

Analysis of Finite Source Traffic Model of Wireless LAN : A Case Study

Indar Surahmat^{*1}, Satria Try Manggala²

^{1,2}Department of Electrical Engineering, Universitas Muhammadiyah Yogyakarta
Kampus Terpadu UMY, Jl Lingkar Selatan Yogyakarta

^{*}Corresponding author, e-mail: indar.surahmat@umy.ac.id

Abstract – *Investigating characteristics of users and their pattern to access networks are parts of important points in traffic engineering. In this paper, those problems, specifically in wireless LAN, were observed and some related theories were explored. The observation was done in UMY, a campus located in Yogyakarta, Indonesia. In this term, we examined the traffic in an Wireless Action Point. Resuming the content of this paper, we found that the traffic had unique characteristics. For instance, during the day of measurement, the peaks of traffic were in the morning and in the afternoon. Average connected clients in all action points in UMY was between 2500 and 3000 users. When we selected an AP, the characteristic of the AP was examined, the number data size are 62.99 MB and 129.22 MB for mean and standard deviation respectively. We found that hourly trend of the usage of APs resembled self-similar model.*

Keywords: *traffic, traffic engineering, wireless LAN, action point.*

I. Introduction

The rapid development of technology in telecommunication sectors, especially wireless communication has led some concerns on how to characterize wireless users in terms of their behaviour in accessing networks. In the past, people could only access network at specific places like their home or their office because they used a fixed connection. Today, mobile users are more popular after some research findings. Then, as an impact, the behaviour on how and when they access networks are changing.

Modelling the traffic characteristics is important to contribute on planning resources, testing the performance and controlling the network admission. These are explained by Xu et al. in [1]. Xu mentioned that there were two methods in modeling traffic; Tracing measurement of networks which is termed as black box and modeling of the traffic generation process which is known as structural traffic modeling.

Previous works and researches on the same topic have been done. Papadopouli et al. in [2] did researches and reported a model of client arrival of wireless networks. The model only discuss the

behaviour on user arrival. Then, Ghosh et al. in [3] published their research on modeling user come to the system. The main parameters investigated are how the patterns of arrival are, what the suitable model of arrival is, how long users connected to the system are and how many users connected in the same time are.

In here, we chose UMY for our location of observation because UMY has more than 20,000 potential users of Wireless LAN. Students, lecturers and staffs are registered in the network. It is clear that the number of users which allowed to access the network is finite.

We took a consideration in the number of sources since users must do a log-in process before they can fully access services from the network. The number of users who can access the network is limited. They need to register first before they use the services. And thus, there are two conditions; random users come to the network is the total of logged users and unlogged users. These condition leads us use the different traffic model.

This paper only focus on determining the arrival model of the network in a specific access point. The selection of access point examined is based on hourly traffic monitoring. Some models which are researched by previous scientists are presented to

investigate the suitable model.

The organization of this paper are as the following. Part I is introduction which explains the background and the related works. In Part II, we discuss some relevant theories to analyse the result of measurement observed. Part III shows the methods we use to gain the data extracted from the wireless LAN network. The data analysis are explored in Part IV. For the last or Part V, we make conclusion of our work and suggest further research to be followed up .

II. Relevant Theories

The following are some theories related to the topic researched in this paper

II.1. User Arrival Modeling

The most popular conventional distribution on describing user arrival is Poisson Arrival especially in telephony. User arrival can be described by using two approaches; number representations and time representations. Both approaches result Poisson distribution.

Some researches on Wireless LAN used different methods to investigate user arrivals. Some of them explored the arrival of packets as in [1] [4]. Xu et al. in [1] made classification based on popularities of traffic models, i.e. common models and special models. Further, they examined some those theories and made analysis. They did not mention any specific model which satisfy some requirements.

Wang and Popstojanova in [4] did a statistical approach using traffic models which fit to the criterion. Their model included test of non-stationarity and Gaussinity. They did a range of tests by using ARMA, FARIMA, ARIMA and FGN based on the fitness between model and measurement.

