The Prediction of Demands for Bandwidth in Computer Network Through Fuzzy Time Series Song-Chissom Method

Mohamed Ahmed

Abstract

The increasing of utilization and development of internet network infrastructure in higher educational institutions influences highly on the increasing of demands for bandwidth in computer network, hence it is necessary to have an efficient, reliable, yet economical service system. It can occur by conducting an appropriate and good system planning. Prediction of demands for bandwidth is one of supporting factors to plan an efficient and reliable internet servive system. This study implemented fuzzy time series Song-Chissom method adjusted the academic calendar. The forecasting system of fuzzy time series captures the pattern of prior (past) data to project the following (future) data. From the testing results obtained by using Song-Chissom method, it was found the mean error of 7.669% for course condition. From the result of prediction of demands for bandwidth, it is proposed 16 Mbps for course condition in State Polytechnic of Sriwijaya.

Key Words: Prediction/Forecasting, *Bandwidth, Fuzzy time series*

1. Introduction

Prediction of demands for bandwidth in computer network in higher educational institutions is needed to identify the demands for bandwidth in internet connection network in order that it can improve the quality of academic services and estimate how much the expense to spend for bandwidth leasing. If bandwidth is higher than the true demands, it will waste the bandwidth. In contrast, if the bandwidth is lower than the true demands, the access will be slower and it gives the user (consumer) adverse consequence. The internet use collectively can affect network performance as the number of users is increasing. The network performance has an important role to arrange the demands for bandwidth for each various internet application service. The availability of network bandwidth is an important factor to select web service (Foster and Kesselman, 2003).

The development of computing technology has come toward soft computing technologies (another term of smart computer) (Mulyadi et al, 2006). One component of soft computing is fuzzy logic which has been widely applied in various fields of life. The most important application is to assist people to make decisions. The truth that the problems in the real world in various fields are complex and involve many variables and require quick decisions, open a big opportunity for fuzzy logic to be applied. Fuzzy logic can provide a value from zero continuously up to one (Liman and Johansah, 2005).

Many forecasting methods develop as many fields require more accurate forecasting. Time series modeling by using artificial intelligence affords to learn the behavior of data to gain more accurate forecasting. Time series modeling with fuzzy time series is one of the artificial intelligence methods which are growing better. Forecasting system with fuzzy time series is able to capture the pattern of prior (past) data to project the

following (future) data (Robandi, 2006). The process also does not require a learning system of a complex system, hence fuzzy time series is easier to use.

This study developed fuzzy time series system to predict demands for bandwidth in computer network. The data source needed was the daily traffic data to implement in the software system designed. Borland C# program is used as supporting software.

2. Literature Review

2.1 Theory of Forecasting

In general term, forecasting is interpretation. However, by using particular techniques, forecasting is not only interpretation. Forecasting is defined as the use of statistical techniques in form of the future vision based on the processing of historical numbers (Buffa et al., 1996).

2.2 Fuzzy Logic

Commonly, fuzzy logic is introduced as an appropriate way to map an input space into an output space, which has continuous values and fuzzy logic is indicated in degree of membership and degree of truth (Kusumadewi, 2002).

Some reasons to use fuzzy logic among others (Kusumadewi and Purnomo, 2004) are:

1. The concept of fuzzy logic is easy to understand, because fuzzy logic includes simple and easily-understood mathematical concepts which underlie fuzzy reasoning.
2. Fuzzy logic is very flexible, which is able to adapt to alterations and uncertainties belong to the problem.
3. Fuzzy logic can tolerate inaccurate data.
4. Fuzzy logic is able to model very complex nonlinear functions.
5. Fuzzy logic can work with the control techniques conventionally.
6. Fuzzy logic is based on natural language. It uses daily language that is easy to understand.
7. Fuzzy logic can establish and apply the experiences of experts directly without training process.

Fuzzy Discourse

Fuzzy discourse is discourses to discuss in a variable in the fuzzy system (Kusumadewi and Hartati, 2006). It is used to anticipate uncertain values. In crips discourse, the membership value of an ítem in a discourse may have two pissibilities; one (1) which means an ítem belongs to a discourse, or zero (0) which means an ítem does not belong to the discourse. Whereas, fuzzy discourse has membership values in range of 0 – 1, which means that fuzzy discourse can represent interpretation of each value based on its opinion or decision, or probability.

