MPRI-JPAS | INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE
VOL. 11 NO.9 JUN-2020 ISSN: 1660-5332

STUDY OF THE DEGREE OF TRANSPORT CONTROL PROTOCOL SELF-SIMILARITY ON A WIRELESS LOCAL AREA NETWORK

[1]ROMANUS I.O, [2]MARIA DINGARI AND SUNDAY AYIGUN[3]

[1&3]Department of Applied Physics, Federal Polytechnic Mubi, Adamawa State. [2]Department of Science Laboratory Technology, Federal Polytechnic Mubi, Adamawa State.

Abstract

The magnitude of self-similarity on a wireless network determines the QoS available to the end users, it becomes very much necessary to have an insight into the network traffic self-similarity characteristics and its measurement. This research took a look at the degree of Transport Control Protocol (TCP) self-similarity on a wireless local area network and its measurement. A total of 1,204,215 TCP traffic volumes were captured, out of which 178,563, 176,127, 163,556, 225,594 231,728, and 228,647 traffic volumes were captured on 1th, 5th, 10th, 13th, 18th, and 21st of February 2018 respectively. It was observed that the ACF takes longer range to decay in lag of (-200,000 to 200,000) on 13th, 18th and 21st February, 2018 with the rest having the lag of (-150,000 to 150,000). The research confirmed the presence of LRD traffic by applying autocorrelation test on the captured live traffic and it showed an autocorrelation function that decays hyperbolically. Four Hurst estimators were

[redacted]

employed so as to complement each other's shortfall. An average daily Hurst parameter of 0.748, depicting a persistent characteristics. The research recommends the use of neuro-fuzzy principles to x-ray the cause of self-similarity on a wireless network.

KEYWORDS: *Wireless, Transport, Protocol, Self-Similarity, Local Area Network.*

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]

INTERNATIONAL JOURNAL – PAS
VOL. 4 NO.1 MAR-2018 ISSN: 1660-5332

INTRODUCTION

In wireless network communication (WNC), transport layer is fully described in internet protocol stack (IPS) or in open system interconnection (OSI) reference model, the protocol usually used in transport layer is popularly known as transport control protocol (TCP). This protocol provides service or performance grantee such as end to end connection between source and destination, reliability, controls flow of packet and perms multiplexing. Hence, regular monitoring of activities of a wireless network on the transport layer based on TCP becomes paramount. (Romanus, and Ali, 2018)

Transmission Control Protocol (TCP) is developed to initiate reliable end-to end delivery of data over unreliable networks. In theory, TCP should be independent of the technology of the underlying infrastructure. In particular, TCP should not care whether the Internet Protocol (IP) is running over wired or wireless connections. In practice, it does matter because most TCP deployments have been carefully designed based on assumptions that are specific to wired networks. Ignoring the properties of wireless transmission can lead to TCP implementations with poor performance (Guo. et al., 2001)

In wireless networks, the principal problem of TCP lies in performing congestion control in case of losses that are not induced by network congestion. Since bit error rates are very low in wired network, nearly all TCP versions nowadays assume that packet losses are due to congestion. Consequently, when a packet is detected to be lost, either by timeout or by multiple duplicated ACKs, TCP slows down the sending rate by adjusting its congestion window. Unfortunately, wireless networks suffer from seal types of losses that are not related to congestion, making TCP not adapted to this environment. Numerous enhancements and optimizations have been proposed over the last few years to improve TCP performance over one-hop wireless (not necessarily ad hoc) networks (Figueredo 2002).

Transmission Control Protocol (TCP) is the de facto standard for the Internet. Using the connection-less, unreliable data delivery services of Internet Protocol (IP), it provides end-to-end, in-order, reliable data delivery services to various applications like File Transfer Protocol (FTP), Telnet, and Hypertext Transfer Protocol (HTTP). Initially, TCP did not

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]
mediterraneanresearchpub@gmail.com

incorporate end-to-end congestion control . The latter was first installed in TCP 3 around 1988 in a successful attempt to combat the occurrence of congestion collapse, and has since become an essential element of TCP, central to ensuring the stable, efficient operation of the present Internet. TCP is connection-oriented. A TCP connection is set up between two end systems when they agree on a few parameter settings and each reserve resources like memory space for the transmission via TCP, after which the transmission begins. The transmission ends when the connection is closed and the resources reserved by end systems will be released (Guo. et al., 2001)

The TCP connection is duplex. Both end systems can transmit and receive data. Nevertheless, we can conceptually focus on one-way transmission over a TCP connection to avoid unnecessary complications. Moreover, in reality the bulk of data transfer over a TCP connection takes place in one way, say, from a web server to a web client. Correspondingly, the end systems that transmits and receives data are referred to as TCP sender and TCP receiver, respectively. A TCP sender consecutively assigns byte numbers to its transmitted data bytes, and a TCP receiver expects the number of received data bytes to be consecutively ordered. When receiving a data packet, a TCP receiver notifies the sender of successfully delivered data via an acknowledgment (ACK) packet that contains acknowledgment number, namely the number of next byte it expects. The time elapsed between when a TCP sender transmits a data packet and when the packet is acknowledged is known as round-trip time (RTT).

