, } P \sim \text{Pareto}\,(\hat{x}_{min}, \hat{\alpha}) \end{cases}$$

Once we have a synthetic collection of data S, we can submit it to the same partition process to find a minimum value $\hat{s}_{min}$ for the tail and compare the resulting KS statistic $\mathrm{KS}\left(S_{\hat{s}_{min}}\right)$ of the best fitting partition with the one of the original dataset. The ratio of synthetic datasets with greater KS statistic over total synthetic datasets gives us a p-value that will be used to validate the resulting power law model. Given a set of synthetic datasets D, the calculation of the p-value can be expressed as:

$$p = \frac{\left|\left\{S \in D : \mathrm{KS}\left(S_{\hat{s}_{min}}\right) > \mathrm{KS}\left(X_{\hat{x}_{min}}\right)\right\}\right|}{|D|}$$

Higher p-values provide strong evidence of the resulting power law model's validity.

For the purpose of this work we generated 1000 synthetic datasets, so we consider a threshold of 0.2 as the minimum p-value to accept that a sample of data behaves according to a power law. For example, Figure 12 shows the distribution of RTT measurements corresponding to an Internet connection captured in the commercial dataset. It also shows the power law fit for the "tail" of the distribution which we then validate by generating 1,00 synthetic datasets; the resulting p-value is 0.998 which is more than the minimum p-value of 0.1 as recommended by [6][1]. To fit the distribution of the lower RTT values, we use the shifted Gamma distribution as suggested in [15].

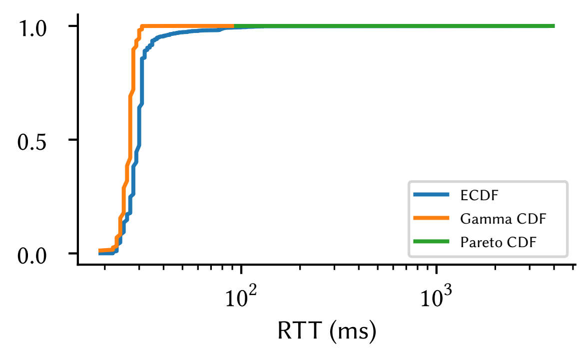

Figure 12: Sample fit of the RTT distribution corresponding to a connection from the commercial dataset.

Figure 13 shows the result of testing the power law goodness of fit for the two original datasets we analyzed and for the data generated using the ns-3 network simulator [18] for different topologies (see Section 5 for more details on our simulation experiments and

[1]According to[6], the recommended minimum p-value is 0.1 using 2,500 synthetic datasets; in our case, since we used 1,000 synthetic datasets to cut down on computational time, we raised the minimum p-value to 0.2 as a way to compensate for the loss of precision.

results). We plot the distribution of the p-values that we found in box-plots so we can see the range of values found. We can observe that both CAIDA and the commercial dataset show the presence of power law behavior, albeit the higher variance of the commercial dataset's p-values. The synthetic datasets resulting from our simulation experiments shown in Figure 13 are further discussed in Section 5.

4.4 Summary

We summarize our observations of RTT behavior based on the CAIDA and commercial datasets below. As expected, both the CAIDA and commercial datasets show differences in the general distribution of RTTs according to the location and time of day of the measurements. More specifically:

- Figures 2 through 7 and 10 confirm the dependence of RTT behavior on where RTTS are measured (i.e., the network location of the application client);
- The pattern of the RTT "peaks" observed over multiple days from the CAIDA dataset also shows how daily patterns influence network performance (Figures 1 and 9);
- RTT power law behavior is still prevalent on the Internet (Figures 12 and 13);
- Both datasets confirm the presence of Cauchy behavior in RTT jitter (Figures 8, 11 and 14) which is important for Internet real-time applications.

The continued presence of power law and Cauchy behavior in the Internet reinforces the need for RTT prediction methods which account for RTT's heavy tails.

