Bandwidth Control Algorithm on Youtube Video Traffic in Broadband Network

Murizah Kassim, Aini Azmi, Ruhani Ab.Rahman, Mat Ikram Yusof, Roslina Mohamad, Azlina Idris
Faculty of Electrical Engineering, Universiti Teknologi MARA,
40450 UiTM Shah Alam, Selangor, Malaysia.
murizah@salam.uitm.edu.my

***Abstract*—This paper presents an analysis of YouTube video traffic and fitted to best distribution traffic model to control bandwidth usage in a broadband network. The study scope comprised of collections of inbound YouTube video traffic for 7 days with the time-interval of each day is 3 hours. The broadband network is supported at 10Gbps line speed to Wide Area Network (WAN). The objective of this research is to characterize YouTube video traffic on broadband network, to fit the original traffic to best traffic model and bandwidth control algorithm called Policing and Shaping is developed based on time based threshold for 0.5Gbps at night and 1.0Gbps in day time. Performance shows the bandwidth controlled as bandwidth save, reduced traffic burst and processing time. Results present benefits of the developed algorithms where enhancement in processing time is 25.25% and the bandwidth is saved about 7.1668Mbps with Policing algorithms. Shaping algorithm process presents performance of processing time is increased up to 55.26% and the bandwidth is saved for about 25.548Mbps. Results also present best Cumulative Distribution Functions (CDF) traffic model using Maximum Likelihood Estimator (MLE) technique four best traffic models is identified which are Extreme Value, Weibull, Normal and Rician traffic model. Among the four, Weibull shown as the best fitted model that presents value of MLE=-1178.4 with the Scale α=9.49411e+08 and Shape β=2.81324 for 2 parameters traffic modeling. Research benefits in the development of design algorithm for Network Quality of Services (QoS) especially for bandwidth control and performance.**

***Index Terms*—Bandwidth Control; Cumulative Distribution Function; Eeibull; Extreme Value; Maximum Likelihood Estimator; Video Traffic Model; Youtube.**

I. INTRODUCTION

Internet played a very important role in this era. Nowadays, most of activities in this world need to use internet to complete each of task or to communicate among each other either for a short distance or long distance communication. There are many applications built from the use of internet such as social networking sites, online streaming sites and the latest is cloud computing. The application of video traffic is widely use either for individual, company or commercial such as YouTube, television websites or individual websites which provide the free drama online streaming. A research presented YouTube accounted more than 50% of the peak download traffic on North American fixed networks in 2015. Video streaming represents a significant source of Internet traffic. Survey has looks at emerging research into the application of client-side, server-side and in-network rate adaptation techniques to support DASH-based content delivery. It motivated on the application techniques which reviews the important of notable video traffic measurement and characterization studies [1]. The use of video traffic usually involved high speed of bandwidth data to avoid buffering and delayed to the user. Because the number of users is increasing year by year, the demand to supply the internet bandwidth are getting high. Internet Service Provider (ISP) need to find a way to allocate the bandwidth to the customer as per requested to avoid traffic delay or bad service happened [2]. The application of video traffic consists of two components which are sound and image [3]. Both components are varying based on the length of the video itself. In this situation, the long the length of the video, the more bandwidth is used and the process of sending the video or uploading the video to the internet is longer time. Here is one of the applications by the user which caused to high bandwidth utilization. In order to control the bandwidth usage, there are many modeling and algorithms that can be applied in the video traffic network to make sure customer satisfied with the service. Traffic shaping is a process of controlling the bandwidth by passing the burst data to the available bucket data below the committed rate with a bit time of delay. It can be applied at the interface of the networks to manage the incoming and outgoing traffic. Traffic shaping can be used to avoid saturation from happening and it can control and manage network latency [4]. Therefore, the Demand Side Management algorithm was purposed to reduce the delay and enhanced the performance of the service.

