IPASJ International Journal of Electrical Engineering (IIJEE)
Web Site: http://www.ipasj.org/IIJEE/IIJEE.htm
A Publisher for Research Motivation.....
Email: editoriijee@ipasj.org
Volume 4, Issue 2, February 2016
ISSN 2321-600X

USING H-PARAMETER TO REDUCE SELF-SIMILARITY INTCP NETWORKS

Joseph Kweku Arthur[1], Emmanuel Kwame Mensah[2]

[1]Faculty of engineering, Ghana Technology University College
PMB 100, Tesano- Accra, Ghana

[2]Faculty of engineering, Ghana Technology University College
PMB 100, Tesano- Accra, Ghana

ABSTRACT

In the twenty first century, many people and machines are talking to each other using several types of communication systems. The greater the number of communication systems, the more varied and complex, the means of having them interconnected. Telecommunication operators need to provide quality of service (QoS) to their subscribers and to achieve that quality of service, operators need to focus on the control and the management of traffic. However, it has been observed that the traffic in a network communication system exhibits a problem called "Self-Similarity". This is a phenomenon, which occurs when the system is subject to larger queuing delays, higher drop rates and extended period of congestion. Therefore,the objectives of this study is to propose a method of eliminating self-similarity in the network communication system. The procedure used to solve this problem of Self-Similarity is as follows: the H-parameter was used to estimate the degree of self-similarity in the network communication system. To compute this H-parameter, the VTP (Variance Time Plot) of packets sniffed using Wire shark was analysed. Then, the principle of RED (Random Earlier Detection) was used to write an algorithm to reduce the level of self-similarity in TCP traffic. The results showed a reduction in the value of the H-parameter from 0.75 to 0.5 using RED. H-parameter value of 0.5 means that the network is good as the degree of self-similarity has been reduced. This implies that the RED application has improved upon the self-similarity in the network communication system

Keywords: QoS, self-similarity, H-parameter, variance time plot, random earlier detection,TCP

1. INTRODUCTION

In the twenty-first century, communication systems need to carry voice, data, video (V/D/V) and control information. Many people and machines are talking to each other using all kinds of systems than ever before in the history of the world. The greater the number of systems, the more varied and complex, the means of having them interconnected. The systems need to talk to each other over a connected information highway that can deliver information at high speed and without distortion. There is no doubt that in the coming years there will be a continuing demand for the ever-increasing quantities of data at ever-increasing speeds.

The quality of service available from wireline systems is of significant interest to both users and providers of those systems. Some telecommunication operators have to make commercial agreements with users to guarantee a particular quality of service (QOS)and to achieve that QOS, operators need to focus on one of the key aspect in telecommunication, the control and the management of the traffic.

The complexity of the traffic in a multimedia network is a result of integrating over a single communication channel, different types of traffic such as video, voice, and data that significantly differ in their traffic patterns, as well as their performance requirements. Specifically, "Bursty" traffic patterns generated by data sources and Variable Bit Rate (VBR) real-time applications such as compressed video and audio exhibits a certain degree of correlation between arrivals, and portray Long-Range Dependence in time (Self-Similar Traffic) [1]. The question that arises here are how prevalent such traffic patterns are and under what conditions are the performance analysis of the network, critically dependent on the self-similarity of the traffic in the network. The self-similar and multifractal nature of traffic can translate into a number of undesirable effects like high buffer overflow rates, large delays and persistent periods of

congestion [2]. Therefore, the objectives of this study is to propose a method of eliminating self-similarity in the network communication system.

2. LITERATURE REVIEW

2.1 Long-Range Dependence and Heavy-Tailed Distributions

Long-range dependent processes are characterized by an autocorrelation function, which decays hyperbolically [3]. This implies that the auto-correlation is non-summable, unlike the more conventional short-range dependent processes, which have auto-correlation functions that decay exponentially [4].

In addition, this distribution follows a power law. A random variable that follows a heavy-tailed distribution can take on extremely large values with non-negligible probability [5]. Heavy-tailed distributions can be used to characterize probability densities such as packet interarrival times and burst length.

2.2 The Hurst Parameter: A measure of Self-Similarity

The Hurst parameter H is a measure of the level of self-similarity of a time series. H-parameter takes values from 0.5 to 1. In order to determine if a given series exhibits self-similarity, an analytical method is needed to estimate the H-parameter for a given series. Currently, there are 4 approaches to doing that:

I. Analysis of the variances of the aggregated processes $X^{(m)}$
II. Analysis of the rescaled range (R/S) statistic for different block sizes
III. Periodogram
IV. A Whittle estimator.