Basically, there are two approaches used to analyze the arrival of the users in wireless LAN. After Paxson et al. in [5] found mismatches between Poisson arrival model and real data arrival in internet traffic, some researches introduced new model namely self-similar traffic model as in [6],[7] and [8]. And also Lee et al. in [9] summarized previous researches and stated that traffic model used in data flows is different with traffic model in the user of internet. The researches inferred that Poisson arrival model is suitable to investigate user or session arrival as well as self-similar is suitable for modelling data or packet arrival.

II.2. Poisson arrival model

The most common conventional model for modelling between an user arrival and the next user arrival is Poisson arrival model. If we consider a stochastic Poisson process with the following notation

$$\{N(t), t \geq 0\} \quad (1)$$

and the rate of interarrival times of random variable X_1, X_2, \dots is λ , the cumulative probability function of exponential distribution can be written

$$P\{X_n(t) \leq t\} = 1 - e^{-\lambda t}, \lambda > 0, t \geq 0 \quad (2)$$

The random variable has value of mean and variance as the following

$$m_1 = \frac{1}{\lambda} \quad (3)$$

$$\sigma^2 = \frac{1}{\lambda^2} \quad (4)$$

where m_1 and σ^2 are mean and variance respectively.

II.3. Arrival and departure of users

The next term related to this paper is arrival and departure of users and also termed as birth-date process. As mentioned in the introduction, the number of clients who access the network are limited. In here, Binomial model is utilized. The process that users come to the network, access services and leave the system can be drawn as the following (Fig. 1).

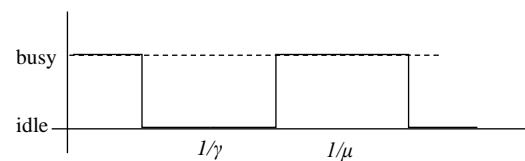


Fig.1 User behaviour in Binomial model

Consider the number of limited sources is S , the intensity of idle sources are exponentially distributed with value γ , and the departure rate is μ , state transition diagram can be illustrated in Fig. 2. The number of available channels depends on the setting of AP, denoted as n . The setting of maximum number of clients granted to connect to a specific AP influences performance of the AP as mentioned by Couta da Silva et al. in [10]. Obviously, the number of S is greater than or equal

the number of n , $S \geq n$. This condition fits Engset distribution.

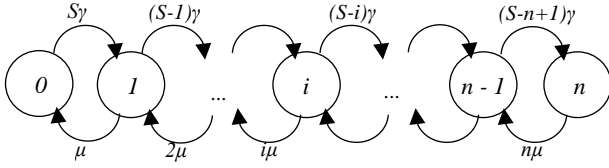


Fig.2 State transition diagram of Engset distribution

Then, let the offered traffic of single idle source β which the value is $\beta = \gamma/\mu$. Probability of the state i can be determined

$$p(i) = \frac{\binom{S}{i} \cdot \beta^i}{\sum_{j=0}^n \binom{S}{j} \cdot \beta^j}, 0 \leq i \leq n \quad (5)$$

II.4. Self-similar model

Self-similar model assumes a stochastic process $X(t)$ determined for $t = (0,1,2, \dots)$. The random variable $X(t)$, which has mean m_1 and variance σ^2 , represents the length or volume of data packet sent through wireless LAN. The characteristic properties of self-similar model are autocorrelation function $r(k)$ and autocovariance function $\gamma(k)$. For the value $k \geq 0$, those properties are defined as the following.

$$r(k) = E\{(X(t) - m_1)(X(t+k) - m_1)\} \quad (6)$$

$$\gamma(k) = \sigma^2 r(k) \quad (7)$$

Then, the popular term related to self-similar model is Hurst parameter which produces some conditions. As mentioned in [9], the condition said to be exactly second-order self-similar is fulfilled if $\gamma(k)$ of $X(t)$ has a value

$$\gamma(k) = \frac{\sigma^2}{2} ((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}) \quad (8)$$

with $0.5 < H < 1$. For the weaker situation where $X(t)$ said to be asymptotically second-order self-similar, then,

$$\lim_{s \rightarrow \infty} \gamma^s(k) = \frac{\sigma^2}{2} ((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}) \quad (9)$$

has the valid value. The value $\gamma^s(k)$ is come from autocovariance of $X^s(k)$

$$X^s(k) = (X(ks - s + 1) + \dots + X(ks))/s \quad (10)$$

with value $k=1,2,\dots$

Next, the important step associated with self similar model is predicting Hurst parameter value. Lee et al. in [9] used variance method to estimate Hurst parameter. Different with Lee, Lu et al. in [11] provides three methods for predicting the value of Hurst parameter; i.e variance method, R/S method, and Whittle method.