Fuzzy discourse has two attributes, namely:

- Linguistic, is naming a group representing a particular condition by using natural language such as: NEAR, MEDIUM, FAR.
- Numerical, is a value (number) indicating size of a variable such as: 40, 25, 50 etc.

Fuzzification is a process to transform an input variable in form of crips into a linguistic variable in the form of fuzzy discourses with each membership functions (Wahyudi, 2005).

Definitions of *Fuzzy Time Series* by Song-Chissom

There are several definitions and theories about *fuzzy time series* of method introduced by Song and Chissom as follow:

Definition 1: For example Y(t) (t =...., 0, 1, 2,....), a subset of R^1, becomes *universe discourse* with *fuzzy set* $f_i(t)$ (i = 1, 2,....) defined and F(t) is a set of $f_i(t)$, $f_2(t)$,....then F(t) is known as *fuzzy time series* defined in Y(t) (t = ..., 0, 1, 2,....). From this definition, F(t)is known as linguistics variables and $f_i(t)$ (i = 1, 2,...) as linguistic possibility values F(t). Because in different time, F(t) values can be different, F(t) as *fuzzy set,* is function of time t. And *universe discourse* can be different in each time. Therefore, Y(t) is used for that t.

Definition 2: For example, F(t) is only caused by F(t-1), that is $F(t-1) \rightarrow F(t)$. Then this relationship can be represented as $F(t) = F(t-1) \circ R(t, t-1)$ where R(t, t-1) as fuzzy relationship between F(t – 1) and F(t), and $F(t) = F(t-1) \circ R(t, t-1)$ which is stated as the first order model of F(t).

Definition 3: For example, R(t, t – 1) is the first order model of F(t). If for each t, R(t, t – 1) is independent t, that is for each t, R(t, t – 1) = R(t - 1, t - 2), then F(t) is called as time- variant fuzzy time series. Conversely, if for each t, R(t, t – 1) is dependent t, that is for each t, R(t, t – 1) = R(t - 1, t - 2), then F(t) is known as time- invariant fuzzy time series.

Theory 1: For example F(t) is fuzzy time series. If for each t, F(t) = F(t – 1) and F(t) only has finite element, then F(t) is time-variant fuzzy time series. $f_i(t)$ (i = 1, 2,...n.) is determined subjectively. It defines that time-invariant or time –variant fuzzy time series are in the same universe. Nevertheless, since it is not arbitrary to determine fuzzy membership of $f_i(t)$, so the concept of time-invariant or time –variant fuzzy time series should be highly significant. Theory 2: If F(t) is fuzzy time series, F(t) = F(t – 1) for each t and F(t) only has finite element $f_i(t)$ (i = 1.....n), so:

$R(t,t-1) = f_{i1}(t-1) \times f_{j1}(t) \cup f_{i2}(t-2) \quad f_{j2}(t-1) \cup .....\cup f_{im}(t-m) \times f_{jm}(t-m+1)$

With m > 0 and all fuzzy sets are different.

Definition 4: if F(t) is caused by less than several fuzzy sets F(t – n), F(t - n+1), …F(t – 1), then its fuzzy relationship represented as:

$$A_{i1}, A_{i2}, \ldots A_{in} \rightarrow A_{ij} \qquad (1)$$

Where $F(t-n) = Ai1$, $F(t-n+1) = A_{i2}$,…..$F(t-1) = A_{in}$, this relationship is called nth-order fuzzy time series model.

Definition 5: For example F(t) is caused by a F(t – 1), F(t – 2),…., and F(t – m) (m > 0) simultaneously and and the relationship is time variant. F(t) is known as time-variant fuzzy time series and this relation is illustrated as fuzzy relation equation formulated as follows:

$$F(t) = F(t-1) \text{ o } R^{w}(t, t-1) \quad (2)$$

Where w > 1 as the time parameter to influence forecasting f(t). Diverse computation methods scarcely provided to compute related to $R^{w}(t, t-1)$.

