

A Hybrid Algorithm to Forecast Enrolment Based on Genetic Algorithms and Fuzzy Time Series

Haneen Al-Wazan¹, Kais Ibraheem², and Abdul-Ghafoor Salim¹

¹Department of Mathematics, Mosul University, Iraq

²Department of Computers Sciences, Mosul University, Iraq

Abstract: In this paper, we proposed a hybrid algorithm to forecast enrolment based on fuzzy time series and genetic algorithms, the proposed algorithms presents a good forecasting result with higher accuracy rate. Historical enrolment of the University of Alabama from year 1948 to 2010 are used in this study to illustrate the forecasting process.

Keywords: Fuzzy time series forecasting, genetic algorithms prediction.

Received July 25, 2012; accepted January 15, 2013; published online April 4, 2013

1. Introduction

Forecasting is one of the important activities in business, enrolment, finance, etc., that helps in decision making. There are different models out of which researches had shown that genetic algorithms with fuzzy logic is one of the most powerful non-linear models for forecasting for any time series, The classical time series methods can not deal with forecasting problems in which the values of time series are linguistic terms represented by fuzzy sets.

Chena and Hsu [4] proposed a new method to forecast enrolments using fuzzy time series. Timma [13] studied forecasting using neural networks and genetic algorithms. Sah and Degtiarev [10] presented a novel improvement of forecasting approach based on using time-invariant fuzzy time series. Chen and Chung [5] investigated the forecasting enrolments of students by using fuzzy time series and genetic algorithms. They presented a new method to deal with forecasting problems based on fuzzy time series and genetic algorithms. Tahseen *et al.* [2] presented multivariate high order fuzzy time series. Jilani *et al.* [12] studied multivariate high order fuzzy time series forecasting for car road accidents. Mandal *et al.* [8] studied soft computing approach in prediction of time series data. Lee *et al.* [7] proposed the adoption of the weighted and the difference between actual data toward midpoint interval based on fuzzy time series, by using the enrolment of the University of Alabama and University Technology Malaysia (UTM). Zaharim *et al.* [15] proposed Genetic algorithm in time series fatigue analysis. Aydin *et al.* [3] studied the prediction algorithm based on fuzzy logic using time series data mining method. Memmedli and Ozdemir [9] investigated an application of fuzzy time series to improve ISE forecasting. Alina and Elena [1] presented a hybrid approach for modeling financial time series.

Haneen *et al.* [12] proposed a novel algorithm to forecast enrolment based on fuzzy time series. In this paper we proposed a hybrid algorithm to forecast enrolment based on fuzzy time series and genetic algorithms. We used the Historical data of the University of Alabama to apply the algorithm on it.

2. Basics of Fuzzy Time Series

In This section we briefly summarizes basic fuzzy and fuzzy time series concepts. The main difference of fuzzy time series and traditional time series is that the values of fuzzy time series are represented by fuzzy sets rather than real values. Let U be the universe of discourse, where $U = \{u_1, u_2, \dots, u_n\}$. A fuzzy set defined in the universe of discourse U can be represented by equation 1:

$$A = f_A(u_1)/u_1 + f_A(u_2)/u_2 + \dots + f_A(u_n)/u_n \quad (1)$$

Where f_A denotes the membership function of the fuzzy set A , $f_A: U \rightarrow [0, 1]$, and $f_A(u_i)$ denotes the degree of membership of u_i belonging to the fuzzy set A , and $f_A(u_i) \in [0, 1]$ and $1 \leq i \leq n$.

- *Definition 1:* Assume $Y(t) \subset \mathbb{R}$ (real line), $t = \dots, 0, 1, 2, \dots, n$ to be a universe of discourse defined by the fuzzy set $f_i(t)$. $F(t)$ consisting of $f_i(t)$, $i = 1, 2, \dots, n$ is defined as a fuzzy time series on $Y(t)$. At that, $F(t)$ can be understood as a linguistic variable, whereas $f_i(t)$, $i = 1, 2, \dots, n$, are possible linguistic values of $F(t)$ [5, 9].
- *Definition 2:* If $F(t)$ is caused by $F(t-1)$, denoted by $F(t-1) \rightarrow F(t)$, then this relationship can be represented by $F(t) = F(t-1) \circ R(t, t-1)$, where the symbol " \circ " is an operator; $R(t, t-1)$ is a fuzzy relation between $F(t)$ and $F(t-1)$ and is called the first-order model of $F(t)$ [4, 14].