5 GENERATING AND VALIDATING SYNTHETIC RTT TRACES

As previously pointed out, access to RTT datasets that are representative of real Internet activity is quite limited and pose significant obstacles to the study of the Internet, its protocols and applications. Using network simulation platforms is an attractive alternative, however it is critical to ensure that simulated environments are accurate representations of reality. In this section we propose the use of power law properties of Internet RTT distributions as a way to validate and calibrate network simulation environments' fidelity. We begin by presenting our initial simulation environments and experiments.

5.1 Simulation Setup

We used the simulator ns-3 [18] and experimented with two different types of network topologies, namely: topologies generated using the BRITE topology generator [16] and topologies based on the dumbbell model. Both topologies differ only in how the data sources and sinks are created, they use the same distribution, the Zipf, to model the distribution of connections as described by Adamic and Huberman [3]. The models for the behavior of applications and their proportions is also shared between topologies using the models given in the simulation specifications developed by the 3rd generation Partnership Project 2 (3GPP2) [1]. Those provide three models of data using the TCP protocol, which are based on FTP, HTTP, or video usage, and one for voice, which allow for loss. The different TCP usage models differ in the size of files that are transmitted, interval between requests, and patterns of request: individual, in groups, or constant.

BRITE Topologies. Our original implementation using the BRITE topology was configured to have ten ASs, each one with a thousand nodes. External bandwidth (between ASs) is on a uniform range of 10 to 1024 Mbps, while internal bandwidth follows a heavy-tailed with same limits. ASs were placed according to Barabasi methodology, while nodes inside each AS were placed according to Waxman methodology [16]. Each AS was configured to have five data sources and a hundred possible data sinks, connected to the nodes of the topology generated by BRITE using a 5Mbps connection. Connections between sources and sinks were determined randomly according to the Zipf distribution.

Dumbbell Topologies. We also utilized in out experiments the dumbbell topologies, which have been widely used as simplified topology models of the internet. Dumbbell topologies consist of a bottleneck link connecting two network nodes, which in turn connect a group of end users. In out dumbbell topology experiments, we use a 2Mbps bottleneck link with 2ms propagation delay, to which nodes connect through 5Mbps connections. We created five servers and a hundred clients connected to the opposite ends of teh dumbbell, so that all the servers will have to cross the common link to reach their clients. We used the same number of sources and sinks and distributed connections and applications according to the same distributions to have the same style of traffic load.

5.2 Simulation Validation

Here we use the same power law fitting process as we described in Section 4.3 to generate power law fits for the datasets obtained from the simulation experiments. Our goal is two-fold, namely: (1) verify the fidelity of synthetically-generated RTT traces; and (2) fine-tune simulation environments in order to generate realistic RTT traces. In Figure 13 we also plot the p-values for RTT traces obtained from simulations using the BRITE and dumbbell topologies. We observe that while the BRITE simulations exhibit some degree of power law behavior, the p-values for the dumbbell topology experiments indicate that their power law features are quite weak.

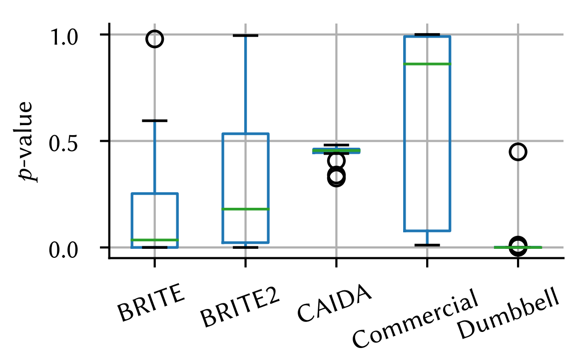

Figure 13: Distribution of p-values from power law fits for the real RTT datasets as well as datasets from simulations.

Additionally, to achieve our second goal of trying to calibrate the fidelity of synthetically generated RTT traces, we generated a slightly different BRITE topology, named BRITE2 in the graph, and show that, based on its p-values, it yields a higher degree of RTT power law behavior. The only modification in this new topology was a different distribution of the neds , which are now arranged in 100 ASs of 200 nodes each. All other parameters were kept the same, although since the number of sources was defined by AS that means we have more active sources of traffic. this suggests that we can use the p−-value to fine-tune simulation environments in order to generate more realistic RTT distributions and consequently increase the simulation's fidelity.

