The Delay Measurement and Analysis of Unreachable Hosts of Internet

Ali Gezer
Electronic and Telecommunication Technology, Kayseri University, Turkey

Abstract: *Delay related metrics are significant quality of service criteria for the performance evaluation of networks. Almost all delay related measurement and analysis studies take into consideration the reachable sources of Internet. However, unreachable sources might also shed light upon some problems such as worm propagation. In this study, we carry out a delay measurement study of unreachable destinations and analyse the delay dynamics of unreachable nodes. 2. Internet Control Message Protocol (ICMP) destination unreachable Internet Control Message Protocol-Destination Unreachable (ICMP T3) packets are considered for the delay measurement according to their code types which shows network un reach ability, host un reach ability, port un reach ability, etc., Measurement results show that unreachable sources exhibit totally different delay behaviour compared to reachable IP hosts. A significant part of the unreachable hosts experiences extra 3 seconds Round Trip Time (RTT) delay compared to accessible hosts mostly due to host un reach ability. It is also seen that, approximately 79% of destination un reach ability causes from host un reach ability. Obtained Hurst parameter estimation results reveal that unreachable host RTTs show lower Hurst degree compared to reachable hosts which is approximately a random behaviour. Unreachable sources exhibit totally different distributional characteristic compared to accessible ones which is best fitted with Phased Bi-Exponential distribution.*

Keywords: *RTT distributions, ICMP-T3, ICMP, self-similarity, worm propagation, Hurst parameter.*

Received February 3,2020; accepted April 21, 2021
https://doi.org/10.34028/iajit/19/1/8

1. Introduction

Network measurements play a vital role for the evaluation of networks, and regularly taken measurements are a necessity for optimal network management [26]. Delay is a significant quality of service metric for telecommunication systems. In a packet switched network, packet delay depends on many factors such as propagation time, processing time, queuing time, and transmission time. Many more factors could add variations to delay such as network congestion, route changes, routing loops etc., Packet delay presages many anomalies during the operation of a network such as link failure, routing anomalies, service degradations, worm affection, and congestion [13, 16]. Especially in Transmission Control Protocol/Internet Protocol (TCP/IP) communication such as Internet, delay affects many mid-steps in the communication process of hosts [29]. Retransmission Time Out (RTO) mechanism in TCP communication needs Round-Trip Time (RTT) to determine waiting time in the absence of any feedback from the remote data receiver [13]. TCP congestion avoidance mechanism also uses RTT as a metric to perform congestion control [4, 19, 23].

Almost all delay analysis studies take into account the reachable sources over the Internet for revealing delay behaviour. Delay-based studies have developed new systems and designed new architectures to reveal network problems via considering different delay quantities such as RTT delay [6, 20], one-way delay [5], queuing delay [15] and processing delay [27, 28]. An algorithm was proposed to estimate one-way delay to a client by cooperating with two other servers, requiring neither clock synchronization nor client trustworthiness [1]. A delay-based verification technique Client Presence Verification (CPV) was designed to verify an assertion about a device's presence inside a prescribed geographic region [2]. In [3], a new dynamic delay-based congestion control algorithm for background traffic, Eclipse was developed. Eclipse could dynamically adapt to the network characteristics for minimizing the additional network delay and maximizing the utilization of spare network capacity. A lightweight delay measurement system was built, and a new robust method was devised to calculate the per-packet delay in a large-scale wireless sensor network. Spatial and temporal characteristics of delay were determined, and a delay model was proposed to capture those factors [32]. Choi *et al*. [13] conducted a study about how to measure and report delay in a meaningful way for an Internet Service Provider (ISP) and how to monitor it efficiently.

Typically, performance measurement studies over the Internet take into consideration the quantities of accessible hosts for early determination of network problems. However, in some manners the behaviour of unreachable sources might be more significant, give

some clues about the current state of network and presages some abnormalities in advance that might cause serious damages in the future. In [9], an early warning system was proposed that uses Internet Control Message Protocol-Destination Unreachable (ICMP-T3) messages to identify the random scanning behaviour of Internet worms. Random scanning of Internet IPs causes many vacant Internet addresses being probed. In the designed system, participating routers send copies of all their locally generated ICMP-T3 messages to a central collection point and these messages are evaluated for the detection of worm activity. In [11], a distributed anti-worm architecture was proposed that automatically slows down or even halts during the worm propagation. New architecture exploits the higher connection failure rate of infected hosts. The connection request fails if the destination host does not exist, or TCP reset packet is returned. Considering TCP resets and host unreachable packets, connection failure rate is obtained and exploited for worm evaluation. In [22], a hybrid method is proposed to detect internet worms by analysing ICMP-T3 messages and worm characteristic matching. Blenn *et al*. [10] introduced a method which leverages backscattered ICMP-T3 packets to quantify whether a server falls over or not. By monitoring ICMP-T3 packets, the authors can estimate the attack size needed to successfully Distributed Denial of Service (DDoS) a server. Although existence of ICMP-T3 messages have been evaluated for attack detection, delay characteristics of unreachable sources have not been considered in the mentioned studies.

