DOI: 10.21917/ijct.2015.0177

PERFORMANCE ANALYSIS OF IPv4 AND IPv6 INTERNET TRAFFIC

Rupesh Jaiswal[1], Shashikant Lokhande[2], Anuj Bakre[3] and Kishor Gutte[4]

[1,3,4] *Department of Electronics and Telecommunication Engineering, Pune Institute of Computer Technology, India*
E-mail: rcjaiswal@pict.edu[1], anujbakre@gmail.com[3], kishor11793@gmail.com[4]
[2]*Department of Electronics and Telecommunication Engineering, Sinhgad College of Engineering, India*
E-mail: sdlokhande.scoe@sinhgad.edu

Abstract

The gigantic growth of the internet communication technology has illustrated its value and benefits to private businesses, government organizations, worldwide professionals, academic institutes and individuals over the past few years. The size and range of computing devices connected to the internet, substantially increased because of IPv6 and offers the potential to establish a much more powerful internet compared to the IPv4. IPv6 developed by the IETF to deal with a shortage of IP addresses under IPv4. New features of IPv6 enhance packet processing speeds over routers, switches and end systems. These improved features will have different traffic characteristics than IPv4. The internet traffic which was earlier assumed as Poisson is now shown to have fractal characteristics as; heavy tailedness, self-similarity and long range dependency. Internet traffic showing above characteristics are found to have burstiness at multiple timescales. This behavior impacts network performance and degrades it substantially. It also increases complexity for network design and create difficulties to maintain desired QoS. IPv4 traffic has been well established as self-similar traffic. Nowadays, IPv6 forming a larger share of the internet traffic and it is pivotal to asses IPv6 with regards to fractal behavior. This will enable network designers to do necessary changes in the existing network to reconcile with IPv6. In this paper we compared IPv4 and IPv6 with respect to fractal behavioral characteristics. It is found that IPv6 shows higher degree of heavy tailedness, higher values of Hurst parameter values, higher fractal dimension values i.e. it is more self-similar, greater autocorrelation achieved even at larger lag and thus showing more burstiness.

Keywords:

IPv4, IPv6, Self-Similarity, Long Range Dependence, Heavy Tailedness, Burstiness

1. INTRODUCTION

Nature of internet traffic must be correctly understood in order to design computer networks and network services. Earlier, the modelling of internet traffic was same as that of telephony, hence Poisson model was commonly used. Study of internet packet traffic has challenged this approach. For the traffic following Poisson process, the property of burstiness would eventually die off over long time scales. Instead, it is observed that considerable burstiness is present even over long enough time scales [1, 2, 6].

The line of thinking was changed by the research paper on long range dependence by Leland et al. [3]. It waived off the Poisson model adapted for telephony. It clearly stated that new approach was required for modelling the data packet traffic. The paper emphasized that internet traffic exhibits self-similarity or fractal characteristics, long range dependent (LRD) behavior and Heavy tails (Power laws) [4].

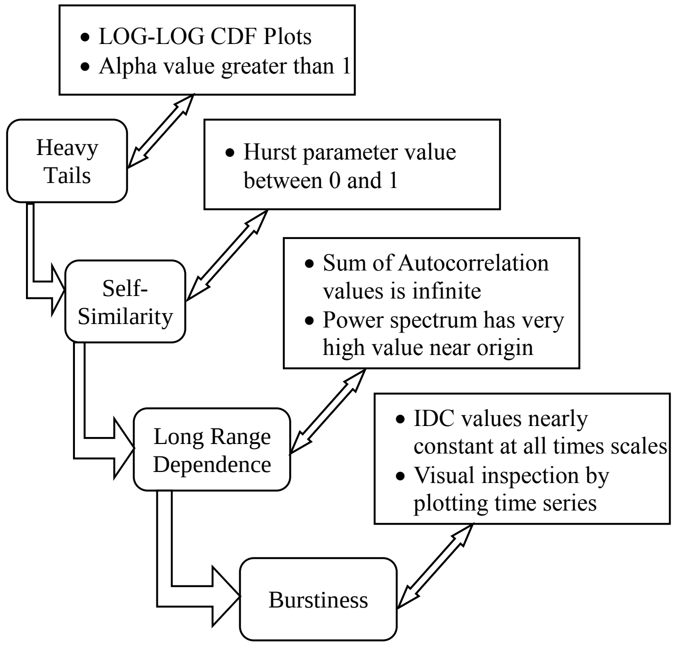

Fig.1. Flow of the internet traffic analysis. Sub points indicates the methods and logic used for analysis

In this paper, we analyzed internet traffic with respect to heavy tailed behavior, self-similarity, long range dependency and burstiness for both IPv4 and IPv6 [9, 10, 11] traffic. It is seen that the three characteristics like heavy tailed behavior, self-similarity and long range dependency are responsible for burstiness in traffic. The sequence correctly illustrates the interrelationships as shown in Fig.1.

Heavy tailedness is a significant property of internet traffic which follows power law. The distribution of file sizes follows power law and therefore very large file transfer can be expected with non-negligible probability. Many such files in a network environment together causes self-similarity. Thus self-similar traffic results from heavy tailed behavior of file sizes being transferred over the network [5].

Self-similarity means the property exhibited by a fractal object which appears unchanged at any viewing scale [7]. Thus, Internet packet traffic showing burstiness at various time scales can be labeled as self-similar traffic. However, the traffic must exhibit long range dependence for above mentioned phenomenon to occur [1].

In this study, our goal is to characterize IPv6 and IPv4 packet traffic based on heavy tails, self-similarity, long range dependence, auto correlation and power spectral density (PSD) analysis for the parameters like packet length and packet interarrival time. We also carried out probability distribution fitting for IPv4 and IPv6 packets for above parameters. In order to generate accurate data traffic for network simulation, it is essential to use correct probability distribution models.

Organization of paper is as follows. Firstly IPv4 and IPv6 protocols are explained in detail in section 2. Information about data traces used and analysis perspective is explained in section 3. In section 4, we perform comparison of IPv4 and IPv6 traffic for the parameters like inter-arrival time and packet length in terms of probability distribution fitting, heavy tailedness, self-similarity, autocorrelation, power spectral density analysis and burstiness. Finally section 5 shows conclusion of research work.

2. BACKGROUND WORK

2.1 IPv4 AND IPv6 PROTOCOL

The main limitation of IPv4 (with 32 bit address) is its lack of address space and provides up to 4 billion IP addresses and was never designed to have all 4 Billion IP addresses used simultaneously. The current world population is of 7 Billion [22] and the number of computing devices connected to the Internet are much more than number of humans [23]. IPv6 uses 128 bit address (approximately 3.4 × 1038 addresses) and can connect massive number of devices at a time. Internet Protocol was comprehensively developed for long-term growth of the internet. The IPv4 header contains of 20-bytes. The maximum length of the IPv4 header is 60 bytes and it has 13 fields to recognize control settings. The IPv6 header is a fixed header of 40-bytes, with only 8 fields. For optional information an extension header is used after IPv6 header.

