Mining Time Series Data with Fuzzy Association Rules

Pao-Ta Yu[*] *Chung-Ming Own*

Department of Computer Science and Information Engineering*

National Chung Cheng University

160, San-Hsing, Ming-Hsiung

Chiayi, Taiwan 62107, R.O.C

Tel[*]: 886 5 2720411 ext. 6015

Email[*]: csipty@ccunix.ccu.edu.tw

ABSTRACT

A novel algorithm for forecasting the relationship of time series is proposed to address two important issues: how to combine the fuzzy time series and association rule, and how to improve efficiency and simplicity without losing the accuracy. A complete procedure is proposed and the results of this algorithm show more efficient than those of the methods proposed in previous studies on the application of forecasting enrollments.

1.INTRODUCTION

The problems of forecasting time series have been recognized as an important subject in our daily life such as forecasting the weather, earthquake, stock fluctuation, students' performance and any phenomenon that can be modeled by a collection of random variables indexed on time [1]. Numerous investigations have been proposed to solve these problems by the Moving Average, Integrated Moving Average and Autoregressive Integrated Moving Average [2][3]. Further, a breakthrough of computational intelligence paradigms was proposed in this area in the 1980s [4]. However, most of these research always focus on analyzing the raw data of historical records, instead of meta data which can be described by linguistic values such that some implicitly messages aren't lost [5] [6]. Since these linguistic values can be translated into fuzzy sets, fuzzy time series have been developed to solve these problems.

The fuzzy set theory has fruitful achievements in both theory and applications since Zadeh proposed from 1965 [7]. The motivation of fuzzy set theory attempts to capture expert's knowledge, cope with the vagueness of human's reasoning and rule the implementation. Using the fuzzy set theory, the trend of the time series can be implemented and forecasted. Song and Chissom first presented the definition of fuzzy time series and modeled its fuzzy relationship from historical enrollments data [7]. Then, Wu proposed the fuzzy-based method to predict future trends and had a better result than traditional forecasting methods on forecasting teacher numbers, government expenditure and so on [9]. Besides, Hwang *et al*. also presented a method that focus on the variation between the enrollments of current year and the trend of past years [1]. Hwang *et al*.'s model is more efficient and simpler than most of other approaches, although its accuracy is quite limited. In addition, the relational equation that they proposed is clearer.

From above specification, fuzzy time series have successfully applied to solve the fuzzy forecasting problem. However, the limitation of these methods fails to mine the hidden patterns, trends and relationship from the historical data. Thus, data mining, the efficient discovery of hidden patterns from large collection of data, is implemented as a novel method to improve the forecasting time series. Due to the combination of fuzzy time series and association rule of data mining, this is one of the most important contributions in this paper.

Data mining recently has been recognized as a new field to find hidden patterns, trends and relationships on data. Owning to the applicability of information management, policy decision, fraud detection, marketing strategy, financial forecasting, processing control and many applications, data mining has attracted tremendous attention from statistics, computer science, and artificial intelligence research [10]. Besides, the task how to accomplish and the data how to retrieve is the focal point of data mining techniques. In the supermarket, analysis of past transaction data is a commonly used approach to give insight into the merchandise. The transaction data often consists of records with items bought by the customers. Therefore, identifying rules and relationships of patterns from the transaction databases means to mine association rules from records. The association rules are the most commonly methods to seek hidden patterns introduced in [10].

We propose the fuzzy association rules as a novel approach to deal with forecasting problems where the historical data are translated into linguistic values. This approach integrates the fuzzy set theory and the association rules to find "associations" among the items with linguistic values. The advantage of this approach is to eliminate the deficiency of traditional fuzzy time series approaches and satisfy the accuracy.

In Section 0, we review the concept of association rules and the fuzzy time series, while the method of fuzzy association rules is described in Section 0. Performance analysis and some conclusions are given in Sections 0 and 0, respectively.