- **Definition 3:** Denoting $F(t-1)$ by A_i and $F(t)$ by A_j the relationship between $F(t-1)$ and $F(t)$ can be defined by a logical relationship $A_i \rightarrow A_j$ [2, 9, 14].
- **Definition 4:** Let $R(t,t-1)$ be a first-order model of $F(t)$. If for any t , $R(t,t-1)=R(t-1,t-2)$, then $F(t)$ is called a time-invariant fuzzy time series. Otherwise, it is called a time-variant fuzzy time series [5].
- **Definition 5:** Fuzzy logical relationships, which have the same left-hand sides, can be grouped together into fuzzy logical relationship groups. For example, for the identical left-hand side A_i such grouping can be depicted using equation 2 [9]:

$$\left. \begin{array}{l} A_i \rightarrow A_{j1} \\ A_i \rightarrow A_{j2} \\ \vdots \end{array} \right\} \Rightarrow A_i \rightarrow A_{j1}, A_{j2}, \dots \quad (2)$$

3. Basics of Genetic Algorithms

Genetic algorithms (Gas) based on the principles of natural evolution. Due to its ease of applicability, numerous application of genetic algorithms are found in the area of business, scenic, engineering, and forecasting problems. Now we will mention some basics of genetic algorithms.

- **Chromosome**

All living organisms consist of cells. In each cell there is the same set of chromosomes. Chromosomes are strings of DNA and serves as a model for the whole organisms. In genetic algorithms terminology these are the point in the search space and represented by string of coded genes [13].

- **Population**

A population a collection of chromosomes. A population consists of number of chromosomes being tested. The two important aspects of population used in genetic algorithms are:

1. The initial population generation.
2. The population size [11].

- **Fitness Function**

The fitness of a chromosome in genetic algorithms is the value of an objective function. The fitness not only indicates how good the solution is, but also corresponds to how close the chromosome is to the optimal one. There is several types of fitness function but in time series forecasting mean square error, absolute error, root mean square error average error etc., are good convergence criterions to be used in the forecasting process [11].

- **Genetic Algorithms Operators**

A genetic algorithms is composed of three operators:

1. Initialize population.

2. Encoding.
3. Reproduction/Selection.
4. Cross over.
5. Mutation [11].

- **Best Chromosome**

A best chromosome is the one with the minimum fitness [11].

4. A Hybrid Algorithm to Forecast Enrolment Based on Fuzzy Time Series and Genetic Algorithms

In this section we proposed a hybrid algorithm to forecast enrolment of the University of Alabama based on fuzzy time series and genetic algorithms .The forecasting method consists of two ways, the first way is used to determined the best interval with less fitness function we use in this way only steps1-7. The second way forecasting is the find the final forecasting results we use in this steps 4, 5 and step 8 with the 3 rules. The steps of the algorithm are given below:

- **Step 1:** Collect the data.
- **Step 2:** Determine the maximum and the minimum of the interval $U = [D_{min} - D_1, D_{max} + D_2]$ where D_1 and D_2 are constant to define the universe of discourse $U = [a, b]$.
- **Step 3:** Divide the universe of discourse U into m intervals, u_1, u_2, \dots, u_m where $u_1 = [a, x_1], \dots, u_m = [x_{m-1}, b]$. Where x_1, \dots, x_{m-1} are integer constant and $x_1 < \dots < x_{m-1}$ as follows:

Define each chromosome consisting of $m-1$ genes as shown below: x_1, x_2, \dots, x_{m-1} .

Initially, the values of x_1, x_2, \dots, x_{m-1} of each chromosome are randomly generated by the system, and the system generates N chromosomes as an initial population. Where N is the numbers of chromosomes.

- **Step 4:** Fuzzify the historical enrolments with the chromosomes in the population. Determine the fuzzy logical sets by using equation 3:

$$A_i = (d_{i-1}, d_i, d_{i+1}, d_{i+2}) \quad (3)$$

Starting by $A_1 = (d_0, d_1, d_2, d_3)$ and ending by $A_m = (d_{m-1}, d_m, d_{m+1}, d_{m+2})$.

Where $d_0 = D_{min}-100, d_{m+2} = D_{max}$ and fuzzify the historical enrolments shown in Table 1, where the fuzzy set A_i denotes a linguistic value of the enrolments represented by a fuzzy set, and $1 \leq i \leq m$.