Self-Similarity on a Wireless Network

The presence of self-similar features in real network traffic has sparked a new research discipline that attempts to analyze these features and model and reproduce them in synthetic traffic. Since it is not always possible and practical to obtain large traces of real network traffic, mathematical models of self-similar traffic are the only viable alternative for obtaining input traces. Recent research in the area has provided very strong empirical evidence that self- similar models are much better than Poisson processes at capturing crucial network traffic characteristics such as burstiness and correlations over large lags (Leland *et al.,* 1995, Paxson and Floyd, 1997). These properties have profound

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]
mediterraneanresearchpub@gmail.com

INTERNATIONAL JOURNAL – PAS
VOL. 4 NO.1 MAR-2018 ISSN: 1660-5332

effects on network queuing performance (Erramilli *et al.,* 1996). While theoretical frameworks are currently being developed to estimate the performance of such systems, simulation will remain a valuable tool for validating these theoretical models, and providing insight into systems that are too complicated to resolve analytically (Roughan, 1999).
However, research on self-similarity being in its infancy, there are no known simulation tools that encompass a self-similar generator. Thus, if one wishes to simulate the behavior of a network under self-similar conditions, it is necessary to use mathematical models of self-similar traffic to generate the traces that are then input to a simulation tool. This not only highlights the importance of proposing a suitable mathematical model (one which realistically models true network traffic) but also the need to have robust and accurate tools to effectively test a series for self-similarity.

MATERIALS AND METHODS

This research presents the materials and methods that were employed for this research, it gave an insight on how the traffic was captured and filtered so as to extract the transport control protocol whose time series length was used for the analysis.

Materials

The materials used for this research work are as follows:

- A laptop
- A gsm mobile phone.
- An excel software
- Selqos (A self-similarity software)
- Wireshark (A network protocol analyzer)

Method of Data Collection

Data for this research was collected from a WLAN network set up using protocol analyzing software called wireshark for the six hours per day for six days. The network traffic so captured was filtered to obtain the transport control protocol (TCP) which was used for the analysis.

Method of Data Analysis

The captured protocol was exported to excel spread sheet where the time series data was extracted and was further exported to a self-

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]
mediterraneanresearchpub@gmail.com

INTERNATIONAL JOURNAL – PAS
VOL. 4 NO.1 MAR-2018 ISSN: 1660-5332

similarity tool SELQOS, where it was used to determine the Autocorrelation function and the Hurst Parameter for two days per week for the period of six months which was used to evaluate the degree of fractals and also observe quality of service (QoS) on the wireless network for the said period. This is as presented in Figure1.