In variance method, self-similar in terms of second-order process of the stochastic process $X(t)$ is defined as

$$r(k) \approx H(2H - 1)k^{2H-2}, k \rightarrow \infty \quad (11)$$

$$r(k) \approx H(2H - 1)k^{2H-2} \approx ck^{-\beta}, \quad 0 < \beta < 1, k \rightarrow \infty \quad (12)$$

$$\text{Var}(X^s) \approx s^{-\beta}, \beta = (2 - 2H) \quad (13)$$

Variance series plot is gathered from values of $\log(\text{Var}(X^s))$ in the series of $\log(s)$. Then, Hurst parameter itself is gained from formula $H = 1 - \beta/2$.

Meanwhile, in R/S method, the estimation process of H is derived from adjustment of R/S statistic by rescaling range for observation sets of $X(t)$ which is

$$\{X(t); t = 1,2,3, \dots, n\} \quad (14)$$

with mean $\bar{X}(n)$ and standard deviation $S(n)$. The expected value of R/S statistic especially for long-range dependent

$$E\left(\frac{R(n)}{S(n)}\right) \approx cn^H, 0.5 < H < 1 \quad (15)$$

where $R(n)$ is defined

$$R(n) = \max(0, W_1, W_2, \dots, W_n) - \min(0, W_1, W_2, \dots, W_n) \quad (16)$$

and W_i is defined

$$W_i = (X_1 + X_2 + \dots + X_n) - i\bar{X}(n) \quad (17)$$

III. Research Methodology

We paid attention to some aspects when we explored or took data from wireless LAN networks. As mentioned before, we examined wireless LAN in terms of wireless access points (AP).

In this research, we used CISCO 5500 Wireless Controller to extract data from all APs in UMY. CISCO 5500 Wireless Controller is featured with

statistic tools. We recorded measurements in certain periods based on the following considerations.

III.1. Time of Observations

Time evaluation is one of important notes in taking data from not only in wireless LAN but also the other telecommunication networks. Finding the busiest time over the period of observations is the main goal. To commit this purpose, we did weekly, daily and hourly monitoring.

The aforementioned time have specific purposes. Weekly monitoring investigates the busiest day over one week period, while daily measurement shows the busiest hour over the day of the observation. Taking measurement in hour based have a slight different goal. Hour measurement produces information on how users or clients access networks.

III.2. Network Selection

At first, all APs are monitored over some weeks to gain traffic trends. Statistical methods are explored here to extract information from the data. Networks or APs, which are the most relevance for gathering samples, are selected among all APs. Next, exploration of the data in the busiest day and in the busiest hour are conducted. In this research, we run evaluation of the measurement of a specific access points which satisfy the requirements over the busiest hour.

III.3. Analysis Data

In analysing the data, some parameters need to be fulfilled to decide a good approach. For selecting a distribution model especially in arrival of user, once again, statistical processes are used. Some models of distribution have been mentioned in the previous part of this paper.

IV. Analysis

As first step of the monitoring, we measured all action points in UMY. Since working hour is from 7AM – 5PM, we only collect measurement from 8AM to 4PM. The traffic outside that time range is not interesting to be investigated.

Fig. 3 shows the number of users of all APs in UMY. The result exhibits a similar pattern between one day and another. From the figure, maximum number of clients can be reached at 10.00 – 11.00

AM. From the information, we selected the hour for further measurement as the busiest-hour monitoring time. Whilst, Fig. 4 illustrates the average of connected clients to all APs in UMY within one week period. On Saturday, as example on 11th March, averaged clients is in the lowest number. Normally, it is around 2500 – 3000 active clients for each measurement time.

As example of single AP monitoring, we chose an AP among 118 APs in UMY. The measurement over the busiest hour illustrated in Fig. 5 and Table 1.

