The Context-Aware Calculating Method in Language Environment Based on Hedge Algebras Approach to Improve Result of Forecasting Time Series

Minh Loc Vu[1(✉)], Hoang Dung Vu[2], and The Yen Pham[1]

[1] Giadinh University, Ho Chi Minh City, Vietnam
vuminhloc@gmail.com, yen637@gmail.com
[2] National University of Singapore, Singapore, Singapore
hoangdung@u.nus.edu

Abstract. During the recent years, many different methods of using fuzzy time series for forecasting have been published. However, computation in the linguistic environment one term has two parallel semantics, one represented by fuzzy sets it human-imposed and the rest is due to the context of the problem. Hedge algebra is the algebraic approach to the semantic order structure of domain of the linguistic variables that unifies the mentioned above two semantics of each term and therefore, there is a Context-Aware calculating method in the language environment. That is the core of the new approach we mentioned in this article to increase accuracy of solve the time series forecasting problem. The experimental results, forecasting enrollments at the University of Alabama and forecasting TAIEX Index, demonstrate that the proposed method significantly outperforms the published ones. The experimental results, forecasting enrollments at the University of Alabama, demonstrate that the proposed method outperforms the others listed methods.

Keywords: Forecasting · Fuzzy time series · Hedge algebras · Enrollments Intervals · Context-Aware

1 Introduction

Fuzzy time series originally created in 1993 by Song and Chissom [1] and applied to forecast the enrollments 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 the new study direction of using fuzzy time series to forecast time series. In this study, Chen suggested the idea of utilize the intervals in the formula of computing the forecasting values only using the arithmetic operators. Sine then, clearly, it seem to be that Phase 1 stronger affects the forecasting accuracy rate. We can see that the step of partitioning the universe of discourse belong to Phase 1.

P. Cong Vinh et al. (Eds.): ICCASA 2017/ICTCC 2017, LNICST 217, pp. 110–123, 2018.
https://doi.org/10.1007/978-3-319-77818-1_11

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 are the intervals as the source that provide 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. Normally, the method of partition the universe of discourse can be divided into two types through resulted intervals, equal or not the sized intervals. From the empirical results in the list, apply the second type gives the forecasting accurate rate better than the other. Thus, the recent researches focus on the second method.

There have been pretty much method of partitioning the universe of discourse such as paper [5] is the first research confirmed the important role of partitioning the universe of discourse, this employed distribution and average based length as the way to solve the problem. In turn, [6] proposed frequency density, [7] suggested the ratios and [8] use modified genetic algorithm as basis to improve quality of intervals. Information granules are applied in [9–11] to get good intervals on the universe of discourse. By the hedge algebras approach [12, 13] presented the 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 perform. The rest of this paper is organized as follows: Sect. 2, briefly introduce some basis concepts of HA; Sect. 3 presents the proposed method; Sect. 4 presents empirical results on forecasting enrollments at University of Alabama; Sect. 5 is the conclusion of this paper.

2 Preliminaries

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

2.1 Fuzzy Time Series

Fuzzy time series are first introduced by Song and Chissom in 1993 [1], it is considered as the set of linguistic values that is observed by the time. Linguistic values are also called linguistic terms. 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 are 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 $f_i(t)\,(i = 1, 2, \ldots)$ are defined and $F(t)$ is the collection of $f_i(t)\,(i = 1, 2, \ldots)$. 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 $f_i(t-1) \circ F(t-1)$ such that there exists a fuzzy relation $R_{ij}(t, t-1)$ and $f_j(t) = f_i(t-1) \circ R_{ij}(t, t-1)$ where '$\circ$' is the max-min composition, then $F(t)$ is said to be caused by $F(t-1)$ only. Denote this as

$$f_i(t-1) \to f_j(t)$$

or equivalently $F(t-1) \to F(t)$.

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.

2.2 Some Basis Concepts of Hedge Algebras

In this section we refer to paper [14] to briefly introduce some basis concepts in HA, these concepts are employed as basis to build our proposed method. HA are created by Ho 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 $AX = (X, G, C, H, \leq)$, 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, "$\leq$" is a *semantically ordering relation* on X. For each $x \in X$ in HA, $H(x)$ is the set of hedge $u \in X$ 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 vise versa with negative hedges. Without loss of generality, it can be assumed that $H^- = \{h_{-1} < h_{-2} < \ldots < h_{-q}\}$ and $H^+ = \{h_1 < h_2 < \ldots < h_p\}$.

