International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

Time Series Forecasting Using Fuzzy Time Series With Hedge Algebras Approach

Minh Loc Vu[1], The Yen Pham[2], Thanh Nghia Huynh Pham[3]

[1,2] *Faculty of Computer, Gia Dinh University, Vietnam*
[3] *Faculty of Basic Sciences, Baria-Vungtau University, Vietnam*

***Abstract*— During the recent years, many different methods of using fuzzy time series for forecasting have been published. This paper presents a novel technique based on the hedge algebras (HA) approach. Based upon the HA, the fuzziness intervals are used to quantify the values of fuzzy time series. The intervals are determined through the fuzziness intervals and the adjusted fuzzy logical relationships. The experimental results, forecasting enrolments at the University of Alabama, demonstrate that the proposed method significantly outperforms the published ones.**

***Keywords*— Forecasting, Fuzzy time series, Hedge algebras, Enrolments, Intervals.**

I. Introduction

Fuzzy time series was originally proposed by Song and Chissom[1] and it has been applied to forecast the enrolments at University of Alabama [2, 3]. All steps in the procedure of using fuzzy time series to forecast time series fall into three phases, Phase 1: fuzzifying historical values, Phase 2: mining the fuzzy logical relationships, Phase 3: defuzzifying the output to get the forecasting values. In 1996, Chen [4] opened a new study direction of using fuzzy time series to forecast time series. In this study, Chen suggested an idea of utilizing the intervals in the formula of computing the forecasting values by only using the arithmetic operators. Since then, clearly, it seems to be that Phase 1 strongly affects the forecasting accuracy rate. We can see that the step of partitioning the universe of discourse belongs to Phase1.

Partitioning the universe of discourse is the essential issue in the method of using fuzzy time series as a tool for forecasting time series. Indeed, product of partitioning the universe of discourse is the intervals as the source that provides the values in the future of time series. So, the better method to partition the universe of discourse we have, the better forecasting values we get. Commonly, the method of partitioning the universe of discourse can be divided into two types through the resulted intervals, equal or not the sized intervals.

From the empirical results in the list, applying the second type gives the better forecasting accuracy rate than others. Thus, recent researches focus on the second method.

There have been many methods of partitioning the universe of discourse such as [5] which is the first research confirmed the important role of partitioning the universe of discourse, this employed distribution and average based length as a way to solve the problem. In turn, Jilani et al. [6] proposed frequency density, Huarng and Yu [7] suggested the ratios and Bas et al. [8] used modified genetic algorithm as basis to improve quality of intervals. Information granules were applied in [9, 10, 11] to get good intervals on the universe of discourse. By the hedge algebras approach Ho et al. [12] presented a method of partitioning the universe of discourse. According to this approach, fuzziness intervals are used to quantify the values of fuzzy time series that are linguistic terms. These fuzziness intervals are employed as intervals on the universe of discourse.

Based upon the fuzziness intervals of values of fuzzy time series, distribution of historical values of time series and adjusted fuzzy logical relationships, we can get the intervals on the universe of discourse. This is the way that the proposed method works.

The rest of this paper is organized as follows: Section 2 briefly introduces some basis concepts of HA; Section 3 presents the proposed method; Section 4 presents empirical results on forecasting enrollments at University of Alabama, Forecasting TAIEX Index and comment; Section 5 concludes the paper.

II. Preliminaries

In this section, we briefly recall some concepts associated with fuzzy time series and hedge algebras.

A. Fuzzy time series

Fuzzy time series was first introduced by Song and Chissom [1], it is considered as the set of linguistic values that is observed by the time. Linguistic values are also called linguistic terms.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

It can be seen that conventional time series are quantitative view about a random variable because they are the collection of real numbers. In contrast to this, as the collection of linguistic terms, fuzzy time series are qualitative view about a random variable. There are two types of fuzzy time series, time-invariant and time-variant fuzzy time series. Because of practicality, the former is the main subject which many of researchers focus on. In most of literature, the linguistic terms are quantified by fuzzy sets. Formally, fuzzy time series are defined as following definition

Definition 1. Let $Y(t)$ $(t = \ldots,0,1,2,\ldots)$, a subset of R^1, be the universe of discourse on which $f_i(t)$ (i = 1,2,...) are defined and F(t) is the collection of $f_i(t)$ (i = 1,2,...). Then F(t) is called fuzzy time series on $Y(t)$ $(t = \ldots,0,1,2,\ldots)$.