In the first method, the variance time plot relies on a slowly decaying variance of a self-similar series. The variance of $X^{(m)}$ is plotted against m on a log-log plot. Then a straight line with a slope $(-\beta)$ greater than $-1$is indicative of self-similarity and the parameter .

In the second method, the R/S plot, uses the fact that for self-similar data, the rescaled range or R/S statistic grows according to a power law with exponent H as a function of the number of points including, n. Thus, the plot of R/S against n on a log-log plot has a slope which is an estimation of the $H-\text{parameter}$.

Periodogramis used to determine whether the generated sequence is a Long Range Dependent process or not. If the autocorrelations are summable, then near the origin of the frequency domain, the periodogram should be scattered round a constant level. If in the other case, the autocorrelation is not-summable, the point of the sequence are scattered around a straight line with a negative slope. The periodogramis obtained by plotting log (periodogram) againstlog (frequency). The slope of the graph is β and the Hurst Parameter is definedas:

$$H = (1-\beta)/2 \quad (1)$$

While the preceding two graphical methods are useful to estimate H, they may be biased for large values of H. The fourth method, a Whittle estimator, does provide a confidence interval. This technique uses the property that any long-range dependent process approaches fractional Gaussian noise (fGN) when aggregated to a certain level, and so should be coupled with a test of the marginal distribution of the aggregated observation to ensure that it has converged to the normal distribution. As m increases, short-range dependences are averaged out of the data set. If the value of H remains relatively constant, it is almost certain that this H value measures the true level of self-similarity of the data set.

2.3 Causes of Self-Similarity

Since self-similarity is believed to have a significant impact on network performance, understanding the causes of self-similarity in the traffic is important. Research has revealed that traffic generated by World Wide Web transfers shows self-similar characteristics [4]. Comparing the distributions of ON and OFF times, they found that ON time distribution was heavier-tailed than the OFF time distribution. The distribution of file sizes in the web might be the primary determiner of web traffic self-similarity. It has been presented that, the transfer of files whose sizes are drawn from a heavy-tailed distribution is sufficient to generate self-similarity in the network traffic. The ON and OFF periods do not need to have the same distribution. The results suggest that the self-similarity of web traffic is not a machine-induced artifact; in particular, changes in protocol processing and document display are not likely to remove the self-similarity of the web traffic [6].

In a realistic client/server network environment, the degree of which file sizes are heavy-tailed can directly determine the degree of traffic self-similarity at the link level. This causal relation is proven to be robust with respect to changes in network resources, network topology, the influence of cross-traffic, and the distribution of interarrival times. Specifically, measuring self-similarity via the Hurst parameter H and the file size distribution by its power law exponent α, has shown that, there is a linear relationship between H and α over a wide range of network conditions.

2.4 Network performance

2.4.1 Description

Well-defined metrics of delays, packet loss, flow capacity, and availability are fundamental ways for measurement and comparison of path and network performance. In general, users are mostly interested in metrics that provide an indication of the likelihood that their packets will get toit's destination in a timely manner. Therefore, estimates of past and expected performance for traffic across specific internet paths, not simply measures of current performance, are important. Users are also increasingly concerned about path availability information, particularly as it affects the quality of multimedia applications requiring higher bandwidth and lower latency, such as internet phone and videoconferencing. Availability of such data could help in scheduling online events such as internet-based distance education seminars, and influence user willingness to purchase higher service quality and associated service guarantees.

Given the ubiquity of scale-invariant burstiness observed across diverse networking contexts, finding effective traffic control algorithms capable of detecting and managing self-similar traffic has become an important problem.

The control of self-similar traffic involves modulating the traffic flow in such a way that the resulting performance is optimized. Scale-invariant burstiness introduces new complexities into optimization of network performance and makes the task of providing QoS together with achieving high utilization difficult.

2.4.2 The effects of Self-Similarity on Network Performance

Many analytical studies have shown that self-similar network traffic can have a detrimental impact on the network performance, including amplified queuing delay and packet loss rate. On the other hand, it has been found that long-range dependence was unimportant for buffer occupancy when there were strong short-range dependence and the Hurst parameter was not very large$(H < 0.7)$. However, they did not touch the case where there was strong long-range dependence with a larger Hurst parameter. [6]

One practical effect of self-similarity is that the buffers needed at switches and multiplexers must be bigger than those predicted by traditional queuing analysis and simulations. These larger buffers create greater delays in individual streams than were originally anticipated. The delay-bandwidth networks and QoS issues stemming from support of real-time multimedia communication have added further complexities to the problem of optimizing performance.

