Fuzzy Logic Tsukamoto for SARIMA

On Automation of Bandwidth Allocation

Isna Alfi Bustoni, M.Eng
Department of Computer Science and Electronics
Universitas Gadjah Mada, Yogyakarta, Indonesia

Adhistya Erna Permanasari, Ph.D
Department of Electrical Engineering and Information Technology,
Universitas Gadjah Mada, Yogyakarta, Indonesia

Indriana Hidayah, M.T.
Department of Electrical Engineering and Information Technology,
Universitas Gadjah Mada, Yogyakarta, Indonesia

Indra Hidayatulloh, M.T.
Software Engineering Study Program, Institut Teknologi Telkom Puwokerto, Purwokerto, Indonesia

***Abstract*—The wireless network is used in different fields to enhance information transfer between remote areas. In the education area, it can support knowledge transfer among academic member including lecturers, students, and staffs. In order to achieve this purpose, the wireless network is supposed to be well managed to accommodate all users. Department of Electrical Engineering and Information Technology UGM sets wireless network for its daily campus activity manually and monitor data traffic at a time then share it to the user. Thus, it makes bandwidth sharing becomes less effective. This study, build a dynamic bandwidth allocation management system which automatically determines bandwidth allocation based on the prediction of future bandwidth using by implementing Seasonal Autoregressive Integrated Moving Average (SARIMA) with the addition of outlier detection since the result more accurate. Moreover, the determination of fixed bandwidth allocation was done using Fuzzy Logic with Tsukamoto Inference Method. The results demonstrate that bandwidth allocations can be classified into 3 fuzzy classes from quantitative forecasting results. Furthermore, manual and automatic bandwidth allocation was compared. The result on manual allocation MAPE was 70,76% with average false positive value 56 MB, compared to dynamic allocation using Fuzzy Logic and SARIMA which has MAPE 38,9% and average false positive value around 13,84 MB. In conclusion, the dynamic allocation was more effective in bandwidth allocation than manual allocation.**

Keywords*—Bandwidth allocation management; dynamic allocation; fuzzy logic; Tsukamoto inference method; SARIMA

I. Introduction

Computer network becomes the important aspect in data communication. Thousands of user are using the network to transfer information to the remote area. The wireless network is common of computer network technology that is widely used in many different institutions to achieve different activity, including educational institution [1]. Therefore, network bandwidth in campus is supposed to be managed to meet user needs.

Bandwidth is channel capacity or the maximum throughput of a physical or logical communication path in a digital communication system [2]. The higher bandwidth consumption needs a good management. Bandwidth management is a way to achieve optimum usage with limited available bandwidth [1].

Universitas Gadjah Mada (UGM) as an educational institution is also implementing Wireless Local Area Network (LAN) or Wi-Fi on their campus. This Wi-Fi is installed at all faculties in UGM. Every UGM staffs and students are provided with their own username and password to connect to the internet through the Wi-Fi. The number of users who connects to access point tends to rise. Thus, it is necessary to manage bandwidth efficiently.

The earlier study, especially in Electrical Engineering and Information Technology presented that bandwidth usage had a seasonal pattern. Seasonal Autoregressive and Moving Average (SARIMA) was applied to predict bandwidth usage. The results showed that bandwidth traffic tends to rise on Tuesday and get down on Sunday. Traffic fluctuations indicate the weekly bandwidth allocation should be changed [3].

Indra Hidayatulloh et al, added Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) to reduce heteroscedasticity effect on bandwidth prediction using SARIMA. SARIMA- EGARCH increased prediction accuracy around 19,15% compared to stand alone SARIMA. Unfortunately, the auto-correlation effect happened during prediction on time series bandwidth data [4].