Song and Chissom employed fuzzy relational equations as model of fuzzy time series. Specifically, we have following definition:

Definition 2. If for any $f_j(t) \in F(t)$, there exists an $f_i(t\text{-}1) \in F(t\text{-}1)$ such that there exists a fuzzy relation $R_{ij}(t,t\text{-}1)$ and $f_j(t) = f_i(t\text{-}1) \circ R_{ij}(t,t\text{-}1)$ where "∘" is the max-min composition, then F(t) is said to be caused by F(t-1) only.

Denote this as

$$f_i(t\text{-}1) \rightarrow f_j(t)$$

or equivalently

$$F(t\text{-}1) \rightarrow F(t)$$

In [2, 3], Song and Chissom proposed the method which use fuzzy time series to forecast time series. Based upon their works, there are many studies focus on this field.

B. Some basis concepts of Hedge Algebras

In this section, we briefly introduce some basis concepts in HA, these concepts are employed as basis to build our proposed method. HA are created by Ho Cat Nguyen et al. in 1990. This theory is a new approach to quantify the linguistic terms differing from the fuzzy set approach. The HA denoted by AX = (X,G,C,H,≤), where,

G = $\{c^+, c^-\}$ is the set of primary generators, in which c^+ and c^- are, respectively, the negative primary term and the positive one of a linguistic variable X, C = $\{0,1,w\}$ a set of constants, which are distinguished with elements in X,H is the set of hedges, "≤" is a semantically ordering relation on X.

For each x ∈ X in HA, H(x) is the set of hedge u ∈ X that generated from x by applying the hedges of H and denoted $u = h_n,\ldots,h_1\,x$, with $h_n,\ldots,h_1 \in H$. $H = H^+ \cup H^-$, in which H^- is the set of all negative hedges and H^+ is the set of all positive ones of X. The positive hedges increase semantic tendency and vice versa with negative hedges. Without loss of generality, it can be assumed that.

$H^- = \{h_{-1} < h_{-2} < \cdots < h_{-q}\}$ and $H^+ = \{h_1 < h_2 < \cdots < h_p\}$.

If X and H are linearly ordered sets,

then AX = (X,G,C,H,≤) is called linear hedge algebra, furthermore, if AX is equipped with additional operations Σ and Φ that are, respectively, the infimum and supremum of H(x), then it is called complete linear hedge algebra (ClinHA) and denoted AX = (X,G,C,H,Σ,Φ,≤).

Fuzziness of vague terms and fuzziness intervals are two concepts that are difficult to define. However, HA can reasonably define these ones. Concretely, elements of H(x) still express a certain meaning stemming from x, so we can interpret the set H(x) as a model of the fuzziness of the term x. With fuzziness intervals can be formally defined by following definition.

Definition 3. Let AX = (X, G, C, H, ≤) be a ClinHA. An $fm: X \rightarrow [0,1]$ is said to be a fuzziness interval of terms in X if:

(1). $fm(c^-)+fm(c^+) = 1$ and $\sum_{h\in H} fm(hu) = fm(u)$, for $\forall u \in \underline{X}$; in this case fm is called complete;

(2). For the constants 0, W and 1, $fm(\boldsymbol{0}) = fm(\boldsymbol{W}) = fm(\boldsymbol{1}) = 0$;

(3). For $\forall x, y \in X$, $\forall h \in H$, $\frac{fm(hx)}{fm(x)} = \frac{fm(hy)}{fm(y)}$, that is this proportion does not depend on specific elements and, hence, it is called fuzziness measure of the hedge h and denoted by $\mu(h)$.

Proposition 3. For each fuzziness interval fm on X the following statements hold:

(1). $fm(hx) = \mu(h)fm(x)$, *for every* $x \in X$;

(2). $fm(c^-) + fm(c^+) = 1$;

(3). $\sum_{-q\le i\le p, i\ne 0} fm(h_i c) = fm(c)$, $c \in \{c^-, c^+\}$;

IJETAE
EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

(4). $\sum_{-q \le i \le p, i \ne 0} fm(h_i x) = fm(x)$;

(5). $\sum_{-q \le i \le -1} \mu(h_i) = \alpha$ *and* $\sum_{1 \le i \le p} \mu(h_i) = \beta$, *where* $\alpha, \beta > 0$ *and* $\alpha + \beta = 1$.

HA build the method of quantifying the semantic of linguistic terms based on the fuzziness intervals and hedges through υ mapping that fit to the conditions in following definition.

