Turk J Elec Eng & Comp Sci, Vol.20, No.5, 2012, © TÜBİTAK
doi:10.3906/elk-1008-696

Packet traffic features of IPv6 and IPv4 protocol traffic

Cebrail ÇİFLİKLİ, Ali GEZER*, Abdullah Tuncay ÖZŞAHİN
Kayseri Vocational College, Erciyes University, Kayseri-TURKEY
e-mails: cebrailc@erciyes.edu.tr, aligezerr@hotmail.com, atozsahin@erciyes.edu.tr

Received: 25.08.2010

Abstract

Nowadays, the IPv6 protocol is in a transition phase in operational networks. The ratio of its traffic volume is increasing day by day. The many provided facilities for IPv6 connection increase the total IPv6 traffic load. IPv6-over-IPv4 tunnels, pilot programs to provide IPv6 connections, IPv6/IPv4 dual stack operating systems, and free IPv6 tunnel brokers cause the IPv6 protocol to expand quickly. For efficient resource utilization, the characteristics of network traffic should be determined accurately. Many traffic characterization studies regarding IPv4 have demonstrated that most of the network traffic is self-similar. Self-similarity causes significant impacts on network performance. With the increasing volume of IPv6 traffic, the characteristics of IPv6 traffic and differences between IPv4 traffic in terms of characterization should be explicitly revealed.

In this study, we investigate the characteristics of IPv6 packet traffic and the differences between IPv6 and IPv4 packet traffic in terms of spectral density, autocorrelation, distribution, and self-similarity of packet interarrival time and packet size. The results obviously show that IPv6 traffic exhibits totally different properties in comparison to that of IPv4. Distribution fittings prove that packet interarrival time and packet size have different distributions in the 2 traffic types. While beta distribution models the empirical cumulative distribution of IPv4 packet size, log-logistic distribution gives more efficient results for IPv6 packet size. Furthermore, a significant difference is observed in self-similarity degrees. IPv6 protocol traffic gives greater self-similarity than that of IPv4. Results show that IPv6 traffic would cause greater performance degradations in computer networks in comparison to IPv4 due to high self-similarity.

Key Words: *IPv6/IPv4 protocol, self-similarity, distributions, traffic analysis*

1. Introduction

Due to the limitations of IPv4, it was inevitable that a new internet protocol would be designed. IPv6 was considered instead of IPv4 in 1996 to provide solutions to a number of shortcomings. The new protocol has the advantage of many more addressable nodes on the Internet, and it extends 32-bit address sizes to 128 bits. IPv6 provides a more simple and extensible header structure. Some IPv4 header fields are dropped or made optional to reduce the processing cost of packet handling and to limit the bandwidth cost [1]. Flow labeling is added to address the packets belonging to a particular traffic flow.

*Corresponding author: Kayseri Vocational College, Erciyes University, Kayseri-TURKEY

Although there are many provided facilities, the widespread adoption of IPv6 protocol has been delayed for some reasons. It also seems difficult to completely abandon IPv4 in a short period of time. Limited use of IPv6 protocol in operational networks has kept the traffic load very low in comparison to IPv4 traffic. Until now, IPv6-relevant studies have been limited to protocol design, connections, routing mechanisms, and transition mechanisms [2-8]. Only a few studies have dealt with traffic measurements. Those studies mostly concentrated on packet loss rate, round-trip time differences, anomaly IPv6 traffic detection, mobile IPv6 handover analysis, and delay and jitter for specific applications [8,9]. A new IPv6-based measurement technique to assess the performance experienced by real user traffic was introduced in [10]. In [11], system-level application traffic was observed on a globally used 6to4 relay router to gather information about IPv6 deployment. An IPv6 traffic measurement study was carried out to evaluate the performance of real IPv6 traffic, especially from the perspective of end users. In the study, a total of 585,560 packet-level traces collected from 936 IPv4/IPv6 dual-stack Web servers located in 44 countries were analyzed to compare IPv4 and IPv6 in terms of connectivity, packet loss rate, and round-trip time [8]. However, none of these studies were concerned with the traffic characteristics of IPv6 in terms of self-similarity, distributions, and modeling.

IPv4-related traffic studies have generally proven that most of the IPv4 network traffic is self-similar [12-17]. Self-similar network traffic causes serious performance degradations in network performance. If a self-similar structure is not taken into consideration in network studies, efficient network management would be impossible with wrong traffic models [16,18].

Nowadays, the IPv6 traffic load is increasing in operational networks with provided facilities such as tunneling, increasing the number of operating systems that support the IPv4/IPv6 dual stack structure. Many internet service providers use IPv6 pilot programs to provide IPv6 connectivity for their customers. Free IPv6 tunnel brokers make it possible for someone to get IPv6 connection [12]. Increasing facilities to provide an IPv6 internet connection make its total traffic load much greater in comparison to that of the recent past. Thus, it is necessary to analyze IPv6 traffic and mixed traffic in terms of characterization. How they are different from IPv4 traffic and the effects of this differentiation on network performance should be addressed, especially if there is significant difference between the 2 traffic types.

We focused on the characterization differences of the 2 protocol traffics in our recent studies. In [19,20], BitTorrent packet traffic features for IPv6 and IPv4 were examined. BitTorrent is a popular P2P application that is increasing its traffic rate on the Internet. Particularly, we focused on BitTorrent traffic to observe its characterization features for IPv6 and IPv4. In [21], a basic packet traffic characterization study was carried out for IPv6 and IPv4 protocol traffic. Only cumulative distribution, power spectral density, and autocorrelation of packet size and packet interarrival time were taken into account to compare the 2 traffic types. The observed differences inspired us to carry out a more detailed characterization study for IPv6 and IPv4 protocol traffic. For this reason, we propose a characterization study of IPv6 packet traffic in terms of self-similarity, long-range dependency, and spectral density for packet interarrival time and packet size. Detailed analyses are performed to extract statistical differences between IPv6 and IPv4 protocol traffic. We also design 2 scenarios to elaborately examine both protocol traffics. First, we analyze aggregated incoming traffic at different time scales to examine its burstiness. Second, interarrival time points with different amounts of traffic are analyzed to investigate traffic density. To model the packet size and packet interarrival time of IPv4 and IPv6 packets, distribution fitting is also carried out in this study.

This paper is organized as follows. Some background information about IPv6 protocol is given in the next section. Measurement details of inspected internet traces are given in Section 3. Packet interarrival time and packet size comparisons of IPv6 and IPv4 traffic in terms of autocorrelation, power spectral density, and cumulative distribution are given in Section 4. Two scenarios for detailed analysis of IPv6 and IPv4 protocol traffic are then given. Finally, we perform a self-similarity analysis of the obtained time series and distribution fitting, and we conclude the paper.

2. Basics of IPv6

IPv6 is the new version of internet protocol. It was designed by the Internet Engineering Task Force in 1996 [22-24]. IPv6 provides a number of enhancements over the existing internet protocol [25]. The most challenging problem of the old protocol was its limited addressing capacity. IPv4 protocol could only address 2^{32} nodes on the Internet due to its 32-bit addressing structure. IPv6 increases the address length from 32 bits to 128 bits. The anycast address type was introduced with the new protocol to send a packet to any one of a group of nodes. The IPv4 header varies between 20 octets and 60 octets. It comprises a mandatory 20 octets and could optionally increase by 40 octets for option and padding fields. While the IPv4 header consists of 13 fields, there are only 8 fields in the IPv6 header. There is no packet fragmentation and reassembly field in the IPv6 packet header. These tasks are carried out only by source and destination nodes. Therefore, packet processing speeds increase in intermediate routers. Furthermore, header format simplification limits the bandwidth cost of the IPv6 header [19].

3. Internet traces

We analyzed packet-level internet traces captured on an IPv6 line connected to WIDE-6Bone in Japan. The traces were taken from the MAWI Working Group traffic archive [26]. For analysis of IPv6 protocol traffic, all daily traces of May 2008 and 10 days of randomly selected traces from 2008 were used. Each day, capturing started at 1800 hours and continued until 2 million packets had been captured. It takes nearly 2 h to capture 2 million packets. Analyses were carried out on a daily basis due to the daily capturing. The captured traffic was logged in tcpdump format. All daily traces comprised different kinds of application traffic, including P2P, http, ftp, and smtp. The protocol breakdowns of May 2008 for IPv6 and of 15 consecutive days for IPv4 are given in Table 1.

A packet sniffer application was used to analyze packet-level traffic traces and extract necessary information from traces. In total, over 70 million IPv6 packets were analyzed. To extract differences between IPv4 and IPv6, we also analyzed traffic characteristics of real IPv4 traffic. For IPv4 traffic, we analyzed a consecutive 15 days (from 1 August to 15 August 2008) and 10 randomly chosen days of traffic from 2008 captured on a 150-Mbps transpacific link between the United States and Japan. IPv4 traces were also taken from the MAWI Working Group traffic archive [26]. Due to the abundance of packet counts in IPv4 traffic, the last 2 million packets were analyzed in the daily traces. Analyses were carried out with respect to packet size and packet interarrival time.

4. Analysis details

Network simulations should be based on true traffic models to give efficient performance estimation results. Recent traffic analyses obviously prove that most of the network traffic types exhibit self-similarity. Spectral density, autocorrelation, and cumulative distribution give significant details about the self-similar nature of the traffic. Comparative power spectral density and autocorrelation analysis were carried out for IPv6 and IPv4 traffic at various time scales. We also determined distributions for packet interarrival time and packet size. To simulate self-similar traffic, the distribution functions that represent heavy-tailed behavior were used. Empirical cumulative distribution plots were used to compare the 2 protocol traffic types. Packet interarrival time and packet size were modeled with known distribution functions.

Table 1. Packet breakdown of IPv6 and IPv4 packet traffic.

Protocol packets (%)	IPv6	IPv4
Ip	100	99.55
Tcp	74.89	80.57
http(s)	17.26	37.73
http(c)	36.3	19.86
squid	0.043	1.30
smtp	0.042	3.46
ftp	0.1	0.98
pop3	0.08	0.16
ssh	0.255	1.55
other	18.87	12.96
udp	20.21	13.00
dns	18.785	6.22
icmp	4.35	4.88

4.1. Autocorrelation analysis

Long-range dependent packet traffic causes serious performance degradations in computer networks. Network packet traffic keeps its correlated structure during long time lags. That is, autocorrelations of packets do not decay exponentially. For long-range dependent time series, the autocorrelations follow:

$$r_x(k) \approx |k|^{-\beta} \; as \; k \to \infty, \; 0 < \beta < 1, \tag{1}$$

where k represents the time lag and β is closely related with the Hurst parameter. Another important property of autocorrelations for long-range dependent processes is nonsummability [19]. That is:

$$\sum\nolimits_k r_x(k) = \infty. \tag{2}$$

According to the obtained autocorrelation values, IPv6 traffic exhibits greater correlations than IPv4 in terms of packet size and interarrival time. The correlation values of IPv4 packet size and interarrival time decrease to low levels in a very short span of time. However, corresponding results for IPv6 packets show much more long-range dependency than those of their IPv4 counterparts. The plots of 1 representative day for IPv6 and IPv4 are given in Figures 1 and 2. Similar autocorrelation plots were obtained for other days.