If X and H are linearly ordered sets, then $AX = (X, G, C, H, \leq)$ is called *linear hedge algebra*, furthermore, if AX is equipped with additional operations $\boldsymbol{\Sigma}$ and $\boldsymbol{\Phi}$ that are, respectively, infimum and supremum of $H(x)$, then it is called *complete linear hedge algebra* (ClinHA) and denoted $AX = (X, G, C, H, \Sigma, \Phi, \leq)$.

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 $AX = (\underline{X}, G, C, H, \leq)$ be a ClinHA. An fm: $X \to [0, 1]$ 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 $\boldsymbol{0}$, $\boldsymbol{W}$ and $\boldsymbol{1}$, $fm(\boldsymbol{0}) = fm(\boldsymbol{W}) = fm(\boldsymbol{1}) = 0$;
(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 $\mu(h)$.

The condition (1) means that the primary terms and hedges under consideration are complete for modeling the semantics of the whole real interval of a physical variable. That is, except the primary terms and hedges under consideration, there are no more primary terms and hedges. (2) is intuitively evident. (3) seems also to be natural in the sense that applying a hedge h to different vague concepts, the relative modification effect of h is the same, i.e. this proportion does not depend on terms they apply to.

The properties of fuzziness intervals are made clearly through following proposition.

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). $\sum_{-q \le i \le p, i \ne 0} fm(h_i c) = fm(c)$, $c \in \{c^-, c^+\}$;
(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 $AX = (X, G, C, H, \boldsymbol{\Sigma}, \boldsymbol{\Phi}, \le)$ 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 $\mathbf{0}, \mathbf{1} \in C$;
(2). Continuity:
$\forall x \in X$, $\upsilon(\boldsymbol{\Phi} x) = \textit{infimum}\, \upsilon(H(x))$ and $\upsilon(\boldsymbol{\Sigma} x) = \textit{supremum}\, \upsilon(H(x))$.

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

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

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

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 $h'hx \neq hx$ and h' is negative w.r.t. h;
$Sign(h'hx) = +Sign(hx)$, if $h'hx \neq hx$ and h' is positive w.r.t. h.
(4). $Sign(h'hx) = 0$ if $h'hx = hx$.

3 Proposed Method

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

3.1 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:

With A_m is the linguistic term that we are considering. If $Rer\upsilon(A_m)$ is the semantically quantifying mapping of A_m 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 $Rer\upsilon(A_m)$, that mean they together reflex the meaning of A_m. If a is the value that belong to $A_{m(+)1}$ and $|Rer\upsilon(A_m) - \mathrm{a}| > |Rer\upsilon(A_{m-(+)1}) - \mathrm{a}|$, then a is more close semantic with $A_{m-(+)1}$ than A_m. So, we can extend $fm(A_{m(+)1})$ cover up a.

3.2 Method for Partitioning the Universe of Discourse

We name the proposed method is VL

Step 1:
Determine the U, the universe of discourse of fuzzy time series $F(t)$. $U = [\mathrm{Min}.F(t) - D1, \mathrm{Max}.F(t) + D2]$, where *D1* and *D2* are two proper positive numbers. Setting n is the number of intervals that we would like to divide on the universe of discourse.

Step 2:
Building the Clin HA 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 mean determining the set of parameters of *AX*. 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 $h_{-1}Ai$ and $h_{+1}Ai$.

Step 3:

Calculating the average of values of $F(t)$, $\bar{F}(t)$;Calculating *W* and *μ(L)* as follows:

$$\mu(h_{+1}) = W = \frac{\bar{F}(t) - Min.F(t)}{Max.F(t) - Min.F(t)}(*),$$

Where $Min.F(t)$ and $Max.F(t)$, respectively, are the max and min of historical values of F(t).

Step 4:

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

3.3 Algorithm for Forecasting

Step 1:
Apply VL to partition the universe of discourse.

Step 2:
Mine the fuzzy logical relationships: $A_p \rightarrow A_q$, where A_p and A_q, respectively, are the linguistic values of *F(t)* and *F(t+1)*.
Set the group of fuzzy logical relationships having the same left side: $At \rightarrow Au(\mathrm{m}) \ldots Av(n)$, $m, \ldots, n$ are the number of iterations of fuzzy logical relationship $At \rightarrow Au$ and $At \rightarrow Av$.
Adjust the fuzzy logical relationships following rule 3.1.