Definition 4. Let $AX = (X, G, C, H, \Sigma, \Phi, \le)$ be a CLinHA. A mapping $\upsilon : X \to [0,1]$ is said to be an semantically quantifying mapping of AX, provided that the following conditions hold:

(1). υ is a one-to-one mapping from X into [0,1] and preserves the order on X, i.e. for all $x, y \in X$, $x < y \Rightarrow \upsilon(x) < \upsilon(y)$ and $\upsilon(\mathbf{0}) = 0$, $\upsilon(\mathbf{1}) = 1$, where $0, 1 \in C$;

(2). Continuity: $\forall x \in X$, $\upsilon(\Phi x) = infimum\ \upsilon(H(x))$ and $\upsilon(\Sigma x) = supremum\ \upsilon(H(x))$.

Semantically quantifying mapping υ is determined concretely as follows.

Definition 5. Let *fm* be a fuzziness interval on X. A mapping $\upsilon : X \to [0,1]$, which is induced by *fm* on X, is defined as follows:

(1). $\upsilon(\boldsymbol{W}) = \theta = fm(c^-)$, $\upsilon(c^-) = \theta - \alpha fm(c^-) = \beta fm(c^-)$, $\upsilon(c^+) = \theta + \alpha fm(c^+)$;

(2). $\upsilon(h_j x) = \upsilon(x) + Sign(h_j x)\{\sum_{i=Sign(j)}^{j} fm(h_i x) - \omega(h_j x)\, fm(h_j x)\}$,

where $j \in \{j: -q \le j \le p\ \&\ j \ne 0\} = [-q \wedge p]$ and

$\omega(h_j x) = \frac{1}{2}[1 + Sign(h_j x) Sign(h_p h_j x)(\beta - \alpha)] \in \{\alpha, \beta\}$;

(3). $\upsilon(\Phi c^-) = 0$, $\upsilon(\Sigma c^-) = \theta = \upsilon(\Phi c^+)$, $\upsilon(\Sigma c^+) = 1$, and for $j \in [-q \wedge p]$,

$$\upsilon(\Phi h_j x) = \upsilon(x) + Sign(h_j x)\{\sum_{i=sign(j)}^{j-sign(j)} \mu(h_i)\, fm(x)\} - \frac{1}{2}(1 - Sign(h_j x))\mu(h_j) fm(x),$$

$$\upsilon(\Sigma h_j x) = \varphi(x) + Sign(h_j x)\{\sum_{i=sign(j)}^{j-sign(j)} \mu(h_i)\, fm(x)\} + \frac{1}{2}(1 + Sign(h_j x))\mu(h_j) fm(x).$$

The Sign function is determined in the following

Definition 6. A function *Sign*: $X \to \{-1, 0, 1\}$ is a mapping which is defined recursively as follows, for $h, h' \in H$ and $c \in \{c^-, c^+\}$:

(1). $Sign(c^-) = -1$, $Sign(c^+) = +1$;

(2). $Sign(hc) = -Sign(c)$, if h is negative w.r.t. c; $Sign(hc) = +Sign(c)$, if h is positive w.r.t.c;

(3). $Sign(h'hx) = -Sign(hx)$, if $h'hx \ne hx$ and h' is negative w.r.t. h;

$Sign(h'hx) = +Sign(hx)$, if $h'hx \ne hx$ and h' is positive w.r.t. h.

(4). $Sign(h'hx) = 0$ if $h'hx = hx$.

Definition 7. [12] Given $AX^2, (k \ge 1)$, the similar fuzzy space of set $X_{(k)}$ denoted $\zeta_{(k)}$ is a set of similar fuzzy space of all grades from $X_{(k)}$ for $\forall x \in X_{(k)}$, $\Im$g(x) $\in \zeta_{(k)}, g + l(x) = k$ unchanged (ie $\forall x \in X_{(k)}$, $\Im$g(x) made up of the same fuzzy space of level k^*) and $\zeta_{(k)}$ is a partition of [0,1].

Definition 8. [12] Given $AX^2, k \ge 1, \forall x \in X_{(k)}$ identify the similar fuzzy space $\Im$g(x) $\in \zeta_{(k)}$ definition of the compatibility level $g = k + 2 - l(x)$ of quantitative value ν for Grade x to be a mapping $s_g : [0,1] \times X \to [0,1]$ determined based on the distance from ν to $\upsilon(x)$ and two similar fuzzy space close to $\Im$g(x) as follows:

$$s_g(\nu, x) = \max\left(\min\left(\frac{\nu - \upsilon(x)}{\upsilon(x) - \upsilon(y)}, \frac{\upsilon(x) - \nu}{\upsilon(z) - \upsilon(x)}\right), 0\right)$$

Where y, z are two grades defining two similar fuzzy space neighbors left and right of $\Im$g(x).