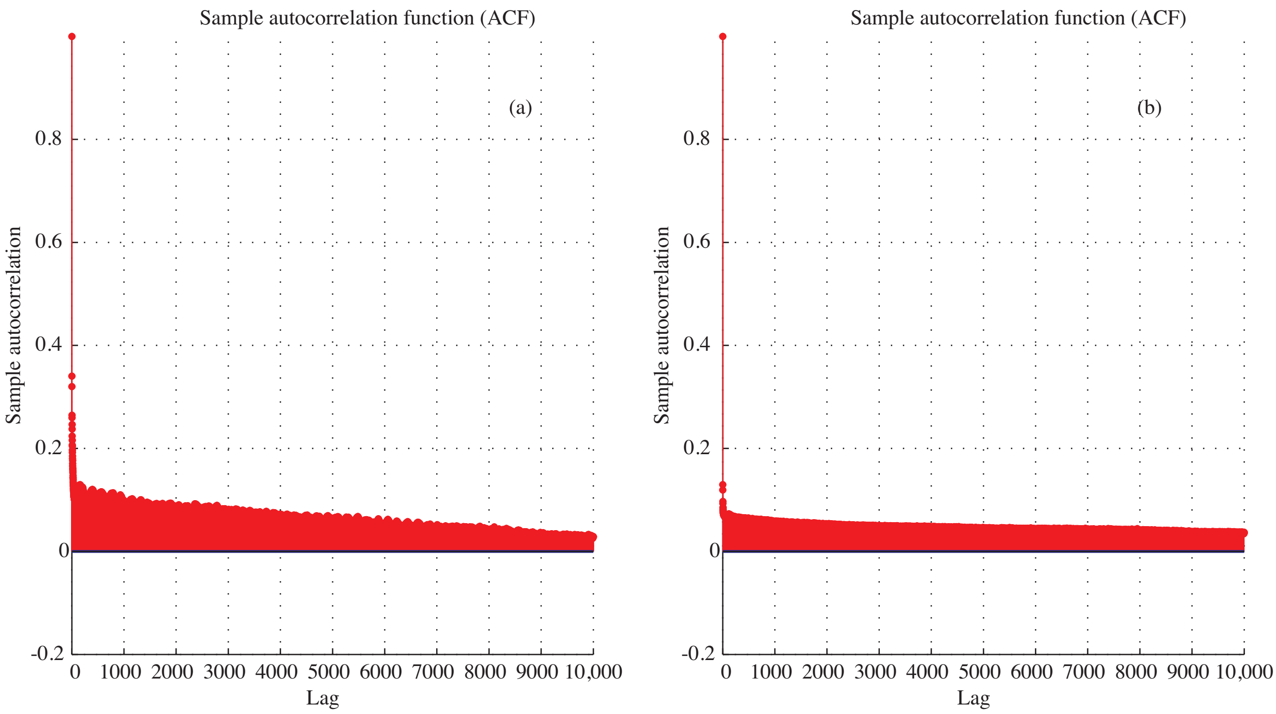

Figure 1. Autocorrelation plots of IPv6 traffic on 3 May 2008: a) packet size, b) interarrival time.

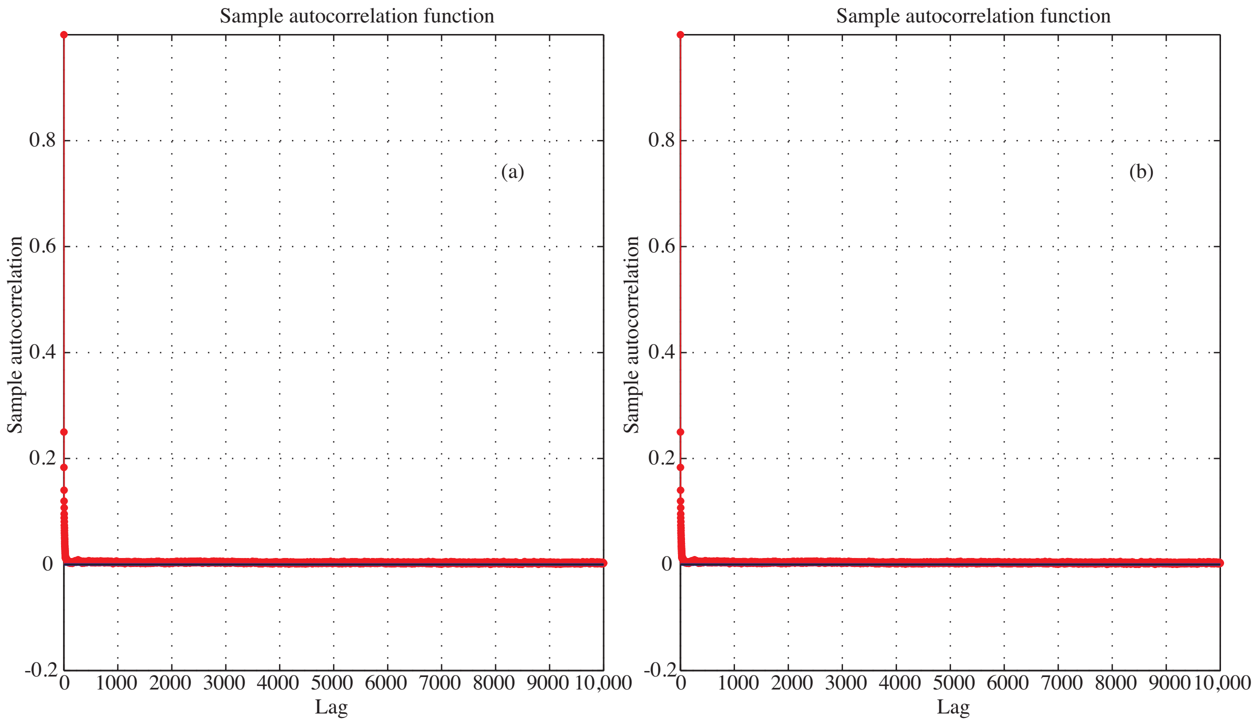

Figure 2. Autocorrelation plots of IPv4 traffic on 13 August 2008: a) packet size, b) interarrival time.

4.2. Cumulative distribution analysis

The modeling of network traffic via a known distribution function is a significant step in understanding the nature of traffic. If a time series could be defined with a known distribution function, similar time series could be generated easily. Therefore, we observed cumulative distribution differences for IPv6 and IPv4 in terms of packet size and interarrival time. Network traffic could generally be modeled with heavy-tailed distribution functions. That is, very large values also had significant probabilities. If a time series is heavy-tailed, it follows:

$$P[X > x \approx x^{-\alpha}] \text{ as } x \rightarrow \infty, \quad 0 < \alpha < 2, \tag{3}$$

where α is related to the Hurst parameter [19]. The cumulative distribution function of a time series is:

$$F_x(x) = P(x \leq x). \tag{4}$$

Plotted cumulative distribution functions (CDFs) of interarrival time and packet size demonstrated significant differences between IPv4 and IPv6 protocol traffic. CDF plots of IPv4 packet size showed that nearly half of the packets were larger than 1200 bytes. The packets that were smaller than 200 bytes and larger than 1200 bytes formed most of the traffic load. However, the packet-size CDF of IPv6 traffic was different. Nearly 80% of the packets were smaller than 200 bytes. The obtained packet-size CDF plots for both protocols are shown in Figure 3. Similar results were obtained from traces from other days. The mentioned differences between the CDFs of packet size of both protocol traffics obviously show that distribution fitting should be done with different functions.

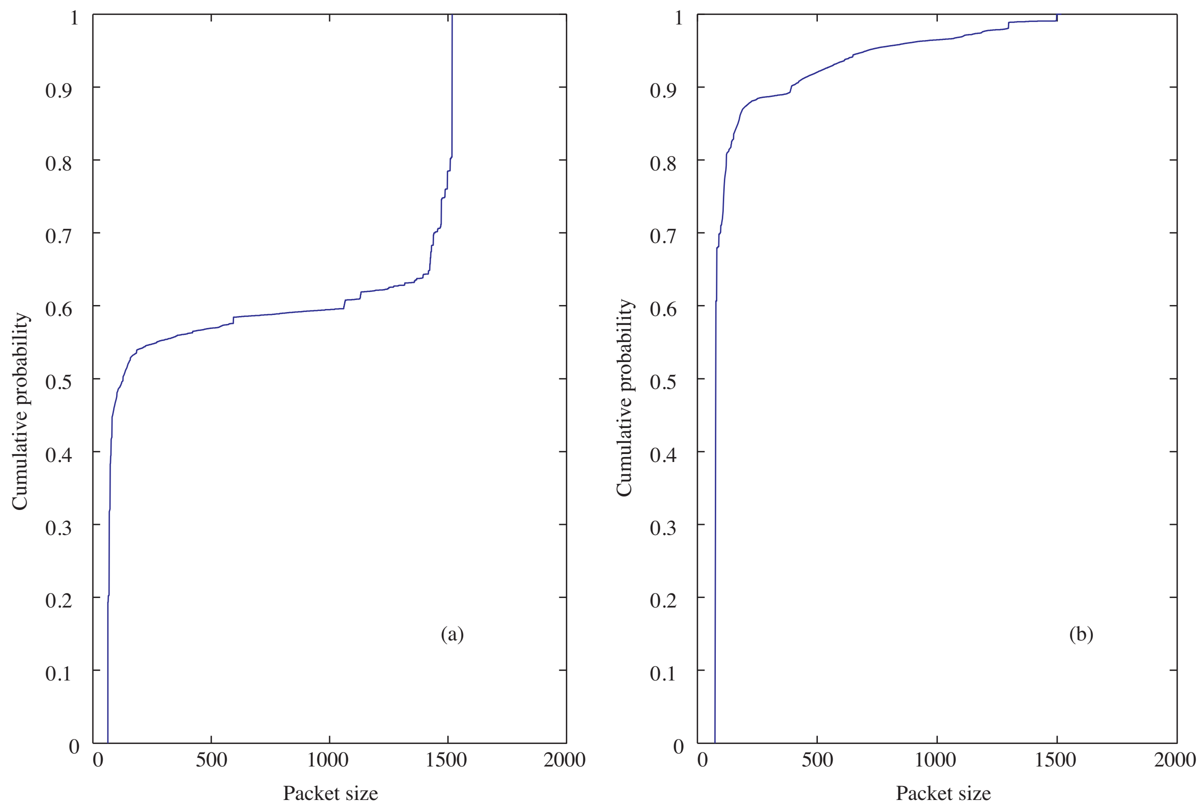

Figure 3. Packet-size CDF of both protocol traffics: a) 11 August 2008, IPv4; b) 2 May 2008, Wide-6Bone IPv6.

There was also an obvious difference in terms of packet interarrival times between IPv4 and IPv6 traffic. The relevant probability density functions of IPv4 represented 3 different peaks. This result could be observed in CDF plots of IPv4 packet interarrival times. Figure 4 demonstrates that the distribution fittings of IPv4 and IPv6 packet interarrival times should also be done with different functions.

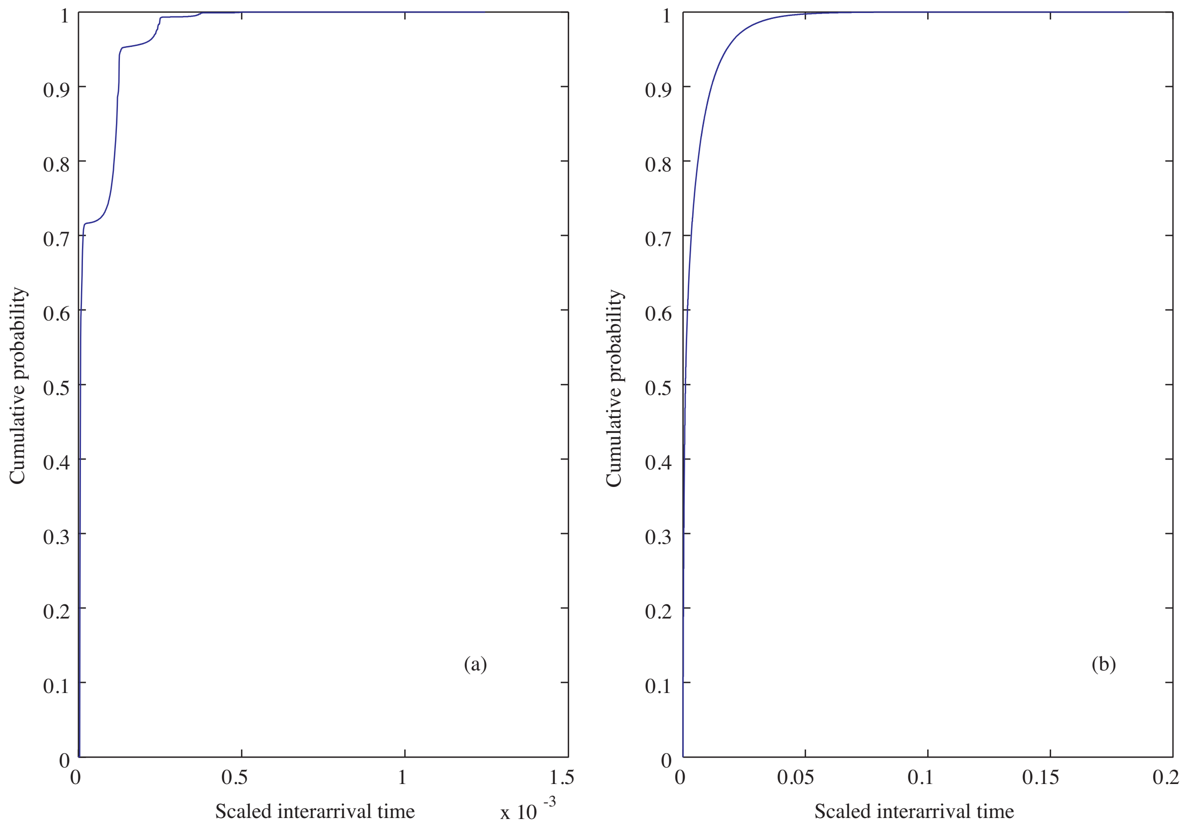

Figure 4. Packet-size CDF of both protocol traffics: a) 11 August 2008, IPv4; b) 2 May 2008, Wide-6Bone IPv6.