In this study, we perform a delay measurement and analysis study via considering unreachable sources of Internet. The RTT values of unreachable destinations are analysed to reveal delay characteristics of inaccessible IP hosts. ICMP protocol is used to determine the reach ability of hosts. If interested source could not be reached at that moment, ICMP-T3 messages are raised by routers or end hosts. According to the code types of ICMP-T3 messages, it could be understood what type of Internet node generates notifying ICMP message. This source might be a transit router, stub router or directly destination host. We investigate the behavioural delay differences of reachable and unreachable hosts over the Internet. Considering more than 1 million ICMP-T3 packets for the analysis, delay dynamics of unreachable hosts could provide valuable information to network administrator and Internet service provider for the evaluation of network traffic in terms of cyber-attacks such as DDoS attacks and worm propagation. We also propose a method to differentiate reachable and unreachable destinations via the estimation of self-similarity parameter.

This paper is organized as follows. Internet Control Message Protocol (ICMP) is introduced in section 2. Background information about self-similarity is given in section 3. The details about the measurement point is given in section 4. Self-similarity and RTT analysis results are shared in sections 5 and 6, respectively. Then, the paper is concluded in section 7.

2. Internet Control Message Protocol (ICMP)

ICMP gives significant information about IP communication problems such as routing anomalies, inaccessibility of end nodes, and RTT delays [6, 25]. ICMP messages are mostly generated by routers, gateways, or destination hosts. Despite the broad usage of ICMP messages, the most common practices are to determine the reach ability of network nodes, and round-trip time between nodes.

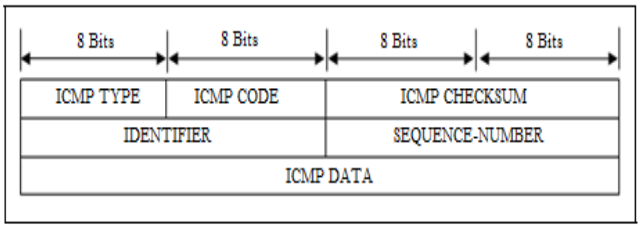


Figure 1. ICMP frame format.

A standard ICMP packet structure is shared in Figure 1. Type and code fields give information about the content of ICMP datagram. Identifier and sequence numbers are used to match associated ICMP request with its reply packet. Some common ICMP types are given in Table 1.

Table 1. ICMP types.

Type	Message
0	Echo Reply
3	Destination Unreachable
5	Redirect
8	Echo Request
11	Time Exceeded
13	Timestamp
14	Timestamp Reply
17	Address Mask Request
18	Address Mask Reply
30	Traceroute

ICMP echo reply, and echo request are exploited for controlling accessibility of IP nodes in TCP/IP networks. Destination unreachable ICMP-T3 messages are generated mostly by routers, in some cases by end hosts. For example, when the node address or network address specified in IP data grams could not be reached, the notifying message are sent by associated routers. If Don't Fragment (DF) bit in the datagram is marked and the packet reached a node of which it must be fragmented, ICMP-T3 packets are generated to notify the source host. Depending on the code type of ICMP-T3 message, the notifying node might be a stub router, transit router, gateway, or destination host. In Table 2, the most common ICMP-T3codesare shared.

Table 2. ICMP unreachable destination codes.

Code	Message
0	Network Unreachable
1	Host Unreachable
3	Port Unreachable
10	Com. With Dest. Host Administratively Prohibited
13	Communication Administratively Prohibited

Network unreachable messages are generated by transit routers to point out there is no way to reach the network or sub-network of that destination IP. When an IP datagram reaches a router on the way, the router looks its routing table and determines its output interface by applying suitable subnet mask. When there is no default gateway option anymore and subnet masks matching with the mentioned IP address, the router discards IP datagram and notifies the source host via sending a network unreachable message. The most common causes for network unreachability are invalid destination address, wrong subnet mask usage for the associated port, or being the link is off at that moment.

Host unreachable messages are generated by stub routers to point out the host unreachability. All the way from the source host to last router is reachable, but stub router could not reach the source host. At this point the router generates ICMP host unreachable message to notify the source host is not reachable. If the router knows Media Access Control (MAC) address of the destination host, it directly sends a packet to the source via using known MAC address. The destination host might be off, or its connection might fail at that moment, then the router generates host unreachable message. But, if router doesn't know the network interface address of interested destination host, the router broadcasts Address Resolution Protocol (ARP) packets to all hosts in that domain to learn the MAC address of the interested host and waits for ARP reply. If the router doesn't get any ARP response after repeated attempts, it sends a host unreachable message to the source host. In this case the RTT communication of unreachable host message takes more time.

Port unreachable messages are generated from directly end destination hosts. In the communication process, last router communicates with the interested host and sends frames to that host. The host accepts frames, puts into its communication buffer, and processes them. All the way to this process is alright but if destination port is not open in the interested host, then port unreachable message is raised. The main causes for port unreachable messages are the interested port address might not run or might be swapped out.

Communication with destination host is administratively prohibited message is raised when the destination device is not allowed to send packets. Communication administratively prohibited is generated if a router cannot forward a packet due to administrative filtering. The cause of filtering might be firewall blockage or something else due to the message content.

3. Self-Similarity

Self-similar processes are emerging as a powerful representation of many physical phenomena such as network traffic. Previously, the modelling of network traffic used to be done with Poisson distribution [24]. Nowadays, self-similarity artifacts have occurred in most of the broadband network traffic [8, 14, 21]. Hurst parameter is a numerical measure of self-similarity. The value of Hurst parameter indicates whether a stochastic process has long range dependency or not. A continuous time stochastic process $\{X(t), t \in R\}$ is strictly self-similar with the Hurst parameter $\{H, 0<H<1\}$ if the following condition is supplied.

$$X(\text{at}) \overset{d}{=} a^H X(t) \tag{1}$$

$X(at)$ is a new process scaled by factor a, and $\overset{d}{=}$ means equal in finite dimensional distributions. When the Hurst value is between 0.5 and 1, the process has long range dependency.