The research of behavior of network traffic has been done and it shows that traffic policing and shaping can only decrease the data rates of the network but the aggregate traffic behavior is not affected [5]. In other way, the bandwidth usage can be control by identifying the model of the distribution which suite to the video traffic data. Some of the research present that the modeling traffic can reduce the traffic burst and at the same time can control the bandwidth use in customer side [6]. It is use the Maximum Likelihood Estimator (MLE) technique to find the best distribution model to be use by fitted the real data to Cumulative Distribution Function (CDF) using MATLAB software. It is also compared four type of distribution to get the best method and the value of the parameter in the distribution is used to control the network traffic flow. There are also some other test being used to compare the distribution such as the application of Goodness-of–Fit statistical toolkit which produced the best algorithm in the internet traffic which is not be used in this research [7]. There are many different research have been done to control the traffic in the network and it depends on the component that involved such as the traffic analysis on the feature like packet size, frequency, inter-arrival time, traffic burst and others [8]. Generalize Extreme Value model being used to model the extreme events by using the block maxima approach. Weibull distribution model was applied to capture the transformation of inter-arrival process and the parameter being used to zoom in from session to flow and to packet level inter-arrivals. Some research use the parameter value of α (scale) to control their traffic but in this research, the value of β (shape) being used[9].

This paper presents a statistical analysis on video traffic network and best fitted distribution for YouTube data. It is supported 10Gbps Committed Access rate speed line. Throughput video traffic is collected for 7 days with 56 numbers of tracers in GByte. The inter-arrival time for each tracer is 3 hours. The throughputs are fitted to CDF using MATLAB software and the values of MLE log are taken. Best four distributions that show four highest MLE log value are identified. Daily and night based YouTube throughput on bandwidth control is developed and presents the network or traffic performance in bandwidth saved, burst control and processing time. YouTube throughput distributions model and its parameters is identified that can be used in next control algorithms based on videos traffic in network.

II. Literature Review

Two important literatures area which are important as taken as guidelines in this research is the existence of bandwidth control algorithms and the traffic distributions model. Previous research is identified as reviews.

A. *Traffic Control Algorithms*

There are many control algorithms being used in video traffic network. It is depends on the subject that the research is focus on such as to reduce the burst, to control the bit rate, delay, packet receive and others. One of the researches was done in controlling the delay by using Leaky-Bucket based rate control algorithm. This algorithm can reduce the burst in the video traffic and the video can be transmitted smoothly without buffer [10]. In other research, it is focusing on the packet size on the video streaming in LTE network. The novel scheduling algorithm is presented to give the better service in LTE network by applying the algorithm at the enhancement layer[11]. This algorithm can improve the packet delivery ratio and serve the better service . Data burst always happened when there is heavy flow of traffic occurred in the network. To overcome this situation, Bit Rate Throttling algorithm was proposed to reduce the burst in RTP which is use to transfer the real time data in the network. This algorithm is focusing to control the burst at the video encoder by separated the RTP data into packet before sending to the destination. This algorithm can avoid the data spike occurred and can reduce buffer in the video traffic network[12] . Recent algorithms developed on internet campus traffic which focus on inbound internet traffic control bandwidth and traffic burst [9, 13]. This research present traffic performance in bandwidth control and processing time. Performance in networking is very important to make sure the service can be delivering in a good condition. In order to complete this performance, the traffic model should be apply with the best defined parameters in each distribution model. Several distribution model has been identified that can be used to measure the traffic.

B. *Youtube Traffic Distribution Model*

A research presented a study on Characteristics of YouTube network traffic at a campus network that measures, models, and analysis its implications. Based on the measurements, traffic video presents duration’s time, data rate of streaming sessions, the popularity of videos, and access patterns for video clips from the clients in the campus network. The analysis of the traffic shows that trace statistics are relatively stable over short-term periods while long-term trends can be observed. It is also analyzed that it is benefits of using alternative distribution infrastructures to improve the performance of a YouTube-like Video on Demand (VoD) service. The results of these simulations show that P2P-based distribution and proxy caching can reduce network traffic significantly and allow for faster access to video clips [14]. Another research presents a study on the relationship between popularity and locality of online YouTube videos. Investigation on whether YouTube videos exhibit geographic locality of interest, with views arising from a confined spatial area rather than from a global one is presented. The analysis is done on a corpus of more than 20 million YouTube videos, uploaded over one year from different regions. It presented that about 50% of the videos have more than 70% of their views in a single region. Generally it concluded that by relating locality to viral the study show that social sharing generally widens the geographic reach of internet videos [15]. Research examined usage patterns, file properties, popularity and referencing characteristics, and transfer behaviors of YouTube, and compare them to traditional Web and media streaming workload characteristics. The research concluded the paper with a discussion of the implications of the observed characteristics example with the traditional Web, caching could improve the end user experience, reduce network bandwidth consumption, and reduce the load on YouTube's core server infrastructure. Unlike traditional Web caching, Web 2.0 provides additional meta-data that should be exploited to improve the effectiveness of strategies like caching [16].