- **Step 5:** Determine the fuzzy logical relations as: $A_j \rightarrow A_i$ and find the fuzzy logical groups.
- **Step 6:** Select two chromosomes randomly from the population to perform the crossover and mutation operations. To perform the mutation operation we need to compare the value of genes i and $i+1; i = 1, 2, \dots, m-1$ if $(gene(i+1)-gene(i)) > 100$ or

$(gene(i+1)-gene(i)) < 200$ then regenerate the $gene(i)$, after the mutation operations, if the chromosomes are not sorted by the values of genes, then the system will sort the values of genes in the chromosomes in an ascending sequence.

• **Step 7:** Based on the corresponding intervals of each chromosome in the generated population the forecasted value at time t is determined by the following rules:

a. If the fuzzy logical relationship group A_j is empty $A_j \rightarrow \emptyset$ then by equation 4:

$$F_{vt} = A_j / 2 \tag{4}$$

b. If the fuzzy logical relationship group A_j is one to one $A_j \rightarrow A_k$ then the interval that contain the forecasted value is A_k then by equation 5:

$$F_{vt} = A_k / 2 \tag{5}$$

c. If the fuzzy logical relationship group A_j one to many $A_j \rightarrow A_{k1}, A_{k2}, \dots, A_{kp}$ then the interval that contain the forecasted value represented by equation 6:

$$F_{vt} = \frac{(A_{k1} + A_{k2} + \dots + A_{kp}) / p}{2} \tag{6}$$

• **Step 8:** The best chromosome will be selected by computing the fitness function which is the Mean Square Error (MSE) for each chromosome the best chromosome is the one that have less men square error and it will be the initial interval to find the new forecasting value we will repeat the steps 4, 5. And we will compute the final forecasting value as follows:

Based on the corresponding intervals of the best chromosome the forecasted value at time t is determined by the following rules:

a. If the fuzzy logical relationship group A_j is empty $A_j \rightarrow \emptyset$ then by equation 7:

$$F_{vt} = A_j / 2 \tag{7}$$

b. If the fuzzy logical relationship group A_j is one to one $A_j \rightarrow A_k$ then the interval that contain the forecasted value is A_k then:

Assume that the fuzzy logical relationship is $A_i \rightarrow A_j$, where A_i denotes the fuzzified enrolment of year $n-1$ and A_j denotes the fuzzified enrolment of year n . Compute Y as follows:

$Y = [(enrolment\ of\ year\ (n) - enrolment\ of\ year\ (n - 1)) - (enrolment\ of\ year\ (n - 1) - enrolment\ of\ year\ (n - 2))]$ then compare the value of Y with zero and apply the following:

1. If $j > i$, and $Y > 0$
 then the trend of the forecasting will go up, and we use the following Rule 2 to forecast the enrolments.

2. If $j > i$, and $Y < 0$
 then the trend of the forecasting will go down, and we use the following Rule 3 to forecast the enrolments.

3. If $j < i$, and $Y > 0$
 then the trend of the forecasting will go up, and we use the following Rule 2 to forecast the enrolments.

4. If $j < i$, and $Y < 0$
 then the trend of the forecasting will go down, and we use the following Rule 3 to forecast the enrolments.

5. If $j = i$, and $Y > 0$
 then the trend of the forecasting will go up, and we use the following Rule 2 to forecast the enrolments.

6. If $j = i$, and $Y < 0$
 then the trend of the forecasting will go down, and we use the following Rule 3 to forecast the enrolments.

c. If the fuzzy logical relationship group A_j one to many $A_j \rightarrow A_{k1}, A_{k2}, \dots, A_{kp}$ then the interval that contain the forecasted value is can be determined as follows:

1. If the difference between any two of $k1, k2, \dots, kp \leq 2$ then the interval that contain the forecasted value is calculated using equation 8:

$$IF_{vt} = \frac{A_{k1} + A_{k2} + \dots + A_{kp}}{p} \tag{8}$$

and the forecasting is the middle of this interval.

2. If the difference between any two of $k1, k2, \dots, kp > 2$ then the forecasted value is the middle of the interval calculated using equation 9:

$$IF_{vt} = \frac{A_{k1} + A_{k2} + \dots + A_{ki-1} + A_{ki+1} + \dots + A_{kp}}{p} = \left(\frac{d_{k1-1} + \dots + d_{(ki-1)-1} + d_{(ki+1)-1} + \dots + d_{kp-1}}{p}, \frac{d_{k1} + \dots + d_{(ki-1)} + d_{(ki+1)} + \dots + d_{kp}}{p}, \frac{d_{k1+1} + \dots + d_{(ki-1)+1} + d_{(ki+1)+1} + \dots + d_{kp+1}}{p}, \frac{d_{k1+2} + \dots + d_{(ki-1)+2} + d_{(ki+1)+2} + \dots + d_{kp+2}}{p} \right) \tag{9}$$

Where A_{ki} is the intervals that have deference > 2 , $i = 1, 2, \dots, p$ and the forecasting of the fuzzy sets A_{ki} is calculated as one to one fuzzy logical relationship by applying step 8-b.