4.3. Power spectral density analysis

Another significant feature in evaluating self-similarity in a time series is its spectral density. If a time series is long-range dependent, its spectral density follows a power law near the origin.

$$\gamma = H - 1 \tag{5}$$

$$s_x(w) \approx |w|^{-\gamma} \text{ as } w \to \infty, \;\; 0 < \gamma < 1 \tag{6}$$

Here, w is the frequency, $s_x(w)$ is the spectral density, and H is the Hurst exponent. The power spectral density of the time series is calculated as:

$$G_m(f) = F[r_m(k)] \tag{7}$$

where $F[r_m(k)]$ indicates the Fourier transform of the autocorrelation function [14].

Power spectral density (PSD) plots of both IPv4 and IPv6 represent similar characteristics and both of them show 1/f-type spectrum behavior. However, the self-similarity of the IPv6 packet interarrival time and packet size appear to be a bit greater than those of IPv4. Obtained plots are shown in Figures 5 and 6.

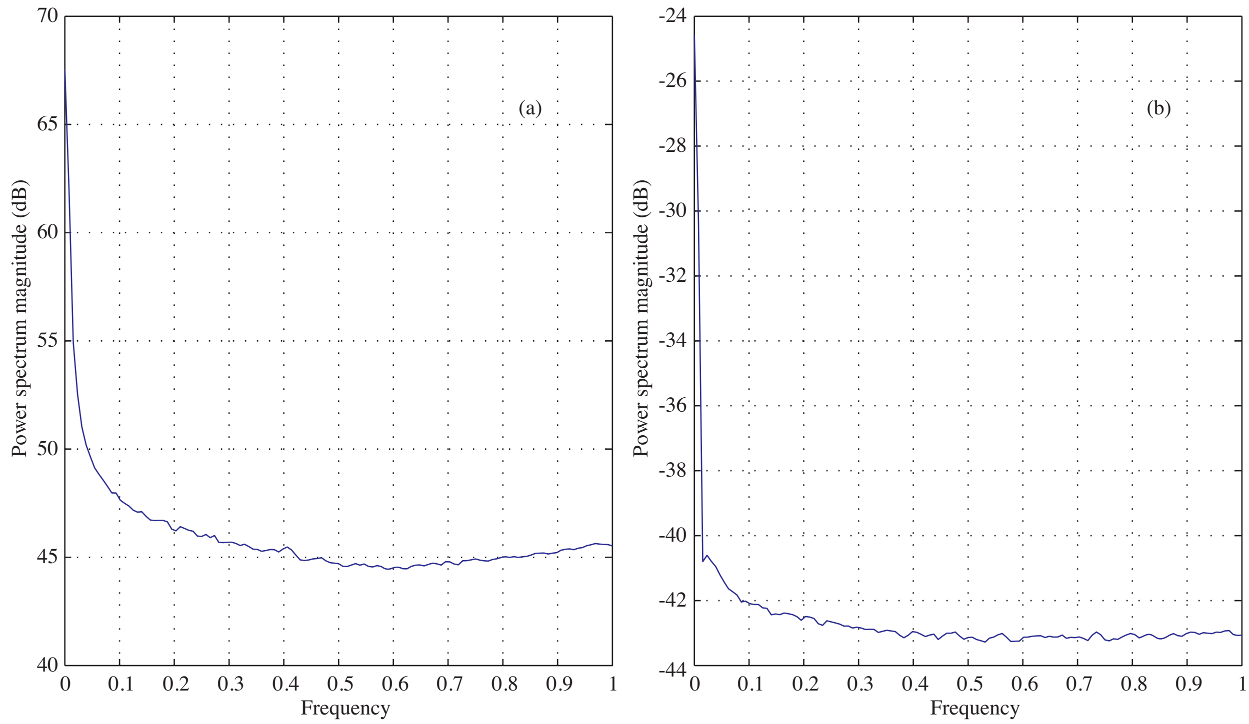

Figure 5. PSD of Wide-6Bone IPv6 traffic of 2 May 2008: a) packet size, b) interarrival time.

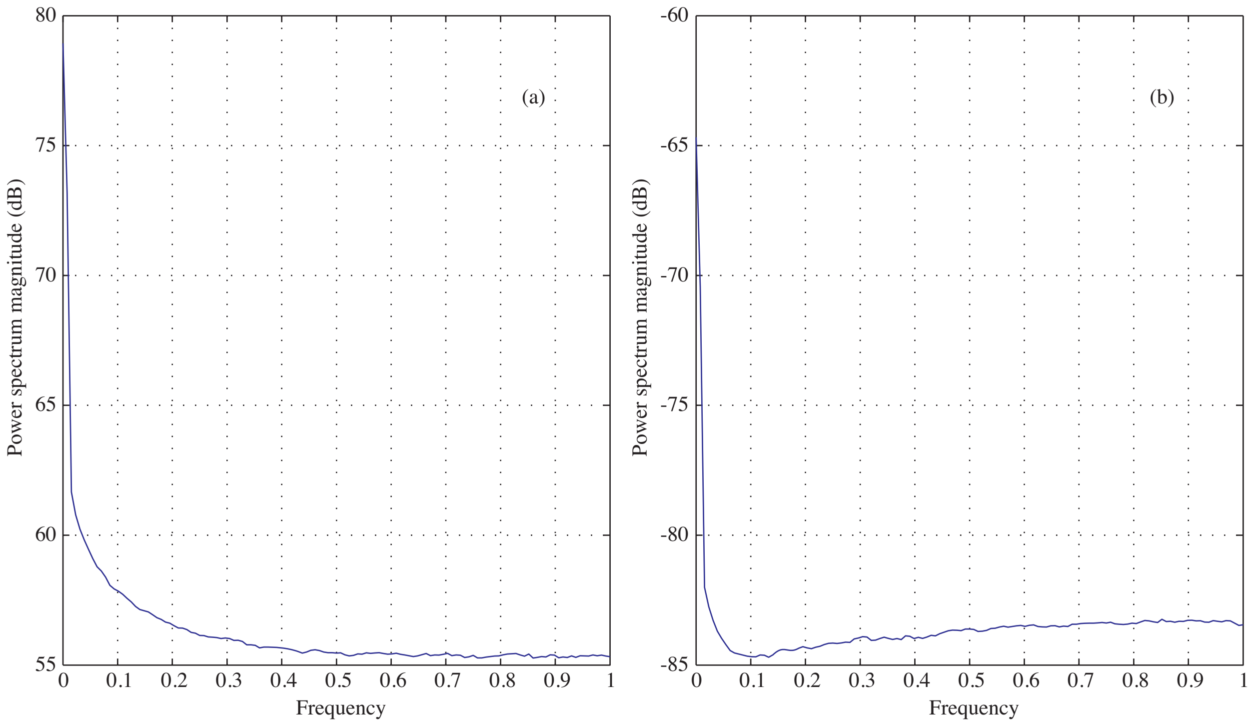

Figure 6. PSD of IPv4 traffic of 11 August 2008: a) packet size, b) interarrival time.

5. Aggregated traffic analysis

Whether there is self-similarity or not could be examined with aggregated traffic at different time scales. While the burstiness of traffic keeps itself even with larger time scales, when the time scale increases up to 10,000 times, the traffic exhibits a bursty and hence self-similar nature. To make the self-similarity differences between the 2 traffic types observable, we performed aggregated traffic analyses at different time scales. IPv4 and IPv6 aggregated traffic were analyzed at different time scales and probability density, autocorrelation, and power spectral density changes were observed. Due to the similar characteristics of each day plot, only graphics belonging to 1 day are given for each time scale.

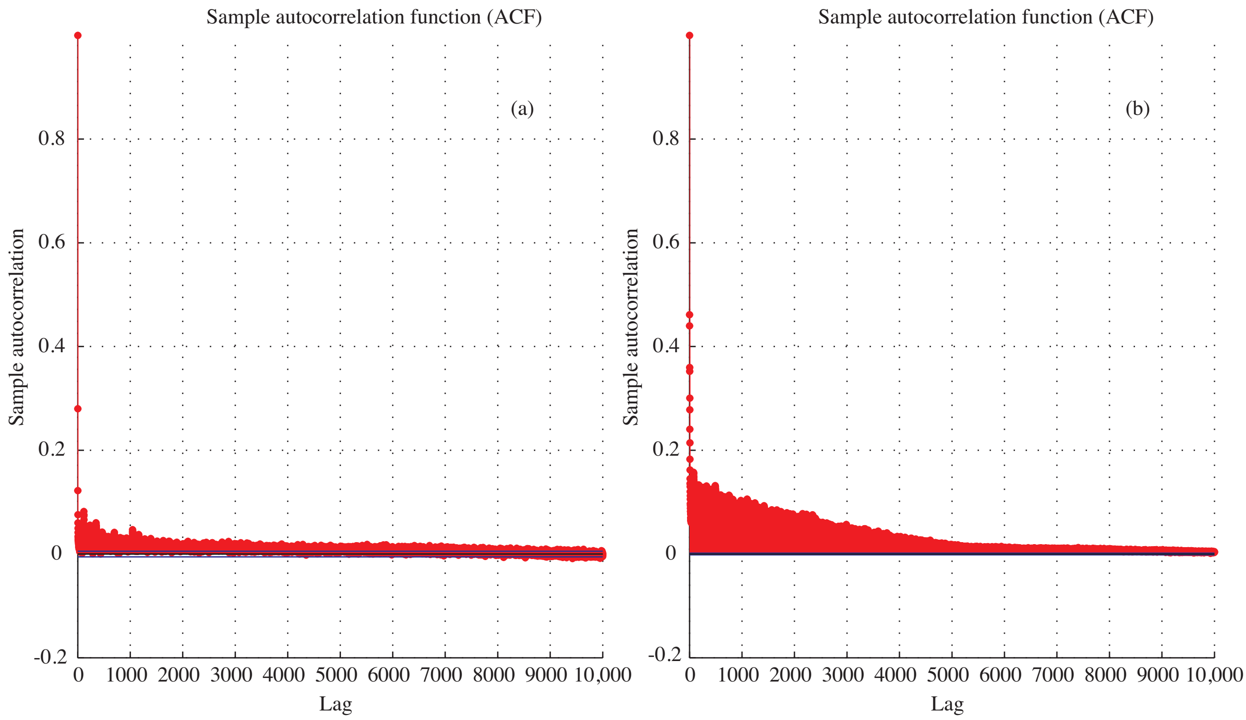

Figure 7. Autocorrelation plot at 0.001 time scale: a) IPv4 traffic of 16 August 2008, b) Wide-6Bone IPv6 traffic of 2 May 2008.

First, we analyzed incoming traffic with a time window of 0.001 s, and each time scale was increased by 10. Results of the 0.001 time scale show different CDF characteristics for IPv6 and IPv4 aggregated traffic, as shown in Figure 7. The CDF of IPv4 traffic looks more like a Gaussian distribution, but its IPv6 counterpart is different. Autocorrelation plots exhibit stronger differences for both protocols. Autocorrelations of IPv6 traffic give greater values than IPv4 at the 0.001 time scale and do not decay during long lags. There is also a significant difference in terms of the power spectral densities of the 2 protocol traffics at the mentioned time scale. IPv6 aggregate traffic at the 0.001 time scale exhibits a stronger 1/f-type power spectrum than IPv4. PSD plots at the 0.001 scale are shown in Figure 8.

Distribution plots are given in Figure 9. IPv4 aggregate traffic has a Gaussian CDF.