As self-similarity is significant in many disciplines, correct estimation of it is a necessity. Powerful properties of wavelet analysis have been an inspiration source for Hurst parameter estimation. The main point which makes wavelet analysis so important for the Hurst parameter estimation is its time-scale dependent working nature. An efficient wavelet-based estimator called Veitch and Arby [31] Daubechies Wavelet Based (DWB) was proposed by Veitch and Arby [31]. Veitch and Arby [31] used Daubechies wavelets as kernel function due to their limited time support which eases handling of border effects.

We employ Veitch and Arby [31] Hurst estimator to get the self-similarity degree of RTT delay for reachable and unreachable destinations. Hurst parameters of the given flows are easily calculated through the relationship between variance of wavelet coefficients and corresponding scale as given in the Equation (2)

$$\log_2(\text{var}(d_j[n])) = (2H-1)j + \text{consant.} \tag{2}$$

Where $d_j[n]$ represents the wavelet coefficients at j scale, and H represents Hurst parameter. The slope between $\log_2(var(d_j[n]))$ and j are in relation with $2H-1$. Therefore, Hurst parameter could be calculated via utilizing a regression line.

4. Measurement Point

Starting in 2003, Analysis of Networ Traffic (ANT) Project Team [18] of Information Science Institute (ISI) from University of Southern California (USC) has been collecting data about IPv4 Internet address space. ANT has been conducting IP address censuses study via probing IPv4 addresses. All allocated IPv4 addresses except multicast, private, loopback, class E addresses, and unallocated addresses are probed

periodically since 2003. California, Colorado, Greece, Japan and Washington DC are chosen to send ping messages to all allocated IPv4 addresses. Via using Measurement and Analysis on the WIDE Internet (MAWI) Widely Integrated Distributed Environment (WIDE) backbone network (AS 2500) [12], ANT is sending ping messages from Fujisawa-Shi, Kanagawa in JAPAN. The probing machine is hosted at Keio University in Fujisawa Campus. The capturing point is located at NTT Otemachi Building (Notemachi) in Otemachi. The packets are captured through WIDE's 150 Mbps incoming/outgoing transit Ethernet link to the upstream ISP. The actual link capacity is 1 Gbps but capped bandwidth of the link is limited with 150 Mbps. Port mirroring technique is utilized for packet capturing and packets are recorded on a web server which is publicly open to the Internet [12].

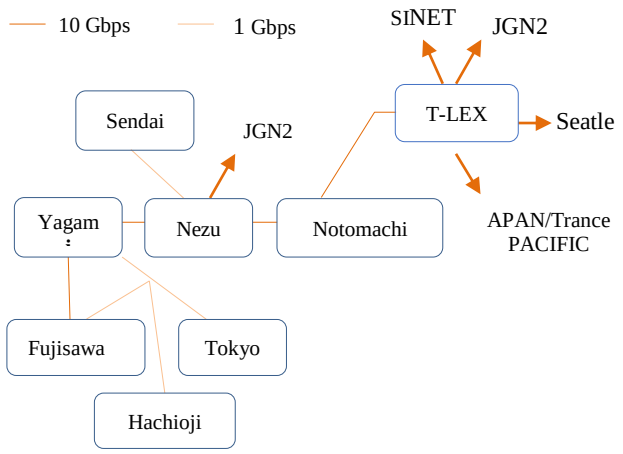


Figure 2. Part of WIDE backbone.

WIDE backbone network includes various speed of links from 2 Mbps CBR ATM up to 10 Gbps Ethernet which spreads all over Japan. A limited section of it which includes probing and capturing points is given in Figure 2.

Captured trace files through 150 Mbps link consist of tens of application traffic including P2P, HTTP, FTP, ICMP, SMTP, etc., and millions of packets. Due to privacy reasons, packet payloads are removed by Tcpdpriv, only first 96 Bytes of packet are captured. MAWI also anonymizes IP addresses with a prefix-preserving method for the privacy of individual IP addresses. Applied prefix mapping is consistent among the traces. Each captured file keeps traffic content of 15 minutes time span and contains more than 100 million packets with a volume of nearly 30 GB.

5. Measurement Data Pre-Processing

The probing host at Keio University sends ICMP echo requests to all IPv4 addresses except some special and unallocated ones. Approximately 15-20% of echo requests take ICMP replies in 15 minutes interval. But most of the probed IPv4 addresses don't give reply due to some reasons such as blocking of firewalls, time exceeded, destination unreachability, etc. Some gateways and transit routers notify source addresses not to transmit their echo requests to the destination. These notifying messages could be caused by hop count limit or destination unreachability at that specific time.

Significant number of hosts, gateways and routers send ICMP-T3messages to point out the destination IP could not be reached. Nearly more than 250.000 ICMP-T3 packets reach the capturing point in 15 minutes interval. The ICMP-T3 packets might be caused by network unreachability, host unreachability, protocol unreachability, etc.

Each trace contains tens of application protocol traffic and millions of ICMP packets. To filter ANT related packets, first the particular IP address associated with ANT should be identified. MAWI's scrambling policy of IP address doesn't make any difference due to IP mapping consistency in all traces. For filtering ANT related ICMP-T3, following filtering is performed in Wireshark protocol analyser.