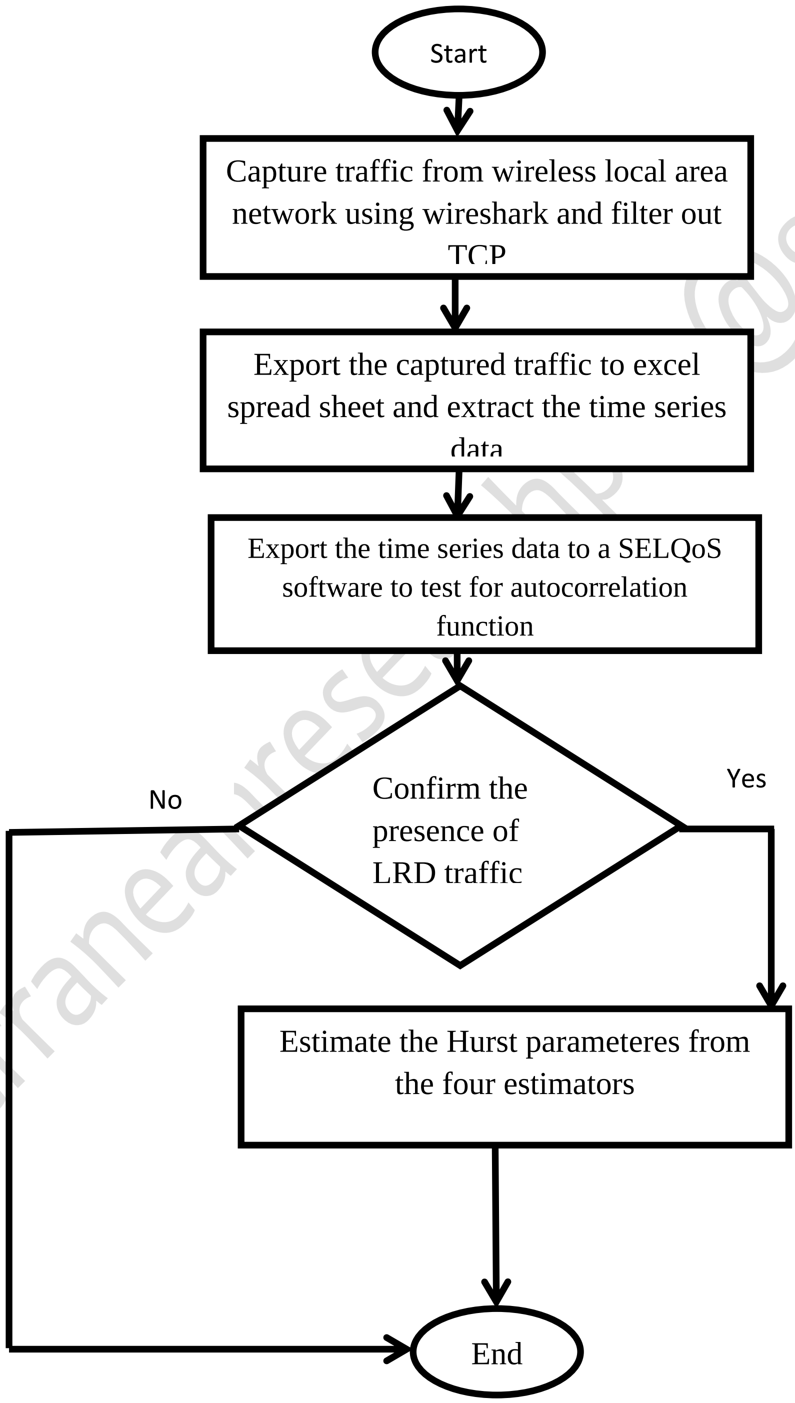

Figure1. The flow chart for the measurement of degree of self-similarity on WLAN

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]

mediterraneanresearchpub@gmail.com

Autocorrelation Test

The autocorrelation test provides a visual representation of the rate of decay of the autocorrelation function for any given time series. One of the most important properties of a time series exhibiting LRD is an autocorrelation function that decays hyperbolically fast. The sample data was analyzed to obtain the autocorrelation function. At lags larger than the "knee", for true self-similar processes, the rate of decay of the autocorrelation function becomes very slow, indicating the presence of LRD. This was tested on all the captured traffics within the period under research.

Estimation of Hurst Parameter (*H*)

It is not possible to use the definition of traffic self-similarity to check whether a finite traffic trace is self-similar or not. Instead different features of self-similarity such as slowly decaying variances are investigated in order to estimate the Hurst parameter *H*. This parameter *H* can take any value between 1/2 and 1 and the higher the value, the higher the degree of self-similarity. For a random series the value is H=0.5, 0<H<0.5 indicates the behavior of an anti-persistent series, 0.5<H<1 indicates the behavior of a persistence series. The methods that were used to test for self-similarity include the rescaled adjusted range plot (R/S plot), Absolute moment, Variance method and the Periodogram plot (Mohammed and Nickolay, 2017)

The R/S method

This method is based on empirical observations Hurst. It estimates *H* based on the R/S statistic, and then indicates (asymptotically) second-order self-similarity. *H* is (roughly) estimated through the slope of linear line in a log-log plot, depicting R/S statistics over the number of points in the aggregated series. For a given set of observations $X = \{X_{n,} n = 1, 2, 3, \ldots.n\}$ with sample mean $\overline{X}(n)$, sample variance $S^2(n)$ and range $R(n)$, the rescaled adjusted range R/S statistic is given by (Park *et al.*, 1997).

$$\frac{R(n)}{S(n)} = \frac{\max(0, \Delta_1, \Delta_2, \ldots.., \Delta_n) - \min(0, \Delta_1, \Delta, \ldots..., \Delta_n)}{S(n)} \tag{1}$$

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]

mediterraneanresearchpub@gmail.com

Where $$\Delta_k = \sum_{i=1}^{k} X_i - k\overline{X},\ for\, k = 1,2,.....n.$$ (2)

Absolute moment

This method is related to the aggregated variance method. The slope *β* of the straight line is in a log-log plot, depicting the first moment of the aggregated block over the block size which provides an estimator for *H*, by *H*= 1+ *β* (Taqqu and Teverosky 1997).

Variance method

Let *Zn* be a vector with the number of packets in the *n*th interval (bin). If for example the bin size has been chosen to 100 ms then *Z*100 is the number of packets that arrived at the first 100 ms. Characteristic of long-range dependent processes is that the variance of the sample mean converges slower to zero than 1/*n* (the reciprocal of the sample size). It can be shown that

$$Var\left(\overline{Z}_n\right) \approx cn^{2H-2}$$ (3)

Where c is constant and c ≥ 0, *n* is the interval size. This is what the variance-time plot method is based on and the actual method to estimate H is as follows:

First the mean of each pair of consecutive, non-overlapping bins are calculated and then the variance of these means is calculated. The logarithm of the variance is plotted against the logarithm of the block size i.e 1. Then the same thing is done for blocks of size 4,8,16,.., *n* bins. The parameter *H* can be estimated by fitting a simple least squares line through the resulting points and using the relation, slope = 2*H* – 2 (Leland *et al.*, 1993). This was evaluated on all the captured traffics within the period of six months.

The Periodogram method

The periodogram estimator, proposed by Daniell, is a graphical method of assessing *H*. The periodogram plot is obtained by plotting log10 (P(λ)) against log10(λ). This shows that if the autocorrelations are assumable, i.e., short range dependence (SRD), then, near the origin, it should be scattered randomly around a constant level. If the autocorrelations are

mediterraneanresearchpub@gmail.com

non-assumable, i.e., long range dependence (LRD) type, the points are scattered around negative slope. This was evaluated on all the captured traffics within the period under research. An estimate of the Hurst parameter is given by $H = (1-\beta)/2$ where β is the slope of a regression line. The periodogram of $X_i = (X_1, X_2, \ldots\ldots\ldots, X_n)$ is defined by a Fourier time series operation over a time period n (Leland *et al*., 1994).

$$p(\lambda) = \frac{1}{2\pi n}\left|\sum_{j=1}^{n} X_j e^{ij\lambda}\right|_2 \qquad (4)$$

Where λ is the frequency and X is the actual time series.