2.BASIC CONCEPTS

2.1 Association Rules

Data mining can be applied to discover the useful patterns and rules by exploring and analyzing a large quantity of data. That is, a collection of data from customer surveys, health studies, market examinations, item banks and other raw data needs further analysis to transform it into useful information. In general, data mining involves the recognition of implicitly patterns that are hard to be analyzed, even though the use of traditional statistical techniques. In addition, an important mining task in this area is to discover association rules [10]. Suppose that basket data consists of items bought by a customer over a period of time, mining a large collection of basket data by association rules is to refine the relation between the sets of items with some specified *confidences* and *support*. The definitions of the association rules are reviewed as follows:

Let D be a database of transactions and $I=\{i_1, i_2, \ldots, i_m\}$ represent the set of m items in a database D. Further, let T be a collection of transactions, where each transaction t has a unique identifier and contains a set of items (itemset is for short) such that $t \subseteq I$. If $x \subseteq t, y \subseteq t,$ and $x \cap y = \varnothing$, then the association rules can be formed as the relation of $x \Rightarrow y$. The *confidence* c is the factor that quantifies how many transactions including itemset x also contains itemset y. Further, the *support* s of $x \Rightarrow y$ is the factor that quantifies how often the itemsets x and y occur together in D. That is, $s = |\{t \in T \mid x \cup y \subseteq t\}|$. The problem of mining association rules is to find all rules that contain *support* and *confidence* greater than the user-specified minimum *support* and minimum *confidence* [11]. Hence, the association rules with unnoted relationship will be refined by the threshold of minimum *support* and minimum *confidence* [12].

Several algorithms have been proposed to mine association rules, such as APRIORI [12], DHP [13], PARTITION [13], DMA [15], AIS [16], SETM [17], etc. However, most of them are based on the APRIORI algorithm, because it is one of the best algorithms for association mining and has excellent scale-up properties [18][19].

2.2 Fuzzy Time Series

Time series that is defined as a collection of random variables indexed on time can be retained to model many occurrences. For example, the temperature changes at time t. People can evaluate the temperature of next day by the information based on the forecasting method and knowledge of experience. In a word, people can prevent and diagnose what situations will happen in advance with the data of forecasting time series. Furthermore, consider the points of view and knowledge of experience; fuzzy set theory has been applied to model the dynamic time series process with linguistic values [7]. This has led to the approach of fuzzy time series [1][6][8][9]. The concept of fuzzy time series is depicted as follows:

Let $\{X_t \in R \mid t = 1, 2, \cdots\}$ be a time series and U be the universe of discourse. Besides, let $\{U_i \mid \bigcup_{i=1}^{n} U_i = U\}$ be a set of ordered partition of U where linguistic variables $\{L_i \mid i = 1, 2, \cdots, n\}$ are given. The $F(t)$ is defined as the collection of memberships of $u_t(1), u_t(2), \cdots, u_t(n)$ corresponding to $L_1, L_2, \cdots, L_n$, for a time instance t, then

$$F(t) = u_t(1)/L_1 + u_t(2)/L_2 + \cdots + u_t(n)/L_n ,$$

Such that '+' denotes the connection, $u_t(i)$ denotes the degree of X_t belonging to L_i for all $i \in \{1, 2, \cdots, n\}$ [20]. For the purpose to make the representation more convenient let the $F(t)$ be represented as the row vector of [1], and show as follows:

$$F(t) = [u_t(i)]_n .$$

There are two categories of quantitative forecasting models: causal models and time-series models. Causal models aim at refining the factors and relationships between the inputs and outputs. Time-series are the other models, which are based on the pattern of the past, and then try to infer the future [21]. Further, fuzzy time series is designed to contain the particulars of two models mentioned above.