Step 3:
Compute the forecasting values: Suppose that the value of the time series at *t−1*, *ft*, if *ft* belong to *Revfm*(*At*), then

The forecasting value at *t* is $\frac{m*Rer\,v(\mathrm{Au}) + \cdots n*Rer\,v(\mathrm{Av})}{m + \cdots n}(**)$

4 Empirical Result

4.1 We Test the Proposed Method on the Time Series that are Enrollments at University of Alabama

This time series have been used in many previous studies.

We apply proposed method for 7 and 17 intervals.

With 7 intervals

Apply the VL to partition the universe of discourse

$$Max.F(t) = 20000, Min.F(t) = 13000, \bar{F}(t) = 16194$$

Building $AX = (X, G, C, H, \Sigma, \Phi, \leq)$, Let $G = \{C^- = \text{Low(Lw)}, C^+ = \text{High(Hi)}\}$, $H = H^- \cup H^+$, $H^- = \{\text{Little(L)}\}$, $H^+ = \{\text{Very(V)}\}$

Follow (*) we have

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

Continue study the data of mathematical, we can choose $\mu(V) = 0.4563$ and we have $fm(V.Lw) = \upsilon(Lw) = 0.20821$.

Based upon 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 υ(Lw), respectively, are left and right border of the linguistic values "LVV.Lw" that mean 'Little-Very-Very-Low'. Similarly, we have:

$$\begin{aligned}
A_2 &= [\upsilon(\text{Lw}), \upsilon(\text{L.Lw}), \upsilon(\text{LVL.Lw}));\\
A_3 &= [\upsilon(\text{LVL.Lw}), \upsilon(\text{VL.Lw}), \upsilon(\text{VVL.Lw})),\\
A_4 &= [\upsilon(\text{VVL.Lw}), \upsilon(\text{VL.Hi}), \upsilon(\text{LLVL.Hi})),\\
A_5 &= [\upsilon(\text{LLVL.Hi}), \upsilon(\text{L.Hi}), \upsilon(\text{LL.Hi})),\\
A_6 &= [\upsilon(\text{LL.Hi}), \upsilon(\text{VLV.Hi}), \upsilon(\text{Hi})),\\
A_7 &= [\upsilon(\text{Hi}), \upsilon(\text{V.Hi}), 1].
\end{aligned}$$

After calculating, we have:

$$\begin{aligned}
&I_1 = [13000, 14457), I_2 = [14457, 15598), I_3 = [15598, 16029),\\
&I_4 = [16029, 16752), I_5 = [16752, 17750), I_6 = [17750, 18263),\\
&I_7 = [18263, 20000]
\end{aligned}$$

The semantically quantifying mappings: $Rer\upsilon(A_1) = 13303$, $Rer\upsilon(A_2) = 15402$, $Rer\upsilon(A_3) = 15833$, $Rer\upsilon(A_4) = 16625$, $Rer\upsilon(A_5) = 17138$, $Rer\upsilon(A_6) = 18029$, $Rer\upsilon(A_7) = 19207$.

From Table 1 we have the group of fuzzy logical relationships that show Table 2 as follows.

Table 1. 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

We have the forecasting result as well as some recent method's as follows:

With 17 intervals:

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

In the field of time series research, RMSE, NE(%) and NNE(%) criterias are alway used to evaluate forecasting quality. RMSE $= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i^{'} - x_i\right)^2}$, NE(%) $= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i^{'} - x_i}{x_i}\right|.\,100$ *and* NNE $= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i^{'} - x_i}{x_{max} - x_{min}}\right|.100$ where x_i' is the forecasting value, x_i is historical value and n is the number of forecasting values. This study is also used one to compare proposed method's forecasting result with some recent method's. Based upon RMSE criteria, we can see that proposed method's RMSE is less than the others. That means proposed method forecasting result give more exactly forecasting resullt than the others (Tables 3 and 4).

Table 2. Group of fuzzy logical relationships

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

Table 3. Compare result of proposed method with some recent method's (with = 7)