The effect of self-similarity on network performance is modulated by the protocols acting at the transport/network layer. An exponential trade-off relationship is observed between queuing delay and packet loss rate.

It is certain that a linear increase in buffer sizes will produce nearly exponential decreases in packet loss, and that an increase in buffer size will result in a proportional increase in the effective usage of transmission capacity [7]. With self-similar traffic, these assumptions do not hold. The decrease in packet loss with buffer size is far less than expected, and as can be seen from Figure 1, the buffer requirements begin to explode at lower levels of utilization for higher degrees of long-range dependence (higher values of H) [8].

It has been showed that for sources with large Hurst parameters, Markov chain models estimated the buffer sizes were not too large (no larger than 10ms for a single source), but these models might not estimate the cell loss rate and mean buffer size accurately for larger buffers. Also, studies has shown that queuing delay exhibited a superlinear dependence on self-similarity when buffer capacity was large. The queue length distribution decayed more slowly for long-range dependent sources than short-range dependent sources [6].

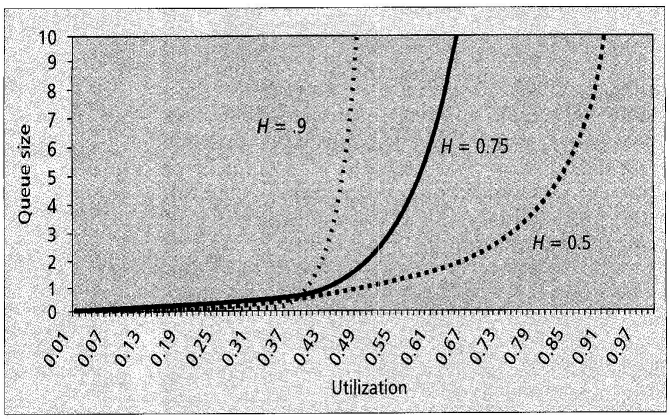


Figure 1 Queue size-utilization trade-off as self-similarity changes defined by the Hurst parameter

2.5 Analysis of traffic flow using WIRESHARK

Wireshark is a network packet analyser, known previously as Ethereal. It lets you examine the network traffic flowing into and out of your Windows or UNIX machine. Network professionals use Wireshark to troubleshoot networking problems, but it is also an excellent way to learn exactly how the network protocols work. For example, it allows you to see the data that your system sends and receives when you type a web address into a web browser (e.g., Internet Explorer or Mozilla's Firefox).

2.6 The problem of congestion

Communication service providers (CSPs) face the threat of churn if they fail to deliver the required quality and speed to their customers [9]. Four options can be proposed for the solution to this problem. The first option being service assurance. One must continually watch the network and ask these questions: How is the traffic flowing, are there the beginnings of potential traffic jam, what could be optimal traffic rerouting strategies? If something goes wrong, one must be able to immediately alert the CSP and, if possible, raise a ticket to resolve the problem. Two, the tools for frequency planning and optimizationmust be used. These are the 'hot' zones with the highest traffic density, and the 'cold' zones where bandwidth is potentially getting wasted? Can the bandwidth be re-distributedso that the capacity is optimized? Are there situations where parameters need to be 'reset'? Would it be possible to tilt or turn the antenna by a bit so that more capacity suddenly becomes available? The third option is to get telecom companies to collaborate more. All too often we encounter situations where one telecom carrier has built huge 'roads' in one geographical region but can't attract the traffic but there is a second carrier in the same region which owns a lot of traffic but doesn't have the roads! Software tools can be built that will facilitate such dynamic buying and selling of capacity. The last option which is being exploited in this paper and arguably the most powerful – is to become an intelligent and voracious user of analytics. Gather data, mediate it, clean it, enrich it, store it, join it, manipulate it and query it. Although telecom companies provide voice and data services to consumers and businesses, the nature of these services differs significantly between the two customer segments. While residential customers mainly use wireless services, businesses use wireline to get high-capacity broadband and advanced communications services [10].

3. METHODOLOGY

This section focuses on the research methodology used for the study. The primary data was obtained from the campus of the Ghana technology university college during peak hours.The VTP (Variance Time Plot) of packets sniffed using Wireshark was analysed to enable the computation of the H-parameter. Then, the principle of RED (Random Earlier Detection) was used to write an algorithm to reduce the level of self-similarity in TCP traffic. The method employed was basically experimental.