A Workload simulator is required for evaluating the methods addressing the problems of high bandwidth usage and scalability of Web 2.0 sites such as YouTube [17]. The distribution models, in particular Zipf-like behavior of YouTube popular video files suggests proxy caching of YouTube popular videos can reduce network traffic and increase scalability of YouTube Web site. YouTube workload characteristics provided in this work enabled us to develop a workload generator to evaluate the effectiveness of this approach. Thus, handling Quality of Services for video traffic is essential. Adaptive of dynamic control has been develop but not particularly concentrated on video or YouTube traffics throughput [18, 19].

III. Methodology

Figure 1 presents the methodology flowchart started form data collections, design algorithm and analysis performance.

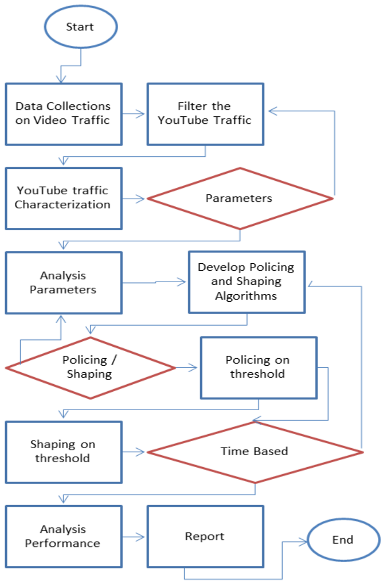

Figure 1: Methodology Flowchart

A set of video traffic data was collected in 7 days. The inter-arrival time for each data was measured in 3 hours each. It is containing of 4 times in a day and 4 times in a night with the total number of tracers are 56 for a week. Inter-arrival data collections for every 3 hours is done base on the limitations of monitoring and collected data by the applications used. Then, statistical analysis is done by comparing the best fitted cumulative distribution function (CDF) with the highest Maximum Likelihood Estimator (MLE) value among models. Collection of YouTube traffic is run with policing and shaping algorithms process by setting up a time based control which is in day and night. Different two threshold values are used which are 0.5Gbps at night and 1.0Gbps in day time to control the receive bandwidth. The policing process is to control the burst bandwidth used by users. Normally, burst happened in video traffic because of the overload of data flows in the network. After policing process, the shaping process is performed to fill the burst in available bucket to prevent the byte lost where some delay happened in processing time. But shaping algorithms will saved the byte and reduce network capacity in time is presented. Policing and shaping using traffic distribution model with identified parameter is done in the second phase in future project. The parameter from the selected model is used to control the bandwidth by adjusting its value.

IV. Analysis and Result

Video data are collected for 7 days with 3 hours interval and it is analyzed. Three hours interval is taken because it is the limitations of the used monitoring and collections traffic applications software.

A. Statistical Analysis

Figure 2 presents the real data of throughput which collected in 7 days with the total number of tracers is 56 tracers. The data is taken in Byte and translated to GByte (GB). The graph shows a heavy throughput in day time which is from 7am to 6pm while for night time from 7pm to 5am, the reading of the throughput is less compared in day time. The minimum traffic flow for YouTube is about 0.2951 GB and maximum value is about 1.320 GB. The mean value is estimated as 0.843 GB.