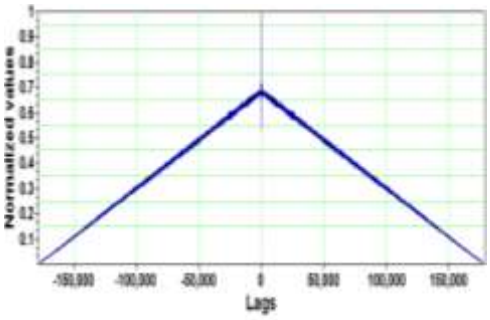

Figure 1 ACF for 1 February, 2018.

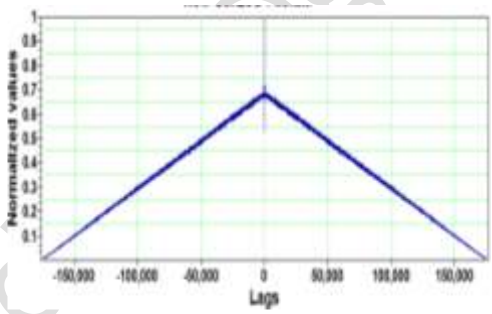

Figure 2 ACF for 5 February, 2018.

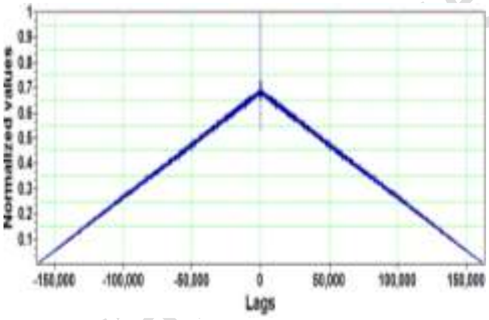

Figure 3 ACF for 10 February, 2018.

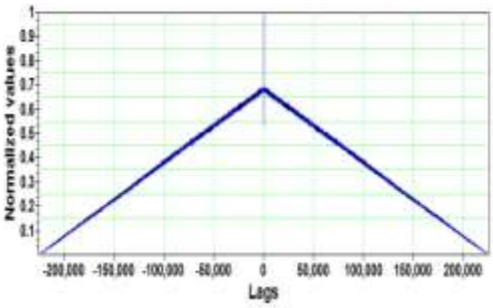

Figure 4 ACF for 13 February, 2018.

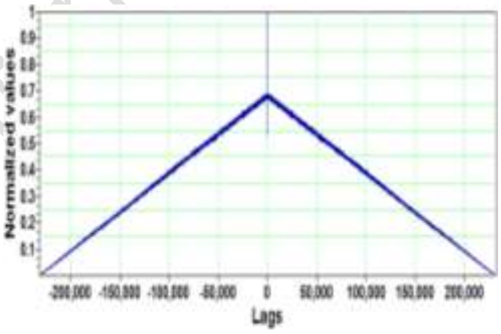

Figure 5 ACF for 18 February, 2018.

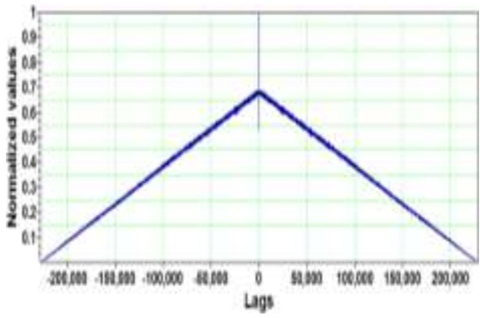

Figure 6 ACF for 21 February, 2018.