3.1 R/S analysis and Variance Time Plot (VTP) for the prediction of self-similarity

The traffic in a wireline system is a major problem in a telecom network whenever the media that is being used to communicate has to be connected by the mean of a wire. Simply because, it is the media that propagates the signal with less loss. However, it faces some constraints that do not allow a signal sent to be received properly. A particular problem that was noticed in the transmission of data through wired cable was self –similarity. Some solutionshave been provided to reduce these effects occurring on the signal, one of them being H-parameter.

Self-similarity occurs when the Hurst parameter is in the interval of [0.5, 1]. Self-similarity can lead to session interarrival times, heavy tailed distributions, larger delays queueing, higher drop rates and extended periods of congestion.

3.1.1 Variance Time Plot (VTP)

Recently, as internet users increase, network congestion happens frequently. Recent studies on packet measurement analysis in various networks, however, have shown that packet traffic exhibits Long Range Dependent (LRD) properties,which means large variance and self-similarity.

On the contrary, network resources more effectively, if networks are controlled considering the self-similarity properties. Only a few proposed network control methods requires that the Hurst Parameter, which is a typical, measure representing self-similarity be constant.

In this study, On-Time VTP (Variance Time Plot) method based on the VTP is proposed. It is the most general Hurst Parameter calculation method, and show its performance.

Variance time plot method,is used to evaluate the self-similarity of traffic by the computation of the Hurst-Parameter. Let $X = (X_\tau : \tau = 0,1,2 \ldots)$ be a covariance-stationary stochastic process with mean $\mu = E(X_\tau)$, variance $\sigma^2 = Var(X_\tau)$ and autocorrelation function $\gamma(\kappa)$. $\gamma(\kappa)$is given by:

$$\gamma(\kappa) = Cov(X_\tau, X_{(\tau+\chi)})/\sigma^2 \ , \ k = 0,1,2 \ldots \quad (2)$$

This process shows that, the packets arrival process X_t is the volume (bytes) of arrival packets at the t-th time slot (time slot =10 ms, for this study). For each $m = 1,2,3,\ldots$, let $X^{(m)} = (X_\kappa^{(m)} : \kappa = 1,2,3\ldots)$ denote the new covariance stationary time series obtained by averaging the original series X non-overlapping blocks of size m. That is, for each $m = 1,2,3,\ldots, X_\kappa^{(m)}$ is given by:

$$X_\kappa^{(m)} = \frac{1}{m}(X_{\kappa m-m+1} + \cdots + X_{\kappa m}), \kappa \geq 1 \quad (3)$$

In this case, for all $m = 1,2,\ldots$, the time series $X^{(m)}$ is the same covariance stationary stochastic process as the time series X, and mean, variance and autocorrelation function are respectively given by [11]:

$$E(X^{(m)}) = \mu, \quad (4)$$

$$\mathrm{Var}(X^{(m)}) = \frac{\sigma^2}{m} + \frac{2\sigma^2}{m}(m-j)\gamma(j), \quad (5)$$

$$\gamma^{(m)}(\kappa) = \frac{\sigma^2}{m^2 \mathrm{Var}(X^{(m)})}\left\{m\gamma(m-\kappa) + 2\sum_{j=1}^{m-1}(m-j)(m,\kappa+j)\right\} \quad (6)$$

If $X^{(m)}$ has the same self-similarity structure as X, then the following expressions hold,

$$\gamma^{(m)}(\kappa) = \gamma(\kappa), \quad (7)$$

$$\mathrm{Var}(X^{(m)}) = \sigma^2 m^{-\beta} \quad (8)$$

This implies that

$$\log(\mathrm{Var}X^{m}) = \log(\sigma^2) - \beta\log(m) \quad (9)$$

Plotting a graph of $\log(\mathrm{Var}X^{m})$ against $\log(m)$ gives a slope of $-\beta$ and an intercept of $\log(\sigma^2)$
In this case, the process X is second-order self-similar with the Hurst-Parameter equals:

$$H = 1 - \frac{\beta}{2} \quad (10)$$

In this study, the Hurst-Parameter adopted as an evaluation index of the self-similarity. If m is large enough, we can define asymptotical self-similarity.
The asymptotical self-similarity is often used in the field of network traffic theory, because it is more practical. In this case, self-similarity is defined as follows:

$$\mathrm{Var}(X^{(m)}) \sim cm^{-\beta}, \text{as } m \to \infty, c: \text{const.} \quad (11)$$

That is, the variance of the sample mean decreases more slowly than the reciprocal of the sample size m.