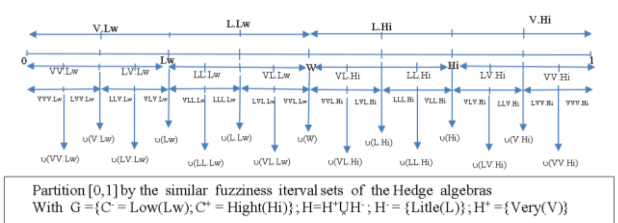

Figure I: Show Partition [0,1] by the similar fuzziness interval sets of the Hedge algebras.

IJETAE
EXPLORING RESEARCH AND INNOVATIONS
International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

III. THE PROPOSED METHOD

For convenience to present proposed method, we name the linguistic values of fuzzy time series as the variables A_i with $i \in N$. Rev$\upsilon(x)$ and Rev $f_m(x)$ are the reversed mapping of $\upsilon(x)$ and $f_m(x)$, respectively, from [0,1] to the universe of discourse of fuzzy time series U. Denote I_k, on U, is the interval corresponding to A_k.

A. Rule for adjusting the fuzzy logical relationships

We can adjust the fuzzy logical relationships to improve forecasting result depending upon the concrete forecasting problem. The rule for adjusting is as follows

Consider sequentially for each group of fuzzy logical relationships. With the group of fuzzy logical relationships considered such as $A_i \rightarrow A_j,...,A_i \rightarrow A_k,...$, and if $|j - k| \geq 1$ then.

If $(j < k)$ then replacing A_k or A_j do not affect any next fuzzy logical relationships, then

If the historical value called aj in fuzziness interval of A_j, I_j, and $|Rev\ \upsilon(A_j) - a_j| > |Rev\ \upsilon(A_{j+1}) - a_j|$, then extend I_{j+1} to cover up a_j. **(r.1)**

If the historical value called b_k in fuzziness interval of A_k,I_k, and $|Rev\ \upsilon(A_{k-1}) - b_k| < |Rev\ \upsilon(A_k) - b_k|$, then extend I_{k-1} to cover up b_k. **(r.2)**

Do the same with the case of $j > k$.

A_j and A_k are as close as possible that is the goal we would like to reach up. We can do **(r.1)** or **(r.2)** because the following reason:

With A_m is the linguistic term that we are considering. If $Rev\upsilon(A_m)$ is the semantically quantifying mapping of Am on the universe of discourse, then one also is the semantic core of A_m. If the other values belonging to the fuzziness interval of A_m, then they are semantically equal to $Rev\upsilon(A_m)$, that means they together reflex the meaning of A_m. If a is the value that belong to $A_{m(+1)}$ and $|\text{Rev}\upsilon(Am) - a| > |\text{Rev}\upsilon A_{(m+)1} - a|$, then a is more close semantic with A_{m+1} than A_m. So, we can extend $f_m(A_{m+1})$ to cover up a

B. Method for partitioning the universe of discourse

The proposed method, named VML, is described as follows

Step 1:

Determine the U, the universe of discourse of fuzzy time series F(t).

U = [min($F(t)$)−D_1,max($F(t)$)+D_2], where D_1 and D_2 are proper positive numbers.

Set n is the number of intervals that we would like to divide on the universe of discourse.

Step 2:

Building the ClinHA with only two hedges, h_{-1} and h_{+1}, $AX = (X,G,H,\Sigma,\Phi,\leq)$ corresponding to linguistic variable that is considered as fuzzy time series F(t). That means determining the set of parameters of AX model needs to be consistent with the context of the problem "forecasting student enrolment number at the Alabama University".

According to the definitions and propositions in Section II.B, parameters of the ClinHA, which contain two hedges, h_{-1} and h_{+1}, have the following properties

$\mu(h_{+1}) = \beta = f_m(c^+)$; $\mu(h_{-1}) = \alpha = f_m(c-)$ where $\alpha,\beta > 0$, $\alpha + \beta=1$.And $\upsilon(W) = f_m(c^-) = \theta\ \upsilon(c^-) = \beta f_m(c^-)$; $\upsilon(c^+) = 1 - \beta f_m(c^-)$ where 0,W,1 are constants and

$f_m(0) = f_m(1) = f_m(W)=0$; $\upsilon(W) = \theta = W$.

