Modelling TCP/IP Traffic of a Convergent Campus Wireless Network

Albert Espinal, Rebeca Estrada, Carlos Monsalve

Abstract— With the deployment of new devices, protocols and applications, network traffic is changing to adapt to these trends. Therefore, it is necessary to analyze the impact over services and resources in data networks. Traffic classification of network is an important requirement to optimize traffic engineering and adequately provision quality of service. In this paper, we propose to analize the traffic in an university campus wireless network, throught the collected data by means of a novel sniffer that ensures the user data privacy. We focus in packet size. The results show that this traffic has a bimodal behaviuor with packets around 60 and 1300 bytes. It is also observed that IPv4 packets represents a big impact over IPv6, mainly TCP packets. And applications such as SSL and HTTP mark this trend. Numerical parameters for poisson distribution are presented in order to compare and simulate such traffic.

Keywords— packet size, sniffer, traffic classification, traffic modeling.

I. Introduction

UNDERTANDING and analyzing the network traffic is an important requisite for planning the network security policies, provide quality of service to applications, and for optimizing the network resources such as bandwidth and delay.

Additionally, it is important to consider factors that influence over traffic, such as the deployment of IPv6 on the Internet, the massive use of applications, and new technologies and devices. The study from Cisco Systems: forecast and trends [1], predicts that by 2022, each person will generate a monthly traffic of 50 GB, compared to 16 GB in 2017. It is expected that the number of network devices will grow from about 18 billion in 2017 to about 28,5 billion in 2022. It is predicted that smart mobile traffic represents 44% compared to 18% in 2017. The traffic from wireless and mobile device will account 71 percent of total IP traffic. Regarding the applications, it is expected that the IP video represents 82% of the global traffic.

In packet-based networks, like the internet or the Wireless Local Area Networks (WLANs), the transmission of information is performed in discrete packets [2]. For analyze and modelling the network traffic, we can to considerate two variables: the packet size and the inter-arrival time [3]. This study is focus on packet size (or packet length). This variable has a stochastically behavior [4][5] which is monitored for the corresponding analysis.

In practice we can measure the traffic network by means of active polling and passive monitoring [6]. The active method generates new traffic, inject it into the network, while passive method consists on monitor, and capture the network traffic. One drawback of the passive method is the privacy of the data to be captured, because the traditional packet sniffers saves the entire packet: headers and payload [7]. We use a novel sniffer that process and save only the header for analysis. The passive measurement can be performed at various levels like byte, packet, flow, and session [8][9]. We centered this study at packet level because is independent of the protocols, and avoid the encrypted payload.

In this work, we proposed to analyze the traffic of a convergent campus wireless network, determine the contribution of protocols and applications, and estimate statistical models that represent and simulate these traffics.

The rest of the paper is organized as follows: section II provides information about related works; in section III we show the data collection, classified by type of traffic, by protocols, and by application, according to the variable packet size. Section IV presents the traffic model that characterize the realistic traffic analyzed. The paper ends with the conclusion in section V.

II. Related Works

Many works have analyzed the network traffic based on packet size, using methods such as statistical analysis, pattern recognition, length of the application messages, packet flows, user behavior, etc. Additionally, these studies had suggested models to simulate the realistic network traffic.

In [10], Zhang et al. presented a state of the art about traffic classification with emphasis in methods based on exact matching, machine learning, and heuristic methods. Around the year 2000, the internet traffic that was tri-modal with packet sizes around 40, 765 and 1.500 bytes [11]. In [12], Sinha et al. observed that the internet traffic was bimodal at packet sizes of 40 and 1500 bytes. Wu et al. in [13] analyzed flow records in an internet service provider and classified this by applications using machine learning. A study for

Albert Espinal is with the Escuela Superior Politecnica del Litoral, ESPOL, Electrical Engineering and Computer Science Faculty & ReDIT Research Group, Guayaquil, Ecuador (phone: 593-958766229; e-mail: aespinal@espol.edu.ec).

Rebeca Estrada, is with the Escuela Superior Politecnica del Litoral, ESPOL, and ReDIT Research Group, Guayaquil, Ecuador (e-mail: restrada@espol.edu.ec).

Carlos Monsalve is with the Escuela Superior Politecnica del Litoral, ESPOL, and Electrical and Computer Engineering Department, Guayaquil, Ecuador (e-mail: monsalve@espol.edu.ec).

main application for this behavior is MDNS (around 80 bytes). Other UDP applications contribute with packets between 60 and 300 bytes in a sparse form.