Assume that the observations at time t are the accumulated results at the previous times, hence, there exists a fuzzy relation $R(t, t-1)$ between $F(t-1)$ and $F(t)$, this relation is called first-order model and shown as [8]:

$$F(t) = F(t-1) \circ R(t, t-1) , \qquad (1)$$

where "$\circ$" is denoted as the relation of composition. The most commonly used composition in the literature is the so-called max-min composition defined by Mamdani [22]. This means that

$$u_t(i) = \max \min_k (u_{t-1}(i), r_{t,t-1}(k,i)) , \qquad (2)$$

where $R(t, t-1) = [r_{t,t-1}(i, j)]_{n \times k}$, which is usually represented as a fuzzy relational matrix, is to extrapolate from $F(t-1)$ to $F(t)$. That is, a matrix of $R(t, t-1)$ whose elements are the membership values corresponding to the pairs of fuzzy relation [23]. This is the first step designed to infer the future by the patterns of the past.

Subsequently, when applying a fuzzy time series model in forecasting, it is assumed that the variable being forecasted is $F(t)$ and the patterns of the past are referring from $F(t-1)$. The relationship in Eq. (1) is analyzed and $R(t, t-1)$ is essential to obtain. Let

$$R(t, t-1) = F(t-1) \times F(t), \quad (3)$$

where " $\times$ " is denoted as the cartesian product. Equivalently, Eq. (10) is represented as follows:

$$r_{t,t-1}(i, j) = \min(u_{t-1}(i), u_t(j)), \quad (4)$$

where uses t-norm as an assessment in [6][8][24]. Similarly, we have

$$r_{t,t-1}(i, j) = u_{t-1}(i) \times u_t(j), \quad (5)$$

where uses the multiplication operation as an assessment in [1][25][27].

In the following discussion, consider that the fuzzy relation may be influenced with some stable time instances. That is, employing the $F(t-1)$, $F(t-2)$, $\cdots$, and $F(t-m)$, then $F(t)$ is obtained as

$$F(t) = (F(t-1) \cap F(t-2) \cap \cdots \cap F(t-m)) \circ R(t, t-m). \quad (6)$$

According to the max-min composition by Mamdani, we have an equivalent definition as follows:

$$u_t(i) = \min_{q=1,\ldots,m} \max_k \min(u_{t-q}(i), r_{t,t-q}(k, i)). \quad (7)$$

This model is also proposed by Song and Chissom [6][8][24]. That is, Let $R(t, t-m)$ be obtained by

$$R(t, t-m) = (F(t-1) \cap F(t-2) \cap \cdots \cap F(t-m)) \times F(t). \quad (8)$$

Equivalently, let the equation is decomposed by the membership values of

$$r_{t,t-m}(i, j) = \min \min_{q=1,\ldots,n} (u_{t-q}(i), u_t(j)), \quad (9)$$

or

$$r_{t,t-m}(i, j) = \min_{q=1,\ldots,n} (u_{t-q}(i) \times u_t(j)). \quad (10)$$

where the operator of Cartesian product is represented as t-norm or multiplication. s

Consider the first order or *m*th order model of fuzzy relation, if for any *t*, $R(t, t-1)$ is independent of *t*, i.e., $\forall R(t, t-1) = R(t-i, t-i-1)$, then the derivation results from Eq.s (1), (6) are called a time invariant fuzzy time series. Otherwise, it is called a time variant fuzzy time series.

3.NEW METHD OF FUZZY ASSOCIATION RULES

Song and Chissom indicate that traditional forecasting methods are not suited to data composed of linguistic values. Fuzzy time series method is the inspiration. On the other hand, fuzzy time series is more and more frequently applied to predict problems about enrollments, because of its simplicity and similarity to human reasoning [1][24]. In addition, association rules algorithm involves the identification of hidden patterns that are not easily acquired by traditional statistical techniques. For the purpose to enhance the forecasting advantage, we propose a novel method to forecast by association rules instead of fuzzy implementation in fuzzy time series.