The rules 1, 2 and 3 used in the algorithm are shown as follows:

• **Rule 1:** When forecasting the enrolment of year 1950, there are no data before the enrolments of year 1947, therefore we are not able to calculate equation 10:

$$Z = ((enrollment(1949) - enrollment(1948)) - enrollment(1948) - enrollment(1947)) \tag{10}$$

Therefore if $|En(1949) - En(1948)| / 2 > \frac{A_j}{2}$ then the trend of the forecasting of this interval will be upward and calculate using equation 11:

$$F_n = 0.75\ of\ A_j \tag{11}$$

If $|En(1949) - En(1948)|/2 = \frac{A_j}{2}$ then the forecasting enrolment falls at the middle value of this interval.

If $|En(1949) - En(1948)|/2 < \frac{A_j}{2}$ then the trend of the forecasting of this interval will be downward, and $F_n = 0.25$ of A_j ; where $En = enrolment$.

- **Rule 2:** If $x = |Y| * 2 + En(n-1) \in A_j$ or $x = En(n-1) - |Y| * 2 \in A_j$ then the trend of the forecasting of this interval will be upward and by using equation 12 we find:

$$F_n = 0.75 \text{ of } A_j$$

$$\text{If } x = \frac{|Y|}{2} + En(n-1) \in A_j \quad (12)$$

$$\text{or } x = En(n-1) - \frac{|Y|}{2} \in A_j$$

Then the trend of the forecasting of this interval will be downward, and calculate by equation 13:

$$F_n = 0.25 \text{ of } A_j \quad (13)$$

If neither is the case, then the forecasting enrolment be the middle value of the interval A_j .

- **Rule 3:** If $x = \frac{|Y|}{2} + En(n-1) \in A_j$ or $x = En(n-1) - \frac{|Y|}{2} \in A_j$ then the trend of the forecasting of this interval will be downward, and we will calculate it using equation 14:

$$F_n = 0.25 \text{ of } A_j$$

$$\text{If } x = |Y| * 2 + En(n-1) \in A_j \quad (14)$$

$$\text{If } x = |Y| * 2 + En(n-1) \in A_j$$

$$\text{or } x = En(n-1) - |Y| * 2 \in A_j$$

Then the trend of the forecasting of this interval will be upward, and $F_n = 0.75$ of A_j . If neither is the case, then we let the forecasting enrolment be the middle value of the interval A_j [5].

5. Results

We will apply the proposed algorithm to forecast the historical enrolments of the University of Alabama that shown in Table 8. The universe set is $U = [5172-172, 27052 + 948]$ i.e., $U = [5000, 28900]$:

1. We will apply our algorithm firstly on enrolment from year 1971 to 1992 so we can compare our method with others, As a start we will divide the universe of discourse $U = [13000, 20000]$ into 11 intervals so each chromosome consisting of 10 genes, the values of each chromosome are randomly generated by the system, and the system will generates 50 chromosomes as an initial population as shown in Table 1.
2. Fuzzify the historical enrolments with the chromosomes in the population. For example for the first chromosome shown in Table 1 we can see that the subintervals of the universe interval $U =$

$[13000, 20000]$ are u_1, \dots, u_{11} , where $u_1 = [13000, 13180]$, $u_2 = [13180, 13226]$, $u_3 = [13226, 13353]$, $u_4 = [13353, 14497]$, $u_5 = [14497, 15177]$, $u_6 = [15177, 15250]$, $u_7 = [15250, 15750]$, $u_8 = [1570, 16434]$, $u_9 = [16434, 18198]$, $u_{10} = [18198, 18210]$, $u_{11} = [18210, 20000]$. The fuzzified enrolment are shown in Table 2.