3.1.2 R/S analysis

H-parameter is the measure of the degree of the self-similarity for IP traffic. The H-parameter expresses the speed of decay of the series autocorrelation function. One of the method used for estimation of the H-parameter is called the R/S method.
In order to calculate the H-parameter, we need to calculate $R(n)/S(n)$ for different values of n. Then, we need to plot a diagram where $\log(E(R(n)/S(n)))$ is plotted on the y-axis and $\log(n)$ is plotted on the x-axis.
First, the expected value of the quality $R(n)/S(n)$ is computed asymptotically, followed by a power law:

$$E\left[\frac{R(n)}{S(n)}\right] \approx cn^H, n \to \infty \quad (12)$$

Where
c is a positive constant,
$R(n)$ is the adjusted range of the samples observed (in this case they were traffic samples sniffed from the GTUC network and expressed in bits),

$S(n)$is the sample standard deviation. Taking log on both sides of $E\left[\frac{R(n)}{S(n)}\right] \approx cn^H, n \to \infty$ (12, we have

$$Log(E(R(n)/S(n)) \approx Log(c) + HLog(n) \quad \textbf{(13)}$$

Plotting a graph of $Log(E(R(n)/S(n))$ against $Log(n)$ gives apositive slope of Hand it is given as:

$$H = \frac{Log[E(R(n)/S(n))]}{Log(n)} \quad \textbf{(14)}$$

$H$is the Hurst parameter with range 0.5 <H <1.
If it is denoted with X_i as the sequence of the samples, then the rescaled adjusted range $R(n)/S(n)$ may be calculated by using :

$$\frac{R(n)}{S(n)} = \frac{\max(0, W_1, W_2, \ldots, W_n) - \min(0, W_1, W_2, \ldots, W_n)}{S(n)} \quad \textbf{(15)}$$

Where

$$W_k = \sum_{i=1}^{n} iX - k\overline{X(n)}, \quad k = 1, 2, \ldots, n \quad \textbf{(16)}$$

The self-similarities inthe network traffic has been established. It has also been shown that this self-similar traffic can lead to larger queueing delays, higher drop rates and extended periods of congestion. In this section, the different aspects of the similaritywas discussed. The methods that must be implied to predict this phenomenon were also discussed.The impact of various buffer management algorithms on the self-similarity of network trafficwas also examined. As TCP is a transport protocol used to minimize the constraints of a heavy traffic in a network system, more specifically, wireline system, the impact of active and passive queue management policies used at the routers on the self-similarity of TCP trafficwas investigated.
The self-similar and multifractal nature of traffic can lead to a number of undesirable effects like high buffer overflow rates, large delays and persistent periods of congestion. The severity of these conditions is directly proportional to the degree of self-similarity or the Hurst parameter. A version of the Random Early Detection (RED) algorithm, which can be used to reduce the degree of self-similarity in TCP traffic, is also presented.
The impact of buffer management policies on the self-similarity of network trafficwas also looked at. In particular, how buffer management policies can affect the TCP related causes of self-similarity. The impact of current active and passive buffer management protocols on the timeouts and exponential backoffwere evaluated. Based on the results, a variation of Random Early Detection (RED) queue management algorithmwas developed. It showed a better performance than both passive tail drop as well as RED queues. The improvement in performance is both in terms of reduced self-similarity as well as throughput, loss rates and percentage of timeouts in the TCP flows. Our method also works even when non-TCP related causes of self-similarity are present, as long as TCP was used as the transport protocol. To support this claim simulationswith the presence of web traffic, heavy-tailed file sizes and typical user behaviour patterns is used. The self-similarity in the traffic is eliminated at low loads and significantly reduced at moderate and high loads if the proposed schemesare implemented.
In the next section, the impact of buffer management scheme on traffic with self-similarity is analysed using the proposed RED algorithm.

4. ANALYSIS OF DATA

4.1 Estimation of the degree of self-similarity

VTP method:

IPASJ International Journal of Electrical Engineering (IIJEE)
Web Site: http://www.ipasj.org/IIJEE/IIJEE.htm
A Publisher for Research Motivation.....
Email: editoriijee@ipasj.org
Volume 4, Issue 2, February 2016
ISSN 2321-600X

In the figure 1, the graph of $(\log(m), \log(\mathrm{Var}(X^{(m)})))$ is plotted. Let $X = (X_t : t = 0,1,2 \ldots)$ be time series of a volume (bytes) of arrival packets. $X^{(m)}$is calculated using to eq. (3) and$f(m) = (\log(m), \log(\mathrm{Var}(X^{(m)})))$is plotted. The value$-\beta = -0.6$is obtained from the plot (figure 2). Finally, the value of the Hurst Parameter using $H = 1 - \frac{\beta}{2}$ (10is obtained as 0.7.