We also validate how well the Cauchy distribution can model RTT jitter, i.e., the differences between consecutive RTT measurements. To this end, we compare the KS statistics of measured jitter against the KS statistics of a Cauchy distribution that uses the sample median as the distribution's location parameter and half of the interquartile range as scale parameter. The resulting KS statistics presented in Figure 14 confirm the presence of the Cauchy behavior in the RTT jitters measured from the real datasets as well as the synthetic ones, though the later contain more occurrences with poorer fit.

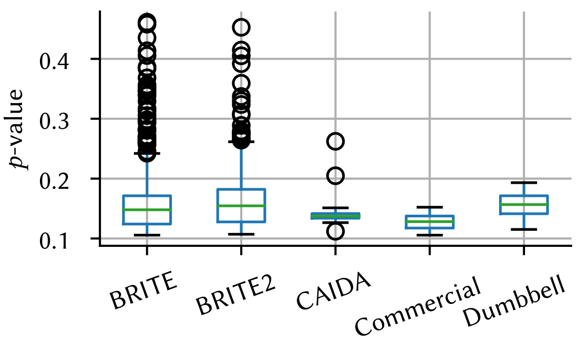

Figure 14: KS statistics for RTT jitter fitting using a Cauchy distribution for the real as well as synthetic datasets.

6 CONCLUSIONS

This paper presented a characterization of current Internet activity by studying recent RTT traces obtained from real Internet traffic as well as from simulated networked environments. Our goal was to validate previously observed RTT behavior as well as uncover new trends. The contributions of our work can be summarized as follows: (1) verifying that temporal and spatial factors affect RTT behavior in repeatable ways, (2) confirming that RTT distributions can be modeled as a power law and RTT jitter follows a Cauchy distribution, (3) using systematic and robust statistical methodology to verify that the power law adequately fits observed RTTs, and (4) employing RTT power law statistics to validate and fine-tune simulation environments.

As future research directions, we plan to expand our study to include additional RTT traces. We also plan on using the insights gained from this study to develop a more accurate RTT predictors as well as develop systematic methodologies to propose and calibrate simulation environments that preserve real-world characteristics and properties.

REFERENCES

[1] 2009. *cdma2000 Evaluation Methodology (Revision B).* Technical Report. 3rd Generation Partnership Project 2 "3GPP2". 75–124 pages.

[2] 2016. *The Zettabyte Era — Trends and Analysis – Cisco.* Technical Report Document ID:1465272001812119. Cisco. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

[3] Lada A. Adamic and Bernardo A. Huberman. 2002. Zipf's law and the Internet. *Glottometrics* 3 (2002), 143–150. http://www.hpl.hp.com/research/idl/papers/ranking/adamicglottometrics.pdf

[4] CAIDA. 2008, 2010, 2012, 2014. The CAIDA UCSD IPv4 Routed /24 Topology Dataset - January, March and September. (2008, 2010, 2012, 2014). http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml

[5] Brent Chun, David Culler, Timothy Roscoe, Andy Bavier, Larry Peterson, Mike Wawrzoniak, and Mic Bowman. 2003. PlanetLab: An Overlay Testbed for Broad-coverage Services. *SIGCOMM Comput. Commun. Rev.* 33, 3 (July 2003), 3–12. https://doi.org/10.1145/956993.956995

[6] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. 2009. Power-Law Distributions in Empirical Data. *SIAM Rev.* 51, 4 (2009), 661–703. https://doi.org/10.1137/070710111 arXiv:http://dx.doi.org/10.1137/070710111

[7] M. E. Crovella and A. Bestavros. 1997. Self-similarity in World Wide Web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking* 5, 6 (Dec 1997), 835–846. https://doi.org/10.1109/90.650143

[8] Peter B. Danzig, Sugih Jamin, Ramón Cáceres, Danny J. Mitzel, and Deborah Estrin. 1992. An Empirical Workload Model for Driving Wide-Area TCP/IP Network Simulations. *Internetworking: Research and Experience* 3 (1992), 1–26.