By having knowledge of W and $\mu(h_{+1})$ or $\mu(h_{11})$, other parameters can be determined. Now let us define two parameters W and $\mu(h)$ that fit the context of the problem based on the historical values and exploit the relationship between them: According to the context, semantics of $\bar{F}(t)$ denotes number of the enrollment students at the medium level and W is the normalization value of , they are calculated according formulas:

$$W = \frac{\bar{F}(t) - \min(F(t))}{\max(F(t)) - \min(F(t))} \qquad (2)$$

where x_i is the historical value of $F(t)$, e.g. $x_1 = F(1971),...,x_{22} = F(1992)$.

The value difference between two adjacent historical values will be the basis for forecasting time series: At time t, it will forecast the trend to increase if the previous historical value is smaller than it, otherwise forecasting trend decreases. In the general case, the average value of the difference will be the basis for forecasting which is defined by the following equation

$$\bar{S} := \frac{1}{n-1}\sum_{i=1}^{n-1}|x_{i+1} - x_i| \qquad (3)$$

To have a better forecasting, we continue to exploit the data relation of historical values based on the relationship between S^- with two values to characterize the increment or decrement of forecasting at maximum difference values

$$S^+ := \left[\max_{(x_{i+1}-x_i)>0}|x_{i+1} - x_i|\right]$$

$$\text{for } 1 \leq i \leq n-1 \qquad (4)$$

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

$$S^- := \left[\max_{(x_{i+1}-x_i)<0} |x_{i+1}-x_i|\right]$$

$$\text{for } 1 \le i \le n-1 \qquad (5)$$

Clearly,

$$0 \le \frac{2\bar{S}}{(S^+ + S^-)} \le 1$$

Let

$$\mu(h) := \frac{2\bar{S}}{(S^+ + S^-)}$$

If $S^+ \ge S^-$ then $h := h_{+1}$ else $h := h_{-1}$.

There is an induced about the trend change in the forecasting in the discourse space into the space of [0, 1] where there is a trend change to the quantitative semantics value due to the impact on the terms of the hedges. That is the basic to we construct the mathematics model for forecasting time series by (HA) approach. Until now, I have finished to construct the ClinHA with only two hedges by method of define two parameters W and $\mu(h)$ based on analyzing data of forecasting.

In this paper, I build $AX = (X,G,H,\Sigma,\Phi,\le)$ as follows:
Let G={C^-=Low(Lo), C^+ = High(Hi)} and

$H = H^+ \cup H^-$.

H^- = {Litlle(L)} and H^+ = {Verry(V)}.

Calculate W using (1) and (2).

Calculate $\mu(h_{+1})$ or $\mu(h_{-1})$ using (3), (4) and (5).

Calculate the remaining parameters according to Section II.B.

Using above HA generate n linguistic terms which use to qualitatively describe time series. The way to determine these linguistic terms as follows: Applying two hedges, h_{-1} and h_{+1}, on the primary generators c^- and c^+, from left to right to generate the linguistic terms.

If the number of linguistic terms are less than, one interval, the number of intervals that we need to divide, then find the interval that contain maximum amount of historical values, assuming that this interval corresponding to the linguistic term A_i. From A_i generating two linguistic term $h_{-1}A_i$ and $h_{+1}A_i$.

Step 3: Based upon the distribution of historical values, put them into the corresponding linguistic term fuzziness interval.

C. Forecasting Algorithm

Step 1:

Apply VML to partition the universe of discourse.

Step 2:

Mine the fuzzy logical relationships: $A_p \to A_q$, where A_p and A_q, respectively, are the linguistic values of $F(t)$ and $F(t + 1)$ respectively.

Set the group of fuzzy logical relationships having the same left side: $A_t \to A_u(m)\cdots A_v(n)$, $m,\cdots,n$ are the number of iterations of fuzzy logical relationship $A_t \to A_u$ and $A_t \to A_v$.

Adjust the fuzzy logical relationships following rule III.A.

Step 3: Suppose that the time series value at $(t - l)$, f_t, belongs to $\mathrm{Rev}f_m(A_t)$ then

$$\hat{x}(t) = \frac{Rev(A_u) + \cdots + Rev(A_v)}{l} \qquad (6)$$

where $\hat{x}(t)$ is the forecasting value at time t and

$l = \mathrm{card}\{u,\cdots,v\}$.

IV. Result And Comment

The proposed approach is applied to forecast the enrolments at the University of Alabama from year 1971 to 1992 (n = 22). The result will then be compared with different published methods. To measure the accuracy of the forecasting methods, the following metrics are used for comparison.