Years	Historical values	Chen et al. 2013	Wang et al. [10]	Lu et al. [11]	Proposed method
1972	13563	14347	13944	14279	14003
1973	13867	14347	13944	14279	14003
1974	14696	14347	13944	14279	14003
1975	15460	15550	15328	15392	15510
1976	15311	15550	15753	15392	15510
1977	15603	15550	15753	15392	15510
1978	15861	15550	15753	16467	15510
1979	16807	16290	16279	16467	17138
1980	16919	17169	17270	17161	17186
1981	16388	17169	17270	17161	17186
1982	15433	16209	16279	14916	15402
1983	15497	15550	15753	15392	15510
1984	15145	15550	15753	15392	15510
1985	15163	15550	15753	15392	15510
1986	15984	15550	15753	15470	15510
1987	16859	16290	16279	16467	17138
1988	18150	17169	17270	17161	17186
1989	18970	18907	19466	19257	19207
1990	19328	18907	18933	19257	19207
1991	19337	18907	18933	19257	19207
1992	18876	18907	18933	19257	19207
RMSE		486.3	506.0	445.2	400.4
NE(%)		2.52	2.68	2.30	1.95
NNE(%)		6.43	6.93	5.88	4.52

4.2 Test of Forecasting TAIEX Index

Chen and Chen [13] have applied their proposed method on the experimental data sets TAIEX Index of November and December 2004. The data set consists of 44 items.

Table 4. Compare result of proposed method with some recent method's

Years	Historical values	Lu et al. [11]	Proposed method
1972	13563	13678	13582
1973	13867	13678	13582
1974	14696	14602	14457
1975	15460	15498	15443
1976	15311	15192	15447
1977	15603	15641	15447
1978	15861	15827	15371
1979	16807	16744	16752
1980	16919	17618	17031
1981	16388	16392	16517
1982	15433	15410	15433
1983	15497	15498	15447
1984	15145	15192	15371
1985	15163	15567	15470
1986	15984	15567	15470
1987	16859	16744	16810
1988	18150	17618	18156
1989	18970	19036	18973
1990	19328	19574	19297
1991	19337	19146	19059
1992	18876	19146	19059
RMSE		256.3	216.1
NE(%)		1.06	0.97
NNE(%)		2.81	2.20

The historical training data of TAIEX is fuzzified into 9 fuzzy sets (h = 9). *The accuracy* metrics of the result:

$$\mathbf{RSME} = \mathbf{56.86; NE(\%)} = \mathbf{0.8; NNE(\%)} = \mathbf{12.44}$$

Our proposed method is applied to the same TAIEX datasets. The process is as follows According to 3.2:

- **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 two proper positive numbers.
- **Step 2.** Building the ClinHA with only two hedges, h_{-1}, h_{+1}, $A = (X, G, H, \Sigma, \leq)$. This means determining the set of parameters of AX model needs to be consistent with the context of the problem "forecasting TAIEX Index" mentioned above.

Let

$$\boldsymbol{G} := \{C^- = Low(Lo),\ C^+ = High(Hi)\};\ \mathrm{H} := H^+ \cup H^-;\ H^- : = \{Little(L)\};\ H^+ := \{Very(V)\}$$

x_1 := 5759.61 (Actual index of day 01/11/2004 … x_{44} := 6139.69 (Actual index of day 31/12/2004).

The following equations are performed

$$S^+ := \left[\max_{(x_{i+1}-x_i)>0} |x_{i+1} - x_i|\right] = 94.74 \qquad S^- := \left[\max_{(x_{i+1}-x_i)<0} |x_{i+1} - x_i|\right] = 138.1$$
$$1 \le i \le 43 \qquad\qquad 1 \le i \le 43$$

$$\bar{F}(t) = \frac{1}{44}\sum_{i=1}^{44} x_i = 5933.51$$

$$W = \frac{F(t) - \min F(t)}{\max F(t) - \min F(t)} = 0.52$$

$$\bar{S} = \frac{1}{43}\sum_{i=1}^{4} 3|x_{i+1} - x_i| = 94.59$$

Because $S^- > S^+$, $\mu(h_{-1}) = \frac{\bar{S}}{S^-}$ hence $\mu(L)$ = 0.71.

Continue to apply the algorithms that we have recommended, the following results are achieved:

***With h* = 7**

The values of I_i are calculated

$$\begin{array}{ll} I_1 = [5700.00, 5835.00) & I_5 = [6051.16, 6097.36) \\ I_2 = [5835.00, 5918.48) & I_6 = [6097.36, 6133.68) \\ I_3 = [5918.48, 5986.20) & I_7 = [6133.68, 6150.00 \\ I_4 = [5986.20, 6051.16) & \end{array}$$

$$\begin{array}{lll} \mathrm{Rev}(A_1) = 5743.76 & \mathrm{Rev}(A_2) = 5\,883.95 & \\ \mathrm{Rev}(A_3) = 5940.14 & \mathrm{Rev}(A_4) = 6001.92 & \\ \mathrm{Rev}(A_5) = 6083.61 & \mathrm{Rev}(A_6) = 6119.36 & \mathrm{Rev}(A_7) = 6119.57 \end{array}$$