Table 1. Variation of the enrollment

Years	Actual Enrollments	Variations
1971	13055	-
1972	13563	508
1973	13867	304
1974	14696	829
1975	15460	764
1976	15311	-149
1977	15603	292
1978	15861	258
1979	16807	946
1980	16919	112
1981	16388	-531
1982	15433	-955
1983	15497	64
1984	15145	-352
1985	15163	18
1986	15984	821
1987	16859	875
1988	18150	1291
1989	18970	820
1990	19328	358
1991	19337	9
1992	18876	-461

In our proposed method, the historical enrollments of the University of Alabama are adopted as our implementation [6][24]. This data are shown in Table 1. For the purpose to represent the relationship of data, we use the variations of the enrollments instead the originals. Hence, we count the variations between any two consecutive years from Table 1. We assume that the enrollment of year t is equal to x and year t–1 is equal to y. Because the variation of the enrollments between years t and t–1 is equal to x–y, we represent year t is equal to x–y.

The procedure of fuzzy association rules is described as follows: First, we define the universe of discourse. Let minimum enrollment be $D_{\min}$ and maximum enrollment be $D_{\max}$, and derive the proper positive values D_1 and D_2 to compute the universe of discourse $U=[D_{\min}-D_1, D_{\max}-D_2]$. Subsequently, the discourse can be divided into even length intervals u_1, u_2, …, u_m, where u_i is the ith divided interval and $1\le i\le m$.

Subsequently, the historical data are applied in the forms of linguistic values, let A_1, A_2, …, A_m be fuzzy sets, all the fuzzy sets will be labeled by possible linguistic values. For example, Song et al., applied the linguistic values to fuzzy sets A_1=(not many), A_2= (not too many), A_3=(many), A_4=(many many), A_5=(very many), A_6=(too many), A_7= (too many many) to forecast the enrollments of the University of Alabama [6][24]. Hence, A_q is defined as

$$A_q = u(u_1)/u_1 + u(u_2)/u_2 + \ldots + u(u_i)/u_i + \ldots + u(u_m)/u_m \,, \quad (11)$$

where u_i is the interval expressed as an element of the fuzzy set. $u(u_i)$ implies the degree of u_i belonging to A_q, and $u(u_i)\in[0,1]$.

Nevertheless, the association rules algorithm can only identify the binary relation among items (In supermarket database, the binary relation means the choice to buy or not to buy items). In this paper, apply the association rule algorithm in our proposed method, we redefine the membership $u(u_i)\in\{0,1\}$. Thus, all the fuzzy sets are defined as the same in Eq. (11) except that

$$u(u_i) = \begin{cases} 1 & \text{if } i = q\text{-}1, q, q+1 \text{ and } 0 < i \le m \\ 0 & \text{otherwise} \end{cases} . \quad (12)$$

Consequently, we fuzzify the values of the historical data by making the judgment of what fuzzy set belong to. As we mentioned above, each fuzzy set A_q is related to a linguistic value and mapped to interval u_q. Hence, the variation of year t can be judged by counting the belonging to what interval, i.e., we assume that the variation of year t is equal to p and $p\in u_q$, then the variation of year t is fuzzified as

$$F(t) = A_q \,, \quad (13)$$

where $F(t)$ is defined as the fuzzified variation of the year t, and $1\le q\le m$.

For simplicity, we express fuzzy set A_q as a row vectors whose entities are corresponding to the membership values in Eq. (11), and it is represented as

$$F(t) = A_q = [u_t(i)]_m \,, \quad (14)$$

where $u_t(i)$ is the membership value of fuzzy set A_q.

Window basis w is defined how many past years' data are used to forecast time series. Besides, the criterion matrix $C(t)$ is defined as the fuzzified variation of the last year, the operation matrix $O^w(t)$ is formed by the fuzzified variations of w past years [1]. They are shown as follows:

$$C(t) = F(t-1) = [u_{t-1}(i)]_m \,, \quad (15)$$

$$O^w(t) = \begin{bmatrix} F(t-2) \\ F(t-3) \\ \vdots \\ F(t-w-1) \end{bmatrix} = \begin{bmatrix} u_{t-2}(1) & u_{t-2}(2) & u_{t-2}(3) & \ldots & u_{t-2}(m) \\ u_{t-3}(1) & u_{t-3}(2) & u_{t-3}(3) & \ldots & u_{t-3}(m) \\ \vdots & \vdots & \vdots & & \vdots \\ u_{t-w-1}(1) & u_{t-w-1}(2) & u_{t-w-1}(3) & \ldots & u_{t-w-1}(m) \end{bmatrix} . \quad (16)$$