The Root-Mean-Square-Error (RMSE) which is defined as

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(\bar{X}_i - X_i)^2}{n}}$$

The Numerical Error (NE) percentage

$$\text{NE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{(\bar{X}_i - X_i)}{X_i}\right| \times 100$$

The Normalized Numerical Error (NNE) percentage

$$\text{NNE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{(\bar{X}_i - X_i)}{\max(x)-\min(x)}\right| \times 100$$

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

where $\hat{x}_i$ are the forecasted value and the actual value at time i respectively, and n is the length of the time series to be forecasted.

A. Result

We apply proposed method for the intervals of 7 and 17.

Interval 7

Apply the VML method to partition the universe of discourse

$\max(F(t)) = 20000$ $\min(F(t)) = 13000$

Follow (1-3.2) and (2-3.2) we have

$$\bar{F}(t) = 16194$$

$$W = \frac{16194 - 13000}{20000 - 13000} = 0.4563$$

According to (3-3.2),(4-3.2) and (5-3.2),

$S^- = 510$ $S^+ = |x_{18} - x_{17}| = |18150 - 16859| = 1291$ $S^- = |x_{12} - x_{11}| = |15433 - 16388| = 955$

As $S^+ > S^-$, according to (6-3.2), $\mu(V) = W = 0.4563$

hence, $f_m(V.\text{Lw}) = \upsilon(\text{Lw}) = 0.20821$

Based on the distribution of historical values we can put the historical values into the following intervals

$A_1 = [0, \upsilon(\text{VVV,Lw}), \upsilon(\text{Lw})]$ where 0 and $\upsilon(\text{Lw})$ are the left and right border of the linguistic values "LLV.Lw" respectively. "LVV.Lw" stands for "Little-Very-Very-Low". Similarly,

$A_2 = [\upsilon(\text{Lw}), \upsilon(\text{L.Lw}), \upsilon(\text{LVL.Lw})]$

$A_3 = [\upsilon(\text{LVL.Lw}), \upsilon(\text{VL.Lw}), \upsilon(\text{VVL.Lw})]$

$A_4 = [\upsilon(\text{VVL.Lw}), \upsilon(\text{VL.Hi}), \upsilon(\text{LLVL.Hi})]$

$A_5 = [\upsilon(\text{LLVL.Hi}), \upsilon(\text{L.Hi}), \upsilon(\text{LL.Hi})]$

$A_6 = [\upsilon(\text{LL.Hi}), \upsilon(\text{VLL.Hi}), \upsilon(\text{Hi})]$

$A_7 = [\upsilon(\text{Hi}), \upsilon(\text{V.Hi})]$

The values of calculated I_i are

$$I_1 = [13000, 14457) \quad I_2 = [14457, 15598)$$
$$I_3 = [15598, 16029) \quad I_4 = [16029, 16752)$$
$$I_5 = [16752, 17750) \quad I_6 = [17750, 18263)$$
$$I_7 = [18263, 20000]$$

The semantically quatifying mappings can also be obtained

$\text{Rev}(A_1) = 13303$ $\text{Rev}(A_2) = 15402$

$\text{Rev}(A_3) = 15833$ $\text{Rev}(A_4) = 16625$

$\text{Rev}(A_5) = 17138$ $\text{Rev}(A_6) = 17795$

$\text{Rev}(A_7) = 19207$

Table I :
Historical and fuzzified values

Years	Enrollments	Fuzzified Values
1971	13055	A_1
1972	13563	A_1
1973	13867	A_1
1974	14696	A_2
1975	15460	A_2
1976	15311	A_2
1977	15603	A_3
1978	15861	A_3
1979	16807	A_5
1980	16919	A_5
1981	16388	A_4
1982	15433	A_2
1983	15497	A_2
1984	15145	A_2
1985	15163	A_2
1986	15984	A3
1987	16859	A_5
1988	18150	A_6
1989	18970	A_7
1990	19328	A_7
1991	19337	A_7
1992	18876	A_7

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 12, December 2017)

Table II:
Group of fuzzy logical relationships

Group 1	$A1 \to A1$ (2), $A1 \to A2$
Group 2	$A2 \to A2$ (5), $A_2 \to A_3$(2)
Group 3	$A3 \to A3$, $A3 \to A5$ (2)
Group 4	$A4 \to A2$
Group 5	$A5 \to A4$, $A5 \to A5$, $A5 \to A6$
Group 6	$A6 \to A7$
Group 7	$A7 \to A7$(3)

Comments

We compare our approach with the method Wei Lu et.al. published in [11] to illustrate our superior efficiency. After commenting on the research publication of other authors, Wei Lu confirmed As noted above, the objective of our research is to develop a novel method of partitioning the universe of discourse deliver some remedy for these shortcoming [11].