2.2 Measurement of Forecasting

In calculating the forecasting error, MAPE (Means Absolute Percentage Error) was used. MAPE is the mean of absolute percentage error of forecasting formulated by equation 2.

$$|X| = \frac{\sum_{t=1}^{n} \frac{|X_t - F_t|}{X_t}}{n} \quad (3)$$

Where:

X_t = Peroid data value t

F_t = Period Forecast t

n = number of data

3. METHOD OF STUDY

3.1. Material of Study

The material needed to conduct this study was historical data of bandwidth use in computer network in State Polytechnic of Sriwijaya.

3.2. Instruments of Study

The instruments used in the study were as follow:

1. Hardware is computer device that can physically be seen. In a computer system, hardware is classified into four parts; the input units, output units, processing units, and storage units. In the development of the system, the minimum specifications of hardware used are: 1GB RAM, Intel Core 2 Duo processor, 320 GB HDD, Keyboard, Monitor, Mouse.
2. Software in computer system is a series of commands with specific rules to operate the hardware. It consists of three parts, namely operation systems, programming language and application programs as contributing factors of computer system. The software used to build this system were: Borland C # and Microsoft Excel.

3.3. Process of Study

The following explanation explains the process of the study:

3.3.1. Procedure of Study

The figure 1 illustrates the procedure of study.

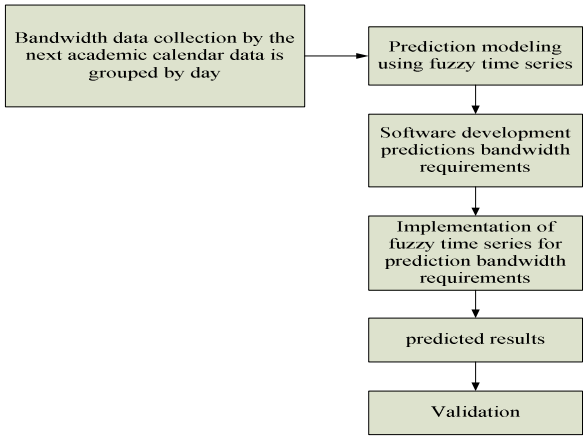

Figure 1 Procedure of study of forecasting of demands for bandwidth on computer network

The procedure of study is elaborted and described in the figures and explanation below.

1. Bandwidth Data Collection

Data of bandwidth use was obtained from computer network administrator in State Polytechnic of Sriwijaya. The data taken was traffic historical data prepared. Furthermore, the data obtained was historical data and grouped daily; Monday, Tuesday, Wednesday, Thursday, Friday and Saturday starting from 07:00 a.m. until 6.00 p.m. The data was used as input data in this study.

For validation, the data obtained was compiled and divided into 2 parts, the first part was used to predict through fuzzy time series method and the second one to examine performance of fuzzy time series whether the method used produced the targeted output.

Based on the academic calendar, the data collected was course data for 7 weeks from November 7, 2011-December 24, 2011, six weeks were used as prediction data and the 7th week as validation data. The more data, the better to resolve the problem.

2. Prediction Modeling through fuzzy time series

These were phases conducted by using fuzzy time series method. They were based on the historical time series, namely:

a. Identifying the universe discourse and compartmentalizing it into intervals with equal length. This phase found out minimum and maximum values of the actual data of bandwidth (U = [min, max]) to be used as actual data of universe discourse. The table 3.1 below shows the values:

Table 1 Bandwidth Data

Bandwidth	*Bandwidth*	*Bandwidth*	The	*Fuzzy*

Data	*max value*	*min value*	Amount of data	set
Collage	26164000	66800	432	401

b. Compartmentalizing the universe discourse into intervals with equal length: u1, u2, ... Um. The number of intervals which would be compatible with the number of linguistics variables (fuzzy sets) A1, A2, ...should be observed. The interval used for course condition was h = 65243.

c. Building the fuzzy sets Ai following the interval in phase 2 and implementing the rules of triangle membership for each interval in each fuzzy set built. In other words, it was defining fuzzy discourse in universe discourse. This phase transformed the compartmentalized universe discourse in discourse of crisp numbers into fuzzy discourse based on the interval. Meanwhile, the value of fuzzy set A1 = 66 800, 66 800 + h = A2, A3 = 2h + 66 800 66 800 + ... A401 = 400h for course condition.