The analysis of IPv6 traffic show that contribute with small packets around 80 bytes with 66.44%, mainly ICMPv6 packets. TCP and UDP traffic over IPv6 are still limited in this university campus wireless network. MDNS over UDP, is the most relevant.

IV. Traffic Modeling

Taking into account the analysis of the network traffic analyzed in the previous section, we estimate some models using the Poisson probability distribution function, based on traffic type, protocols and applications.

For total traffic presented in fig. 2, results a fitted model as a mixture of two Poisson distributions with parameters λ_1 = 93.22, and λ_2 = 1270.11. The probability that the length of a packet belongs to the first distribution is 0.448, while for the second distribution the probability of a packet following that distribution is 0.552. Finally, the model is the result of the sum of two Poisson distributions as in (1):

$$P(X=x) = 0.448 * \frac{e^{-93.22}93.22^x}{x!} + 0.552 * \frac{e^{1270.11}1270.11^x}{x!} \quad (1)$$

Where x is the ocurrence of packet size variable. In fig. 7 we show the simulate model for network traffic total.

For IPv4 network traffic the parameters are λ_1 = 93.42 and λ_2 = 1267.86. The probability that the length of a packet belongs to the first distribution is 0.409, while for the second distribution the probability of a packet following that distribution is 0.591. The model is showed in (2). For IPv6 network traffic, the model is as in (3), with parameters λ_1 = 396.88, and λ_2 = 105.99. The probability that the length of a packet belongs to the first distribution is 0.301, while for the second distribution the probability of a packet following that distribution is 0.699. Figures 8 and 9 show these simulate models.

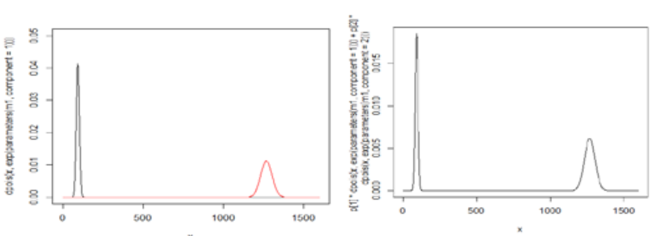

Fig. 7. Poisson model for Traffic Total

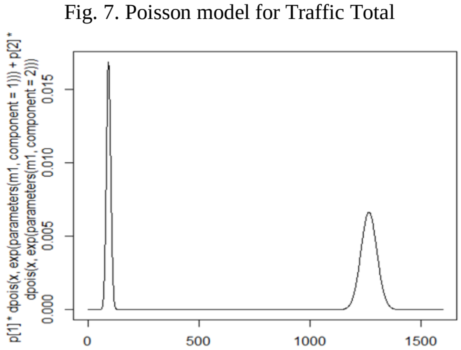

Fig. 8. Poisson model for IPv4 Traffic

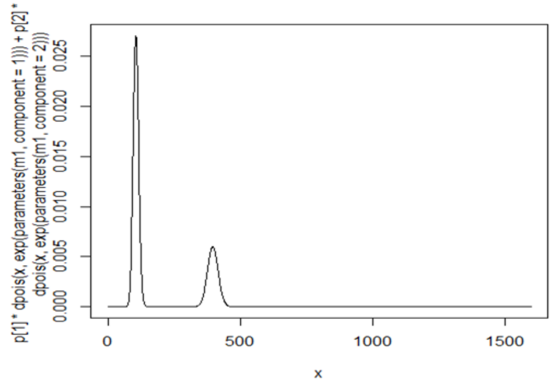

Fig. 9. Poisson model for IPv6 Traffic

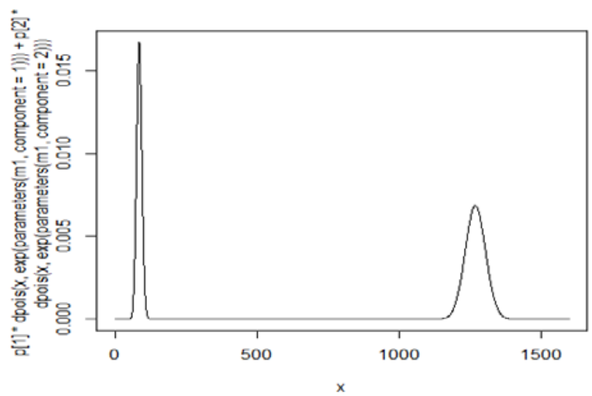

Fig. 10. Poisson models for IPv4 - TCP Traffic

Additionally, we present models for protocols TCP and UDP, over IPv4 and IPv6. Table IV resume the parameters of the models, where λ_1 represent average occurrence in interval 1, λ_2 represent average occurrence in interval 2, P_1 is the probability for a packet following the first distribution, and P_2 is the probability of a packet following the second distribution. For IPv6 only UDP Poisson distribution is necessary for fit the data. Fig 10 show the simulation of these models; and the equations in (4) (5) (6).