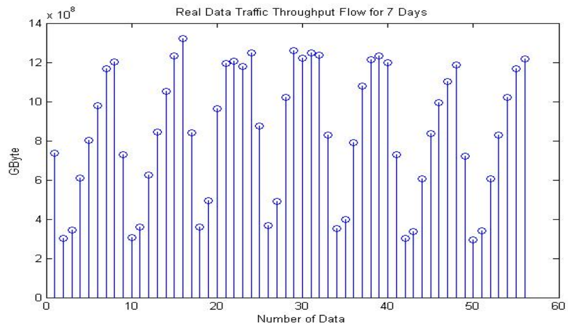

Figure 2: Throughput Real live data collected for 7 days

B. Control Algorithms with Policing

a. Policing at Day Time

Figure 3 presents real traffic flow before policing and after policing in day time. The green line is the Committed Rate Level and Burst traffic shows a 1Gbps rate bandwidth which has been set for YouTube throughput that represents a threshold guide. The red color line shows burst traffic exists in the network. In policing, all the burst bandwidth will be cut off. Policing algorithms is to control burst thus bandwidth can be saved although this process caused byte loss.

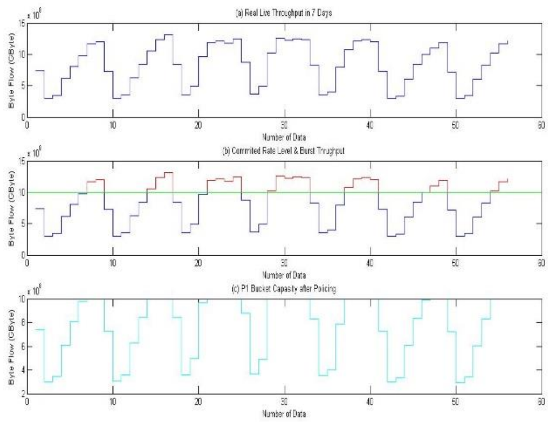

Figure 3: Policing in Day Time

Figure 4 presents the CDF graph for day time before and after applying the policing algorithm. Policing algorithms reduced the throughput according the threshold which present about 0.003 Gbps total bandwidth is saved, 4.211 GB burst is controlled and a difference of 3.12E+11 seconds time is process faster in time for the YouTube traffic. Overall analysis of policing YouTube Traffic at day time is presented as in Table 1.

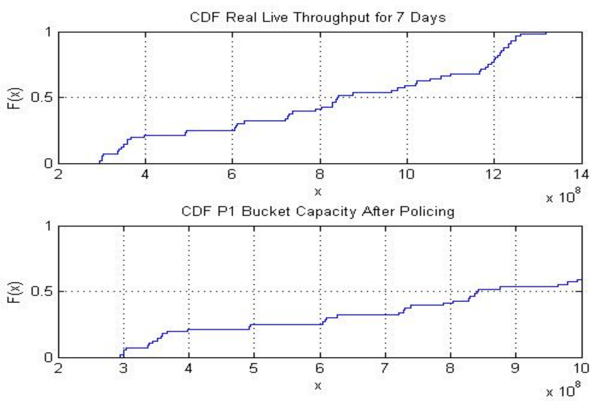

Figure 4: CDF before and After Policing in Day Time

Table 1
Analysis YouTube Throughput on day time

Parameter	Value
Total bandwidth save (bps)	3.1193e+06
Total burst shape (byte)	4.2111e+09
Byte lost (byte)	4.2111e+09
Total Process time (before) (sec)	3.4965e+12
Total Process time (after) (sec)	3.1846e+12
Different Process Time (sec)	3.1193e+11
Minimum value (byte)	2.9507e+08
Maximum value (byte)	1.3194e+09
Mean value (byte)	8.4291e+08

b. Policing at All Time

Figure 5 shows the policing are applied at all time by using different threshold for day and night. This condition can be referred in the CDF graph for policing day and night which there is an increasing value of F(x) at the point of x=5, which is the threshold for night time is 0.5Gbps and day is still set at 1Gbps as in Figure 6.