The IPv4 header uses optional fields for particular processing of packets. These optional fields are not often used and hence might degrade router performance. This is not the case with IPv6. Here extension headers are not processed until the packet reaches the destination node recognized by address field of the IPv6 header. IPv6 provides authentication of the sender packets and encryption of packets thus security is enhanced. Other limitations of the IPv4 protocol are: non-hierarchical addressing, mobility and multi-homing, large routing tables, complex host and router configuration, QoS (Quality of Services), multicasting etc. IPv6 protocol stacks are employed in parallel with IPv4 so that promotion of IPv6 from IPv4 becomes easier. This means that host can work with existing IPv4 network and also process IPv6.

2.2 TRAFFIC CAPTURE AND DATASET PREPARATION

For analysis we downloaded packet level internet traces from MAWI working group traffic archive. We used internet traces captured in the month of May 2009 from IPv6 line connected to WIDE-6Bone. The traffic available is in tcp dump format. All traces are inclusive of application traffic like HTTP, SMTP, P2P, ICMP6 and FTP. For IPv4 traffic, we use the daily traces of trance pacific line (18Mbps CAR on 100Mbps link) throughout the month of April 2006 [8].

Traces with dump file format and pcap file format are processed to calculate the packet length and inter arrival time series for further experiments.

3. RESULT ANALYSIS

3.1 DISTRIBUTION FITTING

The Study of probability distribution functions are pivotal in network traffic analysis. Here we perform probability distribution fitting for time series of parameters like inter-arrival time and packet length. A time series is characterized using known distribution function and hence we study cumulative distribution function [12] for IPv4 and IPv6 traffic in detail.

The pdf and cdf for three-parameter lognormal distribution is given as:

Probability Density Function for three-parameter lognormal distribution is given as:

$$f(x)=\frac{\exp\left(-\frac{1}{2}\left(\frac{\ln(x-y)-\mu}{\sigma}\right)^2\right)}{(x-\gamma)\sigma\sqrt{2\pi}}. \tag{1}$$

Cumulative Distribution Function for three-parameter lognormal distribution is given as:

$$F(x)=\phi\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right) \tag{2}$$

where, ϕ is the Laplace integral.

The pdf and cdf for three-parameter Weibull distribution is given as:

Probability Density Function for three-parameter Weibull distribution:

$$f(x)=\frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right). \tag{3}$$

Cumulative Distribution Function for three-parameter Weibull distribution:

$$F(x)=1-\exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right). \tag{4}$$

The pdf and cdf for three-parameter log-logistic distribution is given as:

Probability Density Function for three-parameter log-logistic distribution:

$$f(x)=\frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}. \tag{5}$$

Cumulative Distribution Function for three-parameter log-logistic distribution:

$$F(x)=\left(1+\left(\frac{\beta}{x-\gamma}\right)^{\alpha}\right)^{-1}. \tag{6}$$

Different probability distribution types are tested for distribution fitting process. As a standard for goodness of fitting, we use Anderson-Darling and chi square test results and decided the appropriate distribution type for packet length and packet inter-arrival time of IPv6 and IPv4 traffic [13, 14, 15].

From the obtained fitting results, we found three parameter log normal distribution as best fit for IPv4 packet length and inter-arrival time. For IPv6 packet inter-arrival time, Weibull distribution gives the best fit whereas for packet length we have log logistic distribution as best fitted distribution.

The range of values of defining parameters of distribution are given below:

Dataset: IPv4 packet length

Best fit: log normal (3P)

Range for σ is 0.01486 to 0.01911, μ is 7.1368 to 7.2083 and γ is -1019.5 to -923.47.

CDF plot for IPv4 Data set for Packet length is best fitted as shown in Fig.2.

Dataset: IPv4 inter-arrival time series

Best fit: log normal (3P)

Range for σ is 2.1863 to 2.34, μ is 2.8457 to 3.5385 and γ is 4.9668 to 4.9834

CDF plot for IPv4 Data set for Packet IAT is best fitted as shown in Fig. 3.

Dataset: IPv6 packet length

Best fit: log logistic (3P)

Range for α is 1.3352 to 1.408, β is 60.801 to 73.216 and γ is 65.991 to 66.0

CDF plot for IPv6 Data set for Packet length is best fitted as shown in Fig.4.

Dataset: IPv6 inter-arrival time series

Best fit: Weibull (3P)

Range for α is 0.52327 to 5.3421, β is 871.89 to 10721 and γ is 3.0 to 4.0

CDF plot for IPv6 Data set for Packet IAT is best fitted as shown in Fig.5.

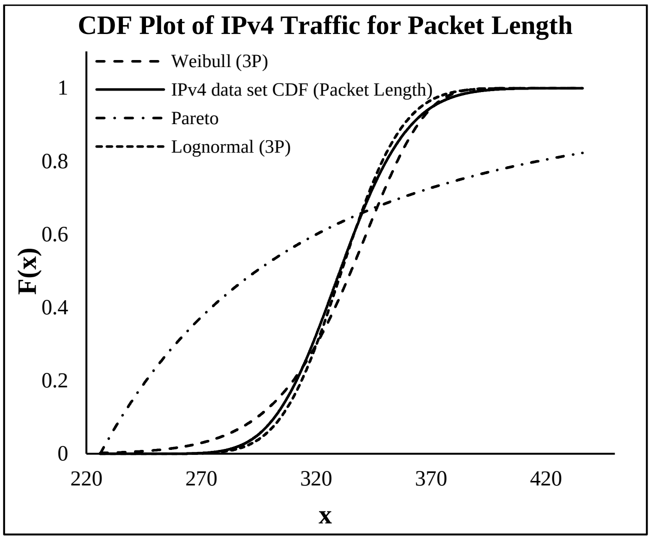

Fig.2. CDF plot for IPv4 Data set (Packet length). Best fit obtained for Lognormal (3P) distribution

CDF plot for packet length and IAT parameters are drawn by finding best fit probability distribution using statistical toolbox.