d. Performing historical data fuzzification and setting logical relationship of fuzzy. This phase identified the value of membership for each fuzzy set of historical data, in value range of 0-1. The membership values were obtained from the membership functions predetermined. Then calculating fuzzy and defining logical relationship of fuzzy by following the rules; If Ai was a fuzzy production of day n-1 and Aj was a production of day n fuzzification, so the logical relationship of fuzzy indicated as Ai → Aj, Ai was the current state and Aj was the next state. *Rule = current state → next state.*

e. Selecting the most appropriate order and basis model (w) and calculating fuzzy operations. This phase determined the value of fuzzy inference based on the basis model (w) in formula:

$$r(MBF)^{W+1} = \frac{MBF_1 + MBF_2 + \cdots + MBF_W}{w} \quad (4)$$

The order used was order 12 for the fuzzy time series Song-Chissom method. While the basis model used was the basis model 6 for bandwidth data in course condition.

f. Defuzzification data
Conducting output defuzzification predicted. This phase determined the predicted results in the form of crisp values through COG defuzzification method (Center of Gravity) for Song-Chissom method in formula:

$$y = \frac{A1 * c1 + A2 * c1}{A1 + A2} \quad (5)$$

$$v = r(NS) * A1 + r(NM) * A2 + \cdots + r(PB) * A7$$

$$Fi = y + v$$

3. The process model used in the software development was a linear sequential model, also known as the waterfall model. It covered activities as shown in the following figure:

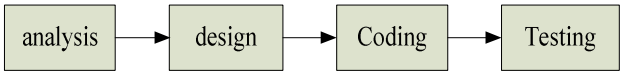

Figure 2 Sequential Linier Model

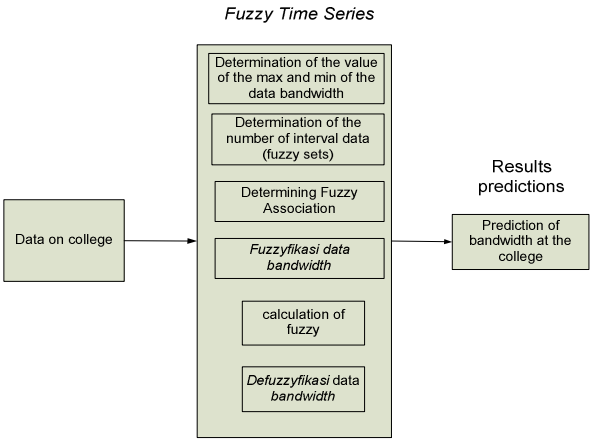

Figure 3 System Scheme (Architecture) of Fuzzy Time Series

4. System implementation is the phase of system setting prepared to operate. The activities conducted followed the plans; data selection and training, hardware and software installation, program testing, system testing. The personnel training was performed to operate the system, including preparing inputs, processing data, operating systems, maintaining and securing the system. The results of prediction of bandwidth obtained were the comparison between the actual data and prediction data.

5. The results of the study was the prediction of bandwidth for each condition of courses, examinations and holidays.

3.4. Software System

In running software system of forecasting of demands for bandwidth, Borland C# program was implemented. The flowchart of software of forecasting through Fuzzy Time Series is illustrated in figure 4:

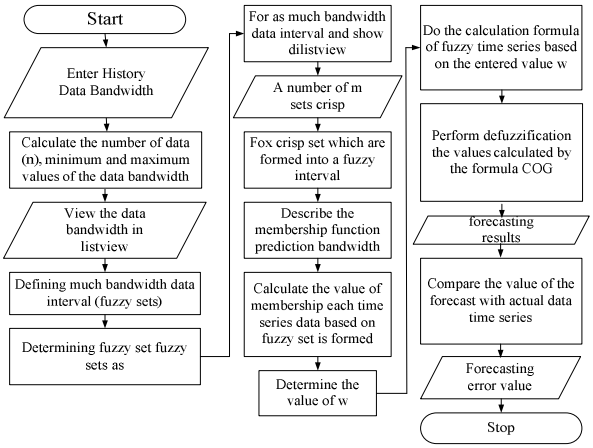

Figure 4 Flowchart of Forecasting software through Fuzzy Time Series

3.5. Interface Design

The interface design is a design of information management in a system. The design was evolved as an overview/object of a program/application which would be established.