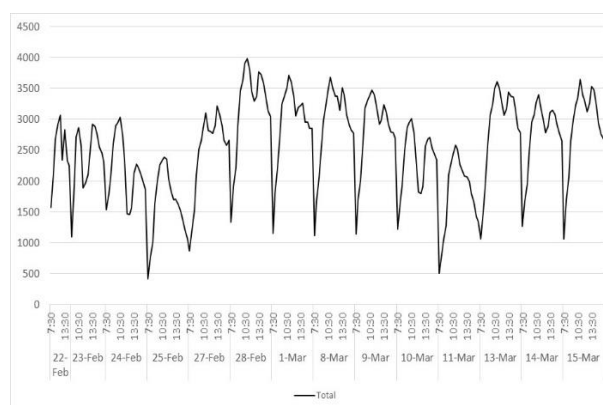


Fig.3 Number of clients measured over two weeks

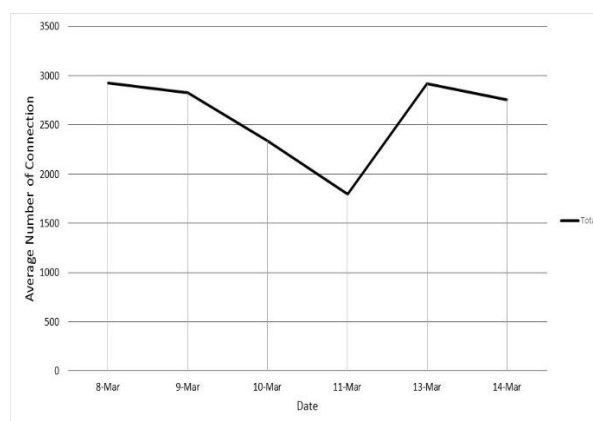


Fig. 4. No of Clients in a typical week in all APs.



Fig.5 Number of clients monitored in the busiest hour and number of arrival and departure clients

After selecting the busiest hour, we monitored and recorded data from a specific AP. The process of choosing an AP was based on the number of clients which was the busiest among the APs. As shown in Fig. 3, the arrival and departure of users were recorded between 10:50 and 11:50. The space between two measurement time was 2 minutes. Table 1 informs number of arrival and departure during the investigation period.

TABLE I

AP'S CLIENT RECORD DURING BUSY HOUR			
Time	Clients	Arrival	Departure
10:50	52	0	3
10:52	49	1	1
10:54	50	2	1
10:56	56	7	4
10:58	53	1	0
11:00	56	3	4
11:02	55	3	4
11:04	53	2	2
11:06	53	2	7
11:08	47	1	1
11:10	48	2	4
11:12	46	2	1
11:14	50	5	5
11:16	46	1	2
11:18	48	4	2
11:20	49	3	2
11:22	50	3	6
11:24	46	2	1
11:26	50	5	8
11:28	45	2	2
11:30	45	2	2
11:32	44	1	1
11:34	43	0	1
11:36	44	2	1
11:38	45	2	2
11:40	43	0	0
11:42	43	0	1
11:44	43	1	0
11:46	45	2	2
11:48	43	0	2
11:50	44	3	0

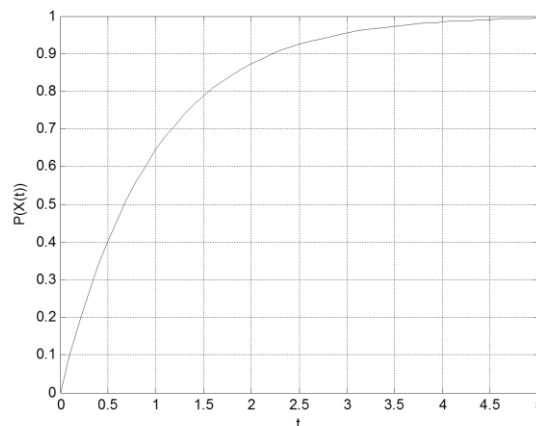


Fig. 6. Poisson distribution of client arrival in the AP

From the data in Table 1, we can calculate some parameters associated with Poisson distribution. First, the main parameter in interarrival distribution is λ . From the data, we got $\lambda = 1.03$ arrivals/minute and the Poisson distribution function can be figured as in Fig. 6