3. Select two chromosomes randomly for the crossover and mutation operations, we choose single point crossover, after the mutation operations, if the chromosomes are not sorted by the values of genes in an ascending sequence, then the system will sort the values of genes in the chromosomes in an ascending sequence. The mutation and crossover operations for the chromosomes are shown in Table 3.
4. The first step forecasting is to determine the best chromosome which is the one with less fitness value. The fitness function used in this paper is the Mean Square Error (MSE) that is given below;

$$MSE = \frac{\sum_{i=1}^n (Actual\ Enrollment_i - Forecasted\ Enrollment_i)^2}{n}$$
5. From 4, we determined the best initial interval for our second step forecasting. The chromosome no.36 is the best with $MSE = 166033$ will be used to find the final forecasting results.
6. In order to get better results we can increase the number of chromosomes genes to 130 genes (i. e., 130 subintervals) this can be shown in Table 4. Table 5 showed that the results will be better by increasing the size of population to 100 chromosomes. Table 6 showed that the proposed algorithm presents better results than the exist methods.
7. In this step we applied the proposed method on the data from 1948 to 2010 as given in Table 7. To do this we needed to generate randomly an initial population consisted of 100 chromosomes and each chromosome contained 365 genes (i. e., each chromosome consisted of 366 subintervals), then the crossover and mutation operators have been done, and we fuzzified enrolment. Then the forecasted enrolment is shown in Table 7.
8. In order to get better results we can increase the number of chromosomes genes to 130 genes (i. e., 130 subintervals) this can be shown in Table 4. Table 5 showed that the results will be better by increasing the size of population to 100 chromosomes. Table 6 showed that the proposed algorithm presents better results than the exist methods.
9. In this step we applied the proposed method on the data from 1948 to 2010 as given in Table 7. To do this we needed to generate randomly an initial population consisted of 100 chromosomes and each chromosome contained 365 genes (i. e., each

chromosome consisted of 366 subintervals), then the crossover and mutation operators have been done,

and we fuzzified enrolment. Then the forecasted enrolment is shown in Table 7.

Table 1. Randomly generated chromosomes.

Chrom No.	Genes Value									
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀
Chrom 1	13108	13226	13353	14497	15177	15250	15750	16434	18198	18210
Chrom 2	13201	13797	14492	16295	16849	17056	17703	19300	19615	19858
Chrom 3	16037	17016	17194	17857	17983	18095	18563	18917	18949	19130
Chrom 4	13870	15178	16429	16537	17400	17482	17486	17848	18625	19684
Chrom 5	13747	14054	15057	16144	16303	16536	16807	17390	17490	18219
Chrom 6	13916	13994	14275	15029	16300	16864	17923	18746	19735	19900
Chrom 7	13946	14697	15953	16046	16428	16877	17105	18772	19500	19571
Chrom 8	13049	13103	13381	16477	16628	16658	17954	18437	18601	18692
Chrom 9	13296	15744	15867	16848	17124	17278	17876	18292	18744	18957
Chrom 10	13669	13738	13986	14036	15788	16144	17012	17274	18624	19221
Chrom 11	13058	14718	14905	15079	15644	15733	15860	15957	16896	17371
Chrom 12	13938	15118	15494	15934	16602	17536	17805	18036	18992	19043
Chrom 13	14205	14284	14461	14884	16109	16276	17937	18070	18585	19732
Chrom 14	13396	13606	14291	15369	16727	17213	18135	18345	18524	18596
Chrom 15	13246	13907	14575	15098	16216	16448	16688	16957	17632	19531
Chrom 16	13206	13327	14276	14736	16261	16677	16870	16925	17424	19300
Chrom 17	15012	15102	15590	17393	17484	17882	17980	18193	18511	19623
Chrom 18	14443	15441	15560	17518	17563	17806	18025	18430	18810	19310
Chrom 19	13247	13891	14537	15630	15969	16583	17778	17951	19415	19455
Chrom 20	13246	14661	15417	15939	16384	17822	17939	18142	18433	19173
Chrom 21	13032	14248	14455	14656	14658	15226	16382	16640	18099	18688
Chrom 22	14960	15368	15697	15912	16424	17519	19281	19756	19818	19821
Chrom 23	14509	16148	16447	16711	16746	17528	18214	19266	19347	19504
Chrom 24	14099	14530	15093	15703	15775	16051	17336	17628	18199	18276
Chrom 25	14051	14470	14926	15052	16362	16639	17017	17325	19101	19401
Chrom 26	14261	14535	14679	15151	15836	16731	16737	17156	17378	19496
Chrom 27	13454	14878	15690	15782	15812	16198	16368	16976	17995	19358
Chrom 28	14946	16312	16862	17211	17219	17579	18325	18646	19139	19145
Chrom 29	13535	13902	14317	14643	16432	17505	17949	18489	19242	19698
Chrom 30	14700	14986	15165	15558	15995	16147	17100	17328	19169	19787
Chrom 31	13908	15300	15618	15777	15916	16028	16071	16288	16482	17096
Chrom 32	13523	14866	15952	16080	16112	16264	16612	17151	18846	19563
Chrom 33	13013	13347	13381	14196	17456	17816	18099	18254	18407	19165
Chrom 34	14445	14470	15061	15313	16422	17593	18152	18633	18708	18781
Chrom 35	13016	14094	14600	14692	14860	15981	16911	18372	18813	19653
Chrom 36	13560	13734	14238	14378	14518	15895	16041	16991	17559	18143
Chrom 37	14533	15362	16100	16193	16423	16601	16849	17483	18658	19208
Chrom 38	13986	14637	15162	15426	17244	17993	18238	18440	18512	19349
Chrom 39	13265	15289	16707	17205	17602	17822	18584	19243	19416	19786
Chrom 40	13188	13425	13535	13996	14710	15234	15896	16334	16869	16872
Chrom 41	13596	13874	14450	15475	16549	17490	17677	18898	19108	19112
Chrom 42	13417	13513	16054	16733	17206	17968	17987	18517	18972	19311
Chrom 43	14308	14570	15056	15376	16314	16877	18261	18364	19002	19794
Chrom 44	14508	15340	15429	15765	15998	16400	17558	18211	18390	19062
Chrom 45	14845	16576	17290	17411	17416	17753	17908	18139	19398	19686
Chrom 46	13657	13949	14504	15915	16341	17705	18058	18406	18845	19283
Chrom 47	14005	14141	14478	14835	15424	15673	16201	17392	18545	19486
Chrom 48	13220	13829	15675	15936	16049	16119	18838	19241	19548	19562
Chrom 49	13225	13977	14518	15304	15674	16288	17681	17804	18772	19433
Chrom 50	13366	13453	13904	14362	15565	15623	15836	16341	16364	18409