One of the methods of bandwidth management is by using system scheduling[5]. Mikrotik outerOS™, that is applied at UGM Hotspot, is a network router based on Linux. Mikrotik is also supported by Windows application (WinBox) to ease its router adjustment. Some scheduling methods are implemented at Mikrotik RouterOS, for example, Simple Queue, Per Connection Queue (PCQ) and Hierarchical Token Bucket (HTB). Simple Queue is used to restrict the number of data for specific addresses or subnet [5]. It is the simplest scheduling since the limitation of maximum upload and download is implemented refer to IP address client. PCQ method can divide bandwidth automatically from the active users. Additionally, this method also has disadvantages, whereas may result in bandwidth leakage or unfair division [6]. HTB method implements link sharing so that the residual bandwidth in a class node can be distributed into another class. It uses Token

Bucket Filter (TBF) as an estimator in the determination of bandwidth allocation. It is more adaptive since TBF provides bandwidth in rate. In HTB method, the admin needs to define the maximum rate in every class node. Unfortunately, it is still doing manually based on admin intuition. Therefore, the bandwidth management is less efficient and less adaptive.

Refer to the problem, there is important to develop an adaptive model for bandwidth management. A model must be able to provide the values of maximum rate automatically. The forecasting results could not be used directly for system input because it has a decimal data type with wide class boundaries that will be hard for a network administrator to manage these quantitative values. Moreover, forecasting results still have more error potential if the data applied directly rather than using a specific range. These data have to be converted into classification form to determine the maximum rate allocation. Fuzzy Logic was implemented to achieve this goal.

Some studies on bandwidth management have investigated. A previous work [7] has explored rate control strategies for real-time multimedia variable bit rate (VBR) services. It was implemented in IEEE 802.16 broadband wireless networks. This study managed bandwidth allocation on max-min fairness queue scheduling using a time constraint condition. Liu, et al [8] has predicted network traffic by using chaos theory and Support Vector Machine (SVM). This research used campus data including wired and wireless. The forecast values could be used to manage the bandwidth. A proposed scheme dynamically reserves and allocates bandwidth based neural network has been studied by Song et al. [9]. It was applied to different types of calls. Lee et al [10] has implemented round-robin schedule to allocate bandwidth. Prediction of Available Bandwidth Estimation with Mobility Management in Ad Hoc Networks has been undertaken by Belbachir et al. [11]. Hierarchical game theory models were also be implemented for bandwidth management [12]. While Maestrelli et al. [13] proposed quantization model for bandwidth adjustment.

The aim of this research was to develop a Fuzzy Logic Tsukamoto in order to support bandwidth allocation decision automation system called BIOMA (Bandwidth Automation Management). Fuzzy Logic Tsukamoto uses monotone membership function with Center Average Deffuzzyfier method. By this defuzzification method, Tsukamoto selects mean from the range given. Meanwhile, Mamdani method selects Minimum or Maximum Value. If the minimum is selected, it might affect bandwidth allocation doesn't meet the network requirement. In the other side, if the maximum is selected, it might cause bandwidth being extravagant. The other method, Sugeno Method, gives consequences value as crisp values using some linear calculation. So this method doesn't meet the system requirement like Fuzzy Logic Tsukamoto.

The remainder of the paper is structured as follows. Section 2 illustrates data input. Section 3 presents the methodology. Section 4 describes experimental results. Finally, Section 5 presents the conclusion of the study.

II. Data Input

Forecasting results of bandwidth usage were used as input data [2]. It was used real downstream dataset (Mbps). Data was collected from monitoring Universitas Gadjah Mada (UGM) portal at http://mon.ugm.ac.id/cacti/weathermap. This study collected time series data for the 20 week period from 09 September 2013 to 27 January 2014. The original data is plotted as presented in Fig. 1. Bandwidth usages were predicted by using SARIMA method with outlier detection.

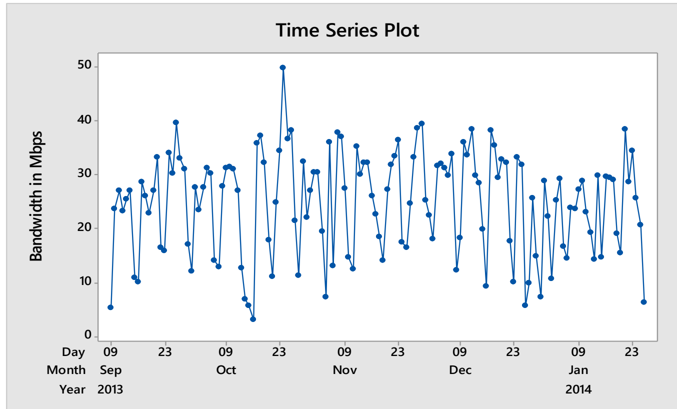

Fig. 1. Bandwidth usage original dataset.