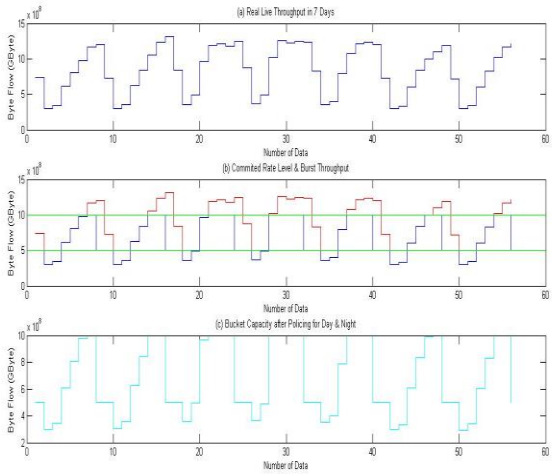

Figure 5: Policing at All Time

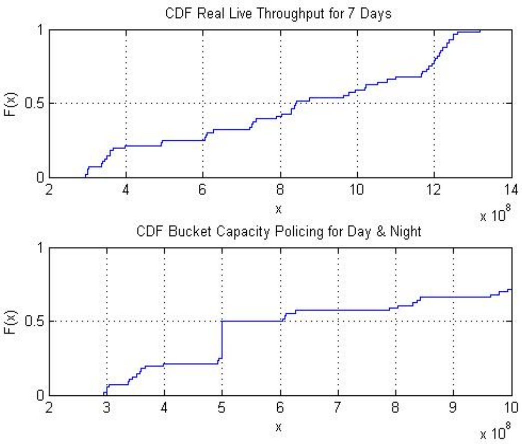

Figure 6: CDF before and after policing at All Time

The readings of the policing graph also being reduce because of the cut off burst in the traffic. The analysis traffic performance is referred in Table 2. The percentage of enhancement in processing time is 25.25% and the bandwidth is saved by applying the policing algorithm is about 7.1668Mbps. Byte loss is happened due to control of bandwidth over threshold. Compared to policing algorithms for day time, it is identified that bandwidth is saved larger when policing is done at all time.

Table 2
Policing All Time

Parameter	Value
Total bandwidth save (bps)	7.1668+06
Total burst shape (byte)	9.6752e+09
Byte lost (byte)	9.6752e+09
Total Process time (before) (sec)	4.9125e+07
Total Process time (after) (sec)	3.6717e+07
Different Process Time (sec)	1.2407e+07
Minimum value (byte)	2.9507e+08
Maximum value (byte)	1.3194e+09
Mean value (byte)	8.4291e+08

c. Shaping YouTube Traffic at All Time

The solution to overcome the byte loss in the policing process is done by applying shaping algorithm. Shaping process will pass the burst data to the next available bucket as the traffic is below the committed rate. Time delay will be happened but bandwidth is saved without dropping the packet data. This is proved by referring to the Figure 7, which is the bucket data in Committed Rate Level and Burst Throughput graph shown that all the burst in the network is pass to the next available bucket to save the bandwidth. The cumulative CDF graph in Figure 8 shows readings which near to consistent value after applying shaping algorithm. Table 3 presents analysis when Shaping is applied where processing time is reduced and the high value of bandwidth used is saved. Results also presents that the higher the bandwidth saved, the speed of the traffic is increased. Traffic delay and buffer capacity also is reduced. Analysis result for shaping presents performance of processing time is increased up to 55.26% and the bandwidth is saved for about 25.548Mbps.