Table 2. Fuzzified enrolment from 1971 to 1992.

Year	Act. En.	Fuz. En.	Year	Act. En.	Fuz. En.	Year	Act. En.	Fuz. En.	Year	Act. En.	Fuz. En.
1971	13055	A ₁	1977	15603	A ₇	1983	15497	A ₇	1989	18970	A ₁₁
1972	13563	A ₄	1978	15861	A ₈	1984	15145	A ₅	1990	19328	A ₁₁
1973	13867	A ₄	1979	16807	A ₉	1985	15163	A ₅	1991	19337	A ₁₁
1974	14696	A ₅	1980	16919	A ₉	1986	15984	A ₈	1992	18876	A ₁₁
1975	15460	A ₇	1981	16388	A ₈	1987	16859	A ₉			
1976	15311	A ₇	1982	15433	A ₇	1988	18150	A ₉			

Table 3. The chromosomes after the crossover and mutation operations and calculating MSE.

X _i Chrom No.	Genes Value										MSE
	X ₁₀	X ₁₀	X ₁₀	X ₁₀	X ₁₀	X ₁₀	X ₁₀	X ₁₀	X ₁₀	X ₁₀	
Chrom **1	13038	13187	13201	14217	16162	16623	16815	17054	17760	19144	411417
Chrom **2	13108	13692	14219	14857	15866	17375	18613	18618	18904	19408	344707
Chrom **3	13035	13687	13870	14223	15950	16374	16616	18680	19243	19352	534125
Chrom **4	13848	14709	15178	16566	17244	17914	18935	19434	19465	19698	293851
Chrom **5	13916	14946	15663	16188	16320	17024	17814	17819	18277	18402	274286
Chrom **6	13698	13747	14281	15736	16524	17358	17577	18976	19450	19470	238009
Chrom **7	13049	14250	16695	16749	17369	17451	18285	19103	19108	19221	492943
Chrom **8	13103	13866	15117	15670	16045	16050	16153	16492	16912	18242	292888
Chrom **9	13669	14383	14761	14999	17007	17552	18436	18557	19313	19625	994932
Chrom **10	13296	13945	13946	14272	14432	16601	16957	17519	18771	19385	385769
Chrom **11	13122	13279	13938	15082	16921	16938	17582	18255	19023	19408	977408
Chrom **12	13058	13885	14440	14511	14683	16377	18426	18793	18862	18866	809108
Chrom **13	13396	14031	14666	14763	17160	18204	18440	18462	18651	19282	825731
Chrom **14	13001	13160	13606	13682	14478	15130	17827	18246	18486	19261	261064
Chrom **15	13206	13972	14369	14536	14846	15770	16641	17376	18076	19081	330278
Chrom **16	13063	13246	13400	13510	13984	14603	16554	16935	18412	18850	41620 3
Chrom **17	13611	13752	14443	15358	16151	17243	18385	18755	19223	19619	359639
Chrom **18	13632	14015	15012	15491	15550	15592	16034	16259	17723	18860	282464
Chrom **19	13246	13977	15337	15908	16145	17043	17430	18078	18732	19005	292395
Chrom **20	13247	13346	13458	13567	15101	15916	17626	18391	18723	19514	403797
Chrom **21	13366	14248	14750	15849	16681	18238	18352	19067	19198	19620	34466
Chrom **22	13012	13032	13051	13134	14150	15007	17870	17898	18813	19371	592839
Chrom **23	13283	13329	14099	14162	14202	16077	16847	17820	18750	19054	459396
Chrom **24	13840	13931	14509	16910	17429	17722	17926	18075	18730	19163	1312138
Chrom **25	14261	14295	16096	17891	18209	18303	18650	19051	19098	19121	458451
Chrom **26	13417	14051	14587	15446	15665	16694	17116	17429	19146	19146	247655
Chrom **27	13692	14468	14878	14893	15314	17027	17202	18149	18573	18908	496517
Chrom **28	13454	13740	14497	15087	15121	15680	18514	18695	19194	19313	607678
Chrom **29	13022	13042	13131	13902	13922	15799	17212	18800	19096	19248	183544
Chrom **30	13340	13535	13557	14777	15465	16246	16880	18813	19170	19337	571821
Chrom **31	13523	13762	14749	15639	16646	17065	17498	17960	18175	19124	244177
Chrom **32	13753	13908	14812	15243	17308	17443	18191	18532	19163	19208	49658