The total number of action points in UMY is 118 APs. Assume that the number of potential user of an AP is distributed evenly, although in real situation, the number of potential clients are different between an AP and another. This calculation is just simplified. Based on the data, from more than 20.000 potential active users in UMY, there are only about 10% to 20% who are active users. From Table 1, during the busiest hour monitoring, the average connected clients is 47.87 (48 clients) and the standard deviation is 4.11. Based on our calculation on the assumption mentioned, the potential users on each AP is about 170 users. From this result, Engset probability of i active users parameters can be illustrated in Fig 7.

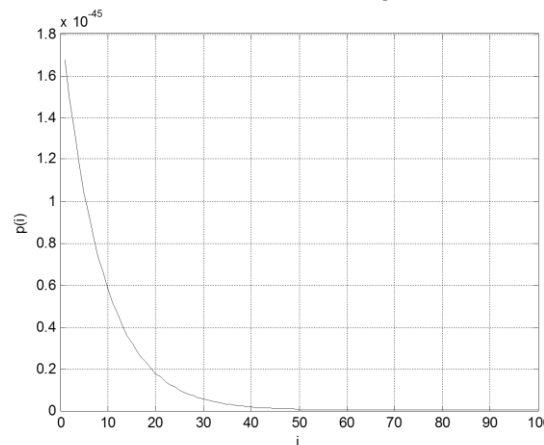


Fig.7 Probability to the state i based on parameter calculation

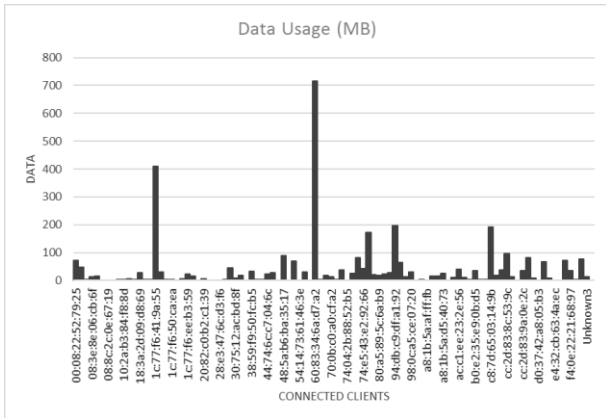


Fig.8 Data usage of connected clients

In terms of data usage, Fig. 8 presents the size of data size committed to the network or an AP for each connected clients over the busiest hour monitoring. We chose the AP with IP address 10.75.0.24. The differences between a client and another is strictly considerable. With simple calculation, the value of mean and standard deviation, respectively, are 62.99 MB and 129.22 MB for data usage over the hour. Furthermore, if we discuss about the cumulative and the growth of data usage in the AP, we summarized it in Fig. 9 as follows

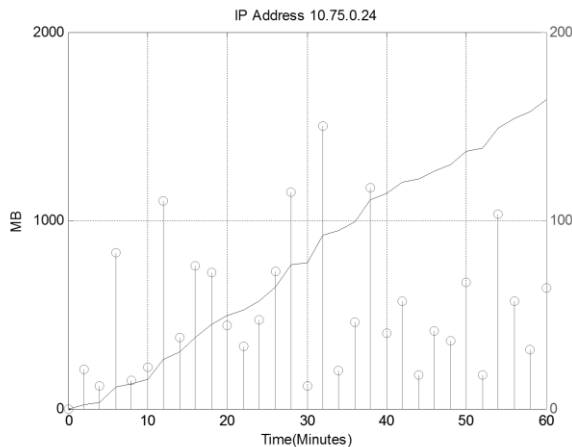


Fig.10 Data usage cumulative and growing pattern of AP with IP address 10.75.0.24

Selecting a user among active users in the AP provides information about behavior of single client. For example, Fig. 10 shows cumulative data usage and its growing pattern. From the selected measurement, the value of mean and standard deviation, where the data before this client is connected to the system is neglected, are 8.57 and

8.29 respectively.