Table 1. Data flow from GTUC network used to estimate the H-Parameter (Time slot = 10ms).

m	n	$X^{(m)}$	Var ($X^{(m)}$)	Log(m)	Log[Var ($X^{(m)}$)]
1	5	6,10,15,5,0.63	29.51138	0	1.47
2	6	9,11,3,16,14,10	20.3	0.301	1.31
3	7	1,3,2,2,2,2,2	0.333333	0.477	-0.477

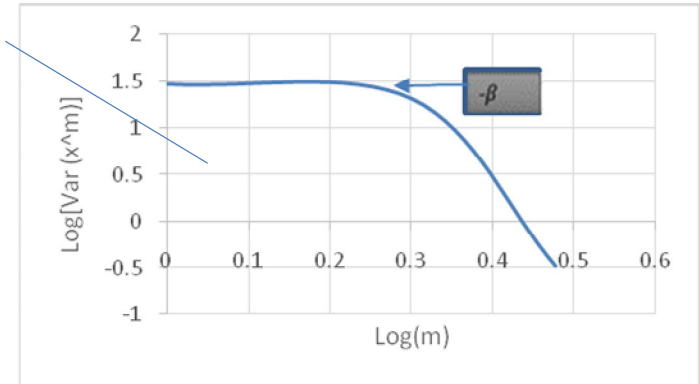


Figure 2. Variance Time Plot

Next, the procedure of calculating the Hurst Parameter by the VTP is explained. Let's assumed that $X^{(1)}$ represents the collected network packets with 10 ms time slot, and shown as follows:

$X^{(1)} = (6,10,15,5,0.63)$.

From $X^{(1)}$,$X^{(2)}$ and $X^{(3)}$is calculated and given as:

$X^{(2)} = (9,11,3,16,14,10)$

$X^{(3)} = (1,3,2,2,22)$.

This phase is called"Phase of making series (P_{ser})". After the derivation of $X^{(k)}$ $(1 \leq k \leq m)$, the following variance is calculated, $\mathrm{Var}(X^{(1)})$,$\mathrm{Var}(X^{(2)})$ and $\mathrm{Var}(X^{(3)})$.This phaseis called "Phase of calculating variance(P_{Var})". Using these two phases $(\log(m)$ and $\log(\mathrm{Var}(X^{(m)})))$is calculated, andthen, the Hurst Parameter is obtained.
With this method of estimating the H-Parameter, the general performance of the networkcan be deduced directly and it will be known if there is self-similarity during the transmission or not. At the same time, the degree of this phenomenon in the systemcan be known. The implication of the estimated value of the H-parameter equals 0.7 is also analysed using the R/S analysis

Table 2 Data flow from GTUC network for estimating H-parameter using R/S analysis**.**

N	Packets	Standard deviation	Max-min= R(s)	R/S	H
7	136	6.29436634	19	3.018	0.57
11	601	22.76081162	71	3.12	0.5
9	275	13.48249894	42	3.115	0.517

Given a sample of n observations $\{X_i\}$, subdivide the whole sample into m_k non-overlapping blocks and compute the rescaled adjusted range for each block. Thus, for a given value of k there are as many as m_k R/S-statistic values. Next, plotting $\log(R(t,k)/S(t,k))$ against $\log k$, the R/S plot is formed. Use the plotted samples to draw a straight line to fit all the data. The slope drawn is the estimated Hurst parameter.

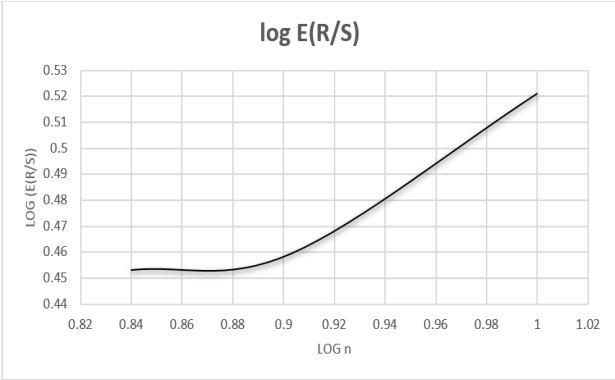


Figure 3 Evolution of the R/S analysis

The Graph above shows the variation of the H-parameter when the network has a problem of congestion.