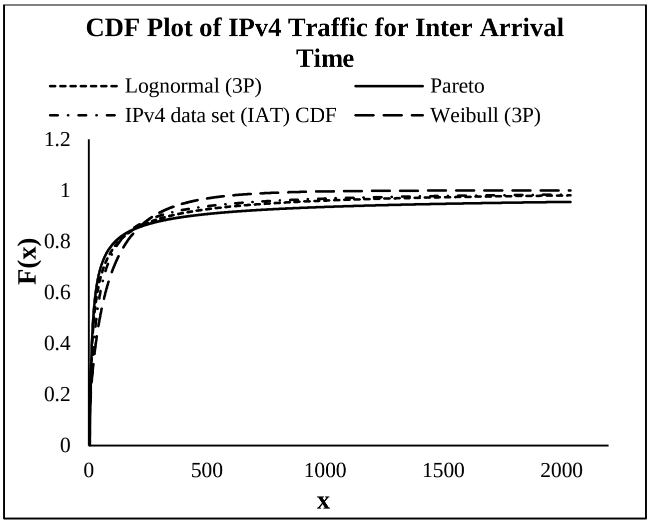

Fig.3. CDF plot for IPv4 Data set (IAT). Best fit obtained for Lognormal (3P) distribution

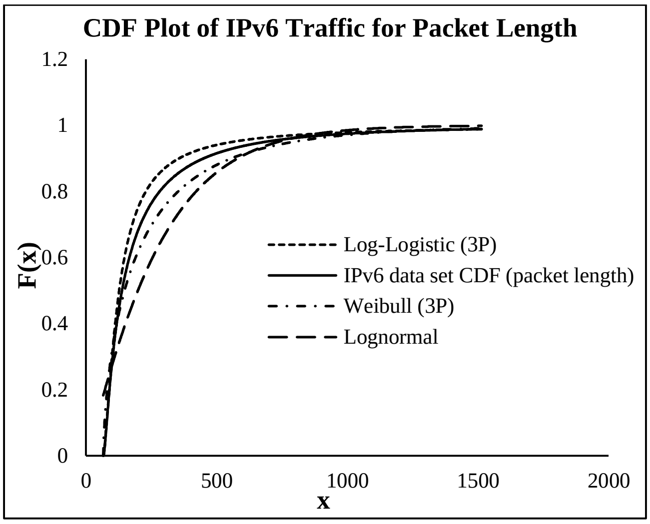

Fig.4. CDF plot for IPv6 Data set (packet length). Best fit obtained for Log-logistic (3P) distribution

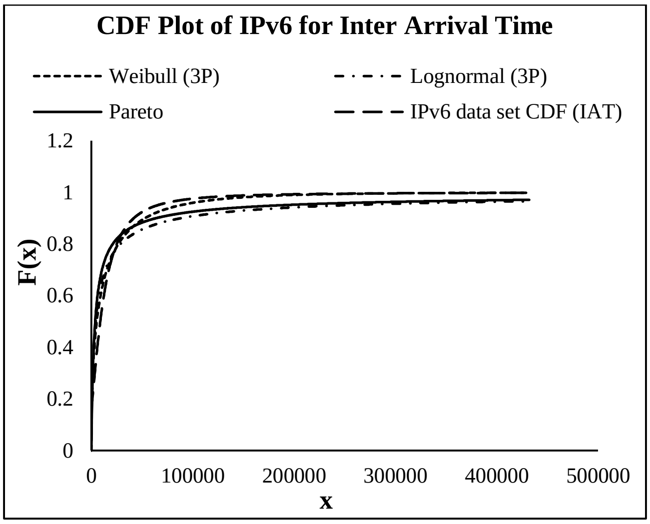

Fig.5. CDF plot for IPv6 Data set (IAT). Best fit obtained for Weibull (3P) distribution

3.2 HEAVY TAIL ANALYSIS

A distribution is heavy-tailed if:

$$P[X > x] \sim x^{-\alpha} \text{as } x \to \infty \tag{7}$$

where, $0 < \alpha < 2$ [1].

That is, the asymptotic shape of the distribution follows a power law.

The properties of Heavy-Tailed Distributions (HTDs) are:

i. CCDF decays slower than the exponential distribution

$$P(X > x) = 1 - F(x) = \bar{F}(x) = e^{-\lambda x} \tag{8}$$

ii. CCDF = Complementary cumulative distribution function

$$\bar{F}(x) = 1 - F(x) \tag{9}$$

iii. For heavy tailed distribution, CCDF is slower by some power of x

$$\bar{F}(x) \to cx^n e^{-\lambda x} \tag{10}$$

The heavy tailed distribution can be evaluated based on shape parameter α. The methods used are:

1. Log-Log Complementary Distribution (CD) plots
2. The Hill estimator

CD plots have the CCDF on log-log-axes. For such a plotted graph, heavy tailed distributions shows

$$\frac{d \log \bar{F}(x)}{d \log x} \sim -\alpha \tag{11}$$

for large value of x.

For visual inspection we plot Log-Log CD plots for both Poisson distribution and the given data trace. If the graph of dataset lies above Poisson's graph, it is heavy tailed [16].

Log-Log CD Graphs for inter arrival time of IPv4 and IPv6 traffic are plotted. They both lie above the log(1-F(x)) plot for Poisson traffic. Hence both show heavy tailed behavior. But more heavy tailedness is observed in case of IPv6 traffic.

Log-Log CD plots for IPv4 dataset for IAT parameter is as shown in Fig.6.

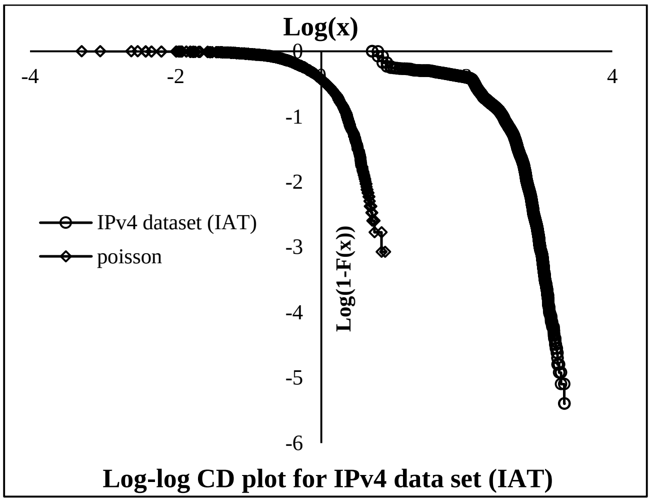

Fig.6. Log-Log CD plots for IPv4 dataset (IAT) and compared with Poisson traffic

Log-Log CD plots for IPv6 dataset for IAT parameter is as shown in Fig.7.