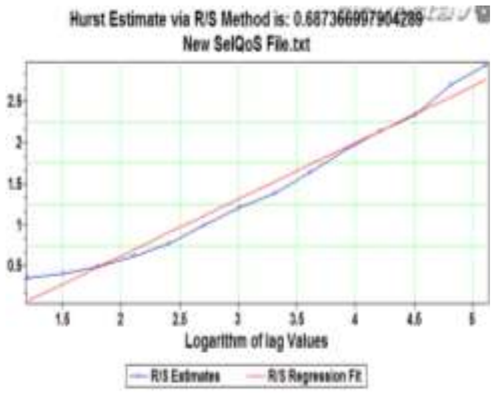

Figure 7 R/S for 1 February, 2018

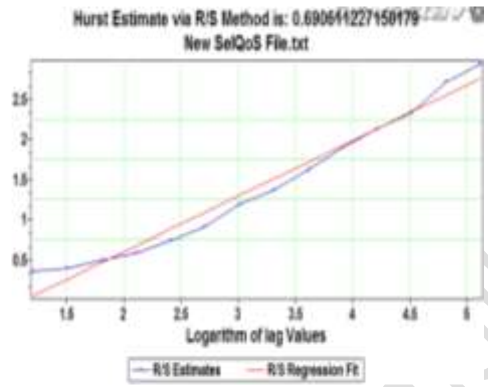

Figure 8 R/S for 5 February, 2018

Figure 9 R/S for 10 February, 2018

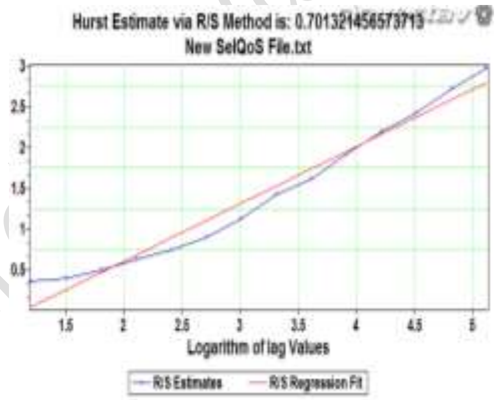

Figure 10 R/S for 13 February, 2018

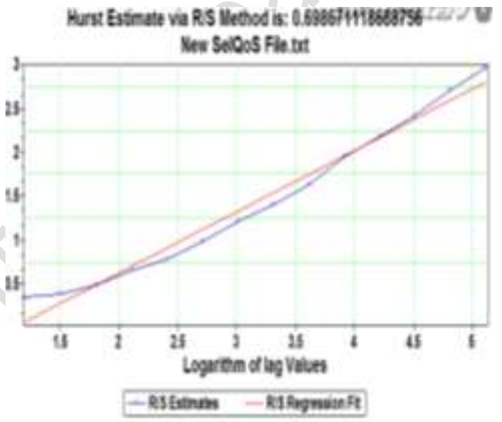

Figure 11 R/S for 18 February, 2018

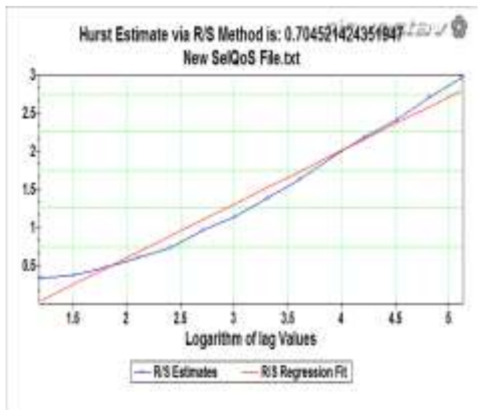

Figure 12 R/S for 21 February, 2018

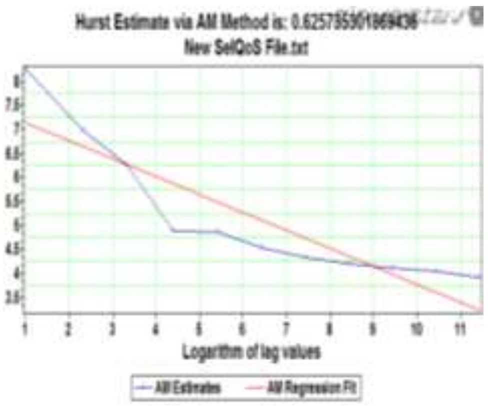

Figure 13 AM for 1 February, 2018

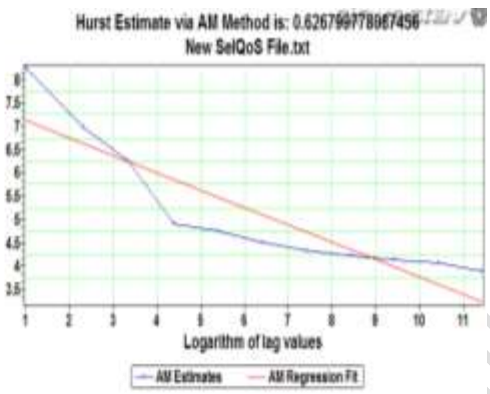

Figure 14 AM for 5 February, 2018

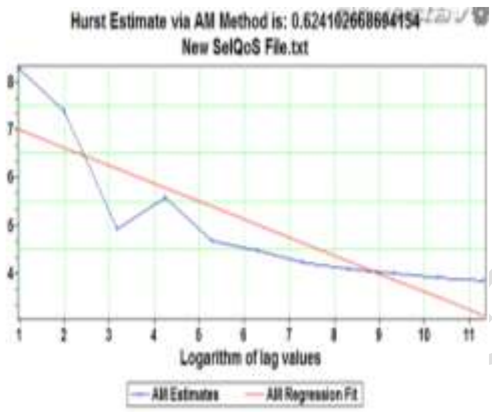

Figure 15 AM for 10 February, 2018

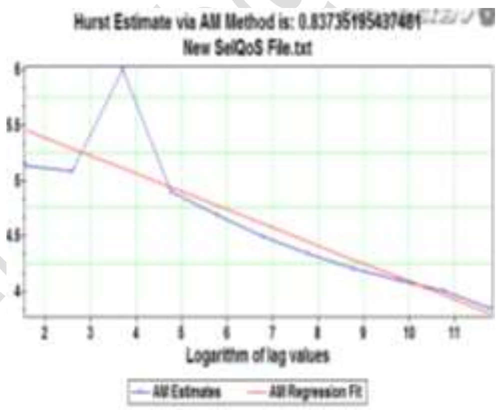

Figure 16 AM for 13 February, 2018

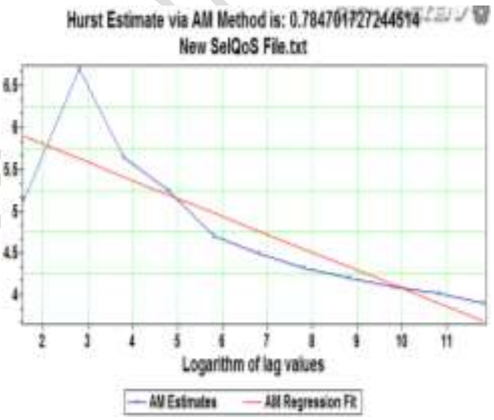

Figure 17 AM for 18 February, 2018

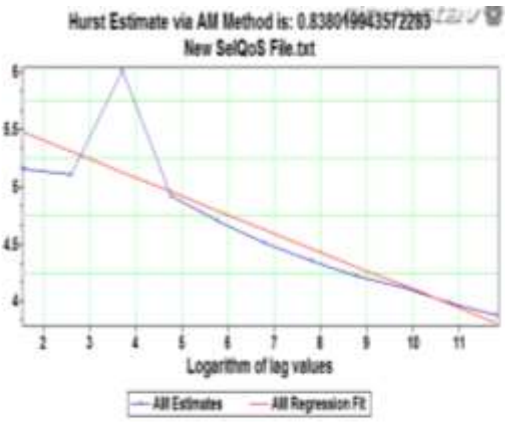

Figure 18 AM for 21 February, 2018

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]