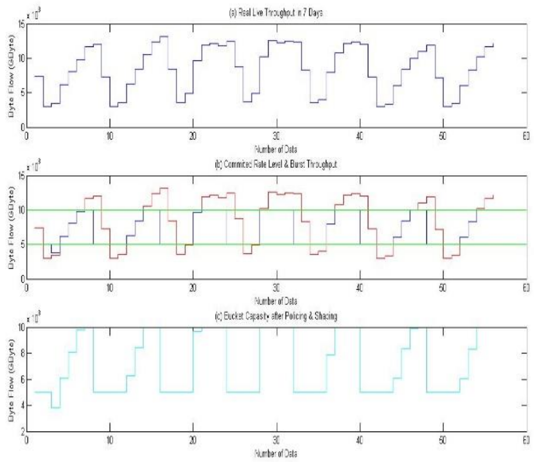

Figure 7: Shaping at All Time

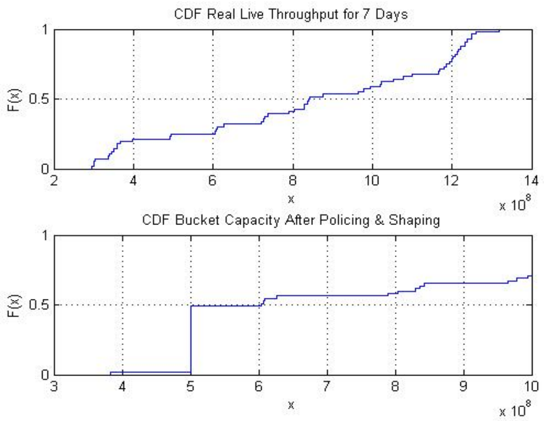

Figure 8: CDF graph for shaping at All time

Table 3
Analysis On Shaping Youtube Throughput

Parameter	Value
Total bandwidth save (bps)	25548e+07
Total burst shape (byte)	34490e+10
Byte lost (byte)	8.3323e+09
Total Process time (before) (sec)	8.6513e+07
Total Process time (after) (sec)	3.8707e+07
Different Process Time (sec)	4.7806e+07
Minimum value (byte)	3.8338e+08
Maximum value (byte)	2.5863e+09
Mean value (byte)	1.4588e+09

d. Fitted Traffic Distribution Model

Figure 9 shows the selected distribution model for the highest four type of distribution that fitted to CDF base on the measurement on MLE log value. The analysis for the graph has shown in Table 1 with the highest MLE log value is -1177.56 for Extreme Value Distribution. Extreme value distribution has 3-parameters used as presents in its mathematical equation. Weibull is identified as a 2-parameters distribution model with the MLE log value -1178.4 that shows its parameter value of α (shape)= 9.49411e+08 and β (scale)= 2.81324.

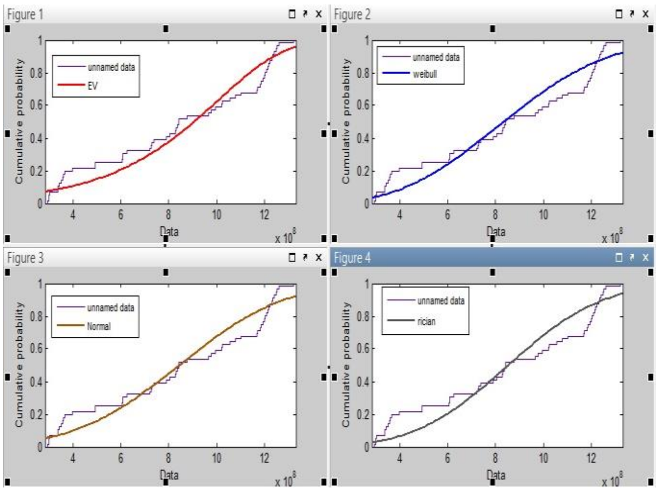

Figure 9 : CDF fit on best MLE throughput

Table 4 shows all fitted identified parameters value for YouTube traffic distribution model. This parameters will be used in next second phase of develop control algorithms using best distributions model and analysis of control algorithms will be presented for its performance value.

Table 4
Best MLE For Youtube Traffic

Traffic Model	MLE	Parameter	
		μ	σ
Extreme Value	-1177.56	8.42914e+08	2.76779e+08
		A	β
Weibull	-1178.4	9.49411e+08	2.81324
		μ	Σ
Normal	-1179.4	8.42514e+08	3.42156e+08
		S	Σ
Rician	-1179.84	7.88109e+08	3.19663e+08

V. Conclusion

The analysis for video traffic flow, develop control bandwidth algorithms and the best modeling in fitted distribution is presented. Result presented the traffic performance of develop algorithms on Policing and Shaping YouTube traffic. It is identified burst is control and bandwidth is saved for both applied algorithms. Results present benefits of the develop algorithms where enhancement in processing time is 25.25% and the bandwidth is saved about 7.1668Mbps with Policing. Shaping algorithm process presents performance of processing time is increased up to 55.26% and the bandwidth is saved for about 25.548Mbps. Traffic distribution model also is presented on YouTube traffic where four best fitted distributions are identified which is Extreme Value, Weibull, Normal and Rician. All distribution model on YouTube traffic is identified and will be used in next develop algorithms such as Weibull distribution is selected as the best video traffic distribution using 2 parameter distribution and Extreme Value Distribution is identified for 3-parameters distribution. Overall result shows the improvement of the traffic in term of speed and bandwidth controlled on YouTube traffic flow. This research is benefits in identifications of tele-traffic engineering algorithms for QoS in a network.