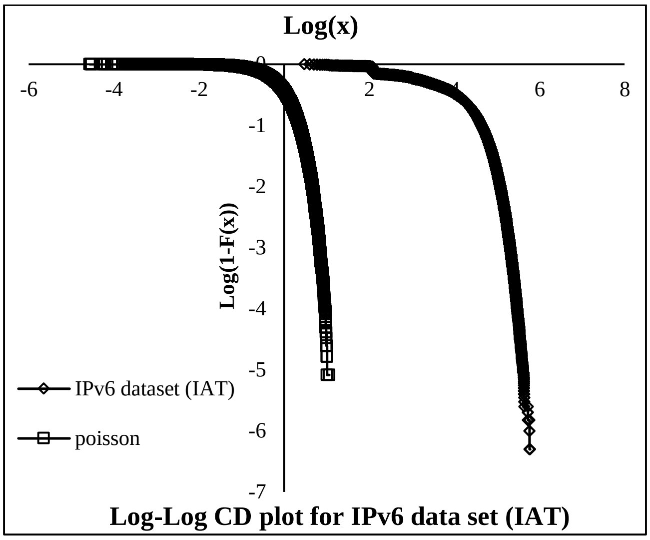

Fig.7. Log-Log CD plots for IPv6 dataset (IAT) and compared with Poisson traffic

3.3 SELF-SIMILARITY

Self-similarity can be simply related to correlation. Insights given by self-similarity in terms of looking an object are valuable. Self-similarity looks at an object at various scales and tells if it looks the same or not. This property is exploited for gaining information from time series of the internet traffic. Computer networks related calculations like resource sharing, routing management and queue management can be done using self-similarity perspective.

Self-similar time series: Instead of continuous process we normally represent time series in network traffic. Then self-similarity can be defined as,

Let $G = \{G(j), j \geq 1\}$ be a stationary sequence.

$$G^{(n)}(k) = \frac{1}{n} \sum_{j=(k-1)n+1}^{kn} G(j), k = 1,2,\ldots, \tag{12}$$

Then, $G^{(n)}(k)$ will be the aggregated sequence. It has aggregation level of m which is obtained by averaging over non-overlapping blocks of size n. So we can state self-similar process, for all integers n as,

$$G \underline{\underline{\text{def}}}\ n^{1-H} G^{(n)} \tag{13}$$

For a process to be randomly scattered it must have Hurst value, $H = 0.5$. When a continuous time stochastic process $\{G(t), t \in R\}$ is called strictly self-similar then it has H value $\{H, 0 < H < 1\}$ for;

$$G(ct) \underline{\underline{\text{def}}}\ C^H G(t) \tag{14}$$

Here c is the scaling factor, so we get a new process $G(ct)$. $\underline{\underline{\text{def}}}$ Means equal in finite dimensional distributions. Hurst parameter values between 0 and 0.5 indicate short range dependency while value of Hurst parameter between 0.5 and 1 shows long range dependency behavior [4].

Table.1. Average Hurst parameter values for IAT and packet length

Data set	R/S method	Abs. Moment method	Time variance method	Difference variance method	Boxed periodogram	Box counting
IPv4 IAT	0.5678	0.5764	0.5777	0.5546	0.5582	1.6867
IPv4 packet length	0.5773	0.5988	0.5907	0.5338	0.5480	1.5992
IPv6 IAT	0.9250	0.8778	0.8535	0.7226	0.6138	1.430
IPv6 packet length	1	0.911	0.911	0.7273	0.6095	1.863

In this section, we carry out self-similarity analysis of IPv6 and IPv4 traffic in terms of packet length and packet inter-arrival time. We calculate Hurst parameter using R/S method, absolute moment method, time variance method, difference variance method and boxed periodogram [17] method. These methods have already been implemented in our previous research work [21]. We also carried out fractal dimension (FD) analysis [18], [21] and demonstrated in this paper using box counting method. The average values obtained for both IPv4 and IPv6 traffic are mentioned below for all above methods.

By comparing the Hurst and fractal dimension values for different methods mentioned above, we can say that IPv4 and IPv6 traffic dataset used are self-similar. But the extent of self-similarity is more in case of IPv6 traffic. This is due to higher heavy tailedness observed in case of IPv6 traffic.

3.4 LONG RANGE DEPENDENCE

3.4.1 Autocorrelation Analysis:

Hurst parameter is a quantitative measure of self-similarity. It articulates the speed of decay of autocorrelation function value for a time series.

Long range dependence is easily observed in LAN and WAN traffic. It degrades the performance as in case of long range dependent traffic where queue length decays slowly. Long-range dependence (LRD) can be shown by a time series. For such a series autocorrelation function can be written as,

$$s(m) \sim m^{-\beta} \text{as } m \to \infty \tag{15}$$

where, the value of β is between 0 and 1.

The relation between Hurst parameter and β is as follows

$$H = 1 - \frac{\beta}{2} \tag{16}$$

Thus for LRD time series, Hurst parameter [5] is given as, $1/2 < H < 1$. As, $H \to 1$. The degree of long-range dependence increases.

A time-series which is self-similar can show long range dependence. For such a process, autocorrelation function is given as, $r(k) \sim k^{-\beta}$ and $k \to \infty$,where value of β is given as, $0 < \beta < 1$. Thus power law behavior is revealed from auto correlation function.

It is known that power law decay slower than exponential. Further we have, $\beta < 1$, hence the sum of autocorrelation values will approach to infinity. Long range dependency can also be depicted in power spectrum. The power spectrum is hyperbolic and rises to infinity at zero frequency.

It is observed from these graphs that for IAT and packet length, low values of auto correlation are obtained. This can be related to the low values of H. We have H values in range of 0.5 to 0.6.

This indicates lesser-long range dependence. IPv4 packet length and packet interarrival time autocorrelation plots shows that IPv4 packet does not show good degree of long range dependence. This is acceptable as Hurst exponent obtained is also less. Similar results were obtained for remaining data traces.

On the other hand autocorrelation values for IPv6 packet length and packet inter-arrival time depict a high amount of long range dependent behavior. This is likely as earlier we obtained high Hurst exponent values. Similar results were obtained for other data traces.

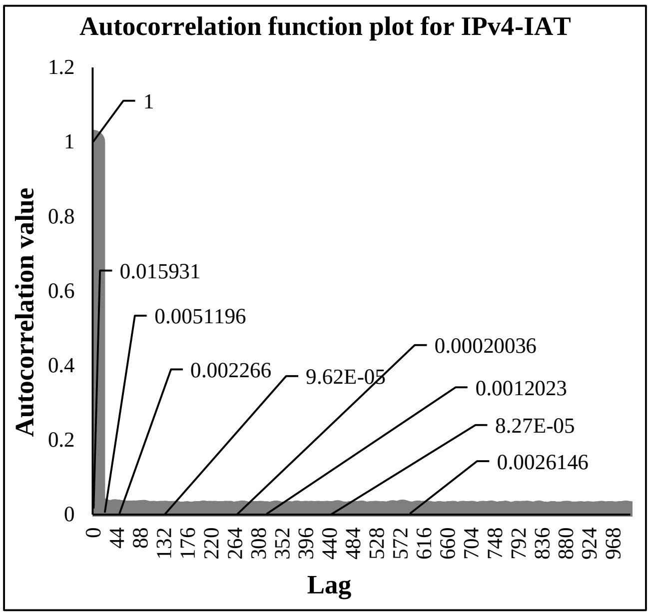

Fig.8. Autocorrelation function plot for IPv4 dataset (IAT)