Contrary to the traditional forecasting on the relationships of matrixes criterion and operation, the association rules used in the fuzzy time series are focused on the discovery of imprecision and uncertainty relationships in the matrix operation. Suppose that database D contains w transactions, $I=\{\boldsymbol{I}_1, \boldsymbol{I}_2, \ldots, \boldsymbol{I}_m\}$ is a set of m distinct items that are corresponding to the entities of fuzzy set A_q. A transaction $\boldsymbol{t}_q$ is defined as a set mapped to the vector of $O^w(t)$, it is shown as:

$$\boldsymbol{t}_{q-1} = \{\boldsymbol{I}_j \mid u_{t-q}(j) = 1 \text{ and } 1 \le j \le m\} \,. \quad (17)$$

Further, apply the w transactions to the APRIORI algorithm mentioned in Section II, and derive the results of forecasted itemsets. Subsequently, let a relation matrix

$$R(t) = [r_t(i,j)]_{r\times m} = [R_1 \;\; \cdots \;\; R_r] \quad (18)$$

be a collection of row vectors translated from the results of itemsets. Note that we assume L_1, …, L_r are the results of itemsets. Each row vector R_i relates to L_i and $R_i = [r_t(i,j)]_{j=1,\ldots,m}$, where

$$r_t(i,j) = \begin{cases} 1 & \text{if } \boldsymbol{I}_j \in L_i \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } j = 1, \ldots, m. \quad (19)$$

While the relation matrix is obtained, we can derive the fuzzified variation $F(t)$ by performing

$$\begin{aligned} F(t) &= R(t)\circ C(t)^{\mathrm{T}} \\ &= [r_t(i,j)]_{h\times m} \circ [u_{t-1}(j)]_{m\times 1} \\ &= \max \begin{bmatrix} r_t(1,1)\times u_{t-1}(1) & r_t(1,2)\times u_{t-1}(2) & \cdots & r_t(1,m)\times u_{t-1}(m) \\ r_t(2,1)\times u_{t-1}(1) & r_t(2,2)\times u_{t-1}(2) & \cdots & r_t(2,m)\times u_{t-1}(m) \\ \vdots & \vdots & \vdots & \vdots \\ r_t(h,1)\times u_{t-1}(1) & c & \cdots & r_t(h,m)\times u_{t-1}(m) \end{bmatrix} \\ &= \begin{Bmatrix} \max[r_t(1,1)\times u_{t-1}(1) \;\; r_t(2,1)\times u_{t-1}(1) \;\; \cdots \;\; r_t(h,1)\times u_{t-1}(1)], \\ \max[r_t(1,2)\times u_{t-1}(2) \;\; r_t(2,2)\times u_{t-1}(2) \;\; \cdots \;\; r_t(h,2)\times u_{t-1}(2)], \\ \vdots \;\; \vdots \\ \max[r_t(1,m)\times u_{t-1}(m) \;\; r_t(2,m)\times u_{t-1}(m) \;\; \cdots \;\; r_t(h,m)\times u_{t-1}(m)] \end{Bmatrix} \\ &= [\dot{r}_1 \; \dot{r}_2 \; \cdots \; \dot{r}_m], \end{aligned} \quad (20)$$

where $\circ$ is the composition of $R(t)$ and $C(t)$ that is defined as the max-product composition, $\times$ is the multiplication operation and $\dot{r}_j$ is the entity of the forecasted result.