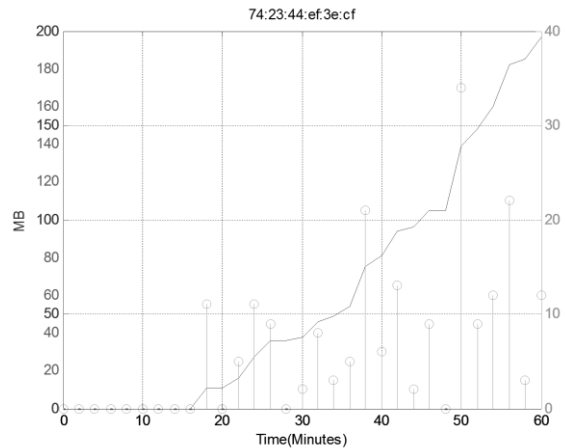


Fig.10 Data usage cumulative and growing pattern of user with MAC 74:23:44:ef:3e:cf

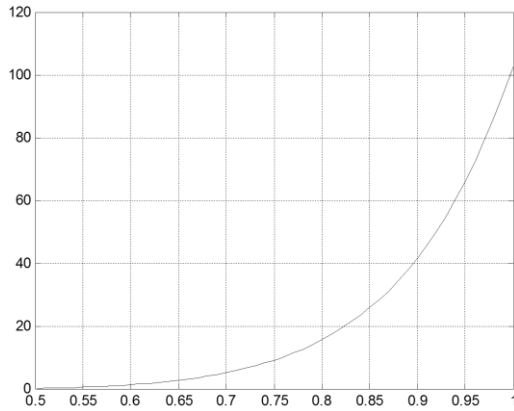
Regarding hourly trend shown in Fig.3, we continued to analyse hourly pattern of a specific access point. Consider, a stable measurement on 27th Feb as reference, table 2 illustrates calculated autocovariance and autocorrelation of the measurement with number $k=18$. The value was same with the number of time measured over a day monitoring. For comparison of value $r(k)$ and $\gamma(k)$ in table 2, we also used moving reference. Regarding the comparison, we got a significant difference between fixed reference and moving reference. In the moving reference, the average of autocorrelation $r(k)$ and autocovariance $\gamma(k)$ are 0.79 and 77.74 respectively, While, in the fixed reference, the value of $r(k)$ and $\gamma(k)$ are 0.57 and 26.85 respectively. In this paper, we used the result of fixed reference for further calculation.

Moreover, from that chosen value, the Hurst parameter is estimated with some ways mentioned in previous section. In hear we were only used iteration. By using iterations, a number satisfied the criteria is found after certain time iterations. The graph in Fig. 11 shows the value of autocovariance calculated from some value of Hurst iteration ($0.5 \leq H \leq 1$). With number $k = 18$, the second-order self-similar is fulfilled with value $H = 0.8$.

TABLE 2
AP'S AUTOCORRELATION AND AUTOCOVARANCE

Fixed Reference		Moving Reference	
$r(k)$	$\gamma(k)$	$r(k)$	$\gamma(k)$
0.68	36.02	0.68	36.02
0.50	25.20	0.91	114.01
0.42	16.84	0.84	79.67
0.59	31.42	0.77	77.19
0.55	33.88	0.92	142.12
0.35	13.23	0.73	78.71

0.68	30.30	0.57	44.18
0.49	18.57	0.80	62.12
0.76	32.81	0.82	62.11
0.66	30.22	0.88	81.25

Fig.11 Autocovariance value from the iteration of H

V. Conclusions and Future Works

The aforementioned analysis stated some points regarding the measurement of APs in UMY.

- We found the busiest time in terms of maximum number of clients was reached at 10 AM. The value at this hour can be used as a consideration to the planning and development plan of the APs in UMY in terms of capacity.
- The arrival of user in the AP, where we selected the investigation, was 1.03 arrival/minute and followed Poisson distribution. This is also the weakness of this research that for calculating the arrival, we found difficulties to differentiate between long-period connected users and short-period connected users
- The data usage between a user and another was strictly different. In regard to the number of variance of data usage, next research need to differentiate between stationary user and mobile user. Actually, this part is also need to be analysed more in terms of behavior of the connected users.
- Regarding pattern of hourly connected clients, self-similar used was used. The result of calculation shows the suitable value of H is 0.8 for the pattern.