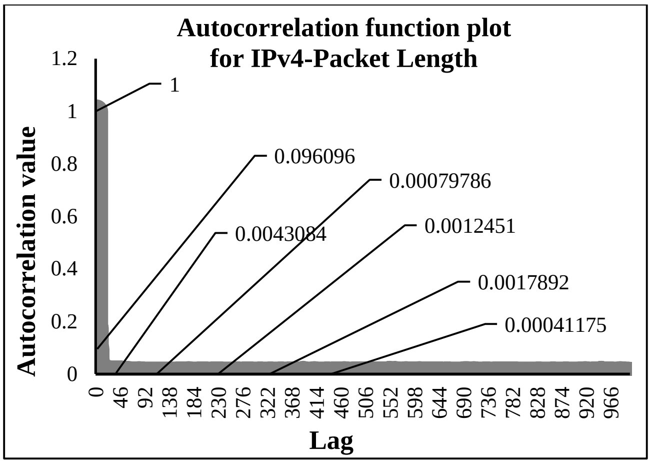

Fig.9. Autocorrelation function plot for IPv4 dataset (Packet length)

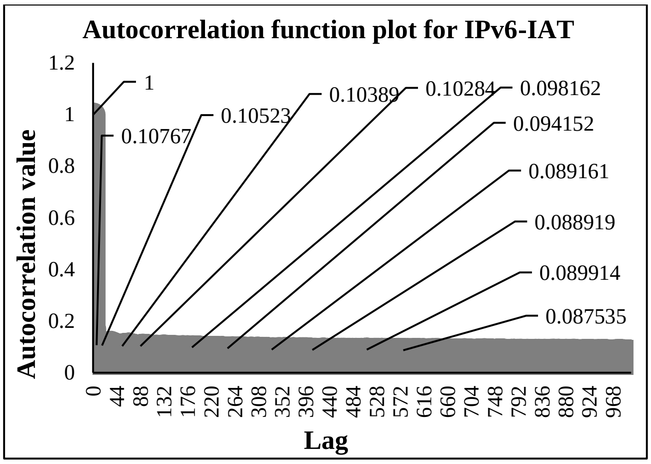

Fig.10. Autocorrelation function plot for IPv6 dataset (IAT)

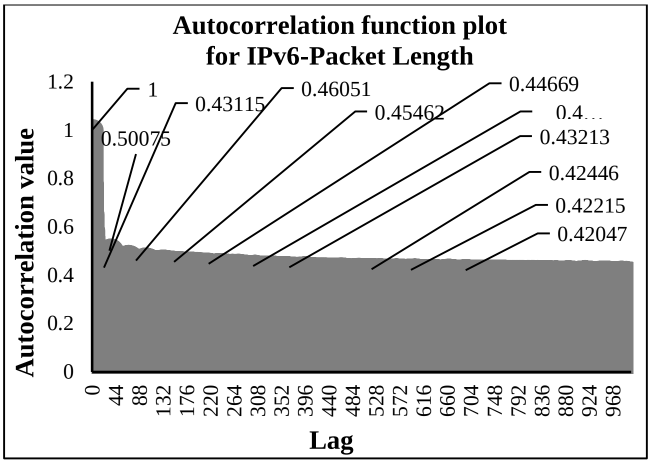

Fig.11. Autocorrelation function plot for IPv6 dataset (packet length)

3.4.2 *Power Spectral Density:*

Self-similar traffic can be characterized by power spectral density (PSD). For LRD time series, PSD follows a power law near origin.

$$Q_x(v) \approx |v|\, as\, v \to \infty,\ 0 < \gamma < 1, \tag{17}$$

where, v is frequency, $Q_x(v)$ is the spectrum density and $\gamma = H - 1$.

They showed analogous characteristics; representing a little $1/f$ type power spectrum behavior. Also Gaussian type power spectra were observed for both of them. Value of PSD is more in case of IPv6 as compared to IPv4. We have shown a plot of PSD [19] for a data trace, but similar plots are obtained for other data traces.

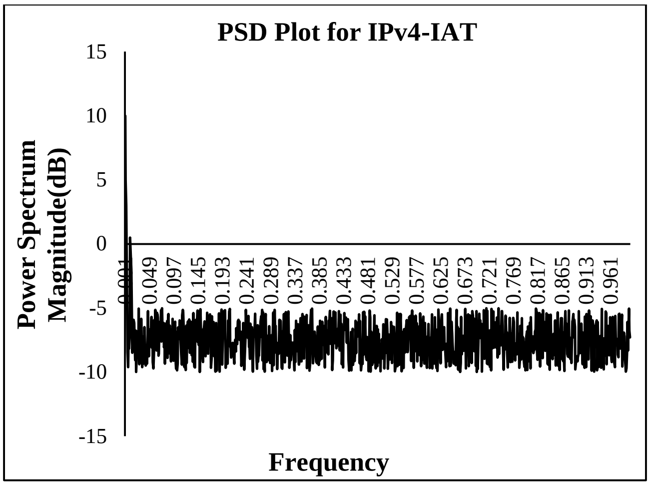

Fig.12. PSD plot for IPv4 dataset (IAT)

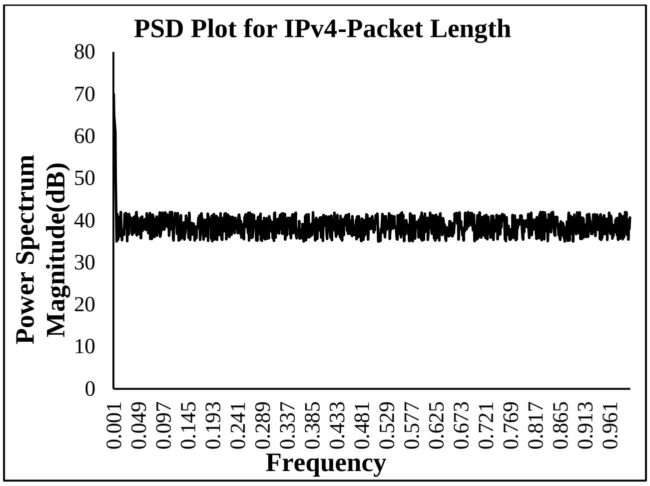

Fig.13. PSD plot for IPv4 dataset (Packet length)