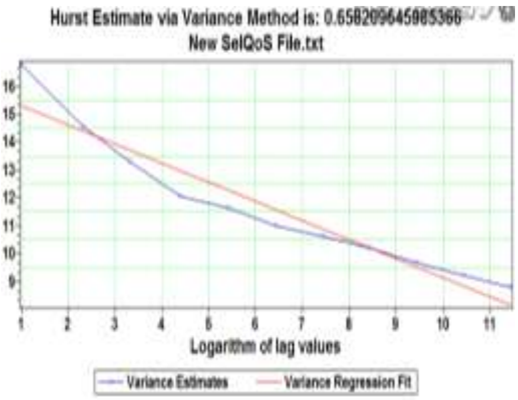

Figure 19 VM for 1 February, 2018

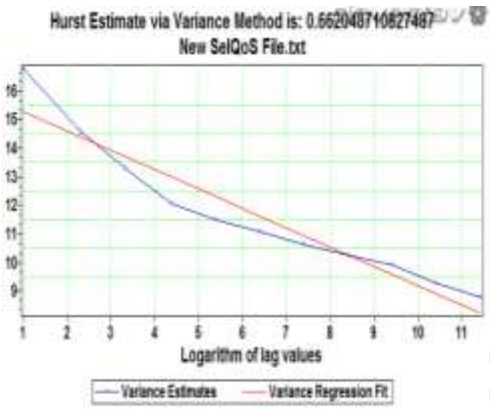

Figure 20VM for 5 February, 2018

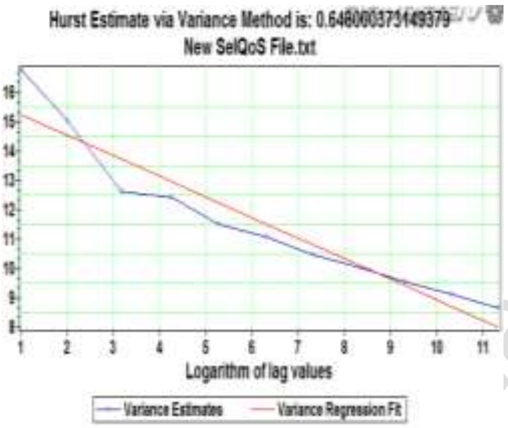

Figure 21 VM for 10 February, 2018

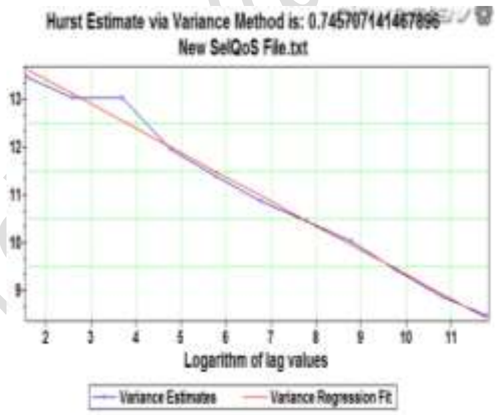

Figure 22 VM for 13 February, 2018

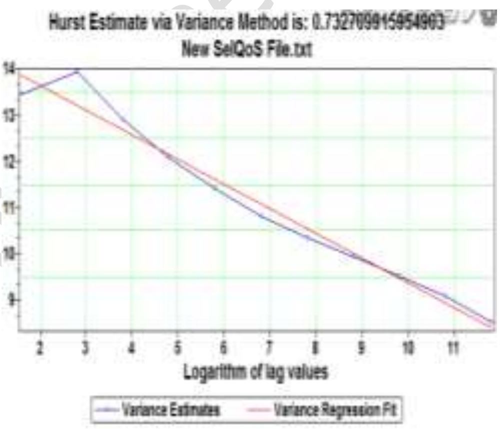

Figure 23 VM for 18 February, 2018

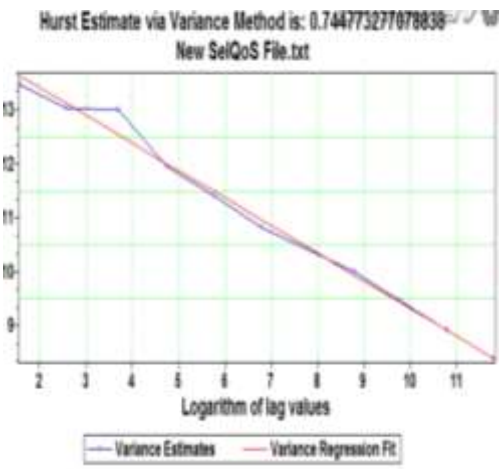

Figure 24 VM for 21 February, 2018

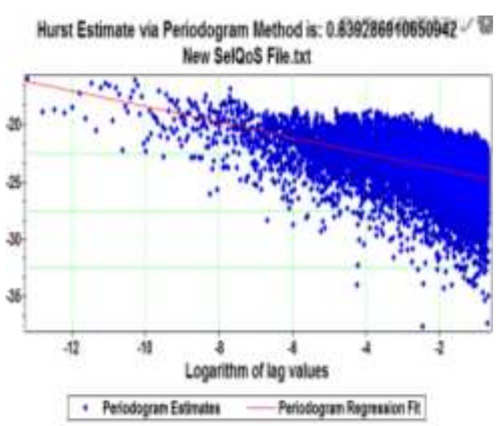