[9] Sally Floyd and Vern Paxson. 2001. Difficulties in Simulating the Internet. *IEEE/ACM Trans. Netw.* 9, 4 (Aug. 2001), 392–403. https://doi.org/10.1109/90.944338

[10] Michele Garetto and Don Towsley. 2003. Modeling, Simulation and Measurements of Queuing Delay Under Long-tail Internet Traffic. *SIGMETRICS Perform. Eval. Rev.* 31, 1 (June 2003), 47–57. https://doi.org/10.1145/885651.781034

[11] Bradley Huffaker, Marina Fomenkov, Daniel J. Plummer, David Moore, and Kimberly Claffy. 2002. Distance Metrics in the Internet. In *in IEEE International Telecommunications Symposium.*

[12] V. Jacobson. 1988. Congestion Avoidance and Control. *SIGCOMM Comput. Commun. Rev.* 18, 4 (Aug. 1988), 314–329. https://doi.org/10.1145/52325.52356

[13] Patrick Loiseau, Paulo Gonçalves, Guillaume Dewaele, Pierre Borgnat, Patrice Abry, and Pascale Vicat-Blanc Primet. 2010. Investigating self-similarity and heavy-tailed distributions on a large-scale experimental facility. *IEEE/ACM Transactions on Networking (TON)* 18, 4 (2010), 1261–1274.

[14] Liangping Ma, G. R. Arce, and K. E. Barner. 2004. TCP retransmission timeout algorithm using weighted medians. *IEEE Signal Processing Letters* 11, 6 (June 2004), 569–572. https://doi.org/10.1109/LSP.2004.827957

[15] Liangping Ma, K. E. Barner, and G. R. Arce. 2006. Statistical analysis of TCP's retransmission timeout algorithm. *IEEE/ACM Transactions on Networking* 14, 2 (April 2006), 383–396. https://doi.org/10.1109/TNET.2006.872577

[16] Alberto Medina, Anukool Lakhina, Ibrahim Matta, and John Byers. 2001. BRITE: An Approach to Universal Topology Generation. In *Proceedings of the Ninth International Symposium in Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS '01).* IEEE Computer Society, Washington, DC, USA, 346–. http://dl.acm.org/citation.cfm?id=882459.882563

[17] Adrian Popescu. 2001. Traffic self-similarity. (2001).

[18] George F. Riley and Thomas R. Henderson. 2010. *The ns-3 Network Simulator.* Springer Berlin Heidelberg, Berlin, Heidelberg, 15–34. https://doi.org/10.1007/978-3-642-12331-3_2

[19] L. Rizo-Dominguez, D. Torres-Roman, D. Munoz-Rodriguez, and C. Vargas-Rosales. 2010. Jitter in IP networks: a cauchy approach. *IEEE Communications Letters* 14, 2 (February 2010), 190–192. https://doi.org/10.1109/LCOMM.2010.02.090702

[20] Srinivas Shakkottai, R. Srikant, Nevil Brownlee, Andre Broido, and Kimberly Claffy. 2004. The RTT distribution of TCP flows in the Internet and its impact on TCP-based flow control. (2004).

[21] Neil Spring, Larry Peterson, Andy Bavier, and Vivek Pai. 2006. Using PlanetLab for Network Research: Myths, Realities, and Best Practices. *SIGOPS Oper. Syst. Rev.* 40, 1 (Jan. 2006), 17–24. https://doi.org/10.1145/1113361.1113368

[22] Kashi Venkatesh Vishwanath and Amin Vahdat. 2006. Realistic and Responsive Network Traffic Generation. *SIGCOMM Comput. Commun. Rev.* 36, 4 (Aug. 2006), 111–122. https://doi.org/10.1145/1151659.1159928