4.1. Results

The use of Borland C# programming language, for application system of prediction of demands for bandwidth in State Polytechnic of Sriwijaya through Fuzzy Time Series Song-Chissom method had been successfully built.

1. Menu of Historical Data of Bandwidth

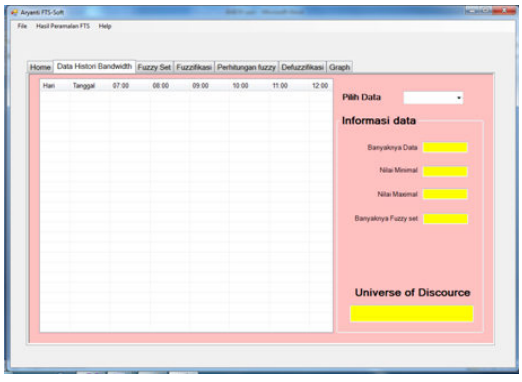

Figure 5. Menu of Historical Data of Bandwidth

The form in figure 5 demonstrates the historical data obtained by using select data column. The column gives three options of data; course data, examination data and holiday data. The figure 5 also presents information about the number of data, the minimum value, the maximum value, the number of fuzzy sets

and the universe discourse for each of data options. The figure 6 shows the menu of historical data of bandwidth.

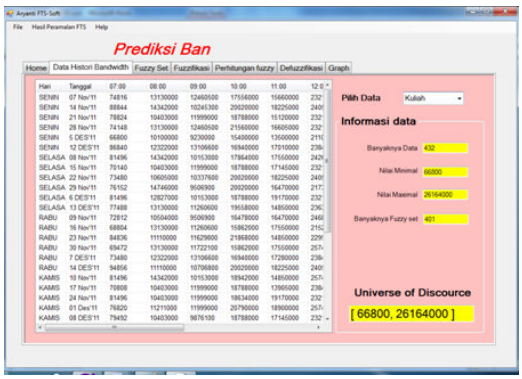

Figure 6. Menu of Historical Data of Bandwidth

2. Menu of Fuzzy Set

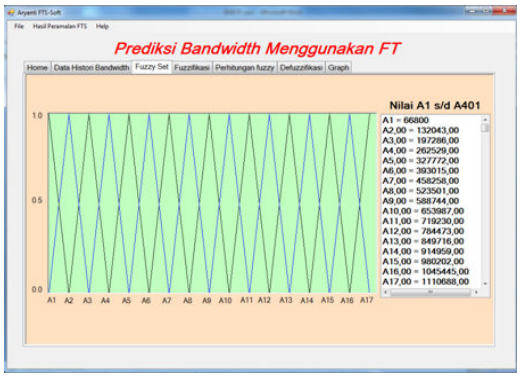

Figure 7. Menu of the Results of Fuzzy Set

The figure 7 illustrates the results of fuzzy sets including the linguistic values and numerical values. The numerical values were represented in numbers of fuzzy set, the linguistic values were from A1 up to A401 which consisted of A1 as the minimum value of the data, A2 as the minimum value of the data added by interval, and forth up to the A401 value as the maximum value of the data.

3. Fuzzification Menu

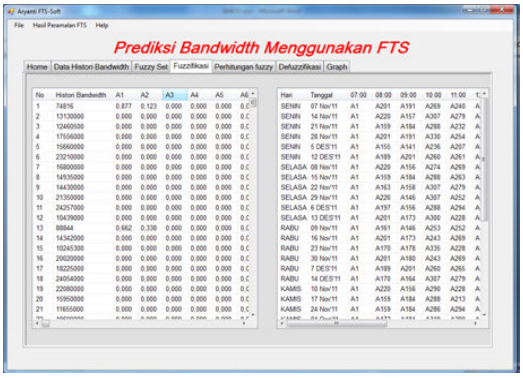

Figure 8 Menu of the Results of Fuzzification

The figure 8 displays the results of fuzzification which were the values of the affiliation function of each historical data based on the fuzzy set configurated. They were categorized in two tables; the first table included fuzzy affiliation values of each historical data, and the second table was fuzzification table.