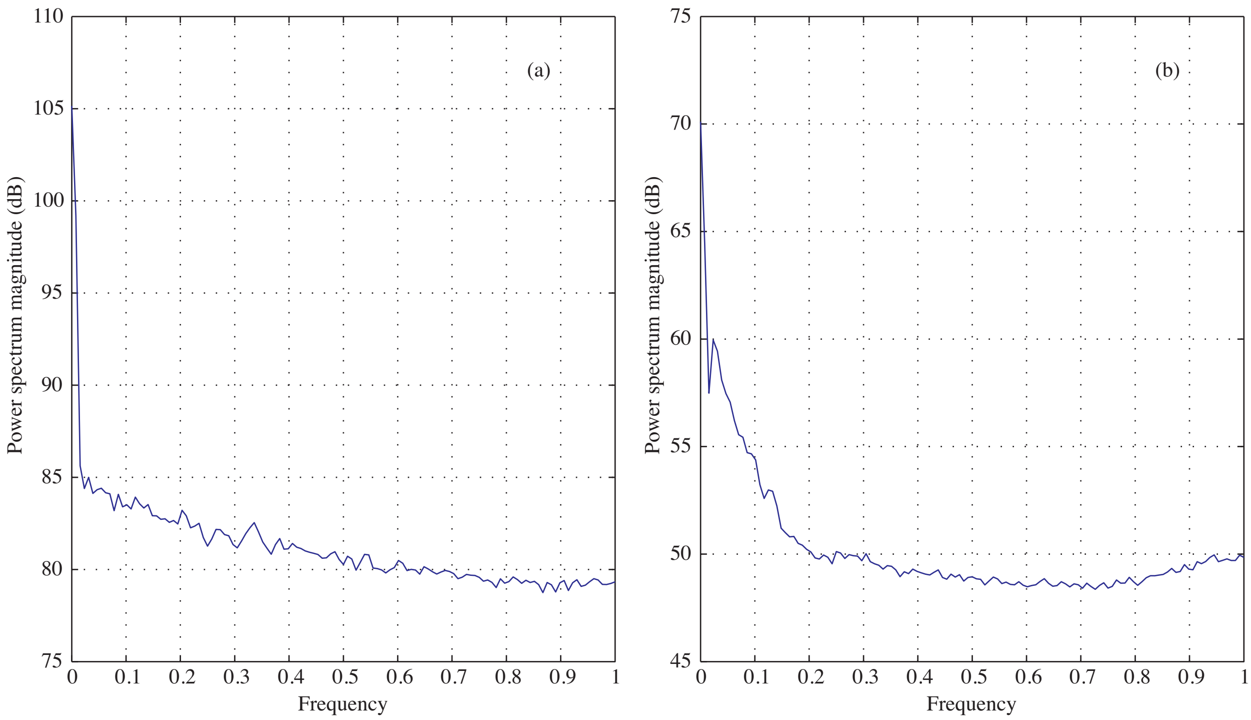

Figure 8. PSD at 0.001 time scale: a) 150 IPv4 traffic of 16 August 2008, b) Wide-6Bone IPv6 traffic of 13 May 2008.

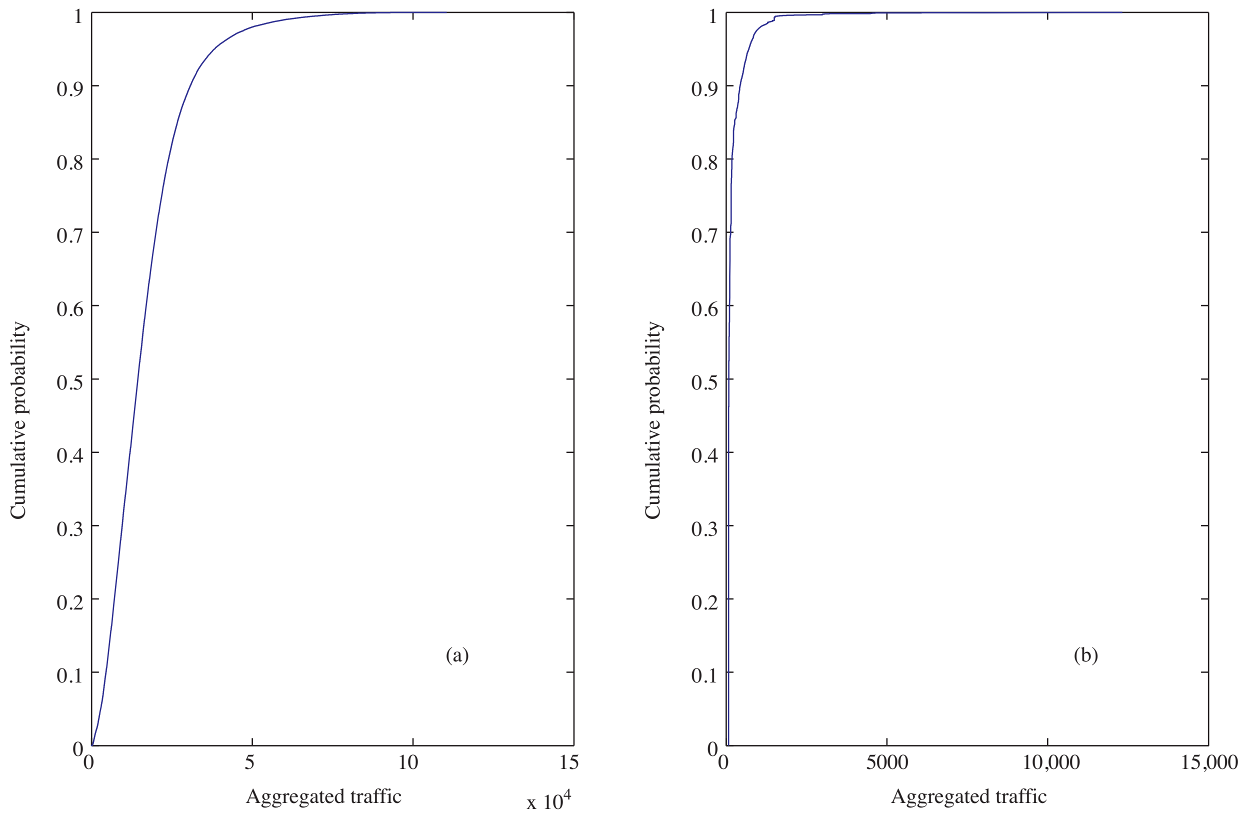

Figure 9. CDF: a) IPv4 traffic, b) IPv6 traffic.

Increasing the time scale by 10 had different effects on each traffic type. Autocorrelations regarding the 0.01-s time scale are shown in Figure 10.

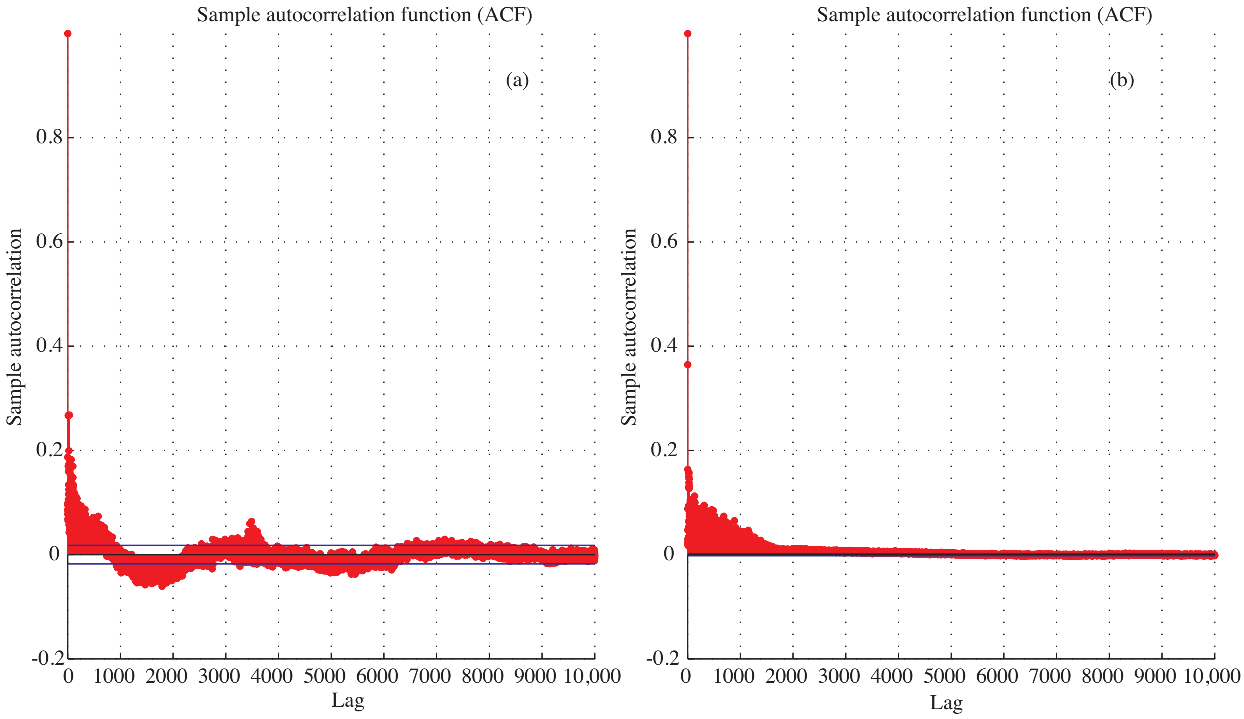

Figure 10. Autocorrelation functions at 0.01 time scale: a) IPv4 traffic, b) IPv6 traffic.

While the autocorrelations of IPv4 aggregate traffic tended to increase, the IPv6 traffic decreased from a scale of 0.001 to 0.01. The IPv6 autocorrelation gives quite high values along lag 4000 on the 0.001 scale. This lag span drops to 2000 at the 0.01 time scale. Increasing the time scale causes a reverse effect for IPv4 aggregate traffic autocorrelations. A remarkable increase was observed in autocorrelation values of IPv4 traffic in comparison to previous scales.

PSD variations at the 0.01 time scale are quite similar for IPv6 and IPv4. The 1/f-type behavior of IPv6 and IPv4 traffic seems to increase. However, the self-similarity degrees of IPv6 traffic seem more powerful than those of IPv4, as shown in Figure 11.

The most interesting results are obtained at the 0.1-s time scale, especially in terms of PSD. For all day traces, very powerful high-frequency components are observed for aggregated IPv6 traffic at 0.1 s. Only plots for 1 day are given in Figure 12. This significant difference is not observed at 0.01 and 1 s, but it repeats itself at the 0.1-s scale for every IPv6 trace. It is not observed at the 0.1 scale for IPv4 aggregated traffic, but the PSD plot of IPv4 traffic at 0.1 shows more 1/f-type spectrum behavior than at the previous scale. Autocorrelation values for IPv6 are also quite high at the 0.1 time scale. Autocorrelation plots for the 0.1 scale are shown in Figure 13.

For IPv4 traffic at the 0.1 time scale, the PSD changes significantly in comparison to previous scales. It obviously represents 1/f-type power spectrum behavior and the correlation values seem to increase, but there are also inverse correlation values.

Lastly, we performed analysis at the 1-s time scale. The obtained graphics show that the autocorrelation values for IPv6 are as good as at the previous scale. This gives an impression about the high self-similarity of aggregated traffic. The power spectral density exhibits a certain 1/f-type spectrum characteristic. This result is also valid for IPv4 traffic on this scale. Autocorrelation values seem quite high, as with the previous scales. Related plots for the 1-s scale are shown in Figures 14 and 15.