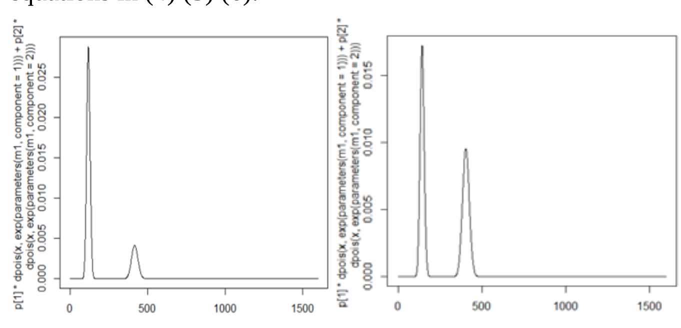

Fig. 11. Poisson models for UDP Traffic – IPv4 and IPv6

Table 4. Poisson Model Parameters for IPv4 / IPv6

Protocol		λ_1	λ_2	P_1	P_2
IPv4	TCP	1270.48	85.67	0.612	0.388
	UDP	418.09	119.06	0.212	0.788
IPv6	TCP	-	-	-	-
	UDP	405.15	143.58	0.482	0.518

$$P(X=x) = 0.409 * \frac{e^{-93.42}93.42^x}{x!} + 0.591 * \frac{e^{1267.86}1267.86^x}{x!} \quad (2)$$

$$P(X=x) = 0.301 * \frac{e^{-396.88}396.88^x}{x!} + 0.699 * \frac{e^{105.99}105.99^x}{x!} \quad (3)$$

$$P(X=x) = 0.612 * \frac{e^{-1270.48}1270.48^x}{x!} + 0.388 * \frac{e^{85.67}85.67^x}{x!} \quad (4)$$

$$P(X=x) = 0.212 * \frac{e^{-418.09}418.09^x}{x!} + 0.788 * \frac{e^{119.06}119.06^x}{x!} \quad (5)$$

$$P(X=x) = 0.482 * \frac{e^{-405.15}405.15^x}{x!} + 0.518 * \frac{e^{143.58}143.58^x}{x!} \quad (6)$$

Finally, table V presents the parameters for the applications that mainly contribute to the total network traffic. As describe, the mainly traffic are SSL and HTTP with 99.15% over TCP, and it´s 95.93% over IPv4, that represents 98.26% of total traffic in this campus wireless network. Fig 12 shows the pattern of SSL packets.

Table 5. Poisson Model Parameters for Applications

Protocol			λ_1	λ_2	P_1	P_2
IPv4	TCP	HTTP	1296.14	85.50	0.685	0.315
		SSL	1267.80	83.61	0.607	0.393
	UDP	DNS	81.94	177.39	0.935	0.065
		MDNS	107.33	353.83	0.670	0.330
		SSDP	362.64	178.61	0.067	0.933
IPv6	UDP	MDNS	405.15	143.58	0.482	0.518
		SSDP	398.54	214.47	0.420	0.580

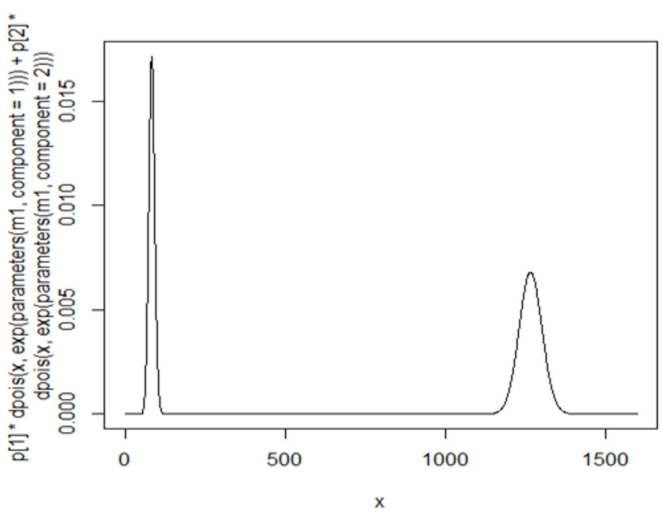

Fig. 12. Poisson models for SSL Traffic – IPv4

V. Conclusions

This paper presents results for stochastic behavior of packet size variable using network traffic measurements in a university campus wireless network. The results show that there is a bimodal traffic distribution with packets around 60 and 1.300 bytes. IPv4 packets represents a big impact in this behavior, mainly TCP packets, and the applications that mark this trend are SSL and HTTP.