(icmp.type == 3 and ip.dst==. . . .) (3)

ICMP-T3 packetincludes original packet content. According to RFC 792 [25], at least the original IP header and 8 bytes of the payload should be included. When an ICMP request packet takes ICMP-T3 message in return, it consists of source/destination address, source/destination ports, and remaining Time to Live (TTL) value of original packet. These fields are exported into a text file and related fields are extracted with a scraper application. To extract required fields from exported text file, we develop a software solution in Microsoft Visual Studio via coding with C#. The extracting operation of related fields takes more than 12 hours for a 15-minute trace file with an Intel Core i7-3770 CPU 3.40 GHz, 16 GM RAM desktop computer.

Associated ICMP request packets for each destination unreachable messages are also required to calculate RTTs. To filter ANT related ICMP requests, we only filter non-reply ICMP requests by applying following filter code in Wireshark.

(icmp.type==8 and !icmp.resp_in and ip.src==. . . .) (4)

At the end of this filtering, nearly 13 million non-reply ICMP packets are obtained for each 15 minutes trace file. We import the obtained fields into a database for processing and matching operations. Matching of ICMP requests and ICMP replies are provided if only IP addresses, the identifier and sequence numbers are all matched. The workflow of pre-processing operation is given in Figure 3.

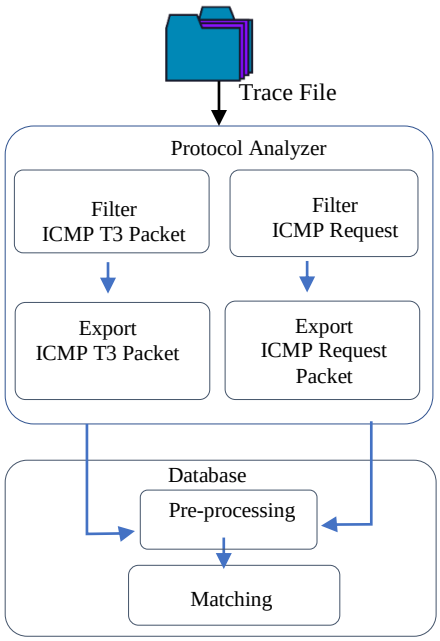


Figure 3. Workflow of ICMP packet pre-processing.

6. RTT Analysis

24.556.763 ICMP packets were captured on December 2nd, 2015 trace between 12.00 and 12.15. Nearly 1% of ICMP packets are ICMP-T3 packets. 248.772 ICMP requests and ICMP-T3 packets are matched considering IP addresses, identifier, and sequence fields. That is, 248.772 ICMP requests to random IP addresses take ICMP-T3 messages in return. Sometimes one ICMP request takes more than one ICMP-T3 messages. For the sake of analysis, we only match associated packets if ICMP echo request is sent one time for one unique IP. At the end of this operation, obtained matching statistics are given in the Unique Destination column of Table 3.

It is observed that sometimes same nodes return many ICMP-T3 messages for separate targeted ICMP requests. It is understood that mentioned sources mostly raise ICMP-T3 messages to ICMP requests whose IP addresses have same prefix. In this manner, average RTT time is calculated to avoid repetition between two ends. Although the destination addresses are being different, eventually this ICMP requests could not reach their destination and same node sends ICMP-T3 messages for these requests. At the end of averaging, obtained case numbers are given in the Distinct Source column of Table 3. We benefit from the RTTs of Distinct Sources in our later analysis.

Table 3. Summary information about ICMP-T3 messages.

Date-Time	Total ICMP packets	Request with Reply	Matching Counts	Unique Dest.	Disinct Source
02.12.2015 00.00-00.15	22.764.400	3.949.096	241.316	236.541	81.795
02.12.2015 00.15-00.30	26.415.27	3.957.438	240.382	236.006	81.470
02.12.2015 12.00-12.15	24.556.763	3.976.507	248.241	243.672	85.294
02.12.2015 18.00-18.15	23.929.802	3.920.760	226.480	221.531	80.343

In Figure 4, RTT delays between ICMP echo requests and their matched ICMP-T3 messages are shown for December 2nd, 2015 12.00-12.15 trace. Each point in the Figure indicates a RTT delay point of a node which notifies the unreachability of targeted IP in the echo request packet. The notifying source of ICMP-T3 message might be gateway, transit router, stub router or targeted host. Via looking the code of returned ICMP-T3 packet, it could be understood, what type of node sent ICMP-T3 message for the targeted IP address. Similar delay characteristics are observed in other trace files for different time periods on the same day.

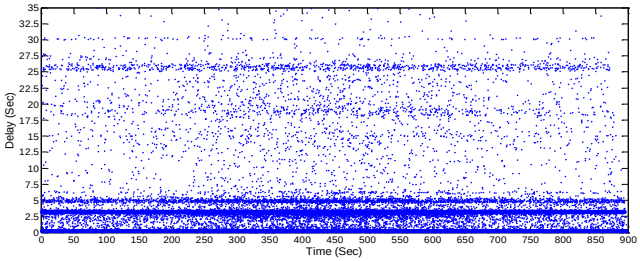


Figure 4. Round trip time (RTT) delay of unreachable destinations (December 2nd, 2015, 12.00-12.15).