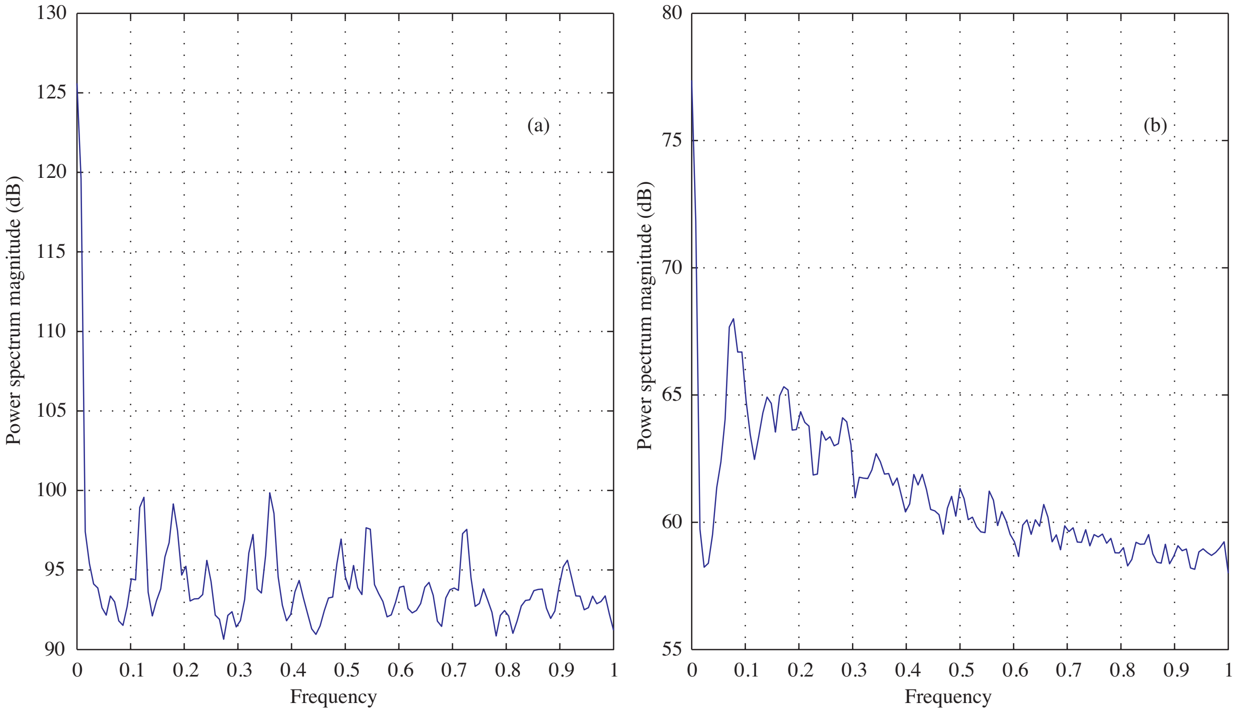

Figure 11. Power spectral density at 0.01 time scale: a) IPv4 traffic, b) IPv6 traffic.

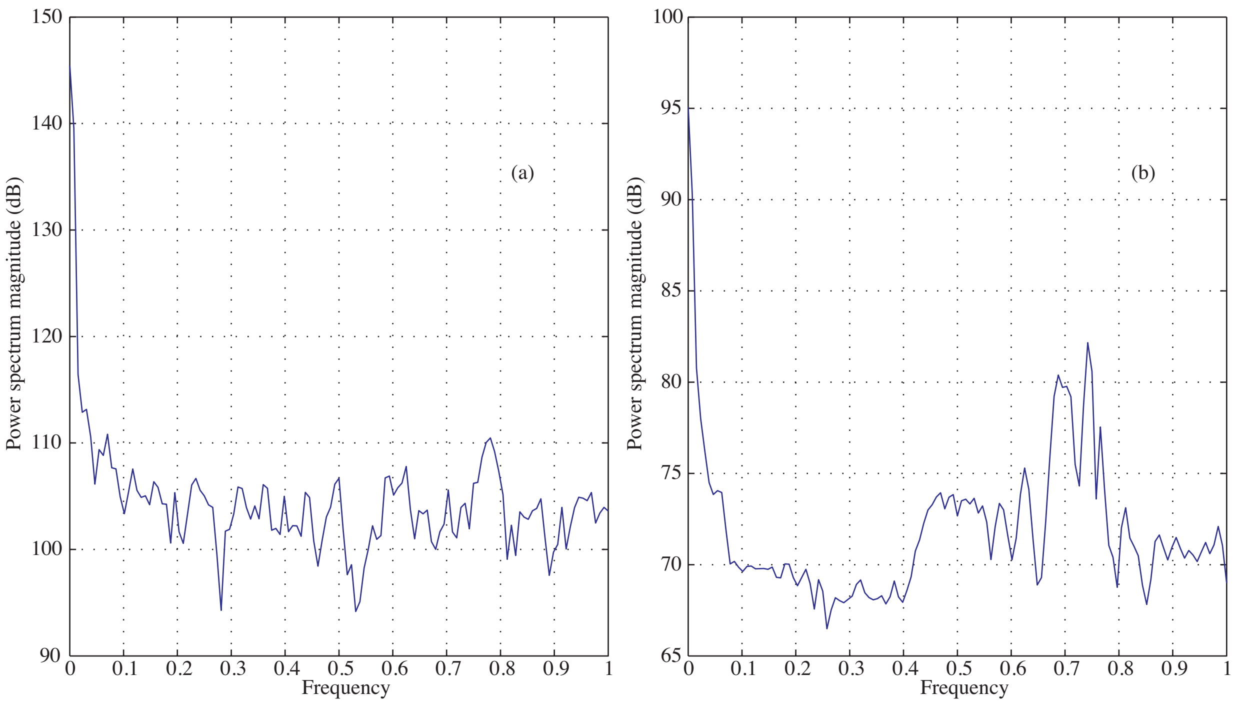

Figure 12. Power spectral density at 0.1 time scale: a) IPv4 traffic, b) IPv6 traffic.

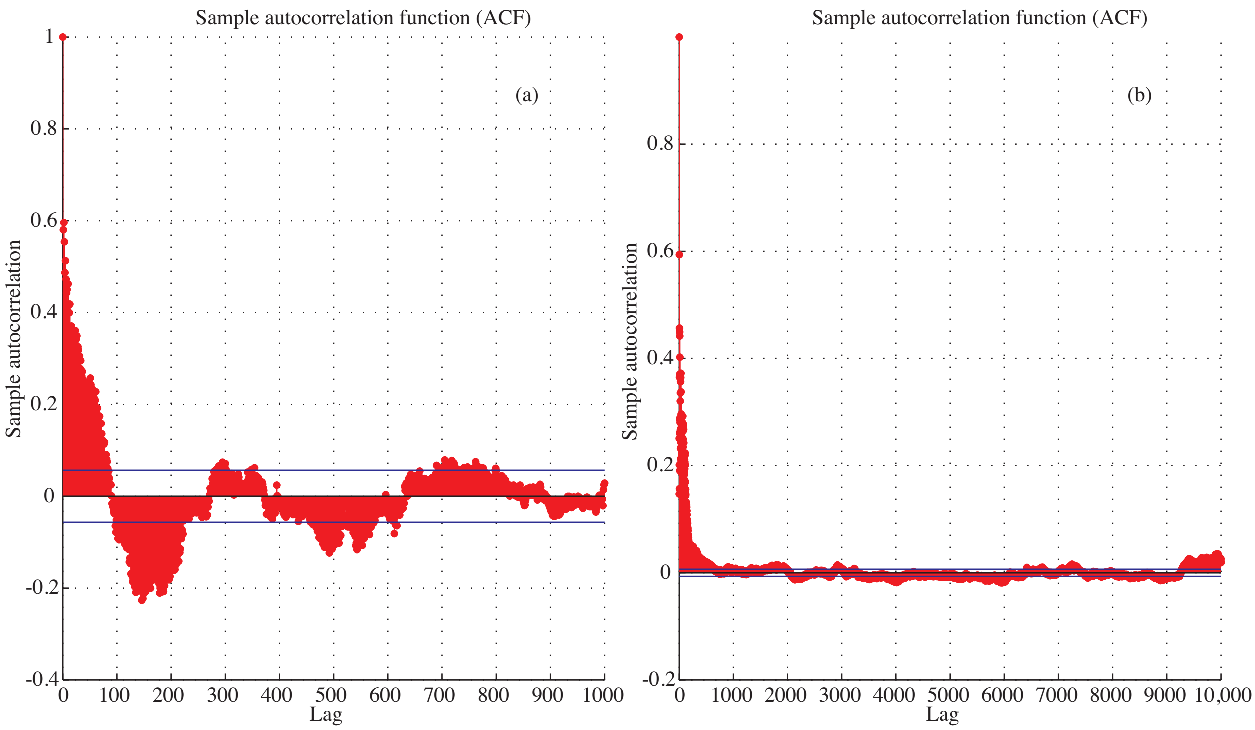

Figure 13. Autocorrelation functions at 0.1 time scale: a) IPv4 traffic, b) IPv6 traffic.

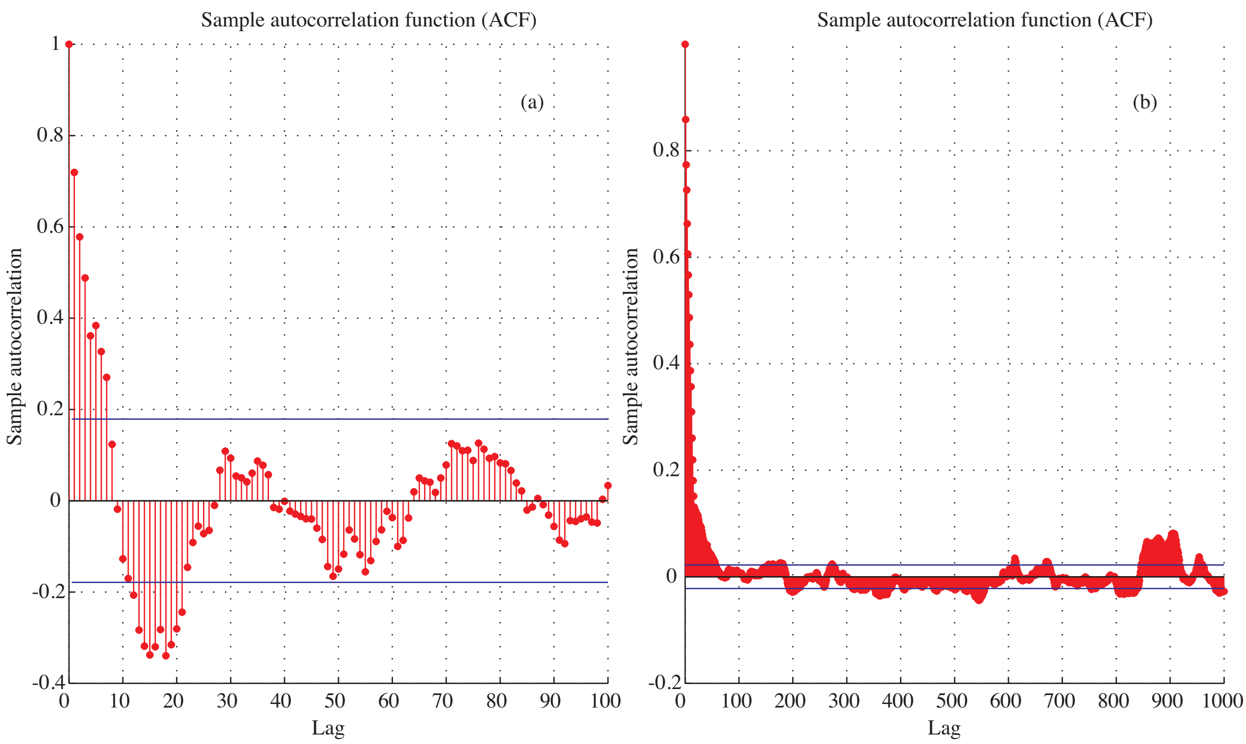

Figure 14. Autocorrelation functions at 1-s time scale: a) IPv4 traffic, b) IPv6 traffic.