Refer to previous research [2], the most appropriate SARIMA model was (0,1,1)(0,1,1)7C from various traces. The observation found that some outliers in data collection influenced forecast accuracy. This problem was solved by including outliers detection to the model. Missing value analysis was done by using mean substitution operation to handle outliers. The approach could reduce forecasting error (MAPE) into 14.49%.

From computation, the parameter results were MA(1) = 0.9519 and SMA(7) = 0.9246 and constant = 0.010686. The estimated parameters were included to form the final model that expressed as backshift model shown in equation (1).

The forecasting results are shown in Table 1 is a month prediction of bandwidth usage in the Department of Electrical Engineering and Information Technology, UGM.

TABLE I. Forecasting Result

Day	Week 1	Week 2	Week 3	Week 4
Monday	9.794235	8.960937	8.037514	7.023966
Tuesday	26.93602	26.08984	25.15355	24.12712
Wednesday	23.88494	23.02589	22.07672	21.03742
Thursday	27.52276	26.65084	25.68879	24.63662
Friday	27.81239	26.9276	25.95267	24.88762
Saturday	25.52054	24.62286	23.63506	22.55714
Sunday	12.64095	11.7304	10.72973	9.63893

III. METHODOLOGY

In order to build adaptive and dynamic bandwidth management, it is necessary to build a system that can model bandwidth needs and give maximum value rate automatically. To do this, SARIMA method was used to predict bandwidth needs and Fuzzy Logic method was used to allocate bandwidth dynamically.

A. SARIMA Model Transformation

The SARIMA model for this research is model (1) which transformed into regular equation form in order to be used as system model.

The model used in the research as follows.

$$z_t = 0.010686 - (1 - 0.9519B - 0.9246B^7 - 0.8801B^8)\alpha_t \quad (1)$$

with,

z_t = prediction value

$\boldsymbol{B}$ = Backshift operator for MA1

$\boldsymbol{B}^7$ = Backshift operator for MA7 (seasonal lag L)

$\boldsymbol{B}^8$ = Backshift operator for MA8 (seasonal lag L)

α_t = forcasting residual.

z_t could be written as follows in order to facilitate computation.

$$z_t = \delta - \theta_1\alpha_{t-1} - \theta_{1,7}\alpha_{t-7} + \alpha_t \quad (2)$$

with,

z_t = Prediction value

δ = constant atmodel

θ_1 = Coefficient of MA1

$\theta_{1,7}$ = coefficient of MA1,7 (seasonal lag L)

α_t = forecasting residual data number t

Based on Four Stationarity Transformation formula by Box [14] the z_t value for SARIMA with the first order of regular and seasonal differencing is as follows.

$$z_t = y^*_t - y^*_{t-1} - y^*_{t-L} + y^*_{t-L-1} \quad (3)$$

with,

z_t = forecasted time series value

y^*_t = time series value at t

y^*_{t-1} = time series value at t-1

y^*_{t-L} = time series value at t-L

y^*_{t-L-1}= time series value at t-L-1

Subsequently, based on the same book [14] a model that has an MA time series character, in both seasonal and regular part, will require one or more multiplicative terms to be combined on the model. Therefore, final model of z_t as follows.

$$z_t = \delta - \theta_1\alpha_{t-1} - \theta_{1,7}\alpha_{t-7} - \theta_1\theta_{1,7}\alpha_{t-8} + \alpha_t \quad (4)$$

with,

z_t = Prediction value

δ = constant atmodel

θ_1 = Coefficient of MA1

$\theta_{1,7}$ = coefficient of MA1,7 (seasonal lag L)