Finally, we interpret and defuzzify the forecasted output. Assume that

$$N_s = \sum_{j \le m} \dot{r}_j \ , \tag{21}$$

where N_s is represented as the number of forecasting items. Since each entity expresses the embedded relationship with interval u_i, the grade of each entity in the fuzzified forecasted vector is $a_i \times midpoint(u_i) \times (N_s)^{-1}$, where $midpoint(u_i)$ is defined as the middle point in interval u_i. Thus, the forecasted variation $DF(t)$ is defined as follows:

$$DF(t) = (\sum_{i \le m} a_i \cdot midpoint(u_i) \cdot \frac{1}{Ns}) / Ns \,. \tag{22}$$

The forecasting method can be presented as the following steps, and we use the enrollments of the University of Alabama as an example [1][6] [24]:

STEP1. Revise the data. From the historical enrollments data, we compute the variation of enrollments between any two continuous years (Table 1). The enrollments of the University of Alabama, we can derive U=[-1000,1400] by $D_{\min}$=-955, $D_{\max}$=1291, D_1=45 and D_2=109.

STEP2. Divide the universe of the discourse. The universe of the enrollments is divided into six even length intervals u_1, u_2, …, u_6, where u_1=[–1000, –600], u_2=[–600, –200], u_3=[–200, 200], u_4=[200, 600], u_5=[600, 1000], u_6=[1000, 1400]. We also defined the middle points of each interval in the Table 2.

Table 2. The middle point of intervals

	u_1	u_2	u_3	u_4	u_5	u_6
$midpoint(u_i)$	-800	-400	0	400	800	1200

STEP3. Define the fuzzy sets of the divided enrollments data. In our example, the enrollments of the University of Alabama can be represented as six fuzzy sets, and the linguistic values are A_1=(big decrease), A_2=(decrease), A_3=(no change), A_4=(increase), A_5=(big increase), A_6=(too big increase). Then the fuzzy sets are defined as:

$A_1=1/u_1+1/u_2+0/u_3+0/u_4+0/u_5+0/u_6$,

$A_2=1/u_1+1/u_2+1/u_3+0/u_4+0/u_5+0/u_6$,

$A_3=0/u_1+1/u_2+1/u_3+1/u_4+0/u_5+0/u_6$,

$A_4=0/u_1+0/u_2+1/u_3+1/u_4+1/u_5+0/u_6$,

$A_5=0/u_1+0/u_2+0/u_3+1/u_4+1/u_5+1/u_6$,

$A_6=0/u_1+0/u_2+0/u_3+0/u_4+1/u_5+1/u_6$.

STEP4. Fuzzify the values of historical data. To fuzzify the variation is to judge what fuzzy set the variation belongs to. If the variation of year t is equal to p and $\exists p \in u_i$, the variation would be fuzzified into fuzzy set A_i. For example, the variation of year 1972 is 508. Therefore, the variation of year 1972 can be fuzzified into A_4. We represent the variation of 1972 as

$$F(1972) = A_4 = [0\ \ 0\ \ 1\ 1\ 1\ \ 0] \,.$$

STEP5. Choose the window basis w, and compute the essential matrixes. For example, we set w=5 and forecast the enrollment of year 1977 as an explanation. In this case, $O^5(1977)$ and $C(1977)$ are shown as follows:

$C(1977)=F(1976)=[0\ 1\ 1\ 1\ 0\ 0]$,

$$O^5(1977) = \begin{bmatrix} F(1975) \\ F(1974) \\ F(1973) \\ F(1972) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

STEP6. Apply the operation matrix $O^w(t)$ to the APRIORI algorithm mentioned in SECTION II. Obtain a series results of forecasted itemsets.