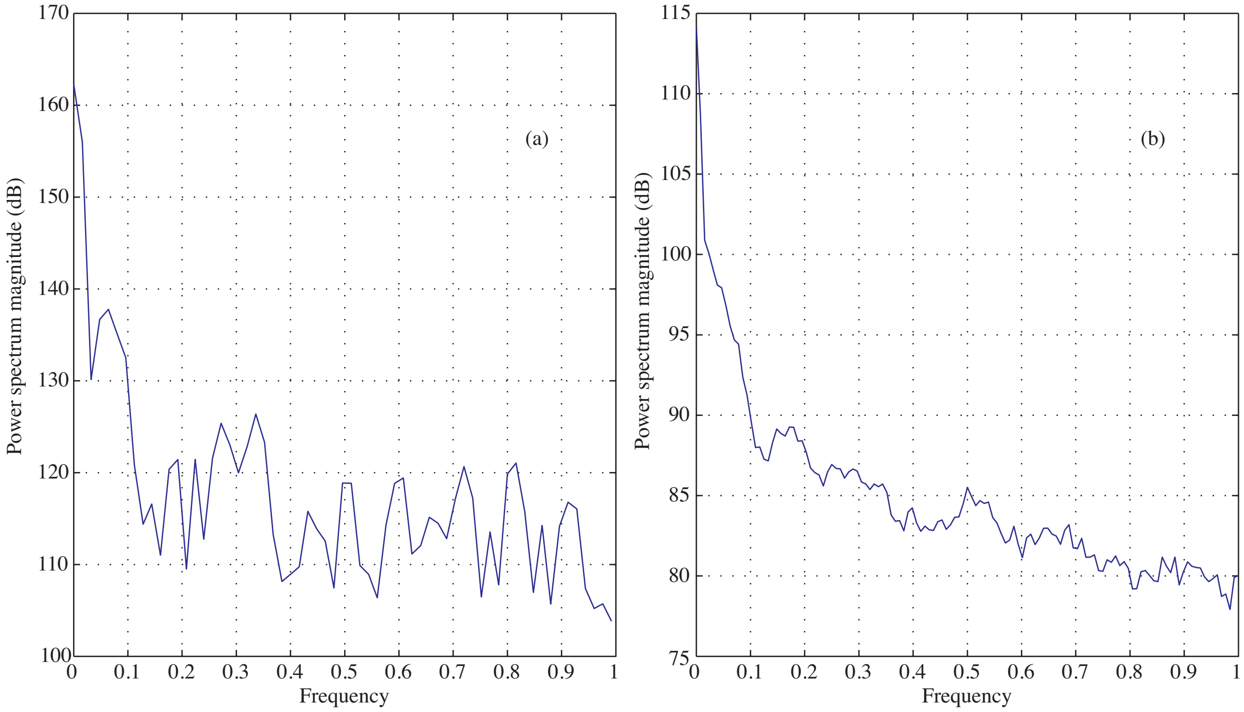

Figure 15. PSD at 1-s time scale: a) IPv4 traffic, b) IPv6 traffic.

6. Interarrival time analysis for aggregated traffic

In this particular analysis, we focused on how IPv6 interarrival times of aggregated traffic are different from their IPv4 counterparts. Is there any significant difference between the 2 protocols in terms of interarrival times when a specific amount of traffic is received? Autocorrelation and power spectral density plots were utilized to find the observable differences.

Interarrival time autocorrelations of IPv6 traffic decrease with increasing traffic load, while there are opposite results for the IPv4 counterparts. This opposite behavior proves that there are certain differences in terms of self-similarity for interarrival times. Self- similarity degrees decrease with increasing traffic load for IPv6 traffic, but they increase for IPv4. Related autocorrelation values for both protocols are represented in Figures 16 and 17.

PSD plots also confirm this case. While the 1/f-type spectrum behavior of IPv6 interarrival time decays with increasing traffic load, the obtained results seem to be the opposite for the IPv4 counterparts. The 1/f-type power spectrum behavior of IPv4 interarrival times increase with incoming traffic load. The mentioned results are represented in Figures 18 and 19.

7. Self-similarity analysis

High self-similarity causes significant performance degradations in computer networks. Therefore, its robust estimation is a necessity. The Hurst parameter is a numerical measure of self-similarity. It denotes whether a stochastic process is long-range dependent or not. A continuous-time stochastic process $\{X(t),\ t \in R\}$is strictly self-similar with a Hurst degree of {H, 0 < H < 1} if the following is true:

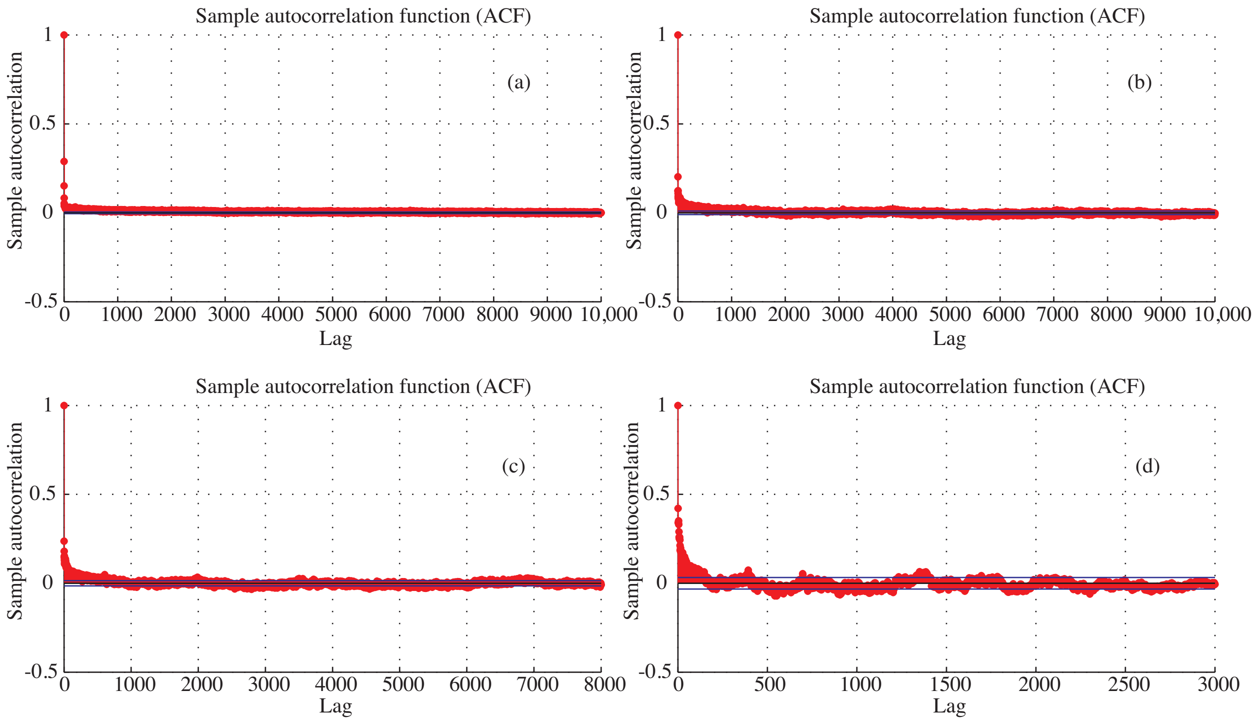

Figure 16. Autocorrelation of IPv4 traffic with reception of: a) 10,000 bytes of traffic, b) 50,000 bytes of traffic, c) 100,000 bytes of traffic, d) 500,000 bytes of traffic.

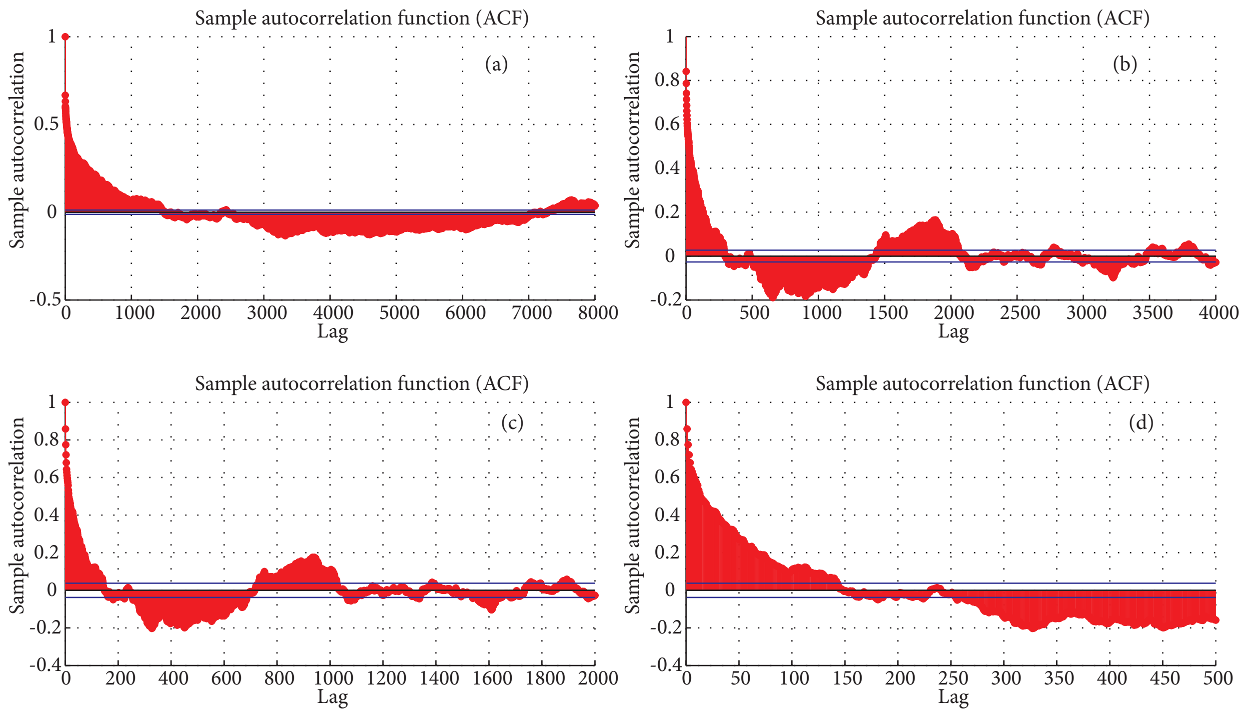

Figure 17. Autocorrelation of IPv6 traffic with reception of: a) 10,000 bytes of traffic, b) 50,000 bytes of traffic, c) 100,000 bytes of traffic, d) 500,000 bytes of traffic.

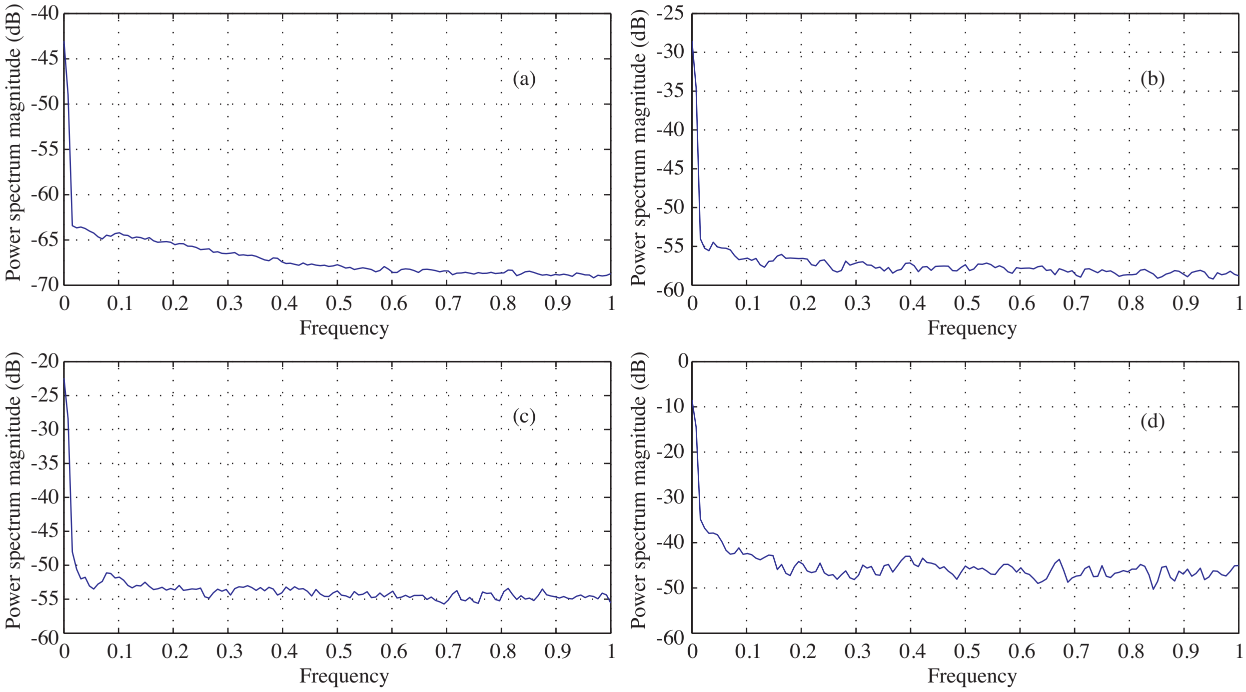

Figure 18. PSD of incoming IPv4 traffic with reception of: a) 10,000 bytes, (b) 50,000 bytes, c) 100,000 bytes, d) 500,000 bytes.