The forecasted values are listed in Table 5, the accuracy metrics are

RSME = 53.87; NE(%) = 0.07; NNE(%) = 10.97

***With h* = 9** The values of I_i are calculated

I_1=[5700.00, 5777.00); I_2=[5777.00, 5835.00). I_3=[5835.00, 5884.00);
I_4=[5884.00, 5992.00). I_5=[5992.00, 6020.00); I_6=[6020.00, 6027.00).
I_7=[6027.00, 6083.00); I_8=[6083.00, 6097.00). I_9=[6097.00, 6150.00].

Table 5. Compare forecasted index result of proposed method with result of Chen

Date	Actual index	Chen' forecasted index	Our forecasted index	
		h = 9	h = 7	H = 9
2/11/2004	5759.61	5674.81	5813.86	5743
3/11/2004	5862.85	5768.14	5813.86	5852
4/11/2004	5860.73	5854.81	5892.44	5876.04
5/11/2004	5931.31	5908.26	5892.44	5876.04
8/11/2004	5937.46	5934.81	5912.05	5912.05
9/11/2004	5945.2	5943.81	5912.05	5912.05
10/11/2004	5948.49	5934.81	5912.05	5912.05
11/11/2004	5874.52	5937.12	5912.05	5912.05
12/11/2004	5917.16	5908.26	5892.44	5919.27
15/11/2004	5906.69	5934.81	5892.44	5919.27
16/12/2004	5910.85	5934.81	5892.44	5919.27
17/11/2004	6028.68	5937.12	5892.44	5919.27
18/11/2004	6049.49	6068.14	5977.41	5979.18
19/11/2004	6026.55	6068.14	5977.41	5979.18
22/11/2004	5838.42	5976.47	5977.41	5979.18
23/11/2004	5851.1	5854.81	5892.44	5876.04
24/11/2004	5911.31	5934.85	5892.44	5876.04
25/11/2004	5855.24	5934.81	5892.44	5919.27
26/11/2004	5778.65	5854.81	5892.44	5876.04
29/11/2004	5785.26	5762.12	5813.86	5797.89
30/11/2004	5844.76	5762.12	5813.86	5852
1/12/2004	5798.62	5834.85	5892.44	5876.04
2/12/2004	5867.95	5803.26	5813.86	5797.89
3/12/2004	5893.27	5854.81	5892.44	5876.04
6/12/2004	5919.17	5854.81	5892.44	5919.27
7/12/2004	5925.28	5937.12	5942.00	5912.05
8/12/2004	5892.51	5876.47	5942.00	5912.05
9/12/2004	5913.97	5854.81	5892.44	5919.27
10/12/2004	5911.63	5934.81	5892.44	5919.27
13/12/2004	5878.89	5937.12	5892.44	5919.27
14/12/2004	5909.65	5854.81	5892.44	5919.27
15/12/2004	6002.58	5934.81	5892.44	5919.27
16/12/2004	6019.23	6068.14	5977.41	5979.18
17/12/2004	6009.32	6062.12	5977.41	5979.18
20.12.2004	5985.94	6062.12	5977.41	5979.18
21/12/2004	5987.85	5937.12	5977.41	5979.18
22/12/2004	6001.52	5934.81	5977.41	5979.18
23/12/2004	5997.67	6068.14	5977.41	5979.18

(*continued*)

Table 5. (*continued*)

Date	Actual index	Chen' forecasted index	Our forecasted index	
		h = 9	h = 7	H = 9
24/12/2004	6019.42	5934.81	5977.41	5979.18
27/12/2004	5985.94	6068.14	5977.41	5979.18
28/12/2004	6000.57	5937.12	5942.00	5979.18
29/12/2004	6088.49	6068.14	5977.41	5979.18
30/12/2004	6100.86	6062.12	6119.36	6119.36
31/12/2004	6139.69	6137.12	6143.57	6143.57
	RSME	**56.86**	**53.87**	**48.02**
	NE(%)	**0.80**	**0.70**	**0.59%**
	NNE(%)	**12.44**	**10.97**	**9.17%**

$\mathrm{Rev}(A_1) = 5740.00$ $\mathrm{Rev}(A_2) = 5829.00$ $\mathrm{Rev}(A_3) = 5869.00$ $\mathrm{Rev}(A_4) = 5940.00$
$\mathrm{Rev}(A_5) = 6002.00$ $\mathrm{Rev}(A_6) = 6026.00$ $\mathrm{Rev}(A_7) = 6051.00$ $\mathrm{Rev}(A_8) = 6085.00$
$\mathrm{Rev}(A_9) = 6119.00$

The forecasted values are listed in Table 5, the accuracy metrics are

$$\mathbf{RSME = 48.02;\ NE(\%) = 0.59;NNE(\%) = 9.17}$$

Compared to the results of [13], our method gives more accurate results and the calculating process is much simpler.