4. Menu of Fuzzy Calculation

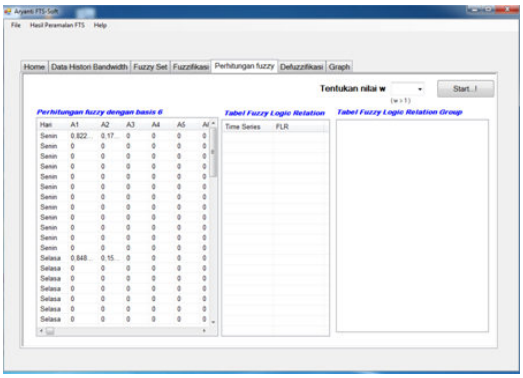

Figure 9 Menu of Fuzzy Calculation

The form in Figure 9 was used for fuzzy calculation, which particularly used Song-Chissom method. Decide the value of the order (12) and press start. The figure 10 below displays the results of fuzzy calculation by using Song Chissom method.

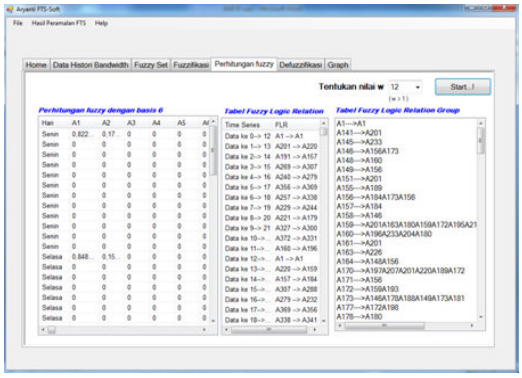

Figure 10 Menu of the Results of Fuzzy Calculation

5. Defuzzification Menu

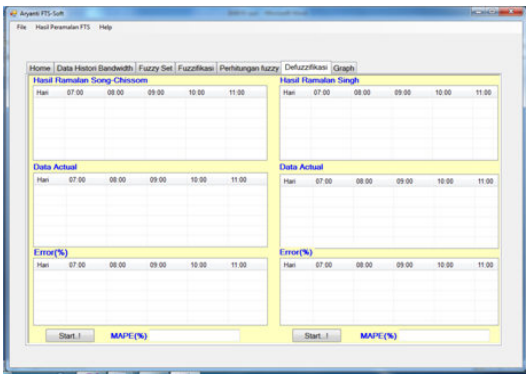

Figure 11 Defuzzification Menu

The figure 11 presents the form of the results of defuzzification which included the results of forecasts, the actual data and error by using Song-Chissom method, the historical data, then select the start column. The figure 12 below demonstrates the results of defuzzification.

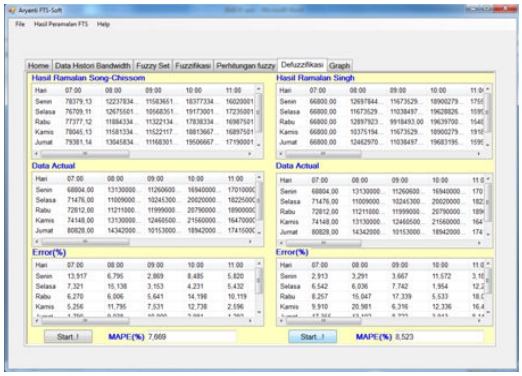

Figure 12 *Menu* of Results of Defuzzification

6. Menu Graph

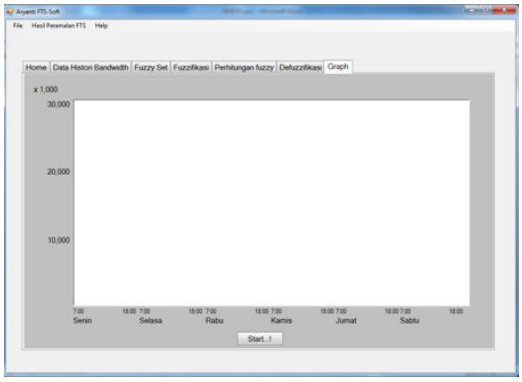

Figure 13 *Menu Graph*

Menu graph is a menu to display graphs of the results of prediction through Song-Chissom method and the actual data from bandwidth. Press the start button to make it displayed. The figure 14 below is the graphic display of the results of prediction of bandwidth.