Table III: Sum up Compare Forecasting Result of Methods (with h=7 and h=17).

Our comparison focuses on two aspects: the calculation convenience and the forecasting accuracy.

First, the convenience in calculations: only with the simple calculations using Method for partitioning the universe of discourse and Algorithm for forecasting", we have obtained results about group of logical relationship like the results from Wei Lu [11]. However, our calculation is much simpler than theirs as shown in Table II.

Second, the forecasting accuracy: Table 3 shows our proposed method is about 10% better in term of accuracy (all metrics) compared to Wei Lu et.al. approach [11].

The reason is because Wei Lu et.al. chose middle point of interval to defuzzifying the forecasting outputs [11] as the basis to calculate the forecasting results hence it carried heavy subjective and is not related to the inherent semantic of information granules (as linguistic values). While with our approach, similarity fuzziness interval (for semantics) is used to fuzzifying the historical data of time series and its quantitative semantic value (as the its semantic core) is the basis for the forecasting.

These values are calculated by determining two parameters using Method for partitioning the universe of discourse mentioned above. It is not decisive subjective and contrast in accordance with the context of forecasting problems, because Hedge algebras are considered as an algebraic approach to the inherent word semantics and word-domains of each individual linguistic variable, which is simply understood as a variable whose values are definite linguistic words of a natural language and establish a formalized foundation to develop quantitative semantics of words, including their fuzzy set based semantics [15]. If Adjusting the fuzzy logical relationships according to Rule III.A mentioned above:

$A_2 \to A_3$, the historical values of I_2 are 14696, 15145, 15163, 15311, 15433, 15460, 15497 and with I_3 are 15603, 15861, 15984. Rev$\upsilon(A_2)$ = 15402, Rev$\upsilon(A_3)$ = 15833.

$|15402 - 15603| = 201 < |15833 - 15603| = 230$, hence $I_2 = [14457,15604)$ and $I_3 = [15604,16029)$ and $A_2 \to A_3$ replaced by $A_2 \to A_2$.

Then the result is even more accurate as RMSE = 387.4, NE(%) = 1.80% and NNE(%) = 4.63%.

This proves Step 2. Find the optimal split point vector $S = p_1,p_2,.p_i..,p_{h-1}p_h$ in the universe of discourse U [11], which be the step is core of our approach[11] has not yet reached optimal.

It is important that in a natural language, semantics of linguistic words be decisive by context. Consequently, That means determining the set of parameters of AX model needs to be consistent with the context of problem of the forecasting student enrollment number at the Alabama University mentioned above, it is interpretation for self-affirmed superior efficiency of our approach

B. Interval of 17

Similarly, apply the proposed method for 17 intervals on the universe of discourse we will have the forecasting result as:

RMSE = 216.1; NE = 0.97%; NNE = 2.20%

Table III. shows that all the metrics: RMSE, NE(%) and NNE(%) of our proposed method is less than of the others

This study is also used one to compare the proposed method forecasting result with some recent methods. This demonstrates the forecast results of the prosed method is the most accurate.

Table III
Sum up Compare Forecasting Result of Methods with h=7 and h=17.