Figure 25 PD for 1 February, 2018

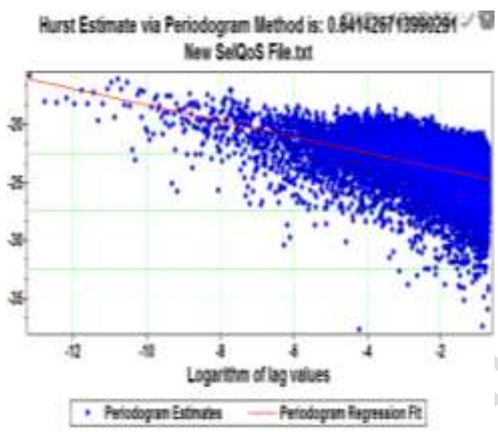

Figure 26 PD for 5 February, 2018

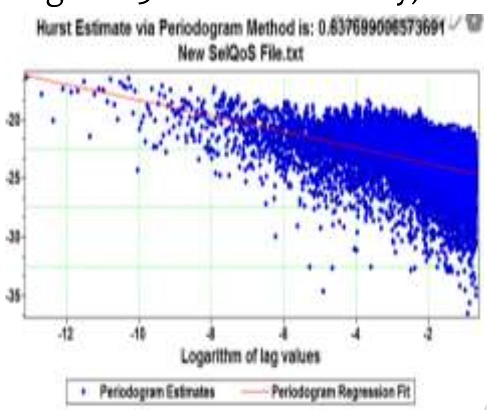

Figure 27 PD for 10 February, 2018

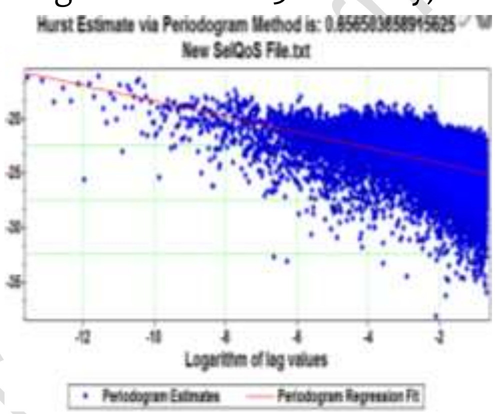

Figure 28 PD for 13 February, 2018

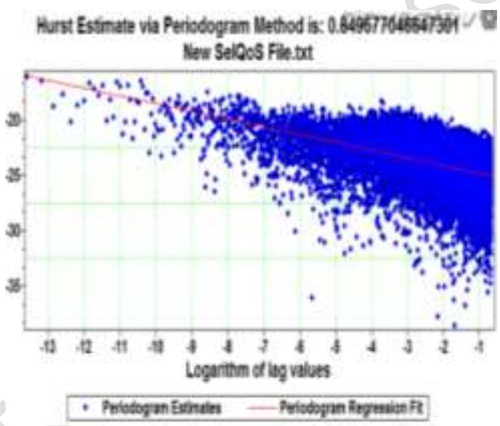

Figure 27 PD for 18 February, 2018

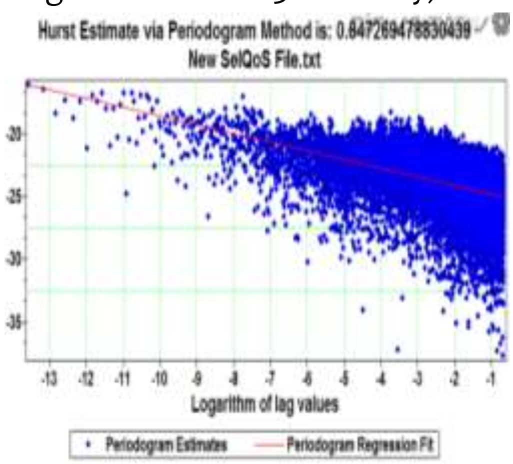

Figure 28 PD for 21 February, 2018

Table 1. Hurst parameters for selected days in the month of February.