A careful examination reveals that there are some delay levels where the number of unreachable destinations increase. Obviously, it seems that these levels are close to 0.5, 3, 5 seconds. The distribution graphic for interested traffic trace proves this outcome more elaborately as could be seen in Figure 5. Distribution graphic exhibits the RTT delays between 0-5.5 seconds with 10 ms bin size. As could be observed, the distribution curve consists of two modals. The first part is between 0-0.5 and the other part is between 3-3.5 seconds. Although the heights of similar peaks are not proportional in two parts, the difference time between the peaks show some resemblances. Furthermore, general shape of each peak and the tail part of the peaks are also in great similarity. The distribution graphic exhibits that some ICMP echo packets experience ICMP-T3 packets 3 seconds later. Figure 6 shows RTTs of unreachable destinations at different time intervals on the same day which is December 2nd, 2015. We also observe 3 seconds latency for unreachable destinations in all traces.

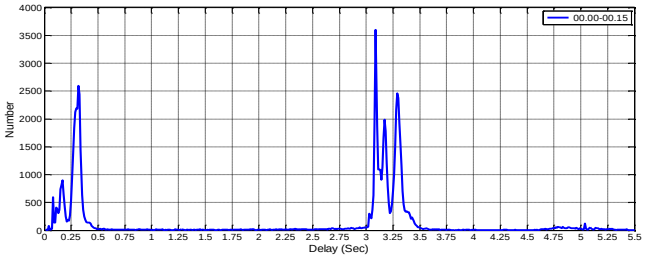


Figure 5. RTT distribution of unreachable destinations with10 msresolution (December 2nd, 2015, 12.00-12.15).

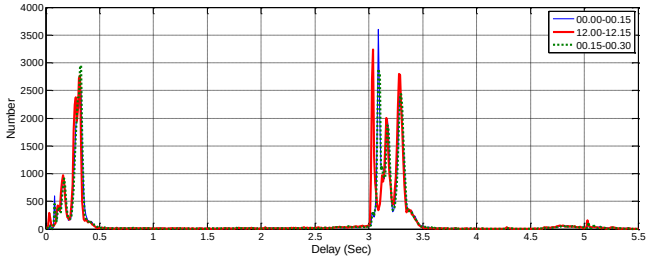


Figure 6. RTT distributions in different time intervals on December 2nd, 2015 with10 ms resolution.

A kind of bimodal behaviour exists in all distribution curves. Considering all distribution figures consists of two main parts, the first peaks in two separate modals mostly represent the RTTs for unreachable hosts close to the capturing point. The second and third peaks mostly shows the far away notifying nodes which could be in other countries, continents or beyond the pacific. The waviness in two modals are caused by clustering of notifying nodes being in different locations in the world [17].

Comparison of reachable and unreachable destination RTTs will give us more information about the delay dynamics of internet hosts. According to the obtained measurement results, in 22 million ICMP requests to randomly chosen IP hosts, 3.948.498 distinct hosts send ICMP reply to ICMP requests between 12.00 and 12.15 on December 2nd, 2015. In the same time interval, 81.795 unique destinations send ICMP-T3messages. The great difference in these numbers makes harder to compare delay dynamics in a linear scale. Therefore, logarithmic scale is chosen for the representation. The logarithmic scale exhibits that there is a certain increase in the number of unreachable destinations beyond 3 second compared to reachable destinations as seen in Figure 7.

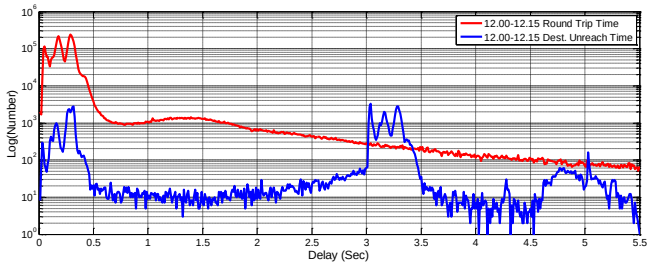


Figure 7. RTT distributions of reachable and unreachable destinations on December 2nd, 2015 with 0.01 Sec. resolution.

To reveal distributional differences, we apply distribution fitting to reachable and unreachable destination RTTs. It proves that reachable and unreachable destinations exhibit totally different distributional characteristics. We test 73 different distribution types for fitting procedure. As selection criteria for goodness of fitting, we employ Smirnov [30] and Anderson and Darling [7] test results. Phased Bi-Exponential gives the best result for unreachable destination RTTs. According to fitting results the parameters of fitted function are found as λ_1=0.33267, γ_1=0, λ_2=5.3652 andγ_2=3.0973. The cumulative density function of Phased Bi-Exponential is as follows,

$$F_x(x) = \begin{cases} \lambda_1 e^{-\lambda_1(x-\gamma_1)} \gamma_1 \le x \le \gamma_1, \\ \lambda_2 e^{-\lambda_2(x-\gamma_2)-\lambda_1(\gamma_2-\gamma_1)} \quad \gamma_2 \le x \le +\infty. \end{cases} \quad (5)$$

Figure 8 shows the cumulative distribution and fitted distribution for unreachable hosts.