STEP7.Convert the result of itemsets into the relation matrix $R(t)$. Therefore, as the above example, we obtain $R(1977)$ as

$$R(1977) = [r_{1977}(i,j)]_{9\times 6} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

STEP8. The fuzzified variation $F(t)$ is obtained following the previous definition. For example, $F(1977)$ is derived as

$$\begin{aligned} F(1977) &= R(1977) \circ C(1977)^{\mathrm{T}} \\ &= [r_{1977}(i,j)]_{9\times 6} \circ [u_{1976}(j)]_{6\times 1} \\ &= \begin{Bmatrix} \max[0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0], \\ \max[0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0], \\ \max[1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0], \\ \max[0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 1], \\ \max[0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0], \\ \max[0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0] \end{Bmatrix} \\ &= [0\ \ 0\ 1\ 1\ 0\ \ 0]. \end{aligned}$$

Table 2. Comparisons of average forecasting errors with different methods (*support*=2)

Average error	*w*=2	*w*=3	*w*=4	*w*=5	*w*=6	*w*=7	*w*=8
Fuzzy association method	2.79%	2.56%	2.62%	2.45%	2.62%	2.90%	2.59%
A method proposed by Hwang *et al*.	2.99%	2.94%	3.12%	2.92%	3.01%	3.08%	2.89%
A method proposed by Song *et al*.	3.15%	3.89%	4.37%	4.41%	4.49%	4.35%	4.45%

Table 3. Comparisons of average forecasting errors with different methods (*support*=2, *w*=4)

Method	Song-Chissom	Song-Chissom with neural network	Chen	Markov	Hwang *et al.s*	Fuzzy association rule
Average error	3.20%	4.37%	3.22%	2.66%	3.12%	2.62%

STEP9. Defuzzify the forecasted output. According to Eq.s (21) and (22), we obtain the summation of forecasted elements Ns =2 and

$$DF(1977)$$

$$= (0+0+1\cdot midpoint(u_3)\cdot\frac{1}{2}+1\cdot midpoint(u_4)\cdot\frac{1}{2})/2$$

$$= (0+0+0\cdot\frac{1}{2}+400\cdot\frac{1}{2})/2=100.$$

Eventually, the forecasted variation of enrollment in 1977 is 100. Hence, the value of forecasted enrollment in 1977 is 15411.

4.PERFORMANCE ANALYSIS

As Song and Chissom point out, traditional forecasting methods are not fitted to solve the problems composed by linguistic values [6]. For reasons to demonstrate the advantage of our proposed method, comparing and analyzing with other forecasting methods are listed in this section.

In Table 2, we compare the average errors with three different methods, our proposed fuzzy association rule method, the methods proposed by Hwang *et al*, and the method proposed by Song *et al* under different window bases [1][24]. The definition of the window basis *w* is the same as the definition of model basis in [24]. The results illustrate that our proposed fuzzy association rule gain from these comparisons. Hence, the method proposed in this paper possesses more efficient results in every window bases.

In the Table 3, we compare the average forecasting error with Song-Chissom method [6], Song-Chissom method with neural net [24], Chen's method [24], Markov method [27], Hwang *et al*.s method [1] and our proposed fuzzy association method. Therefore, the results reflect that our proposed method is more excellent than all the other methods.

5.CONCLUSION

A forecasting method has been developed based on fuzzy time series and the association rules algorithm to solve the forecasting problem which focuses on how to forecast the enrollments of the University of Alabama. In addition, the fuzzy association rules are applied to compare with other methods and the results emphasize that our fuzzy association rules are more efficient than other's method. This indicates that fuzzy association rules are a successful tool for solving fuzzy time series.

Recently, there have been breakthroughs in improving the performance of association rules [10] [13] [15]. About the further research, we can use these methods and the possibility algorithm to find a better fuzzification to forecast the enrollments.

6.REFERENCE

[1] J. R. Hwang, S. M. Chen and C. H. Lee, "Handling forecasting problems using fuzzy time series," Fuzzy Sets and Systems, vol. 100, pp. 217-228, 1998.

[2] George E.P. Box and G. M. Jenkins, "Time series analysis: forecasting and control," Revised Ed., Prentice Hall, 1976.

[3] G. Janacek and L. Swift, Time series forecasting,

simulation, applications, Ellis Horwood, 1993.

[4] A. S. Weigend and N. A. Gershenfeld, Time series prediction: forecasting the future and understanding the past, Addison-Wesley, 1994.