Chrom **33	13318	13347	13348	13624	15476	15480	16364	16428	17910	18715	314741
Chrom **34	13005	13013	13948	14071	14687	15381	15897	16440	16755	18331	453904
Chrom **35	13317	13337	13560	17115	17704	18471	18557	18909	19303	19314	1508675
Chrom **36	13179	13928	14339	14774	15695	16260	16560	17795	18707	19540	166033
Chrom **37	13270	13783	13986	14207	14624	15714	16814	18293	18858	18979	319689
Chrom **38	14335	14493	14533	14831	15139	15841	16423	18435	18569	18899	370970
Chrom **39	13188	13295	13833	14148	15079	15823	18264	19270	19406	19436	602690
Chrom **40	13027	13265	14688	15642	15932	16079	16422	18297	18300	18687	307608
Chrom **41	13398	13417	14336	15409	16503	18153	18762	18971	19363	19675	412638
Chrom **42	13513	15415	15493	16490	16565	16692	16696	16927	18861	19505	427158
Chrom **43	13574	13672	13831	14508	16151	16355	17224	19444	19479	19530	540702
Chrom **44	13585	13677	14308	17065	17913	18004	18418	18636	18662	19100	1077178
Chrom **45	13252	13657	13980	14876	15985	18440	18882	19336	19394	19739	665869
Chrom **46	13154	13289	13949	15185	15314	15805	15981	18357	18883	19342	475012
Chrom **47	13220	15599	16260	16422	16906	17086	17487	18026	18578	19036	307787
Chrom **48	13829	14920	15328	15519	16238	17016	17135	17879	19212	19319	283241
Chrom **49	13083	13366	13832	15156	15959	16779	18993	19027	19099	19133	581187
Chrom **50	13225	13262	14650	16308	16442	16537	17670	17959	18910	19834	444079

Table 4. Forecasted enrolment with population size=50 chromosome and different no. of intervals.

Year	Actual Enrolment	Forecasted Enrolment with Different Numbers of Intervals								
		8	20	40	50	70	90	100	130	
1971	13055	-----	-----	-----	-----	-----	-----	-----	-----	-----
1972	13563	13554	13710	13480	13481	13399	13573	13594	13612	13612
1973	13867	13554	13856	13840	13836	13600	14014	13859	13937	13937
1974	14696	14707	15173	15206	14640	14872	14976	14623	14949	14949
1975	15460	15335	15444	15399	15175	15444	15616	15472	15567	15567
1976	15311	15335	15389	15399	15175	15406	15491	15290	15289	15289
1977	15603	15711	15606	15444	15175	15570	15730	15602	15750	15750
1978	15861	15711	15903	15516	15841	15722	16000	15876	15838	15838
1979	16807	17178	16703	16657	16872	16789	16790	16830	16787	16787
1980	16919	16869	16703	16849	16872	16789	16813	16898	16832	16832
1981	16388	16310	16266	16466	16367	16350	16252	16445	16357	16357
1982	15433	15335	15444	15399	15175	15444	15534	15472	15483	15483
1983	15497	15335	15444	15399	15175	15526	15578	15472	15567	15567
1984	15145	15335	15173	15206	15041	15319	15412	15110	15233	15233
1985	15163	15204	15173	15123	15041	15444	15412	15110	15233	15233
1986	15984	15944	16266	15700	15915	16265	16054	16026	15851	15851
1987	16859	17486	16703	16849	16872	16789	16813	16861	16797	16797
1988	18150	18251	18461	17882	18057	18451	17911	18253	17968	17968
1989	18970	19332	19238	18810	18954	18866	18975	19003	19022	19022
1990	19328	18915	19325	19195	19343	19330	19345	19337	19343	19343
1991	19337	19124	19325	19264	19343	19330	19345	19337	19343	19343
1992	18876	18915	19238	18762	18832	18803	18924	18890	18995	18995
MSE		53420	35102	32726	24428	22780	20914	14908	9816	9816