Network administrators can use these results to design better networks and optimize network traffic in order to give security policies, QoS provisioning, and ensure efficient utilization of resources.

We development models for characterize the network traffic based using mixture Poisson distribution and provide the best statistical fit to the packet size variable of the dataset considered in this paper. These models simulate the data by traffic type, protocols and applications. Research community can use these distribution parameters presented for built traffic models and apply in other studies in the areas of computer networking and traffic engineering.

Acknowledgment

The authors thank to technical staff from Electrical Engineering and Computer Science Faculty of Escuela Superior Politecnica del Litoral, ESPOL, by the facilities for capture of network traffic.

References

[1] CISCO, "Cisco visual networking index: forecast and methodlogy, 2012c2017," 2013. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html.

[2] D. K. Arrowsmith and R. J. Mondrag, "Modelling Network Data Traffic," 2005.

[3] I. Lee and A. Fapojuwo, "Stochastic processes for computer network traffic modeling," *Comput. Commun.*, vol. 29, no. 1, pp. 1–23, 2005.

[4] A. Dainotti, A. Pescapé, and G. Ventre, "A packet-level characterization of network traffic," *2006 11th Int. Work. Comput. Model. Anal. Des. Commun. Links Networks*, vol. 2006, pp. 38–45, 2006.

[5] G. Mansfield, T. K. Roy, and N. Shiratori, "Self-similar and fractal nature of Internet traffic data," *Int. Conf. Inf. Netw.*, vol. 14, no. 2, pp. 227–231, Mar. 2001.

[6] R. Pries, F. Warmer, D. Staehle, K. Heck, and P. Tran-Gia, "Traffic measurement and analysis of a broadband wireless internet access," *IEEE Veh. Technol. Conf.*, 2009.

[7] C. Gandhi, G. Suri, R. P. Golyan, P. Saxena, and B. K. Saxena, "Packet Sniffer – A Comparative Study," *Int. J. Comput. Networks Commun. Secur.*, vol. 2, no. 5, pp. 179–187, 2014.

[8] A. Callado *et al.*, "A survey on internet traffic identification," *IEEE Commun. Surv. Tutorials*, vol. 11, no. 3, pp. 37–52, 2009.

[9] S. Maheshwari, K. Vasu, C. Kumar, and S. Mahapatra, "Measurement and Comparative Analysis of UDP Traffic over Wireless Networks," *Int. Conf. Wirel. Networks*, 2011.

[10] M. Zhang, W. John, K. C. Claffy, N. Brownlee, and U. C. S. Diego, "State of the Art in Traffic Classification : A Research Review," *PAM '09 10th Int. Conf. Passiv. Act. Meas. Student Work.*, pp. 3–4, 2009.

[11] W. John and S. Tafvelin, "Analysis of internet backbone traffic and header anomalies observed," *dl.acm.org*, p. 111, 2007.

[12] R. Sinha, C. Papadopoulos, and J. Heidemann, "Internet Packet Size Distributions : Some Observations," *81*, pp. 1–7, 2007.

[13] X. L. Wu, W. M. Li, F. Liu, and H. Yu, "Packet size distribution of typical Internet applications," *2012 Int. Conf. Wavelet Act. Media Technol. Inf. Process. ICWAMTIP 2012*, pp. 276–281, 2012.

[14] A. Hajjar, J. Khalife, and J. Díaz-Verdejo, "Network traffic application identification based on message size analysis," *J. Netw. Comput. Appl.*, vol. 58, pp. 130–143, 2015.

[15] S. Lee, Y. Won, and D. J. Shin, "On the multi-scale behavior of packet size distribution in internet backbone network," *NOMS 2008 - IEEE/IFIP Netw. Oper. Manag. Symp. Pervasive Manag. Ubiquitous Networks Serv.*, pp. 799–802, 2008.

[16] H. Kim, K. Claffy, M. Fomenkov, D. Barman, M. Faloutsos, and K. Lee, “Internet traffic classification demystified,” in *Proceedings of the 2008 ACM CoNEXT Conference on - CONEXT ’08*, 2008, pp. 1–12.