α_t = forecasting residual data number t

By combining the formulas (4-2) and (4-3) the value of the predicted value y^*_t can be determined as follows.

$$y^*_t = y^*_{t-1} + y^*_{t-7} + y^*_{t-8} - \theta_1\alpha_{t-1} - \theta_{1,7}\alpha_{t-7} + \theta_1\theta_{1,7}\alpha_{t-8} + \alpha_t + \delta \quad (5)$$

B. Fuzzy Logic Method

There is some approach that can be implemented to bandwidth allocation. One of them is a heuristic model. Fuzzy Inference System (FIS) is a heuristic model that widely used [6]. This study implemented Fuzzy Logic Tsukamoto. This method was represented by a fuzzy set with a monotonical membership function. The monotonical reasoning was used when two Fuzzy areas are related with the following implication:

$$IF\ x\ is\ A\ THEN\ y\ is\ B\ or\ transfer\ function \quad (6)$$

$$y = f((x, A), B))$$

The implication function extends monotonical reasoning into:

$$If (x_1\ is\ A_1).(x_2\ is\ A_2).\ ¤ \ldots ¤.(x_n\ is\ A_n)$$
$$THEN\ y\ is\ B\ with\ ¤\ is\ Operator \quad (7)$$

For example, there are two input variable Var-1(x) and Var-2(x), and output variable Var-3(z). Var-1 is divided into two sets: A_1 and A_2. Var-2 is divided into set B_1 and B_2. Whilst, Var-3 is divided into sets: C_1 and C_2, whereas C_1 and C_2 are supposed to be monoton. Therefore, two rules are used:

$$[R_1] IF\ (x\ is\ A_1)\ AND\ (y\ is\ B_2) THEN\ (Z\ is\ C_1)$$

$$[R_2] IF\ (x\ is\ A_2)\ AND\ (y\ is\ B_1) THEN\ (Z\ is\ C_2)$$

The first step is finding membership function for each fuzzy set in its rule. Sets of A_1, B_2, and C_1 are driven from fuzzy rule $[R_1]$ and sets A_2, B_1, and C_2 come from fuzzy rule $[R_2]$. Fuzzy rule R_1 and R_2 may be represented in determining crisp values Z.

Furthermore, the inferred output for each rule is defined as a crisp value induced by the rule's firing strength (α-predicate).

The final output result is taken from the weighted average of the output of each rule [15]. Fig. 2 illustrates steps of Tsukamoto method.

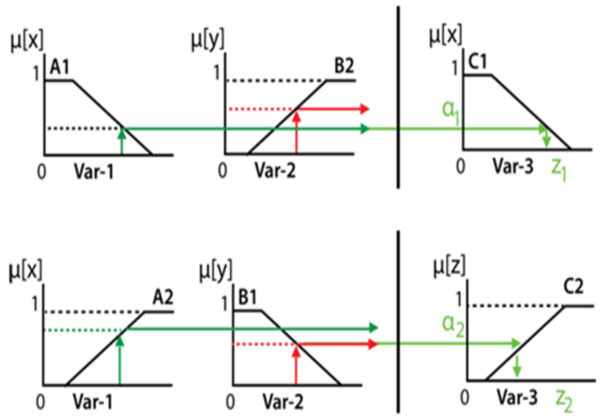

Fig. 2. Tsukamoto Step[15].

IV. Experimental Result

This following section is discussing the result of Tsukamoto steps in order to manage bandwidth allocation on available dataset.

A. Construction of Membership Function

This step is focus on the developing a fuzzy set. There were 2 variables for modeling, "Usage" is for input variable and "Allocation" is for output variable. The Usage has 3 fuzzy sets, they are Small, Medium, and Large. "Usage" is real type variable with its own domain:

Small, domain = [0.00 15.00]

Medium, domain = [12.50 27.50]

Large, domain = [25.00 35.00]

Membership function in this study is represented as triangle curve in Fig. 3.