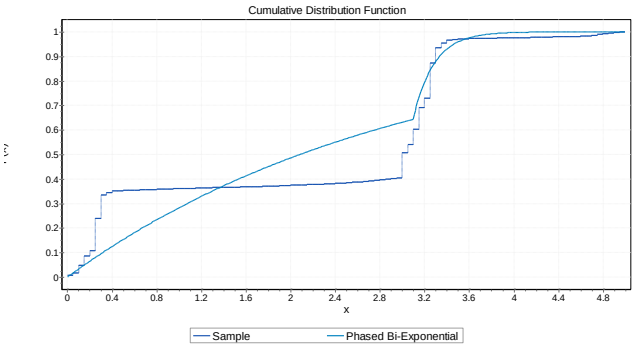


Figure 8. RTTs distribution and fitted distribution of unreachable hosts on December 2nd, 2015 with 0.01 Sec. resolution.

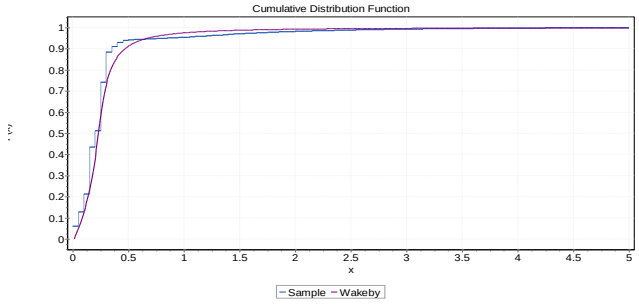


Figure 9. RTTs distribution and fitted distribution of reachable hosts on December 2nd, 2015 with 0.01 Sec. resolution.

Wakeby distribution gives the best fitting result for accessible hosts according to Smirnov [30] and Anderson and Darling [7] test results. Wakeby distribution parameters α, β, γ, δ, and ξare found as 0.77909, 4.2137, 0.0493, 0.67697, and 0.01172 respectively. The cumulative density function of Wake by distribution is as follows,

$$f(x) = \frac{(1-F(x))^{(\delta+1)}}{\alpha t+\gamma} \quad (6)$$

Where F is the cumulative distribution function and

$$t = (1 - F(x))^{(\beta+\delta)} \quad (7)$$

Figure 9 shows the cumulative distribution and fitted distribution for unreachable destinations.

Table 4. Unreachable destination statistics in terms of ICMP codes.

Date-Time	Code 0	Code 1	Code 3	Code 10	Code 13
02.12.2015 00.00-00.15	2438	65.143	9.940	468	3890
02.12.2015 00.15-00.30	2.477	64.301	10.499	453	3.868
02.12.2015 12.00-12.15	2.628	68.321	10.048	429	4.000
02.12.2015 18.00-18.15	2566	64.396	9.300	435	3.788

Considering the codes of ICMP-T3 messages, the distribution curves are obtained. The most common encountered code types are Code 0, 1, 3, 10 and 13 in our measurements. Summary information for ICMP-T3 messages in terms of their codes are given in Table 4. For the reliability of analysis, the repetitions and same source sending more than one ICMP-T3 packet situations are eliminated. According to the obtained results, host unreachability (Code 1) is the most encountered ones in all codes. The distribution curves for each code are obtained and shown with 10 ms resolution in Figure 10. As could be seen, Code 1determines mostly the shape of total distribution curves. We also observe that Code 1 is almost responsible for the distribution curve beyond 0.5 second. Between 0 and 0.5 second all codes contribute the total distribution curve. However, beyond 0.5 second, mostly host reach ability determines the distribution shape.

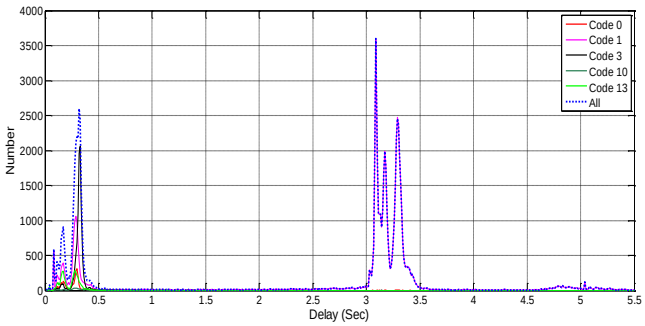
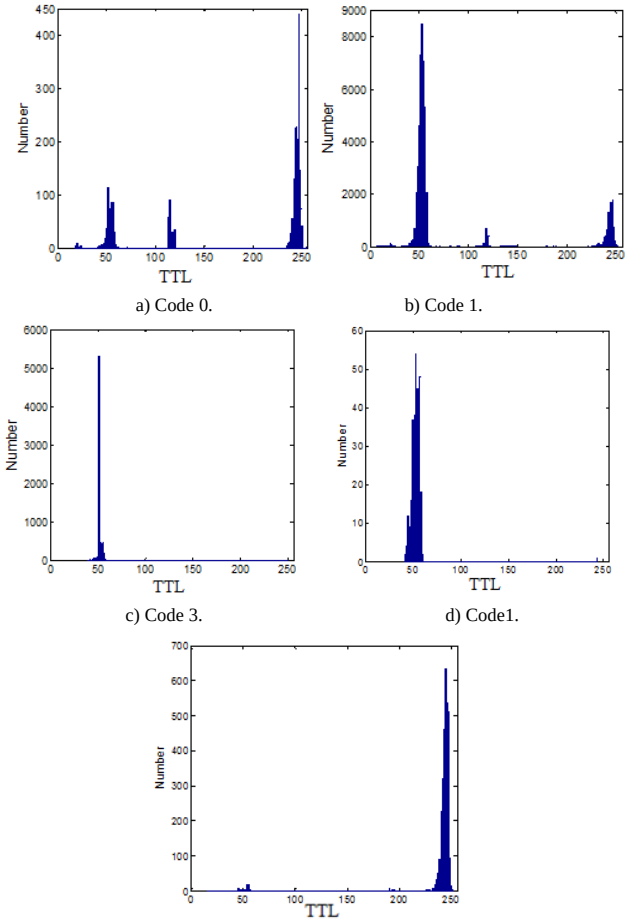


Figure 10. RTT distributions of unreachable destinations in terms of codes (December 2nd, 2015, 12.00-12.15, 0.01 sec. resolution).