[17] M. Zhang, M. Dusi, W. John, and C. Chen, “Analysis of UDP traffic usage on internet backbone links,” *Proc. - 2009 9th Annu. Int. Symp. Appl. Internet, SAINT 2009*, pp. 280–281, 2009.

[18] O. J. Adeyemi, S. I. Popoola, A. A. Atayero, D. G. Afolayan, M. Ariyo, and E. Adetiba, “Exploration of daily Internet data traffic generated in a smart university campus,” *Data Br.*, vol. 20, pp. 30–52, 2018.

[19] J. Cao, W. S. Cleveland, D. Lin, and D. X. Sun, “Internet traffic tends toward Poisson and independent as the load increases,” in *Nonlinear Estimation and Classification*, 2002, pp. 1–18.

[20] L. Bo, D. J. Parish, J. M. Sandford, and P. J. Sandford, “Using TCP Packet Size Distributions for Application Detection,” *7th Annu. Postgrad. Symp. Converg. Telecommun. Netw. Broadcast.*, 2006.

[21] F. Liu, Z. Li, and J. Yu, “P2P applications identification based on the statistics analysis of packet length,” *Proc. - 2009 Int. Symp. Inf. Eng. Electron. Commer. IEEC 2009*, pp. 160–163, 2009.

[22] W. Zhang, “Peer-to-peer traffic anti-identification based on packet size,” *Proc. 2011 Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2011*, vol. 4, pp. 2277–2280, 2011.

[23] N. Vicari, “Modeling of Internet Traffic : Internet Access Influence , User Interference , and TCP Behavior” Norbert Vicari Würzburger Beiträge zur Leistungsbewertung Verteilter Systeme,” 2003.

[24] S. Maheshwari, S. Mahapatra, and K. Cheruvu, “Measurement and Forecasting of Next Generation Wireless Internet Traffic,” Jan. 2018.

[25] I. W. C. Lee and A. O. Fapojuwo, “Analysis and modeling of a campus wireless network TCP/IP traffic,” *Comput. Networks*, vol. 53, no. 15, pp. 2674–2687, 2009.

[26] C. M. Mueller, “On the importance of realistic traffic models for wireless network evaluations,” *COST 2100 12th MCM*, no. 10, pp. 6–13, 2010.

[27] S. A. Mushtaq and A. A. Rizvi, “Statistical analysis and mathematical modeling of network (segment) traffic,” *Proc. - IEEE 2005 Int. Conf. Emerg. Technol. ICET 2005*, vol. 2005, pp. 246–251, 2005.

[28] A. Dainotti, A. Pescapé, and H. C. Kim, “Traffic classification through joint distributions of packet-level statistics,” *GLOBECOM - IEEE Glob. Telecommun. Conf.*, 2011.

[29] E. Castro, M. Alencar, and I. Fonseca, “PROBABILITY DENSITY FUNCTIONS OF THE PACKET LENGTH FOR COMPUTER NETWORKS WITH BIMODAL TRAFFIC,” *Int. J. Comput. Networks Commun.*, vol. 5, no. 3, 2013.

[30] Espinal A., Estrada R., Monsalve C. (2019) Traffic model using a novel sniffer that ensures the user data privacy. Unpublished.

Albert Espinal Santana was born in Montecristi, Ecuador, in 1972. He´s a computer science engineer since 1996, graduated at Escuela Superior Politecnica del Litoral in Guayaquil, Ecuador. He has a magister degree in Information Systems earned at Escuela Superior Politecnica del Litoral in 2001. Additionally, has a master in business administration obtained in 2007 at University of Quebec in Montreal, Canada. Actually, he´s studying a PhD degree in Engineer´s Science at Universidad Nacional de Cuyo, Argentina, since 2016.

He has worked for public and private enterprises. Since 1995 he has been a professor in the networking area of the Engineering in Electricity and Computer Science Faculty. He worked as networking Director of the Escuela Superior Politecnica del Litoral. In 2001 he formed a Cisco networking academy at ESPOL, of which he is its director from 2005 to the present. In 2006 he founded a career in Networks at ESPOL. Actually, he´s working as professor and researcher at ESPOL, in Guayaquil, Ecuador. His research interests are protocol modelling, traffic prediction, and quality of service.

Prof. Espinal is an active member of Internet Society, where has obtained scholarships for assist to Internet Engineering Task Force meetings.