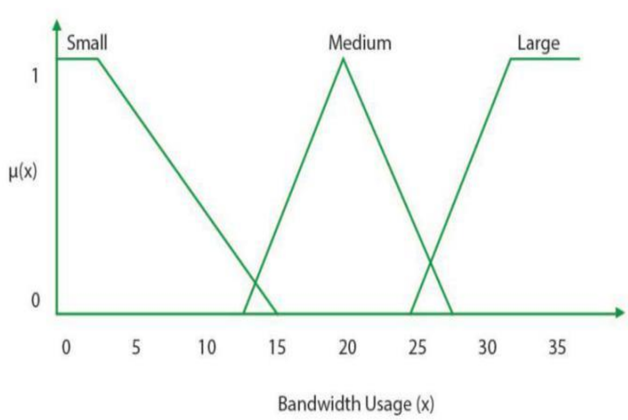

Fig. 3. Triangle curve of data.

Membership function for each set has been calculated as the following:

$$\mu SMALL(x) = \begin{cases} 1, & x \leq 5.00 \\ \frac{15.00-x}{10}, & 5.00 \leq x \leq 15.00 \\ x, & x \geq 15.00 \end{cases} \quad (8)$$

$$\mu MEDIUM(x) = \begin{cases} 0, & x \leq 12.50 \; or \; x \geq 27.50 \\ \frac{x-20.00}{7.50}, & 12.50 \leq x \leq 20.00 \\ \frac{27.50-x}{7.50}, & 20.00 \leq x \leq 27.50 \end{cases} \quad (9)$$

$$\mu LARGE(x) = \begin{cases} 0, & x \leq 25.00 \\ \frac{x-25.00}{10}, & 25.00 \leq x \leq 35.00 \\ 1, & x \geq 35.00 \end{cases} \quad (10)$$

Furthermore, "Allocation" variable was determined as crisp values. This output has 3 sets: SMALL, MEDIUM, and LARGE.

Small, domain = [15.00]

Medium, domain = [27.5]

Large, domain = [35.00]

These values were refer to the highest average from bandwidth allocation in 20 last week.

B. Identify the Headings Rule Formation

Rules are one of FIS requirement [16]. The calculation includes Usage as input variable and Allocation as output variable. It yields the following rule format:

$$\left[R_i\right] IF\ X_i\ is\ A_i ¤.¤\ X_n\ is\ A_n\ THEN\ Allocation\ is\ B_i \quad (11)$$

where:

R_i : fuzzy rule-i (i=1..m).

X_i : weight values of Usage-i

A_i : fuzzy set of Himpunan weight values of Usage-i

¤ : operator

n : number of data

B_i : fuzzy set for allocation variable

C. Weight Calculation and Determination of Bandwidth Allocation

The previous rules were used to determine each data weight. Values of predicate α were found from rule composition (α_i). The predicates are associated differently with the operator.

In AND operator, predicate value of " α is $A_1\ AND\ X_2\ is\ A_2$ " is given as:

$$\alpha_i = \mu_{A_1 \cap A_2} = Min(\mu_{A_1}(x_1), \mu_{A_2}(x_2)) \quad (12)$$

$$\alpha_i = \mu_{A_1 \cap A_2} = Max(\mu_{A_1}(x_1), \mu_{A_2}(x_2)) \quad (13)$$

One consequent values are obtained, y values can be calculated as:

$$y = \frac{\sum_{i=1}^{m} \alpha_i y_i}{\sum_{i=1}^{m} \alpha_i} \tag{14}$$

D. Allocation Calculation

The calculation was started with the determination of membership weight. Membership weight for each data was calculated based on 2 domain, with each domain consists of Small, Medium, and Large. The result is presented in Table 2.