	the number of split intervals is 7 (h=7)							the number of split intervals is 17 (h=17)		
With 7 split points (h=7)								With 17 split points (h = 17)		
Year	Actual enrollment	Chen96	Wang13	Wang14	Chen13	Lu15	Our approah	Huarng	Lu15	Our Approach
1972	13563	14000	13486	13944	14347	14279	14003	13600	13678	13582
1973	13867	14000	14156	13944	14347	14279	14003	14000	13678	13582
1974	14696	14000	15215	13944	14347	14279	14003	14800	14602	14457
1975	15460	15500	15906	15328	15550	15392	15510	15600	15498	15443
1976	15311	16000	15906	15753	15550	15392	15510	15600	15192	15447
1977	15603	16000	15906	15753	15550	15392	15510	15600	15614	15447
1978	15861	16000	15906	15753	15550	16467	15510	15600	15817	15371
1979	16807	16000	16559	16279	16290	16467	17138	16800	16744	16752
1980	16919	16833	16559	17270	17169	17161	17264	17067	17618	17031
1981	16388	16833	16559	17270	17169	17161	17264	17067	16392	16517
1982	15433	16833	16559	16279	16209	14916	15402	15600	15410	15433
1983	15497	16000	15906	15753	15550	15392	15510	15600	15498	15447
1984	15145	16000	15906	15753	15550	15392	15510	156001	15192	15371
1985	15163	16000	15906	15753	15550	15392	15510	15600	15567	15470
1986	15984	16000	15906	15753	15550	15470	15510	15600	15567	15470
1987	16859	16000	16559	16279	16290	16467	17138	16800	16744	16810
1988	18150	16833	16559	17270	17169	17161	17264	17067	17618	18156
1989	18970	19000	19451	19466	18907	19257	19207	18800	19036	18973
1990	19328	19000	18808	18933	18907	19257	19207	19200	19574	19297
1991	19337	19000	18808	18933	18907	19257	19207	19000	19146	19059
1992	18876	19000	18808	18933	18907	19257	19207	19000	19146	19059
	RMSE	638.4	578.3	506.0	486.3	445.2	387.4	353.1	256.3	216.06
	NE(%)	3.11	2.76	2.68	2.52	2.30	1.80	1.51	1.06	0.97
	NNE(%)	7.94	7.17	6.93	6.43	5.88	4.63	3.98	2.81	2.20

V. CONCLUSION

In this paper we presented the proposed method using fuzzy time series with HA approach to forecast enrolments at the University of Alabama. Whereby, based upon the number of intervals that we would like to divide on the universe of discourse, the set of linguistic terms - the values of fuzzy time series are determined by means of HA with only two hedges. Each linguistic term, x, is quantified by semantically quantifying mapping, υ(x), and its fuzziness interval, f_m(x). The distribution of historical values are the basis to decide putting them into the corresponding linguistic terms fuzziness interval. The fuzzy logical relationships can be adjusted to improve the forecasting quality.

Using υ(x) in the formula to calculate the forecasting results are better than using centre point of interval like other studies because υ(x)is the semantic core of x. The experimental results show that the proposed method, with the different number of intervals, owns accuracy rate of forecasting results outperforming the others. We can see that the proposed method can also permit us to forecast on other time series. This is the subject that we focus on the further studies.

REFERENCES

[1] Qiang Song and Brad S Chissom. Fuzzy time series and its models. Fuzzy sets and systems, 54(3): 269-277, 1993.

[2] Qiang Song and Brad S Chissom. Forecasting enrollments with fuzzy time seriespart i. Fuzzy sets and systems, 54(1):1- 9, 1993.

[3] Qiang Song and Brad S Chissom. Forecasting enrollments with fuzzy time seriespart ii. Fuzzy sets and systems, 62(1):1-8, 1994.

[4] Shyi-Ming Chen. Forecasting enrollments based on fuzzy time series. Fuzzy sets and systems, 81(3): 311-319, 1996.

[5] Kunhuang Huarng. Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy sets and systems, 123(3):387-394, 2001.

[6] Tahseen Ahmed Jilani, Syed Muhammad Aqil Burney, and Cemal Ardil. Fuzzy metric approach for fuzzy time series forecasting based on frequency density based partitioning. World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 4(7):1194-1199, 2007.

[7] Kunhuang Huarng and Tiffany Hui-Kuang Yu. Ratio-based lengths of intervals to improve fuzzy time series forecasting. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 36(2): 328-340, 2006.

[8] Eren Bas, Vedide Rezan Uslu, Ufuk Yolcu, and Erol Egrioglu. A modified genetic algorithm for forecasting fuzzy time series. Applied intelligence, 41(2):453-463, 2014.

[9] Lizhu Wang, Xiaodong Liu, and Witold Pedrycz. Effective intervals determined by information granules to improve forecasting in fuzzy time series. Expert Systems with Applications, 40(14):5673-5679, 2013.

[10] Lizhu Wang, Xiaodong Liu, Witold Pedrycz, and Yongyun Shao. Determination of temporal information granules to improve forecasting in fuzzy time series. Expert Systems with Applications, 41(6):3134-3142, 2014. 14.

[11] Wei Lu, Xueyan Chen, Witold Pedrycz, Xiaodong Liu, and Jianhua Yang. Using interval information granules to improve forecasting in fuzzy time series. International Journal of Approximate Reasoning, 57:1-18, 2015.

[12] Nguyen Cat Ho, Tran Thi Son, and Duong Thanh Long. Hedge algebras with limited number of hedges and applied to fuzzy classification problems. Journal of Science and Technology, 48(5), 2012. 15.