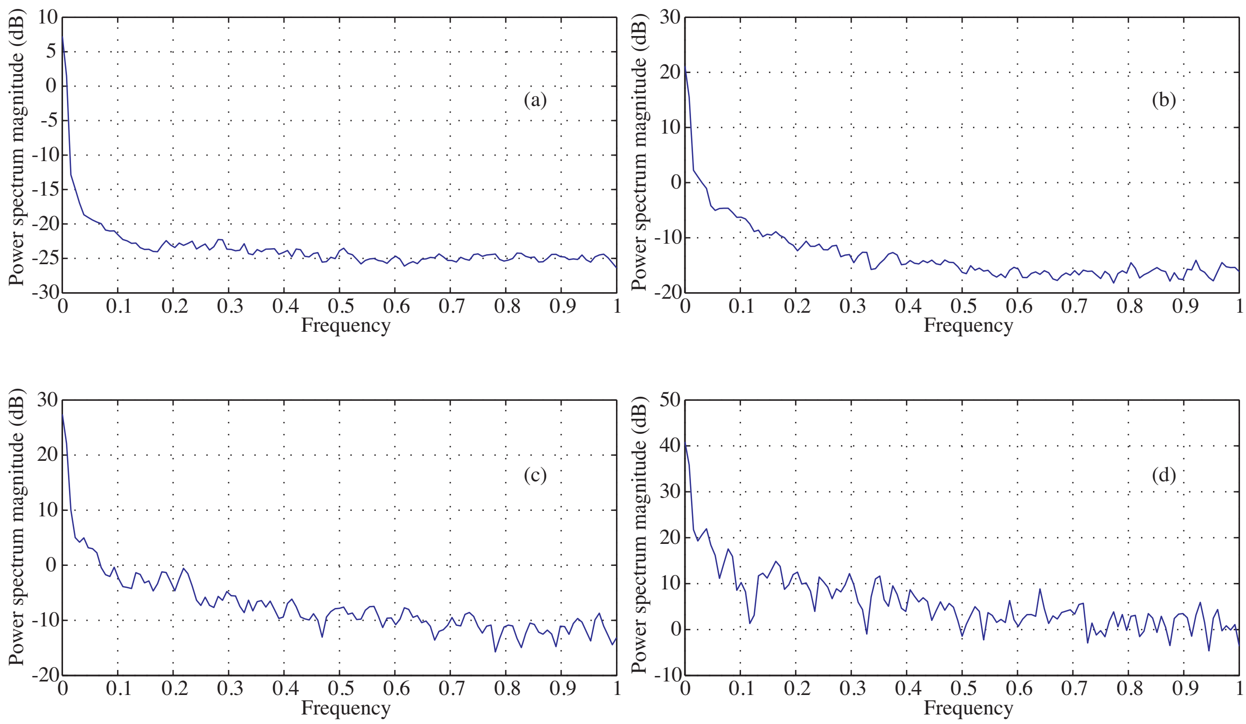

Figure 19. PSD of incoming IPv6 traffic with reception of: a) 10,000 bytes, b) 50,000 bytes, c) 100,000 bytes, d) 500,000 bytes.

$$X(at) \stackrel{d}{=} a^H X(t). \tag{8}$$

$X(at)$ is a new process scaled by a and $\stackrel{d}{=}$ denotes equality in finite-dimensional distributions [19]. Some commonly used time-based estimators are the R/S method [27], the variance time estimator [28], the estimator based on absolute moments [29], and the estimator based on variance of residuals [12]. Frequency-based Hurst estimation methods take into consideration power law behavior of power spectral density. The Daniell periodogram-based estimator [28], Whittle maximum likelihood [27], and local Whittle maximum likelihood estimator [30] are well-known frequency-based estimators. An efficient wavelet-based estimator called the Abry-Veitch Daubechies wavelet-based estimator was proposed by Abry and Veitch in 1998 [31]. Daubechies wavelets are the most studied ones in wavelet-based Hurst estimation [32]. Daubechies 2 and Daubechies 3 give better Hurst estimation results for long-range dependent fractional Gaussian noises [22].

Hurst values of IPv6 and IPv4 traffic in terms of packet interarrival time and packet size were estimated. We also carried out self-similarity analysis of aggregated traffic loads at different time scales and interarrival times when different amounts of traffic load were received. Various Hurst estimation methods were utilized for the sake of estimation trustiness. Average Hurst values for 30 days of daily IPv6 traffic and 15 days of daily IPv4 traffic are given below.

At first, we estimated Hurst degrees of interarrival times per 10,000, 50,000, 100,000, and 500,000 bytes received. The calculated average Hurst values are given in Tables 2 and 3. Self-similarity degrees of IPv4 interarrival time increase with increasing traffic load. On the other hand, the self-similarity degrees of IPv6 interarrival times per 100,000 and 500,000 bytes decrease in comparison to previous scales. Sometimes the estimation methods give Hurst degrees greater than 1. That is not a desirable result for any estimation method. Those values are accepted as quite high self-similarity degrees.

Table 2. Average Hurst values of interarrival time of IPv6 traffic.

Traffic load (bytes), 30-day averages	Absolute value	Differential variance	Aggregated variance	Rescaled range	Wavelet method
10,000	0.892812	0.798774	0.840976	0.999572	1.009182
50,000	0.907677	1.043325	0.895579	0.949949	1.037125
100,000	0.890796	1.041677	0.879513	0.92011	0.991618
500,000	0.784468	0.72281	0.60818	0.851436	0.894279

Table 3. Average Hurst values of interarrival time of IPv4 traffic.

Traffic load (bytes), 15-day averages	Absolute value	Differential variance	Aggregated variance	Rescaled range	Wavelet method
10,000	0.735452	0.564139	0.735806	0.733434	0.610401
50,000	0.811976	0.559586	0.812875	0.806217	0.621486
100,000	0.843958	0.605961	0.84531	0.841827	0.69422
500,000	0.874653	0.894686	0.878519	0.890264	0.946994

Average Hurst values for aggregated traffic at different time scales are given in Tables 4 and 5. IPv6 aggregated traffic at different time scales obviously shows that the self-similarity degrees at all time scales except 1 s are greater than those of IPv4. Moreover, the absolute value and aggregated variance methods give greater Hurst degrees at the 1-s scale for IPv6 traffic loads.

Table 4. Average Hurst values of aggregated IPv6 traffic at different time scales.

Traffic load (bytes), 30-day averages	Absolute value	Differential variance	Aggregated variance	Rescaled range	Wavelet method
0.001	0.904001	0.69713	0.904095	0.902551	0.656181
0.01	0.923599	0.799847	0.923709	0.943639	0.644213
0.1	0.925657	1.122492	0.925943	0.922329	1.120932
1	0.861664	1.10272	0.862754	0.821699	1.040008

Table 5. Average Hurst values of aggregated IPv4 traffic at different time scales.

Traffic load (bytes), 15-day averages	Absolute value	Differential variance	Aggregated variance	Rescaled range	Wavelet method
0.001	0.761289	0.580831	0.761741	0.748904	0.614101
0.01	0.856172	0.655581	0.858048	0.855838	0.783866
0.1	0.84742	0.933738	0.855476	0.891735	0.899996
1	0.724464	1.299874	0.755081	0.988637	1.336524

We compared the self-similarity of the packet interarrival times and packet sizes of both protocol traffic types. We obtained sensible results for the self-similarity of the packet size. Self-similarity degrees of IPv6 packet size are always greater than those of IPv4 according to all estimation methods. Self-similarity estimation of packet interarrival time gives meaningless results. The absolute value, aggregated variance, and rescaled range Hurst estimation methods show that the self-similarity of interarrival times for IPv6 is greater than that for IPv4, but the differential variance and wavelet methods give contrary results. Average values are given in Tables 6 and 7. As mentioned, the IPv6 measurements comprise 30 days and the IPv4 measurements comprise 15 days.

An important conclusion drawn from the self-similarity analysis is that all packet interarrival times and the packet sizes of both protocol packet traffics represent a high degree of self-similarity. The general results for IPv6 show greater self-similarity degrees than those for IPv4.

Table 6. Average Hurst values of packet interarrival time.

Interrival time	Absolute value	Differential variance	Aggregated variance	Rescaled range	Wavelet method
IPv4	0.71935	0.676906	0.719429	0.69677	0.680433
IPv6	0.826074	0.638283	0.826166	0.913019	0.570864

Table 7. Average Hurst values of packet size.

Packet size	Absolute value	Differential variance	Aggregated variance	Rescaled range	Wavelet method
IPv4	0.674958	0.648844	0.675037	0.642257	0.717481
IPv6	0.899921	0.742941	0.900028	1.00554	0.740246

8. Distribution fitting procedure

Obtained CDF plots obviously demonstrate that there are certain differences between empirical CDFs for packet interarrival times and packet sizes of IPv6 and IPv4 protocol traffic. Observed differences prove that the modeling of the mentioned time series should be done with different distribution functions. With efficient packet traffic modeling, true traffic will be generated for both IPv6 and IPv4 in network simulation studies. In the distribution fitting process, we used 48 different distribution functions. For selection criteria of goodness of fitting, we used Anderson-Darling and chi-square test results and determined the best distribution type for packet interarrival time and packet size of IPv6 and IPv4 packet traffic.

Beta distribution gave the best result for IPv4 packet size modeling. The beta probability density function is as follows:

$$f(x) = \frac{1}{B(\alpha_1, \alpha_2)} \frac{(x-a)^{\alpha_1 - 1}(b-x)^{\alpha_2 - 1}}{(b-a)^{\alpha_1 + \alpha_2 - 1}}, \tag{9}$$

and the CDF function is given as:

$$F(x) = \frac{\beta_x(\alpha_1, \alpha_2)}{\beta(\alpha_1, \alpha_2)} = I_x(\alpha_1, \alpha_2). \tag{10}$$

According to test results, the span for beta distribution parameters α_1 and α_2 change between [0.05395, 0.07022] and [0.07061, 0.09236] for the analyzed traces, respectively. The Weibull distribution was selected for interarrival time fitting according to Anderson-Darling and chi-square test results. The PSD function and CDF of the Weibull distribution are as follows:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right), \tag{11}$$

$$F(x) = 1 - \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right). \tag{12}$$

The α and β parameters change between [0.67832, 0.70687] and [3.67e-5, 4.4e-5] for interarrival time, respectively. The distribution fitting plots for IPv4 are given in Figure 20.

The log logistic distribution is the best model for IPv6 packet size according to test statistics. The PSD distribution of the log logistic is:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-2}, \tag{13}$$

and the CDF distribution of the log logistic is:

$$F(x) = \left(1 + \left(\frac{\beta}{x-\gamma}\right)^{\alpha}\right)^{-1}. \tag{14}$$

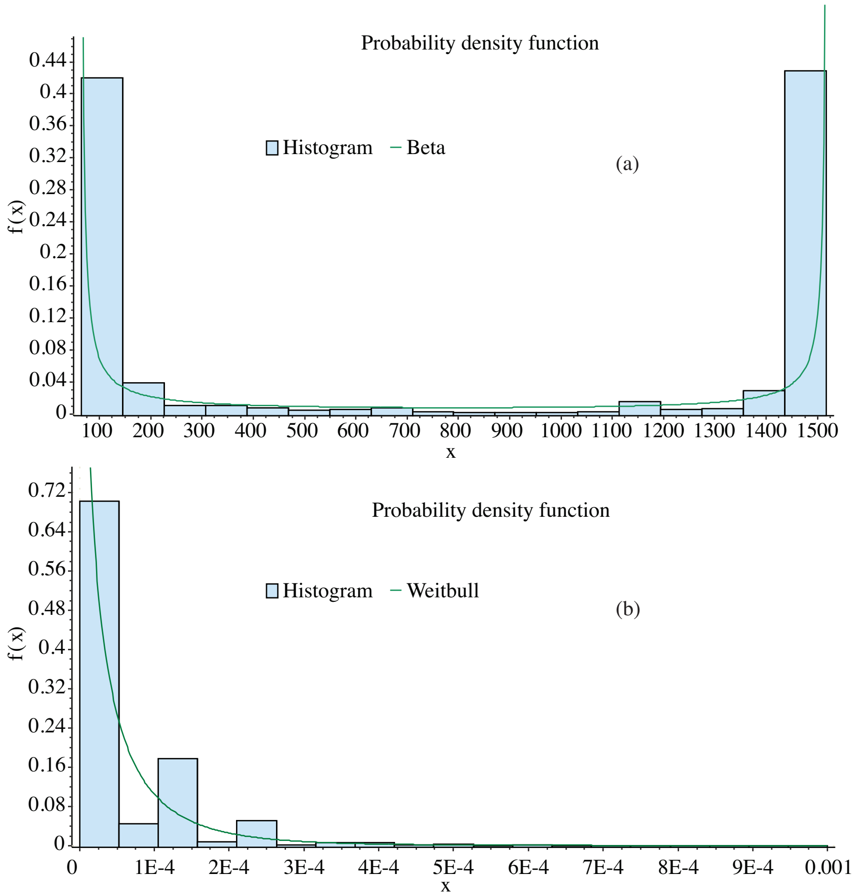

Figure 20. IPv4 distribution for: a) packet size, b) packet interarrival time.