TABLE II. WEIGHT RESULTS

x	SMALL	AVERAGE	LARGE
9.7942	0.5206	0	0
26.9360	0	0	0.1936
23.8849	0	0.2230	0
27.5228	0	0	0.2523
27.8124	0	0	0.2812
25.5205	0	0	0.0521
12.6409	0.2359	0	0
8.9609	0.6039	0	0
26.0898	0	0	0.1090
23.0259	0	0.3948	0
26.6508	0	0	0.1651
26.9276	0	0	0.1928
24.6229	0	0.0754	0
11.7304	0.3270	0	0
8.0375	0.6962	0	0
25.1535	0	0	0.0154
22.0767	0	0.5847	0
25.6888	0	0	0.0689
25.9527	0	0	0.0953
23.6351	0	0.2730	0
10.7297	0.4270	0	0
7.0240	0.7976	0	0
24.1271	0	0.1746	0
21.0374	0	0.7925	0
24.6366	0	0.0727	0
24.8876	0	0.0225	0
22.5571	0	0.4886	0
9.6389	0.5361	0	0

Where, x are time series from SARIMA forecasting results.

Next step was determining y values that represented as allocation values for each time series by using formula (14). The calculation results are shown in Table 3.

TABLE III. RESULTS OF BANDWIDTH ALLOCATION

Day	Week-1	Week-2	Week-3	Week-4
Monday	15	15	15	15
Tuesday	35	27	27	27
Wednesday	27	25	27	27
Thursday	35	35	27	27
Friday	35	35	27	27
Saturday	27	27	27	27
Sunday	15	15	15	15

V. BIOMA SYSTEM IMPLEMENTATION

Bioma system has been implemented with a user friendly interface design using Bootstrap. Fig. 4 shows the interface of dashboard page that contains history of bandwidth allocation and bandwidth prediction the next few days.

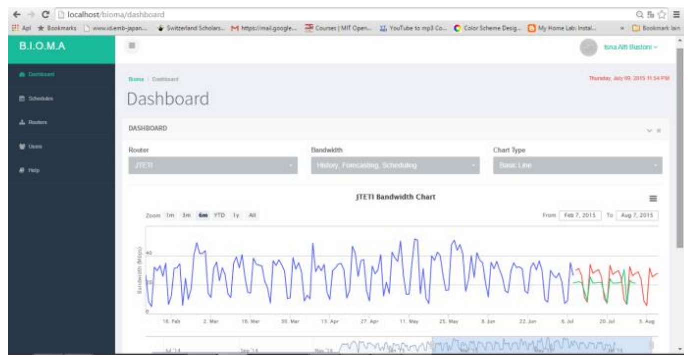

Fig. 4. Dasboard.

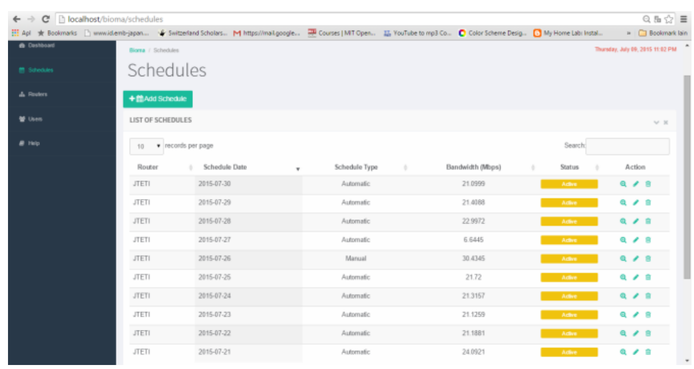

Fig. 5. Schedules.

Fig. 5 as follows is the interface of schedules page that used to set bandwidth allocation on specific date.

Bioma system tested by comparing MAPE between manual/ static and dynamic allocation using Bioma. The test results are shown in Table 4 below:

TABLE IV. RESULTS OF BANDWIDTH ALLOCATION

Method	**MAPE**	**Avg. False Positif Value (MB)**	**Avg. False Negatif Value (MB)**
Manual/ Static	70,77 %	56,61	0
Dynamic (Bioma)	38,90 %	13,84	0

VI. CONCLUSION

Fuzzy Tsukamoto method has been demonstrated its capability in the bandwidth allocation automation at UGM-Hotspot by using forecasting traffic usage data as an input for Fuzzy Logic. The fuzzy approach was able to convert the quantitative values of forecasting results into classification form. Bioma system as dynamic bandwidth allocation system had MAPE amounted to 38.90 %, smaller than the MAPE of manually calculating bandwidth allocation of 70.77%. Besides showing the error, MAPE in this case also shows the huge excess bandwidth allocation or the rest of the unused bandwidth. Thus it can be said that the bandwidth allocation by the Bioma system is more efficient than manually allocation by an administrator under normal conditions.