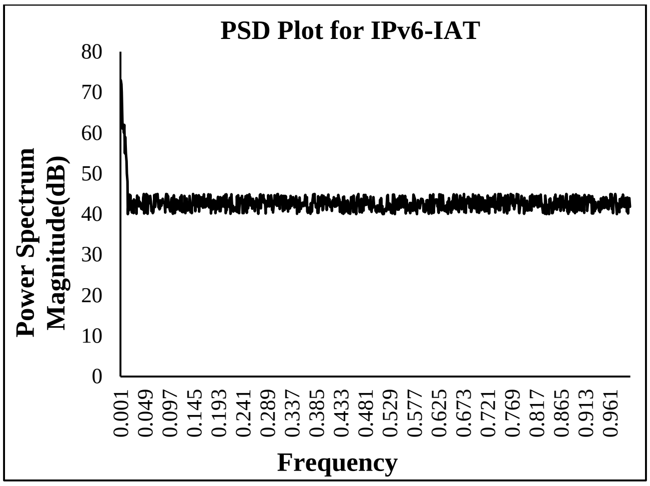

Fig.14. PSD plot for IPv6 dataset (IAT)

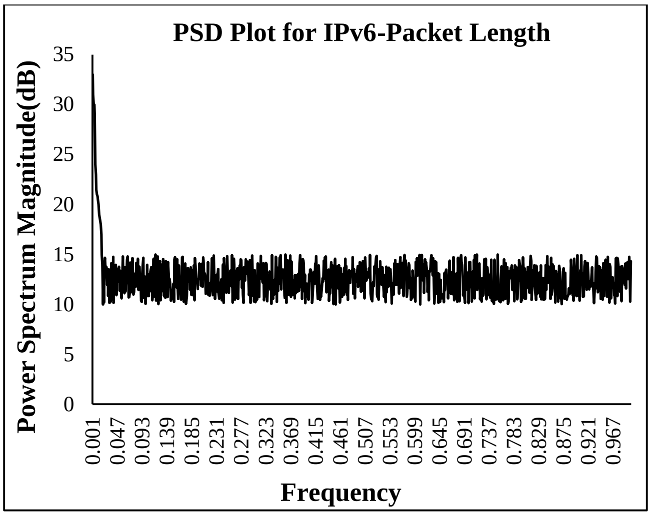

Fig.15. PSD plot for IPv6 dataset (Packet length)

3.5 BURSTINESS

By definition, burstiness is the degree of variation in network traffic. Burstiness can be studied using Peak to mean ratio (PMR) and index of dispersion for counts (IDC).

For time scale = t, IDC can be defined as

$$I_t = \frac{Variance(K_t)}{E(K_t)} \tag{18}$$

where, K_t indicates the number of arrivals in an interval of time t and $E(K_t)$ is mean number of arrivals in time t.

Mathematically, Peak to mean ratio (PMR) can be written as:

$$PMR = \frac{Max(K_t)}{Mean(K_t)}. \tag{19}$$

The value of the IDC is 1 for all 't' for a Poisson process, and also compared with other research works [20].

We check the values for burstiness in terms of IDC and PMR. Value of IDC should be greater than 1 even for larger time scales.

We calculated IDC and PMR for 100 different scales like 1, 10, 25, 50, 75, 100, 125, 150, 175 and 200 seconds. Approximately 2.5 lakh packets containing dataset were analyzed for IPv4 and IPv6 traffic. Higher values of IDC and PMR are obtained for IPv6, even at higher time scales. Even in case of packet length parameter, IDC and PMR values are high as compared to IPv4 traffic. The IDC values obtained are as given in Table.2 and Table.3.

Table.2. Average IDC for IPv4 data traffic

Time duration* (seconds)	1	10	50	100
IPv4 (packet length)	330.5	288.36	221.64	205.6
IPv4 IAT	348.36	280.73	217.9	198.3

*Due to small length traces available for IPv4 IDC could be calculated over short duration only.

Table.3. Average IDC for IPv6 data traffic

Time duration (seconds)	1	10	50	100	200
IPv6 packet length	829.9	857.4	677.3	653.9	628.6
IPv6 IAT	5990	5456	5507	4735	4058

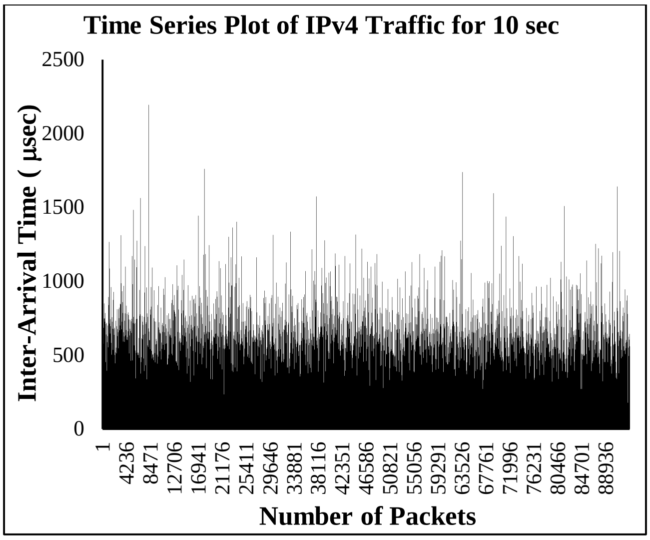

Fig.16. Time series plot for IPv4 data set (IAT) for time scale 10 seconds

Mathematically, an object is said to be self-similar if it looks "roughly" or "exactly" the same on any scale (short, medium or long scale). Self-similar traffic is bursty on many or all timescales (i.e. Variation in the average rate of the traffic stream is same for all time scales). Also for self-similar traffic, as link speed increases, if the traffic is truly bursty at all-time scales, the queuing delay would not decrease with increased traffic aggregation. This is the reason where Poisson model fails to represent the HTTP traffic and well defined by self-similarity process.