[5] R. J. Kuo, and K.C. Xue, “Fuzzy neural networks with application to sales forecasting,” Fuzzy Sets and Systems, vol. 108, pp. 123-143, 1999.

[6] Q. Song, and B. S. Chissom, “Forecasting enrollments with fuzzy time series – part I,” Fuzzy Sets and Systems, vol. 54, pp. 1-9,1993.

[7] L. A. Zadeh, “Fuzzy sets,” Inform and Control, vol. 8, pp. 338-353, 1965.

[8] Q. Song, and B. S. Chissom, “Fuzzy time series and its models,” Fuzzy Sets and Systems, vol. 54, pp. 269-277, 1993.

[9] B. Wu, Fuzzy analysis and forecasting in time series, Thesis of Chengchi University, Taiwan, 1994.

[10] M. J. A. Berry and G. Linoff, Data mining techniques-for marketing, sales, and customer support, John Wiley & Sons, Inc., 1997.

[11] R. Srikant, R. Agrawal, “Mining generalized association rules,” Future Generation Computer Systems, vol. 13, pp. 161-180, 1997.

[12] R. Agrawal, and R. Srikant, “Fast algorithms for mining association rules in large databases,” Proceedings of 20th International Conference on Very Large databases, pp. 478-499, 1995.

[13] M. S. J. Chen and P. S. Yu, “Data mining: an overview from database perspective,” IEEE Transaction on Knowledge and Data Engineering, vol. 8(2), pp. 866-883, 1996.

[14] Savasere, E. Omiecinski, and S. Navathe, “An efficient algorithm for mining association rules in large databases,” Proceedings of 21st International Conference on Very Large database, pp. 432-444, 1995.

[15] D.W. Cheung, V.T. Ng, Ada W. Fu, and Y. Fu, “Efficient mining of association rules in distributed databases,” IEEE Transactions on Knowledge and Data Engineering, vol. 8(6), pp. 911-922, 1996.

[16] R. Agrawal, T. Imielinski, and A. Swami, “Mining association rules between sets of items in large databases,” Proceedings of the ACM SIGMOD Conference on Management of Data, 1993.

[17] M. Houtsma and A. Swami, “Set-oriented mining of association rules,” Research Report RJ 957, 1993.

[18] M. J. Zaki, S. Parthasarathy, W. Li and M. Ogihara, “Evaluation of sampling for data mining of association rules,” Proceedings. Seventh International Workshop on Data Engineering, pp. 42-50, 1997.

[19] Li Shen, H. Shen and L. Cheng, “New algorithms for efficient mining of association rules,” The Seventh Symposium on the Frontiers of Massively Parallel Computation, pp. 234-241, 1999.

[20] B. Wu, “Application of fuzzy time series analysis to change periods detection,” IEEE international fuzzy systems conference proceedings, pp. 697-702, 1999.

[21] P. T. Cheng, “Fuzzy seasonality forecasting,” Fuzzy Sets and Systems, vol. 90, pp. 1-10,1997.

[22] E.H. Mamdani, “Application of fuzzy logic to approximate reasoning using linguistic synthesis,” IEEE Trans. Computing, vol. 26, system 40, pp. 143-202, 1991.

[23] Li Xin Wang, A course in fuzzy systems and control, Prentice-Hall PRT, 1997.

[24] Q. Song and B. S. Chissom, “Forecasting enrollments with fuzzy time series – part II,” Fuzzy Sets and Systems, vol. 62, pp. 1-8,1994.

[25] Q. Song, R. P. Leland and B. S. Chissom, “Fuzzy stochastic fuzzy time series and its models,” Fuzzy Sets and Systems, vol. 88, pp. 333-341,1997.

[26] S. M. Chen, “Forecasting enrollments based on fuzzy time series,” Fuzzy Sets and Systems, vol. 81, pp. 311-319,1996.

[27] J. Sullivan and W. H. Woodall, “ A comparison of fuzzy forecasting and markov modeling,” Fuzzy Sets and Systems, vol. 64, pp. 279-293,1994.