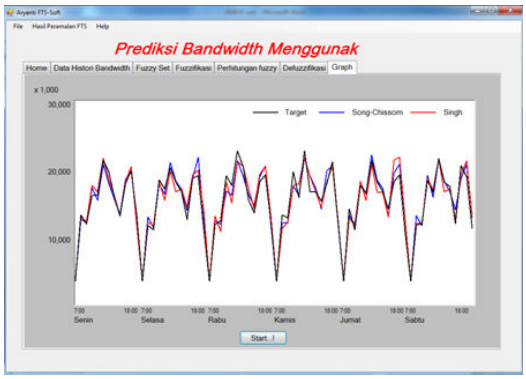

Figure 14 Result *Menu* of prediction of *bandwidth*

4.2. Discussion

4.2.1. *Fuzzification* Process

The historical load data was used in the fuzzification process to gain the values of membership function which would be used for fuzzy calculation to determine the prediction of demands for bandwidth in State Polytechnic of Sriwijaya. The load data used in the fuzzification process was load data in November and December 2011 starting from 7.00 a.m. up to 6.00 p.m. as shown in table 1

.

Table 1 *Fuzzification* of Historical Load Data

Day	Date	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00
Monday	07/11/2011	A1	A201	A191	A269	A240	A356	A257	A229	A221	A327	A372	A160
Monday	14/11/2011	A1	A220	A157	A307	A279	A369	A338	A244	A179	A300	A331	A196
Monday	21/11/2011	A1	A159	A184	A288	A232	A356	A341	A289	A230	A306	A340	A160
Monday	28/11/2011	A1	A201	A191	A330	A254	A356	A257	A249	A221	A308	A334	A233
Monday	05/12/2011	A1	A155	A141	A236	A207	A323	A245	A222	A170	A268	A293	A145
Monday	12/12/2011	A1	A189	A201	A260	A261	A365	A331	A253	A197	A335	A334	A233
Monday	19/12/2011	A1	A201	A173	A260	A261	A365	A331	A253	A197	A295	A334	A198

4.2.1.1 Prediction of Demands for *bandwidth*

The fuzzification data of historical load shown in table 4.1 was used for the fuzzy and defuzzification calculations in Song Chissom method and to determine *fuzzy logic relationship* in order to obtain the output of forecasting of demands for bandwidth through fuzzy time series method proposed by Song-Chissom. The prediction results of data of demands for bandwidth would be analyzed hourly and daily.

4.2.1.2 Analysis of Results of Prediction of Demands for Bandwidth in Course, Examination, and Holiday Conditions

Table 2 demonstrates the results of tested data of demands for bandwidth in course conditions on Monday December 19, 2011 by using Song-Chissom method.

Table 2 Result of Forecasting of *Bandwidth* in Course Conditions

Day	Hour	Forecasting Result	Target	MAPE
		Song-Chissom		Song-Chissom
Monday	7	78379,13	68804,00	13,917
	8	12237834,33	13130000,00	6,795
	9	11583651,00	11260600,00	2,869
	10	18377334,33	16940000,00	8,485
	11	16020001,00	17010000,00	5,82
	12	23104501,00	23843000,00	3,097
	13	19253334,33	21600000,00	10,864
	14	16167501,00	16530000,00	2,193
	15	13246001,00	12876000,00	2,874
	16	20066667,67	19250000,00	4,242
	17	21774001,00	21774000,00	0
	18	12257917,67	12906400,00	5,025
Min		78379,13	68804,00	0
Max		23104501,00	23843000,00	13.917
Mean		15347260,29	15599067,00	5,515

Table 2 shows the results of forecasting in course condition with the maximum error of 13.917%, the minimum error of 0%, and the mean error of 5.515% for Song-Chissom method. The error percentage of forcasting of demands for bandwidth hourly in daily can be seen in Figure 15.