However, False positive happened in both manual and dynamic allocation. This false positive result indicates remaining unused allocation became wasteful. Therefore, further research is needed to be able to automatically transfer the remaining allocation of unused bandwidth so it can be utilized better. In the future research, we should consider other variables such as the number of users, the role of users and the autocorrelation effect in the bandwidth data since bandwidth data are fluctuating.

REFERENCES

[1] Deneen, L. “Bandwidth Management Tools, Strategies, and Issues.” http://net.educause.edu/ir/library/pdf/DEC0202.pdf. 2002.

[2] Comer DE. “Computer Networks and Internets”. 5th ed. Pearson; 2008

[3] Permanasari AE, Hidayah I, Bustoni IA. “Forecasting Model for Hotspot Bandwidth Management at Department of Electrical Engineering and Information Technology UGM”. International Journal of Applied Mathematics and Statistics 2015; 53:227-234.

[4] Hidatulloh, I., & Bustoni, I.A. “SARIMA-EGARCH Model to Reduce Heteroscedasticity Effect in Network Traffic Forecasting”. Journal of Theoretical and Applied Information Technology 2017, vol 95 No. 3.

[5] Nahar A. “Penerapan Sistem Scheduler Pada Manajemen Bandwidth Menggunakan Mikrotik RouterOS”. Politeknik Negeri Lhokseumawe. 2013.

[6] Asmoro AS, Hesti, Widhi AW. “Implementasi Htb (Hierarchical Token Bucket) Untuk Manajemen Bandwidth Pada Router Internet Di Universitas Jenderal Soedirman HTB”. 2011.

[7] Lin CP, Jenhui, Chen HL. “An efficient bandwidth allocation algorithm for real-time VBR stream transmission under IEEE 802.16 wireless networks”. Journal of Network and Computer Applications 2010; 33: 467-476.

[8] Liu X, Fang X, Qin Z, Ye C, Xie M. “A Short-Term Forecasting Algorithm for Network Traffic Based on Chaos Theory and SVM”. Journal of Network and Systems Management 2011; 19: 427-447.

[9] Song Q, Cui Y, Li L. “An Adaptive Bandwidth Management Scheme for Heterogeneous Wireless Networks. Recent Advances in Computer Science and Information Engineering”. Lecture Notes in Electrical Engineering 2012; 127: 431-7.

[10] Lee J, Kang M, Jin Y, Kim H, Kim J. “An Efficient Bandwidth Management Scheme for a Hard Real-Time Fuzzy Control System Based on the Wireless LAN”. Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. Lecture Notes in Computer Science 2005; 3642: 644-653.

[11] Belbachir R, Maaza ZM, Kies A. “Available Bandwidth Estimation with Mobility Management in Ad Hoc Networks”. Advances in Wireless, Mobile Networks and Applications. The series Communications in Computer and Information Science 2011; 154: 304-315.

[12] Rahmes M, Hackett J, Fox K, Pemble R, Lemieux G, Gans H. “A cognitive group hierarchy game theoretic framework for bandwidth management”. Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA) 2014 IEEE International Inter-Disciplinary; 250–6.

[13] Maestrelli R, Almeida L, Coutinho D, Moreno U. “Dynamic bandwidth management in networked control systems using quantization”. Special Issue on the 6th Workshop on Adaptive and Reconfigurable Embedded Systems 2014; 11: 58-61."

[14] Box, E.P. George & Jenkins M. Gwilym. (2004). “Time Series Analisys Forecating and Control”. NewPrentice Hall

[15] Chen B. “Tsukamoto Fuzzy Model”. http://www.bindichen.co.uk/post/AI/tsukamoto-fuzzy-models.html. 2013.

[16] Kusumadewi, S., & Purnomo, H. “Aplikasi Logika Fuzzy untuk Pendukung Keputusan”. Yogyakarta: Graha Ilmu. 2004