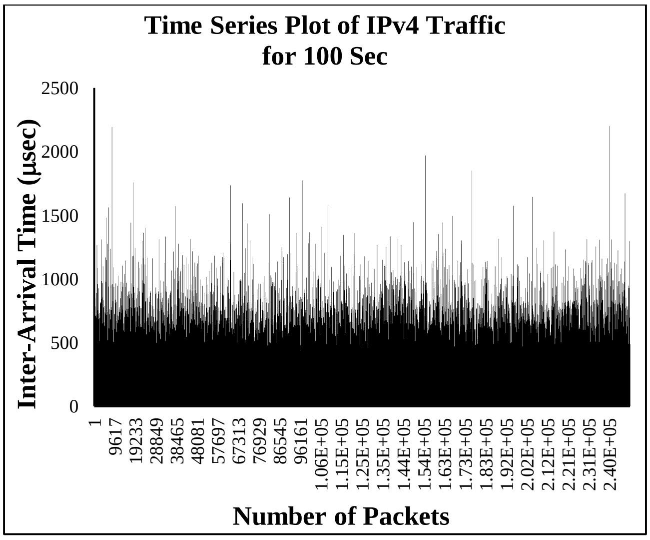

Fig.17. Time series plot for IPv4 data set (IAT) for time scale 100 seconds

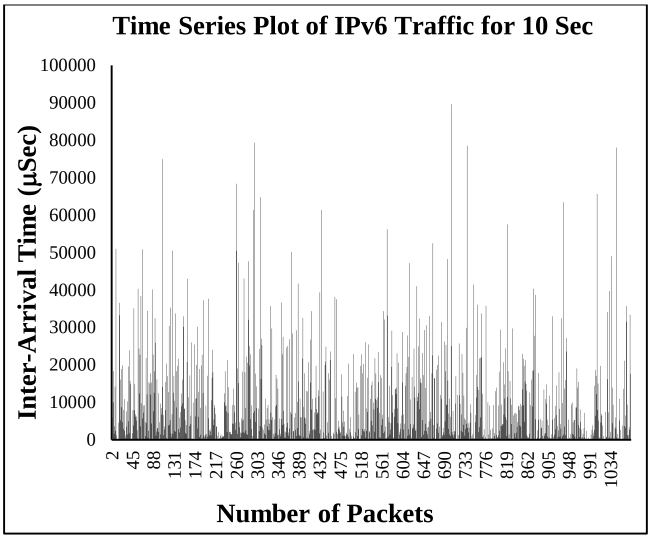

Fig.18. Time series plot for IPv6 data set (IAT) for time scale 10 sec

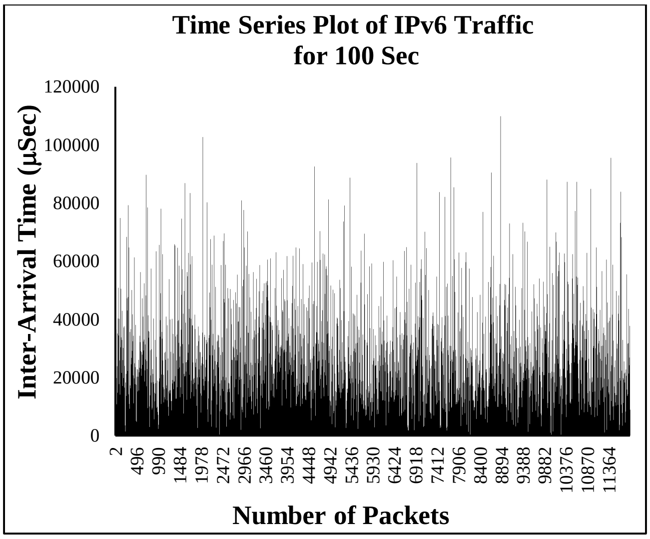

Fig.19. Time series plot for IPv6 data set (IAT) for time scale 100 seconds

Performance analysis is done for IPv4 and IPv6 traffic in this research work whereas it is done for HTTP and VoIP traffic by other researchers. The approach used in this research are based on autocorrelation analysis, fractal dimension analysis and probability distribution fitting analysis etc. Also, the comparison for IPv4 and IPv6 in terms of self-similarity and burstiness is rarely seen in the available literature survey.

In this research work, IPv4 and IPv6 traffic is characterized based on heavy tails, self-similarity, long range dependence, auto correlation, power spectral density (PSD) and burstiness. This analysis is done for the parameters like packet size and packet interarrival time. The Table.4 compares our research work with others.

Table.4. Comparison of our research work with existing research work

Traffic considered	Heavy tailed analysis (HTD), Long range dependence (LRD), autocorrelation analysis (AA), Power spectral density analysis (PSD)	Hurst parameter methods and Fractal dimension calculations (average values given)	Burstiness using index of dispersion for counts (IDC)	Probability Distribution
Our paper IPv4 and IPv6 mixed traffic	HTD: 0<α< 2 LRD:AA: Sum of autocorrelation function tends to be infinity, PSD: Follow power law near origin	R/S: 0.5 (IPv4), 0925 (IPv6) AM: 0.576(IPv4), 0.877(IPv6) VT: 0.577(IPv4), 0.835(IPv6) Diff. variance: 0.55(IPv4), 0.7226(IPv6) Boxed periodogram:0.55(IPv4), 0.61(IPv6) Box counting:1.68(IPv4), 1.43(IPv6)	IDC: 198 IPv4-IAT -100 sec, IDC: 4735 (IPv6- IAT-100 sec)	IPv4 IAT: Lognormal IPv4 packet size: Lognormal IPv6 IAT: Weibull IPv6 Packet size: Log-Logistic
Cebrail: [24] IPv4 and IPv6 mixed traffic	AA: Sum of autocorrelation function tends to be infinity, PSD: Follow power law near origin	R/S: 0.7(IPv4), 0.9 (IPv6), Diff. variance: 0.56(IPv4), 0.79(IPv6) Aggregated variance: 0.73(IPv4), 0.84(IPv6) Absolute value: 0.73(IPv4), 0.89(IPv6) Wavelet Method: 0.61(IPv4), 1(IPv6)	Not Done	Not Done
Jaiswal: [21] HTTP traffic	Not Done	R/S plots: 0.627 V-T plots: 0.62 Absolute Moment: 0.72 Periodogram: 0.68 Correlation integrals: 0.922	Not Done	IAT: Weibull

Ali Gezer: [19] IPv4 and IPv6 Bit Torrent traffic	AA: Sum of autocorrelation function tends to be infinity, PSD: Follow power law near origin	Absolute value: 0.788(IPv4), 0.76(IPv6) Aggregated variance: 0.71(IPv4), 0.77(IPv6) R/S: 0.78(IPv4), 0.75(IPv6) Wavelet method: 0.733(IPv4), 0.688(IPv6)	Not Done	IPv4 packet size: Log logistic, IPv4 IAT: Weibull IPv6 packet size: Pareto IPv6 IAT: Gamma
Crovella: [1] WWW traffic	HTD: $0<\alpha< 2$	V-T plots, R/S Plot, Periodogram, Whittle Estimator: H values are not mentioned but stated that $0.7 < H < 0.8$	Not done	Not Done
Zhu-wang: [25] Web traffic, wireless and wireline IP network	Not done	V-T plots: 0.66 (Wireless), 0.68 (Wireline) R/S plot: 0.66 (Wireless), 0.7 (Wireline)	Not done	Weibull for both Wireless (0.18, 0.67) and Wireline (0.03, 0.53)

4. CONCLUSION

In this paper, we examine characteristics of IPv6 and IPv4 traffic obtained from the standard internet traffic datasets. For analysis, we used packet level internet traces from MAWI working group traffic archive. We used internet traces captured in the month of May 2009 from IPv6 line connected to WIDE-6Bone. For IPv4 traffic, we analyze the daily traces of trance pacific line (18Mbps CAR on 100Mbps link) throughout month of April 2006. Detail analysis is carried out in terms of cdf analysis, self-similarity analysis, autocorrelation analysis, power spectral density analysis, PMR and IDC analysis.