5 Conclusion

Researchers who use information granules as models to predict time series has emphasized the inherent semantics of words, e.g. "Information granulation is inherent to fuzzy time series" and "Information granules are human-centric constructs capturing the semantics of the concepts of interest, which are inherent to all ensuing processes of abstraction" [11]. Information granules are linguistic values (or terms). But the inherent semantics of term is resulted from human knowledge hence depends on context. The context here is 'the high or low level of the annual number of enrolled students at Alabama university'. 'high' and 'low' are the two main words whose semantics are used to describe the information within a context. Other words 'little' and 'very' are the impacting words which have effect on 'high' and 'low' to create mediate semantics to illustrate the 'high', 'low' levels corresponding to the annual number of enrolled students at Alabama university. Hedge algebras is an approach to the inherent semantics of words to represent the semantics of information granules by fuzzy sets with their inherent semantics. With the definitions of 'quantitative semantics mapping', 'fuzzy set based semantics of the words', etc., the hedge algebras have constructed 'a set of weights' which $\in [0, 1]$ hence is a partition of $[0, 1]$. The normalized historical data of time series (from 0 to 1) is distributed within this set hence is the basis for the optimal partitioning. Each partition has a quantitative semantics value which can be considered as "semantics core" meaning all the historical data belongs to this partition will lie around this "semantics core". The fuzzy parameters of the HA are determined based on the analysis of the relationship between the historical values of a given time series. Consequently, the forecasting enrollment has been solved by the Context-Aware approach. The above statements fully explain our approach based on the inherent semantics of term is easy to understand and simplicity in practice to forecasting enrollment in fuzzy time series with remarkable accuracy in comparison with the other approaches has published.

References

1. Song, Q., Chissom, B.S.: Fuzzy time series and its models. Fuzzy Sets Syst. **54**, 269–277 (1993)
2. Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series – part I. Fuzzy Sets Syst. **54**, 1–9 (1993)
3. Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series – part II. Fuzzy Sets Syst. **62**, 1–8 (1994)
4. Chen, S.-M.: Forecasting enrollments based on fuzzy time series. Fuzzy Sets Syst. **81**, 311–319 (1996)
5. Huarng, K.: Efective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets Syst. **123**, 387–394 (2001)
6. Jilani, T.A., Burney, S.M.A., Ardil, C.: Fuzzy metric approach for fuzzy time series forecasting based on frequency density based partitioning. Int. J. Comput. Inf. Syst. Control Eng. **4**, 39–44 (2010)
7. Huarng, K., Yu, T.H.-K.: Ratio-based lengths of intervals to improve fuzzy time series forecasting. IEEE Trans. Syst. Man Cybern.—Part B Cybern. **36**, 328–340 (2006)
8. Bas, E., Uslu, V.R., Yolcu, U., Egrioglu, E.: A modified genetic algorithm for forecasting fuzzy time series. Appl. Intell. **41**, 453–463 (2014)
9. Wang, L., Liu, X., Pedrycz, W.: Effective intervals determined by information granules to improve forecasting in fuzzy time series. Expert Syst. Appl. **40**, 5673–5679 (2013)
10. Wang, L., Liu, X., Pedrycz, W., Shao, Y.: Determination of temporal information granules to improve forecasting in fuzzy time series. Expert Syst. Appl. **41**, 3134–3142 (2014)
11. Lu, W., Chen, X., Pedrycz, W., Liua, X., Yang, J.: Using interval information granules to improve forecasting in fuzzy time series. Int. J. Approx. Reason. **57**, 1–18 (2015)
12. Lan, V.N., Ho, N.C.: Hedge algebras-an order-based structure of terms-domains: an algebraic approach to human reasoning. J. Sci. Technol. **45**, 77–108 (2007)
13. Chen, S.-M., Chen, C.-D.: TAIEX forecasting based on fuzzy time series and fuzzy variation groups. IEEE Trans. Fuzzy Syst. **19**(1), 1–12 (2011)