Acknowledgment

Authors would like to thank Faculty of Electrical Engineering, Universiti Teknologi MARA for the support grant in publishing this research.

References

[1] J. Kua, G. Armitage, and P. Branch, "A Survey of Rate Adaptation Techniques for Dynamic Adaptive Streaming over HTTP," *IEEE Communications Surveys & Tutorials*, 2017.
[2] A. M. AlAdwani, A. Gawanmeh, and S. Nicolas, "A Demand Side Management Traffic Shaping and Scheduling Algorithm," in *Modelling Symposium (AMS), 2012 Sixth Asia*, 2012, pp. 205-210.
[3] K. Salah, F. Al-Haidari, M. Omar, and A. Chaudhry, "Statistical analysis of H. 264 video frame size distribution," *IET communications*, vol. 5, pp. 1978-1986, 2011.
[4] M. F. Alam, M. Atiquzzaman, and M. A. Karim, "Traffic shaping for MPEG video transmission over the next generation internet," *Computer communications*, vol. 23, pp. 1336-1348, 2000.
[5] D. Daniel-Simion and G. Dan-Horia, "Traffic shaping and traffic policing impacts on aggregate traffic behaviour in high speed networks," in *Applied Computational Intelligence and Informatics*

(SACI), 2011 6th IEEE International Symposium on, 2011, pp. 465-467.
[6] M. Kassim, M. A. Abdullah, and M. M. Sani, "Internet traffic analysis with goodness of fit test on campus network," 2006.
[7] S. Guatelli, B. Mascialino, A. Pfeiffer, M. Pia, A. Ribon, and P. Viarengo, "Application of statistical methods for the comparison of data distributions," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, 2004, pp. 2086-2090.
[8] F. Zhang, W. He, and X. Liu, "Defending against traffic analysis in wireless networks through traffic reshaping," in *Distributed Computing Systems (ICDCS), 2011 31st International Conference on*, 2011, pp. 593-602.
[9] M. Kassim, M. Ismail, and M. I. Yusof, "Adaptive throughput policy algorithm with weibull traffic model for campus IP-based network," *Information Technology Journal*, vol. 13, pp. 2632-2644, 2014.
[10] J. Bai, Q. Liao, X. Lin, and Y. Zhang, "A rate control algorithm for VBR video encoding and transmission," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, 2000, pp. 1931-1934.
[11] Y.-C. Lin, W. K. Lai, K.-T. Yang, and M.-T. Lin, "An novel scheduling algorithm for video stream in lte," in *Genetic and Evolutionary Computing (ICGEC), 2012 Sixth International Conference on*, 2012, pp. 107-110.
[12] M. Modi and S. Kasula, "Bit rate throttling algorithm on video over RTP," in *Engineering (NUiCONE), 2013 Nirma University International Conference on*, 2013, pp. 1-4.
[13] M. Kassim, M. Ismail, and M. I. Yusof, "A new adaptive throughput policy algorithm on campus ip-based network internet traffic," *Journal of Theoretical and Applied Information Technology,* vol. 71, pp. 205-214, 2015.
[14] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Characteristics of YouTube network traffic at a campus network–measurements, models, and implications," *Computer Networks,* vol. 53, pp. 501-514, 2009.
[15] A. Brodersen, S. Scellato, and M. Wattenhofer, "Youtube around the world: geographic popularity of videos," in *Proceedings of the 21st international conference on World Wide Web*, 2012, pp. 241-250.
[16] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "Youtube traffic characterization: a view from the edge," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 2007, pp. 15-28.
[17] A. Abhari and M. Soraya, "Workload generation for YouTube," *Multimedia Tools and Applications,* vol. 46, p. 91, 2010.
[18] M. Kassim and N. A. Ayop, "Adaptive Policing Algorithms on inbound internet traffic using Generalized Pareto model," in *Internet Technology and Secured Transactions (ICITST), 2016 11th International Conference for*, 2016, pp. 217-222.
[19] M. Kassim, R. Samsuri, M. I. Yusof, R. Ab Rahman, N. M. Isa, R. Mohamad, and M. Ismail, "Video Traffic Modeling using Kolmogorov Smirnov Analysis in Broadband Network," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC),* vol. 8, pp. 49-53, 2016.