In Figure 11, the Time to Live (TTL) histograms of unreachable hosts are shown. Mostly routers and hosts generate ICMP packets with 64, 128 and 256 TTL values according to their Internetwork Operating Systems (IOS). Concentration of histogram bars close to the mentioned TTL values point proves that the return way includes only a few hops. The most prominent result is that network unreachable packets and communication administratively prohibited ICMP packets have mostly TTL values close to 256 at the capturing point as seen in Figure 11-a) and Figure 11-e). That is, the end host of this message, mostly generate ICMP-T3 message with 256 TTL value. The other types are generated mostly with TTL value close to 64.



a) Code 0. b) Code 1.

c) Code 3. d) Code1.

e) Code 13.

Figure 11. TTLs of destination unreachable messages in terms of codes (December 2nd, 2015, 12.00-12.15).

The Estimated Hurst values for each traffic trace are given in Table 5 for reachable and unreachable destinations.

Table 5. Hurst parameter estimation values of each trace via Wavelet-Based methods.

Date-Time	Unreachable Destinations	Reachable Destinations
02.12.2015 00.00-00.15	0.4992	0.6951
02.12.2015 00.15-00.30	0.5064	0.6851
02.12.2015 12.00-12.15	0.4850	0.5568
02.12.2015 18.00-18.15	0.5050	0.7025

It is observed that reachable destinations give greater Hurst values in comparison to unreachable RTTs which is close to long range dependent behavior. However, unreachable hosts show a random behavior due to their Hurst values are close to 0.5. Obtained results demonstrate that Hurst parameters could be an indicator of abnormal traffic, and it might presage an upcoming cyber-attack in advance.

7. Conclusions

Delay related quality of service metrics are significant ones for the evaluation of Internet quality. The RTT

delays of unreachable destinations over Internet gives us some clues about a healthy Internet and upcoming worm propagation. The comparison of delay behaviour of reachable and unreachable destinations via ICMP protocol have shown us prominent results in terms of case numbers, distributional characteristics, and self-similarity degrees. Approximately, 15-20% of ICMP probing to random IP addresses take ICMP reply in return, only 1% take ICMP-T3 messages. Calculated delay quantities show that while most of the reachable host RTTs are below 0.5second, unreachable sources RTTs concentrated on two points which are below 0.5 second, and between 3-3.5 second. ICMP code-based analysis proves that the difference between the RTTs of reachable and unreachable destinations are mostly caused by host unreachability which is ICMP-T3 Code 1. Mostly, host unreachability adds extra 3 seconds. Distribution fitting results exhibit that while Wakeby distribution models the behaviour of reachable hosts RTTs, Phased Bi-Exponential distribution function models the distribution of unreachable host RTTs. Unreachable hosts RTTs show approximately 0.5 Hurst value which is close to a random behaviour. However unreachable host RTTs shows Hurst degree above 0.6 which is closer to long range dependency. This difference shows that Hurst parameter estimation of RTT delays might be a sign of upcoming worm propagation. In our future studies, we will focus on the detection of cyber-attacks via Hurst parameter estimation.

References

[1] Abdou A., Matrawy A., and Oorschot P., "Accurate One-Way Delay Estimation with Reduced Client Trustworthiness," *IEEE Communications Letters*, vol. 19, no. 5, pp. 735-738, 2015.

[2] Abdou A., Matrawy A., and Oorschot P., "CPV: Delay-based Location Verification for the Internet," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 2, pp. 130-144, 2015.

[3] Adhari H., Dreibholz T., Werner S., and Rathgeb E., "Eclipse: A New Dynamic Delay-based Congestion Control Algorithm for Background Traffic," *in Proceedings of 18th International Conference on Network-Based Information Systems*, Taipei, pp. 115-123, 2015.

[4] Allman M., Paxson V., and Blanton E., "TCP Congestion Control," RFC 5681, 2009.

[5] Almes G., Kalidindi S., and Zekauskas M., "A One-Way Delay Metric for IPPM," RFC 2679, 1999.

[6] Almes G., Kalidindi S., and Zekauskas M., "A Round-Trip Time Delay Metric fo IPPM Round Trip Delay," RFC 2681, 1999.

[7] Anderson T. and Darling D., "Asymptotic Theory of Certain 'Goodness-of-Fit' Criteria Based on Stochastic Processes," *The Annals of Mathematical Statistics*, vol. 23, pp. 193-212, 1952.

[8] Beran J., Sherman R., Taqqu M., and Willinger W., "Long-Range Dependence in Variable-Bit-Rate Video Traffic," *IEEE Transactions Communications*, vol. 43, no. 234, pp. 1566-1579, 1995.

[9] Berk V., Bakos G., and Morris R., "Designing a Framework for Active Worm Detection on Global Networks," *in Proceeding of the 1st IEEE International Workshop on Information Assurance*, Darmstadt Germany, pp. 13-23, 2003.