The α and β parameter spans for all day traces vary between [0.8428, 1.1405] and [7.8066, 12.928], respectively. The generalized Pareto distribution gave the best-fitting result for IPv6 packet interarrival times. The PSD and CDF of the generalized Pareto distribution are as follows:

$$f(x) = \begin{cases} \frac{1}{\sigma}\left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-1-1/k} & k \neq 0 \\ \frac{1}{\sigma}\exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0 \end{cases}, \tag{15}$$

$$F(x) = \begin{cases} 1 - \left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-1/k} & k \neq 0 \\ 1 - \exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0 \end{cases}. \tag{16}$$

The α, β, and k parameters span for all day traces vary between [1.546, 2.1734], [1.24e-4,2.80e-4], and [0.38327, 0.45708], respectively. The distribution fitting plots for IPv4 are given in Figure 21. The obtained distribution results obviously show that IPv6 and IPv4 packet size and packet interarrival time have different probabilities.

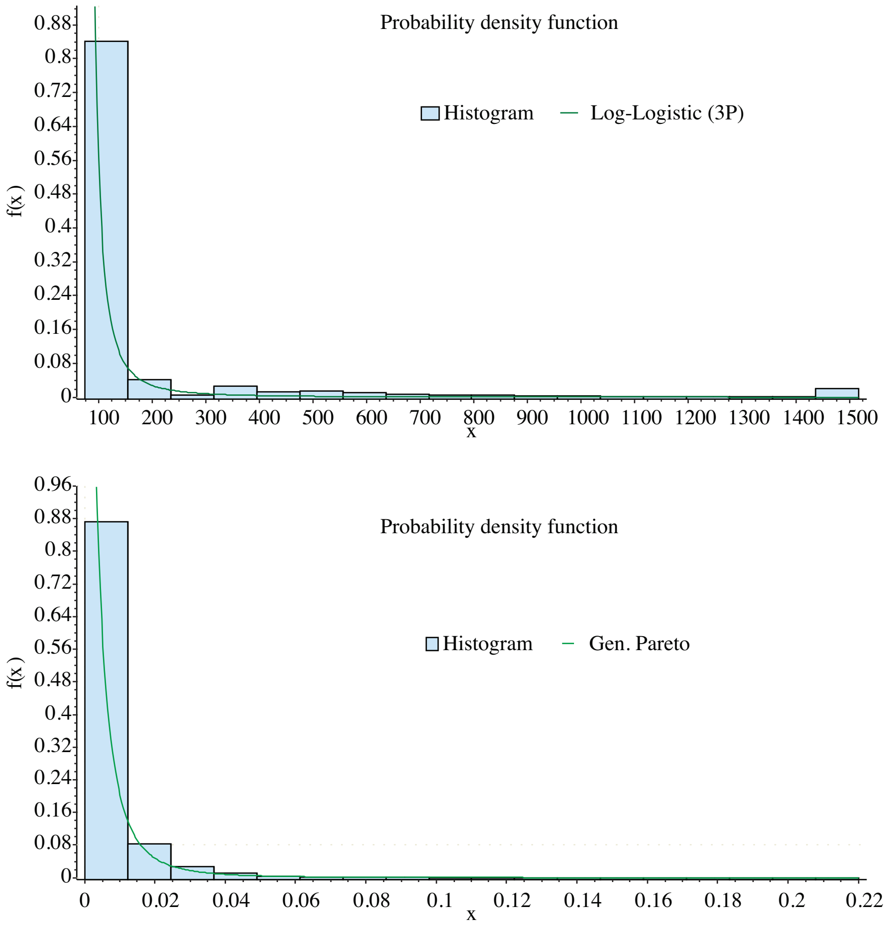

Figure 21. IPv6 distribution for: a) packet size, b) packet interarrival time.

9. Future work

Self-similarity causes performance degradations in computer networks. High self-similarity increases packet loss, delay, jitter, and congestion problems in computer networks. Accurate performance evaluation results can be only obtained with true traffic models in network simulations. With the performed study, we have determined some characteristic features of IPv6 and IPv4 protocol traffic. We will next use self-similarity and distribution results to generate realistic traffic for network simulation studies. By this means, more realistic performance evaluation results for packet loss, packet delay, jitter, buffer usages, and round-trip times could be obtained from network simulators.

10. Conclusion

We analyzed different characteristics of IPv6 and IPv4 protocol traffic generated by various applications. We analyzed IPv6 traces of May 2008 and randomly chose 10 days of traces from 2008 obtained on a link connected to a Wide-6Bone node. Traces were taken from the MAWI Working Group traffic archive. To determine differences between IPv6 and IPv4 traffic, we also analyzed a consecutive 15 days and 10 randomly selected days of IPv4 traces obtained on a 150-Mbps transpacific link between Japan and the United States. Detailed analyses were performed in terms of power spectral density, probability density, autocorrelation, self-similarity, and distribution fitting differences between IPv6 and IPv4 protocol traffic.

Our study demonstrates that there are certain differences in terms of the mentioned traffic characteristics for both IPv6- and IPv4-related traffic. Packet interarrival time and packet size distribution fitting results prove that they should be modeled with different functions. While the beta distribution could model the empirical cumulative distribution of IPv4 packet size, the log logistic distribution gave more efficient results for IPv6 packet size according to chi-square and Anderson-Darling test statistics. Various analysis results showed that the aggregated incoming traffic at different time scales exhibited very different characteristics in terms of power spectral density and autocorrelation. These variations deeply affect the self-similarity degree of aggregated traffic at different time scales for both protocol traffics. Lastly, we analyzed the interarrival times of incoming traffic loads per 10,000, 50,000, 100,000, and 500,000 received bytes in terms of IPv6 and IPv4 protocol traffic. The self-similarity results for the interarrival time series were quite different for the protocols. IPv6 packet traffic exhibited greater self-similarity degrees than IPv4. The results obviously show that IPv6 protocol traffic would cause more performance degradations in computer networks.

References

[1] D. Minoli, Voice Over IPv6: Architectures for Next Generation VoIP Networks, Burlington, MA, USA, Elsevier, 2006, pp. 259-292.

[2] S. Deering, R. Hinden, Internet Protocol, Version 6 (IPv6) Specification, IETF RFC 2460, 1998.

[3] B. Carpenter, K. Moore, Connection of IPv6 Domains via IPv4 Clouds, IETF RFC 3056, 2001.

[4] A. Durand, P. Fasano, I. Guardini, D. Lento, IPv6 Tunnel Broker, IETF RFC 3053, 2001.

[5] R. Rockell, R. Fink, 6Bone Backbone Routing Guidelines, IETF RFC 2772, 2000.

[6] R. Gilligan, E. Nordmark, Transition Mechanisms for IPv6 Hosts and Routers, IETF RFC 2893, 2000.

[7] I. Raicu, S. Zeadally, "Evaluating IPv4 to IPv6 transition mechanisms", IEEE International Conference on Telecommunications, Vol. 2, pp. 1091-1098, 2003.

[8] Y. Wang, S. Ye, X. Li, "Understanding current IPv6 performance: a measurement study", 10th IEEE Symposium on Computers and Communications, pp. 71-76, 2005.

[9] N. Choi, H. Son, Y. Lee, Y. Choi, "Experiences with IPFIX-based traffic measurement for IPv6 networks", Proceedings of ACM IPv6'07, 2007.

[10] D.P. Pezaros, D. Hutchison, F.J. Garcia, R.D. Gardner, J.S. Sventek, "In-line service measurements: an IPv6- based framework for traffic evaluation and network operations", IEEE Network Operations and Management Symposium, Vol. 1, pp. 497-510, 2004.

[11] P. Savola, "Observations of IPv6 traffic on a 6to4 relay", ACM SIGCOMM Computer Communication Review, Vol. 35, pp. 23-28, 2005.

[12] W. Leland, M. Taqqu, W. Willinger, D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)", IEEE ACM Transactions on Networking, Vol. 2, pp. 1-15, 1994.

[13] C.K. Peng, S. Buldyrev, S. Havlin, M. Simons, H. Stanley, A. Goldberger, "Mosaic organization of DNA nucleotides", Physical Review E, Vol. 49, pp. 1685-1689, 1994.

[14] M.E. Crovella, A. Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes", IEEE/ACM Transactions on Networking, Vol. 5, pp. 835-846, 1997.

[15] J. Beran, R. Sherman, M.S. Taqqu, W. Willinger, "Long-range dependence in variable-bit-rate video traffic", IEEE Transactions on Communications, Vol. 43, pp. 1566-1579, 1995.

[16] V. Paxson, S. Floyd, "Wide area traffic: the failure of Poisson modeling", IEEE/ACM Transactions on Networking, Vol. 3, pp. 226- 244, 1995.

[17] W. Willinger, M.S. Taqqu, R. Sherman, 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, pp. 1-86, 1997.

[18] K. Park, W. Willinger, Self-Similar Network Traffic and Performance Evaluation, New York, John Wiley & Sons, 2000, pp. 1-38.

[19] C. Çiflikli, A. Gezer, A.T. Özşahin, Ö. Özkasap, "BitTorrent packet traffic features over IPv6 and IPv4", Simulation Modelling Practice and Theory, Vol. 18, pp. 1214-1124, 2010.

[20] C. Çiflikli, A. Gezer, A.T. Özşahin, Ö. Özkasap, "Comparison of Bittorrent packet traffic characteristics over IPv6 and IPv4", Application of Information and Communication Technologies, Baku, pp. 1-5, 2009.

[21] C. Çiflikli, A. Gezer, A.T. Özşahin, "Traffic measurements of new internet Protocol IPv6", 9th ETAI, Ohrid, 2009.

[22] C. Çiflikli, A. Gezer, "Comparison of Daubechies wavelets for Hurst parameter estimation", Turkish Journal of Electrical Engineering & Computer Sciences, Vol. 18, pp. 117-128, 2010.

[23] S. Deering, R. Hinden, Internet Protocol, Version 6 (IPv6) Specification, IETF RFC 1883, 1995.

[24] M. Goncalves, K. Niles, IPv6 Networks, New York, McGraw-Hill, 1998.

[25] USC Information Sciences Institute, Internet Protocol, IETF RFC 791, 1981.

[26] MAWI Working Group Traffic Archive, available at http://mawi.wide.ad.jp/mawi/.

[27] J. Beran, Statistics for Long-Memory Processes, New York, Chapman & Hall, 1994.

[28] D. Cox, P. Lewis, The Statistical Analysis of Series of Events, London, Chapman & Hall, 1966.

[29] M. Taqqu, V. Teverovsky, W. Willinger, "Estimators for long-range dependence: an empirical study", Fractals, Vol. 3, pp. 785-788, 1995.

[30] P. Robinson, "Gaussian semiparametric estimation of long-range dependence", The Annals of Statistics, Vol. 23, pp. 1630-1661, 1995.

[31] P. Abry, D. Veitch, "Wavelet analysis of long-range-dependent traffic", IEEE Transactions on Information Theory, Vol. 44, pp. 2-15, 1998.

[32] H.J. Jeong, J.R. Lee, D. McNickle, K. Pawlikowski, "Comparison of various estimators in simulated FGN", Simulation Modelling Practice and Theory, Vol. 15, pp. 1173-1191, 2007.