Date	Traffic Captured	R/S.	A.M	VM	P.D
Feb. 1st	178,563	0.687	0.626	0.658	0.839
Feb. 5th	176,127	0.691	0.627	0.622	0.841
Feb. 10th	163,556	0.676	0.624	0.648	0.838

Feb. 13th	225,594	0.701	0.837	0.746	0.857
Feb. 18th	231,728	0.699	0.785	0.733	0.850
Feb. 21st	228,647	0.705	0.838	0.745	0.847

Cumulative average = 0.748

Discussion of Results.

1,204,215 TCP traffic volumes were captured, out of which 178,563, 176,127, 163,556, 225,594 231,728, and 228,647 traffic volumes were captured on 1th, 5th, 10th, 13th, 18th, and 21st of February 2018 respectively. It was observed that the ACF takes longer range to decay in lag of (-200,000 to 200,000) on 13th, 18th and 21st February, 2018 with the rest having the lag of (-150,000 to 150,000). However, it is noticed that, ACF does not depend on traffic volume (Danladi et *al*., 2017), the decaying characteristics is attributed to the presence of fractal traffic on wireless network (Rocha et *al*., 2011).

The selected period had the highest R/S Hurst parameter of 0.705 on the 21 February, with least Hurst parameter of 0.687 on the 1 with an average R/S Hurst parameter of 0.693 showing persistent characteristic in accordance with (Cody and Smith, 1997), the AM estimator had an average Hurst parameter 0.723 exhibiting persistent behavior, the VM had an average Hurst of 0.692 showing persistent characteristics and the PD estimator showed a persistent property all through the selected period with an average Hurst parameter of 0.845, this was in agreement with (Mohammed and Nickolay, 2017).

Conclusion

The magnitude of self-similarity on a wireless network determines the QoS available to the end users, it becomes very much necessary to have an insight into the network traffic self-similarity characteristics and its measurement. This research took a look at the degree of self-similarity on a wireless local area network and its measurement. A total of 1,204,215 TCP traffic volumes were captured, out of which 178,563, 176,127, 163,556, 225,594 231,728, and 228,647 traffic volumes were captured on 1th, 5th, 10th, 13th, 18th, and 21st of February 2018 respectively. It was observed that the ACF takes longer range to decay in lag of (-200,000 to 200,000) on 13th, 18th and 21st February, 2018 with the rest having the lag of (-150,000 to 150,000). The research confirmed the presence of LRD

mediterraneanresearchpub@gmail.com

traffic by applying autocorrelation test on the captured live traffic and it showed an autocorrelation function that decays hyperbolically. Four Hurst estimators were employed so as to complement each other's shortfall. An average daily Hurst parametere of 0.748, depicting a persistent characteristics.

Recommendations

Research has shown that network traffic is self-similar in nature and the self-similarity of traffic leads to lager queuing delays, packet loss, jitter and throughput in telecommunication network. Hence, the research recommends the use of neuro-fuzzy principles to x-ray the cause of self-similarity on a wireless network.

References

Danladi A., and Romanus I.O., (2018). Evaluation of Quality of Service on a Wireless Local Area Network Using Fractal Dimension. International Journal of Pure and Applied Science, 8(1):24-36.

Danladi A., Micheal Y. and Silikwa N.K (2017). Routing Protocol Source os Self-Similaritry on a wireless Network. *Alexandra Engineering journal* Article In Press.

Dependent Packet Traffic, *IEEE/ACM Trans on Networking*, 5 (1): 71 – 96.

Erramilli, Narayan and Willenger (2006). Experimental Queuing Analysis with Long-range

Figueredo D.R.,(2002). Autocorrelation structure of TCP traffic. Computer Networks Journal Special Issue in advances in modelling of Long Range Dependent traffic. 40(3):339-361.

Guo. L, Crovella M and Malta (2001) How TCP generate pseudo-self-similarity. MASCOTS Conference.

Leland W.E, Taqqu M.S, Willinger W and Wilson D.V., (1994) On the self similarity nature of Ethernet traffic, in Proc. *ACM SIGCOMM*: 183-193.

Mohammed A. and Nickolay Z. (2017) Network Link Dimensioning based on Statistical Analysis and Modeling of Real internet Traffic, International Journal of communications,7(3):17-23.

Paxson V, and Floyd D. (1997). Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic, *ACM SIGCOMM Computer Communication Review*, 27: 5-18.

Rocha F.G.C., Santos C.B. and Veira F.H.T. (2011) Traffic modeling in PLC Network using Markov Fluid Model ACF Fitting. *TEMA (Sao Carlos).* 12(4):467-478.

Roughan, Veitch, and Abry,(2011) Real Time Estimation of the Parameters of Long-Range Dependence, IEEE/ACM Transactions on Networking8(4): 467-478.

EDITOR-IN-CHIEF
Prof. G.O. Emerole [NIGERIA]

mediterraneanresearchpub@gmail.com