4.2 Comparison of values of H-parameter without and with the application of dropping algorithm of modified RED

"Modified RED" is a change to RED's dropping policy. The algorithm for the packet dropping is shown below. The idea is not to drop any 2 consecutive packets, which arrive at the queue, unless of course if the queue is full. Since TCP generally sends back packets, ensuring that no 2 consecutive packets are dropped will reduce the probability that multiple packets from the same window are dropped, thereby **reducing** the occurrence of timeouts. Furthermore, an algorithm to reduce these effects on the network was proposed. In the following analysis, the effectiveness of this algorithm will be shown. The process is to take data from GTUC's network, especially at a time where students are browsing (during peak time). The values of the H-parameter with and without the use of RED algorithm is compared. The calculation of the H-parameter is done by the use of the 2 methods: VTP and R/S analysis.

Modified REDAlgorithm:

```
Last_drop_flag ⇐ 0
forEach Packet Arrival do
  iflast drop flag = 1 then
    last_drop_flag= 0;
    gotoenqueue;
else if min_th < avg < max_th then
  with probability d(k), drop the packet
    ifpacket is dropped then
      last_drop_flag= 1;
    end if
  else if max_th < avg then
    Drop the packet;
    last_drop_flag= 1;
  else
    gotoenqueue;
  end if
end for
```

Without RED Algorithm:

- **VTP**

With the data collected from the GTUC's network at peak time, the H-parameter using VTP calculationwas evaluated.The packets were sniffed at 5 different times. The analysis of data was done using the software "Wireshark".

A value of H-parameter approaching 1 was obtained. Meaning that the network exhibits a problem of timeout and congestion. The graph plotted below shows the way "β" was obtained.

Table 3 Data using VTP calculation without RED algorithm

X(m)	Var(X(m))	m	Packets
70,85,19,81,63,58	563.4667	1	376
32,91,54,42	498.6875	2	219
95,39,52,96	646.25	3	282
79,35,12,56	616.25	4	182
11,18,20	14.88889	5	49

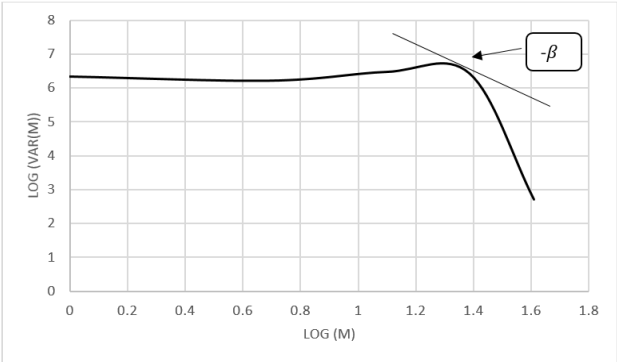


Figure 4H-parameter calculation without RED algorithm

From Figure 4 :-$\beta = -0.5$ and $H = 1 - \frac{\beta}{2} = 0.75$

It was observed that, the value of H is in the interval of 0.75<H<1. Meaning that the GTUC network is facing a problem of self-similarity. The packets were sniffed at a time where there were many students browsing.

<u>With RED Algorithm:</u>

- **VTP**

The idea was to remove the highest sequence of self-similarity form each packet.The behavior of the value of H-parameterwas obtained and compared with the values of the H-parameter without and with the RED algorithm. A value of H-parameter lower than the network value without the application of the RED algorithm was obtained. The table below shows the results obtained with the use of RED algorithm.

Table 4 Data of the VTP calculation with RED algorithm

n	Packets	Sequence X(m)	Var(X(m))	log(Var X(m))
5	291	70,19,81,85,63,58	443.76	6.095283876
3	128	32,54,42,91	80.888889	4.393090207
3	186	95,96,39,52	572.66667	6.350309635
3	103	79,35,12,56	322.88889	5.777311708
2	29	11,18,20	12.25	2.505525937

As for VTP, the graph below shows how the value "β" have been found.