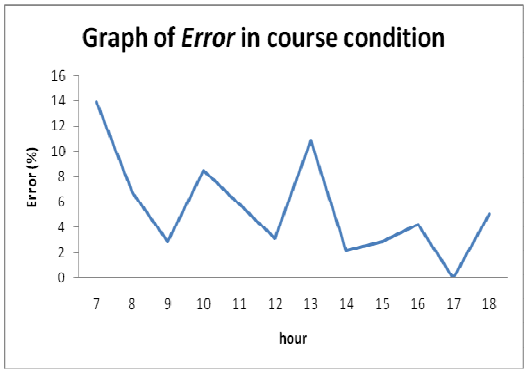

Figure 15 Graph of *Error* in course condition
(19/12/2011)

Figure 15 is the graph of the mean error of prediction of bandwidth in course condition on Monday, December 19, 2011. The graph illustrates that the prediction of bandwidth in course condition had a mean error of 7.669%.

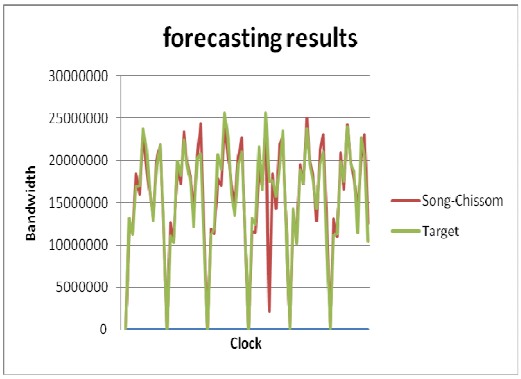

Figure 16 Graph of the result of forecasting in course condition

From the results of forecasting of demands for bandwidth through fuzzy time series Song-Chissom method, it was yielded good forecasting and it came toward the actual data as shown in Figure 16. It indicates that the information system of forecasting of demands for bandwidth by using the fuzzy time series method can be used as an alternative method to forecast the demands for bandwidth.

5. Conclusion

The forecasting of demands for bandwidth through *fuzzy time series* Song-Chissom yielded good forecasting with the error of 7.669% and the result obtained came toward the actual data.

References

Buffa S. Elwood, Rakesh, and K.Sarin, 1996. Modern Production and Operation Management, Eigth Edition, John Willey and Sons Inc., London.

I Foster and C Kesselman, 2003. The Grid: Blueprint for a New Computing Infrastructure, Second edition, Morgan Kaufman.

Kusumadewi S, 2002. Analisis & Desain Sistem Fuzzy Menggunakan Tool Box MathLab Graha Ilmu, Yogyakarta.

Kusumadewi S dan H Purnomo, 2004. Aplikasi Logika Fuzzy untuk Pendukung Keputusan. Graha Ilmu, Yogyakarta.

Kusumadewi S dan S Hartati, 2006. Neuro-Fuzzy: Integrasi Sistem Fuzzy dan Jaringan Syaraf. Graha Ilmu, Yogyakarta.

Liman and Johansah, 2000. Pengenalan Logika *Fuzzy*,h *itp://www I bpkpen ab ur. or. id/kps-jkt/berita/200004/artikel4. htm, 2000,* diakses 26 juni 2011.

Mulyadi Y, Abdullah G.A, dan Hakim L.D, 2006. Aplikasi Logika *Fuzzy* dan Jaringan Syaraf Tiruan

sebagai Metode Alternatif Prediksi Beban Listrik Jangka Pendek

Q Song and B.S. Chissom, 1993. Forecasting Enrollments with Fuzzy Time Series – part I, Fuzzy Sets and Systems 54 (1993), pp. 1-9.

Robandi, I, 2006. Desain Sistem Tenaga Modern, Optimasi, Logika Fuzzy dan algoritma Genetika. Andi. Yogyakarta.

Wahyudi, 2005. Implementasi Fuzzy Logic Controller pada Sistem Pengereman Kereta Api, Transmisi, Vol. 10, No. 2, Desember 2005, 10-13.