Furthermore, the analysis in this paper still need to be explored, especially in the use of data. We found some limitations on extracting the value of data usage from database server.

Acknowledgements

This work was supported by Telecommunication Lab. of Electrical Engineering Department, Universitas Muhammadiyah Yogyakarta.

References

- [1] Xu, S., Chen, X., Wei, G., Wang, D. (2009). Research of Traffic Modeling for Future Wireless Networks. *2009 5th International Conference on Wireless Communications, Networking and Mobile Computing*, page 1 – 6 DOI: 10.1109/WICOM.2009.5301728
- [2] Papadopouli, M., Shen, H., Spanakis, M. (2005). Modeling client arrivals at access points in wireless campus-wide networks. *2005 14th IEEE Workshop on Local & Metropolitan Area Networks*, DOI: 10.1109/LANMAN.2005.1541514
- [3] Ghosh, A. , Jana, R., Ramaswami, V., Roland, J. Shankaranarayanan, N.K. (2011). Modeling and Characterization of Large-Scale Wi-fi Traffic in Public Hot-Spots. *2011 Proceedings IEEE INFOCOM* page 2921–2929 DOI: 10.1109/INFCOM.2011.5935132
- [4] Wang, X., Popstojanova, K.G. (2009). Modeling Web Request and Session Level Arrivals. *2009 International Conference on Advanced Information Networking and Applications*.
- [5] Paxson, V., Floyd, S. (1994) Wide-area traffic: the failure of poisson modeling. *Proceedings of the conference on Communications architectures, protocols and applications*, pages 257–268, New York, NY, USA
- [6] Leland, W.E, Taquq, M.S, Willinger, W., Wilson, D.V. (1994). On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15.
- [7] Willinger, W., Taquq, M. S., Leland, W. E, Wilson, D.V. (1995) Self-similarity in high-speed packet traffic: Analysis and modeling of ethernet traffic measurements. *Statistical Science*, 10(1):67–85.
- [8] Willinger, W., Taquq, M. S., Sherman, R., Wilson, D.V. (1997). Self-similarity through high-variability: statistical analysis of ethernet lan traffic at the source level. *IEEE/ACM Transactions on Networking*, 5(1):71–86.
- [9] Lee, Y, Kim, J. (2008). Characterization of Large-Scale SMTP Traffic: the Coexistence of the Poisson Process and Self-Similarity. *MASCOT 2008. International Symposium on Modelling, Analysis, and Simulation of Computers and Telecommunication Systems*. DOI:10.1109/MASCOT.2008.4770573
- [10] Couto da Silva, A.P, Meo, M., Marsan, M.A.(2012). Energy-performance trade-off in dense WLANs: A queuing study, *Computer Networks* 56 (2012) 2522 – 2537
<http://dx.doi.org/10.1016/j.comnet.2012.03.017>

- [11] Lu, H., Yu, C., Gou, X (2011). Analysis of Traffic Model and Self-similarity for QQ in 3G Mobile Networks. *Proceeding of AIAI 2011. International Conference on Advanced Intelligence and Awareness Internet*. DOI: 10.1049/cp.2011.1442.
- [12] Iversen, V.B (2015) Teletraffic Engineering and Network Planning, *DTU Fotonic*.

Authors' information



Indar Surahmat is currently a researcher in the Department of Electrical Engineering, UMY. He was born in 1982. He got bachelor degree in electrical engineering at UGM, Indonesia (2005) and master degree in telecommunication engineering at ITB, Indonesia (2011).

Previously, he had worked at some multinational companies which provide services in telecommunication sector.

He is interested in two streams of telecommunication. Those are antenna and wave propagation, and teletraffic engineering. He is also a supervisor of some students in the same department majoring in telecommunication.

Mr Surahmat is also the member of the Institution of Engineers Indonesia and APTI, the association of engineering professional Indonesia.



Satria Try Manggala is currently a student in the Department of Electrical Engineering, UMY. He was born in 1994. He is now pursuing bachelor degree in electrical engineering department at UMY, Indonesia.

He is interested in telecommunication. In this research, he gives some supports for collecting the information and data to be analysed.