Figure 5 Estimation of the H-parameter with RED algorithm

$$-\beta = \frac{X_2 - X_1}{Y_2 - Y_1} = \frac{0.5 - 0.4}{6.1 - 6.2} = -1$$

$$H = 1 - \frac{\beta}{2} = 1 - \frac{1}{2} = 0.5$$

From the calculation above, it can be seen that the value of the H-parameter is 0.5. This means that the degree of self-similarity in the network is low, which implies that, the network condition is better.
The main goal of the study was to find a solution to reduce the values of H when the network is congested or timeout. In the analysis, a method for calculating the degree of self-similaritywas compared. It was seen that VTP gave thesame value of H-parameter at all the times the data was picked. For the R/S analysis, different values of H-parameter was obtained at all times. The second part of the analysis evaluated the values of H-parameter without and with the application of the RED algorithm. Using the two methods of calculation, we saw that the VTP brings better results than the R/S analysis.
On theother hand, without the RED algorithm, and in an environment where many students were browsing, we derived values of "H" in the interval of 0.7 to 1. It meant that the network had a problem of self-similarity. With the application of the proposed RED algorithm, the VTP calculation saw a reduction of the H-parameter. Theoretically, if H is close to 0.5, it means that the network is better. The sequence of packets dropped is sent back again to the receiver and the receiver can properly resend these packets.

5. CONCLUSION

The management of telecommunication networks involves many parameters to maintain a guaranteed QoS for the users. In this study, the focus was on network conditions. The H-parameter was used to estimate the degree of self-similarity in the network. To compute this H-parameter, the VTP (Variance Time Plot) of packets sniffed using Wireshark was analyzed. Then, the proposed RED algorithm (Random Earlier Detection) was used to reduce the level of self-similarity in TCP traffic. The results showed a reduction in the value of the H-parameter from 0.75 to 0.5 using the proposed RED algorithm. H-parameter value of 0.5 means that the network is good as the degree of self-similarity has been reduced. This implies that the proposed RED application has improved upon the self-similarity in the network.

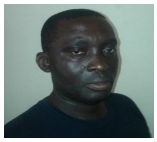
REFERENCES

[1] B. Sikdar and K. S. Vastola, "The Effect of TCP on the Self-Similarity of Network Traffic," in In Proc. of the 35th Conf. on Information Sciences and Systems, Pennsylvania, 2001.
[2] T. Omesh and S. Biplab, "On the Impact of IEEE 802.11 MAC on Traffic Characteristics," IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, vol. 21, no. 2, 2003.
[3] M. Li, W. Zhao and B. Chen, "Heavy-Tailed Prediction Error: A Difficulty in Predicting Biomedical Signals of 1/f Noise Type," Computational and Mathematical Methods in Medicine., New York, 2012.

[4] M. Kulikovs and E. Petersons, "Modelling the On-line Traffic Estimator in OPNET," ELECTRONICS AND ELECTRICAL ENGINEERING, vol. 7, no. 95, 2009.
[5] H. Kettani and J. A. Gubner, " A Novel Approach to the Estimation of the Hurst Parameter in Self-Similar Traffic," in IEEE Conference on Local Computer Networks , LCN, 2002.
[6] Z. Sahinoglu and S. Tekinay, "On multimedia networks: Self-similar traffic and network performance," IEEE explore, vol. 37, no. 1, 2013.
[7] D. Yamkhin and Y. Won, "Modeling and Analysis of Wireless LAN Traffic," JOURNAL OF INFORMATION SCIENCE AND ENGINEERING, vol. 25, 2009.
[8] A. Suarez-Gonzalez, C. Lopez-Ardao and C. Lopez-Garcia, "A batch means procedure for mean value estimation of processes exhibiting long range dependence," in ETSE de Telecomunicacion, Spain, 2002.
[9] S. Bhogle, "How telcos can leverage analytics to solve network congestion," Tele-Talk, India, 2014.
[10] R. Sheffer, "An overview of the US telecom industry," Market Realist, New York, 2016.
[11] University of Washington, Time Series Concepts, Washington: University of Washington, 2015.

AUTHORS

Kweku Arthur is a final year PhD student at the ST. Petersburg State University of Telecommunications. He obtained his BSc and MSc degree in the same university. He lectures at the Ghana Technology university college. His research areas are wireless technology and emerging trends in mobile networking



Emmanuel Kwame Mensah is a lecturer at Ghana Technology University College. Heobtained both BSc and MSc from Kwame Nkrumah University and Science Technology Ghana in the field of Mathematics. He specialized in Mathematical Control Theory. Moreover, he has another degree in Postgraduate Diploma in Education from University of Education, Winneba -Ghana. His research area is stochastic optimal control theory.