Data communication of packets follows heavy tailed distributions in the internet network. Self-similarity can be used to describe this behavior. The processed data traffic shows heavy tailedness and thus follows power law. IPv6 traffic showed higher heavy tailed behavior compared to IPv4. Hurst parameter provides analytical proof of self-similarity using various techniques. Higher degree of heavy tailedness results in higher degree of self- similarity. This is evident from the H values obtained for IPv4 and IPv6 data traces. The H values for IPv4 are close to 0.6 and H values for IPv6 are close to 1 for both parameters like inter arrival time and packet length. Thus, it is evident that IPv6 traffic is more self-similar compared to IPv4.

Long range dependency was tested using auto correlation and power spectral density plots. The results obtained using this analysis, correctly matched with heavy tailed distribution and self-similarity analysis. Lower H values indicates lower long range dependence. This is reflected by auto correlation plots of IPv4 for both inter-arrival time and packet length. IPv6 packet traffic showed higher auto correlation values for greater lags and decayed slowly. Especially, for packet length parameter, auto correlation graphs showed very high auto co-relation values and decayed very slowly with respect to time. This clearly indicates higher self-similarity in case of IPv6 traffic. Burstiness is analyzed using IDC, PMR and plotting time series graphs for different time scales. IPv4 packet traffic shows less burstiness and it decays over higher time scales, whereas IPv6 traffic maintains bursty nature even for higher time scales.

With rigorous experimentations, it is assured that, IPv6 is burstier as it showed higher degree of heavy tailedness. Accordingly, increased self-similarity is noticed. Upon further investigation, it showed higher values of auto correlation function and results large power spectral density value near the origin. This all contributed to increased bursty nature of IPv6 traffic than IPv4. This analysis is very much important and can be used by internet service providers for network design and management to ensure smooth functioning of network flow for IPv4 and IPv6 traffic in future.

REFERENCES

[1] Mark E. Crovella and Azer Bestavros, "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes", *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, pp. 835-846, 1997.

[2] W. Willinger, M.S. Taqqu, R. Sherman and D.V. Wilson. "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level", IEEE/*ACM Transactions on Networking*, Vol. 5, No. 1, pp. 71-86, 2002.

[3] W.E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson, "On the Self-Similar Nature of Ethernet Traffic (extended version)", *IEEE/ACM Transactions on Networking*, Vol. 2, No. 1, pp. 1-15, 1994.

[4] Wei-Bo Gong, Yong Liu, Vishal Misra and Don Towsley, "Self-Similarity and Long Range Dependence on the Internet: A Second Look at the Evidence, Origins and Implications", *Computer Networks*, Vol. 48, No. 3, pp. 377-399, 2005.

[5] Kihong Park, Gitae Kim and Mark Crovella, "On the Relationship between File Sizes, Transport Protocols, and Self-Similar Network Traffic", *Proceedings of the 4th International Conference on Network Protocols*, pp. 171-180, 1996.

[6] Vern Paxson and Sally Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling", *IEEE/ACM Transactions on Networking*, Vol. 3, No. 3, pp. 226-244, 1995.

[7] Benoit B. Mandelbrot, "*The Fractal Geometry of Nature*", W.H. Freeman and Company, 1982.

[8] http://mawi.wide.ad.jp/mawi/.

[9] Internet Protocol, Version 6 (IPv6) Specification, 1995, Available at https://tools.ietf.org/html/rfc1883

[10] M. Goncalves and K. Niles, "*IPv6 Networks*", New York: McGraw-Hill, 1998.

[11] Internet Protocol Version 4 (IPv4) Specification, IETF RFC 791, 1981, Available at https://tools.ietf.org/html/rfc791

[12] Fayez Gebali, "*Analysis of Computer and Communication Networks*", Springer US, 2008.

[13] T.W. Anderson and D.A. Darling, "Asymptotic Theory of Certain 'Goodness-of-Fit' Criteria based on Stochastic Processes", *The Annals of Mathematical Statistics,* Vol. 23, No. 2, pp. 193-212, 1952.

[14] Chi squared test, Available at, http://mathworld.wolfram.com/Chi-SquaredTest.html.

[15] Balakrishnan Chandrasekaran, "Survey of Network Traffic Models", Student Project Report, Department of Computer Science and Engineering, Washington University, St. Louis, 2015.

[16] Tarek Sobh, Khaled Elleithy, Ausif Mahmood and Mohamed Karim, Eds., "*Innovative Algorithms and Techniques in Automation, Industrial Electronics and Telecommunications*", Springer, 2007.

[17] Thomas Karagiannis and Michalis Faloutsos, "SELFIS: A Tool for Self-Similarity and Long-Range Dependence Analysis", *Proceedings of 1st Workshop on Fractals and Self-Similarity in Data Mining: Issues and Approaches*, 2002.

[18] F. Cervantes-De la Torre, J.I. González-Trejo, C.A. Real-Ramírez and L.F. Hoyos-Reyes, "Fractal Dimension Algorithms and their Application to Time Series Associated with Natural Phenomena", *Journal of Physics: Conference Series*, Vol. 475, No. 1, pp. 1-10, 2013.

[19] Cebrail Çiflikli, Ali Gezer, A. Tuncay Özsahin, and Öznur Özkasap, "BitTorrent Packet Traffic Features over IPv6 and IPv4, Simulation Modelling Practice and Theory", *Simulation Modelling Practice and Theory*, Vol. 18, No. 9, pp. 1214-1224, 2010.

[20] Kuan-Ta Chen, Polly Huang and Chin-Laung Lei, "Game Traffic Analysis: An MMORPG Perspective", *Computer Networks*, Vol. 50, No. 16, pp. 3002-3023, 2006.

[21] C. Jaiswal Rupesh and D. Lokhande Shashikant, "Measurement, Modelling and Analysis of HTTP Web Traffic", *International Conference on Communication and Computing*, 2014.

[22] http://www.census.gov/main/www/popclock.html

[23] http://blogs.cisco.com/news/the-internet-of-things-infographic/

[24] Cebrail Çiflikli, Ali Gezer and A. Tuncay Özsahin, "Packet Traffic Features of IPv6 and IPv4 Protocol Traffic", *Turkish Journal of Electrical Engineering & Computer Sciences*, Vol. 20, No. 5, pp. 727-749, 2012.

[25] Chunmei Zhu, Yu Wang, Yan Zhang and Weiling Wu, "Different Behavioral Characteristics of Web Traffic between Wireless and Wire IP Network", *Proceedings of International Conference on Communication Technology*, Vol. 1, pp. 267-271, 2003.