[10] Blenn N., Ghiette V., and Doerr C., "Quantifying the Spectrum of Denial of-Service Attacks Through Internet Backscatter," *in Proceedings of the 12th International Conference on Availability, Reliability and Security*, Reggio Calabria Italy, pp. 1-10, 2017.

[11] Chen S. and Tangi Y., "Slowing Down Internet Worms," *in Proceedings of the 24th International Conference on Distributed Computing Systems*, Tokyo, pp. 312-319, 2014.

[12] Cho K., Mitsuya K., and Kata A., "Traffic Data Repository at the WIDE Project," *in Proceedings USENIX 2000 Annual Technical Conference: FREENIX Track, USENIX Association*, San Dieog, pp. 263-270, 2000.

[13] Choi B., Moon S., Cruz R., Zhang Z., and Diot C., "Quantile Sampling for Practical Delay Monitoring in Internet Backbone Networks," *Computer Networks*, vol. 51, no. 10, pp. 2701-2716, 2007.

[14] Crovella M. Bestavros A., "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835-846, 1997.

[15] Csoma A., Toka L., and Gulyas A., "On Lower Estimating Internet Queuing Delay," *in Proceedings 38th International Conference on Telecommunications and Signal Processing*, Prague, pp. 299-303, 2015.

[16] Gezer A. and Warner G., "Exploitation of ICMP Time Exceeded Packets for A Large-Scale Router Delay Analysis," *The International Arab Journal of Information Technology*, vol. 16, no. 6, pp. 1090-1097, 2019.

[17] Gezer A., "Large-Scale Round-Trip Delay Analysis of Ipv4 Hosts Around the Globe," *Turkish Journal of Electrical Engineering and Computer Science*, vol. 27, no. 3, pp. 1998-2009, 2019.

[18] Heidemann J., Pradkin Y., Govindan R., Papadopoulos C., Bartlett G., and Bannister J., "Census and Survey of the Visible Internet," *in*

Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, New York, pp. 169-182, 2008.

[19] Jacobson V., "Congestion Avoidance and Control," *in Proceedings of ACM SIGCOMM'88*, Stanford, pp. 314-329, 1988.

[20] Karn P. and Partridge C., "Improving Round-Trip Time Estimates in Reliable Transport Protocols," *ACM SIGCOMM Computer communication Review*, vol. 17, no. 5, pp. 2-7, 1987.

[21] Leland W., Taqqu M., Willinger W., and Wilson D., "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1-15, 1994.

[22] Nagata T., Su W., and Lee J., "Using Hybrid Method to Detect Internet Worms by Analyzing ICMP Type 3 Messages and Worm Characteristic Matching," *International Information Institute (Tokyo) Information*, vol. 23, no.1, pp. 21-28, 2021.

[23] Paxon V., Allman M., Chu J., Sargent M., "Computing TCP's Retransmission Timer," RFC 6298, 2011.

[24] Paxson V. and Floyd S., "Wide Area Traffic: the Failure of Poisson Modeling," *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 226-244, 1995.

[25] Postel J., "Internet Control Message Protocol," RFC 792, 1981.

[26] Roy A., Pachuau J., and Saha A., "An Overview of Queuing Delay and Various Delay Based Algorithms in Networks," *Computing*, pp. 1-39, 2021.

[27] Salehin K., Cessa R., Lin C., Dong Z., and Kijkanjanarat T., "Scheme to Measure Packet Processing Time of a Remote Host through Estimation of End-Link Capacity," *IEEE Transactions on Computers*, vol. 64, no. 1, pp. 205-218, 2015.

[28] Salehin K., Rojas-Cessa R., and Ziavras S., "A Method to Measure Packet Processing Time of Hosts Using High-Speed Transmission Lines," *IEEE Systems Journal*, vol. 9, no. 4, 2015.

[29] Sebopetse N., Burger C., Mofolo M., and Lysko A., "Measuring with JPerf and PsPing: Throughput and Estimated Packet Delivery Delay vs TCP Window Size and Parallel Streams," *in Proceedings 7th International Conference on Advanced Computing and Communication Systems*, Coimbatore, pp. 838-832, 2021.

[30] Smirnov N., "Table for Estimating the Goodness of Fit of Empirical Distributions," *The Annals of Mathematical Statistics*, vol. 19, no. 2, pp. 279-281, 1948.

[31] Veitch D. and Abry P., "A Wavelet-Based Joint Estimator for the Parameters of LRD," *Special Issue on Multiscale Statistical Signal Analysis and its Applications IEEE Transactions on Informatics Theory*, vol. 45, no. 3, pp. 878-897, 1999.

[32] Wang J., Dong W., Cao Z., and Liu Y., "On the Delay Performance in a Large-Scale Wireless Sensor Network: Measurement, Analysis, and Implications," *IEEE/ACM Transactions on Networking*, vol. 23, no. 1, pp. 186-197, 2015.

Ali Gezer was born in Kayseri City, Turkey, in 1976. He received the B.S. degree in Electronic and Computer Education from Marmara University in 1999 and M.S. degree in computer engineering from Erciyes University in 2004, and the Ph.D. degree in electronic engineering from Erciyes University, Kayseri, TURKEY, in 2011.

He is an Associated Professor with the Electronic and Communication Technology in Kayseri University. His research interests include internet traffic measurement, self-similarity, traffic modelling, network security and cyber-attack detection.