Table 5. Forecasted enrolment with population size=100 chromosome and 130 intervals.

Year	Act.	Forec.	Year	Act.	Forec.	Year	Act.	Forec.	Year	Act.	Forec.
1971	13055		1977	15603	15626	1983	15497	15498	1989	18970	18958
1972	13563	13568	1978	15861	15849	1984	15145	15125	1990	19328	19306
1973	13867	13869	1979	16807	16796	1985	15163	15160	1991	19366	19356
1974	14696	14686	1980	16919	16953	1986	15984	15988	1992	18804	18903
1975	15460	15463	1981	16388	16395	1987	16859	16859			
1976	15311	15297	1982	15433	15416	1988	18150	18130	MSE	246	

Table 6. A comparison of the MSE of the forecasted enrolment for different forecasting methods.

Method	MSE
Chen and Hsu [4]	5353
Lee <i>et al.</i> [7]	16248.7
Chen and Chung [5]	35324
Huang [4]	226611
Jilani <i>et al.</i> [6]	41426
The proposed method	246

Table 7. Forecasted enrolments of the University of Alabama from year 1948 to 2010.

Year	Actual Enrolment	Forecasted Enrolment	Year	Actual Enrolment	Forecasted Enrolment
1948	8916	----	1980	16919	
1949	7974	7979.00	1981	16388	16394.00
1950	6293	6260.00	1982	15433	15425.00
1951	5269	5260.00	1983	15497	15518.00
1952	5172	5165.00	1984	15145	15120.00
1953	5652	5618.00	1985	15163	15178.00
1954	6111	6068.00	1986	15984	16014.00
1955	7038	7040.00	1987	16859	16869.00
1956	7112	7106.00	1988	18150	18146.00
1957	7032	7040.00	1989	18970	18947.00
1958	7089	7090.00	1990	19328	19318.00
1959	7407	7424.00	1991	19366	19352.00
1960	7848	7816.00	1992	18804	18886.00
1961	8257	8267.00	1993	18909	18886.00
1962	8560	8569.00	1994	18707	18713.00
1963	8879	8859.00	1995	18561	18583.00
1964	9724	9717.00	1996	17572	17573.00
1965	10938	10933.00	1997	17877	17866.00
1966	11975	11974.00	1998	17929	17935.00
1967	12251	12255.00	1999	18267	18653.00
1968	12816	12836.00	2000	18859	18840.00
1969	13035	13057.00	2001	18735	18713.00
1970	13017	13057.00	2002	19181	19167.00
1971	13055	13057.00	2003	19828	19846.00
1972	13563	13548.00	2004	20512	20533.00
1973	13867	13863.00	2005	20969	20948.00
1974	14696	14668.00	2006	21835	21842.00
1975	15460	15468.00	2007	23878	23867.00
1976	15311	15307.00	2008	25580	25586.00
1977	15603	15574.00	2009	27052	27055.00
1978	15861	15824.00	2010	28807	28818.00
1979	16807	16798.00		MSE	323.3

6. Conclusions

In this paper we presented a hybrid algorithm to forecast enrolment based on fuzzy time series and genetic algorithms. Historical enrolments of the University of Alabama are used in this study to illustrate the forecasting process. From Table 5 we can see that when the number of intervals is bigger, then the accuracy rate will be higher, From Tables 5 and 6 we can see that the other effective element is the population size, when the population size was 50 the MSE=9816 but when the population size was 100 the MSE=246, with the same number of intervals. Figure 1 shows a comparison between actual and forecasted enrolment, We also compare our method with other methods, Table 7 shows that our method presents better forecasting results and can get a higher forecasting accuracy rate than the existing methods.

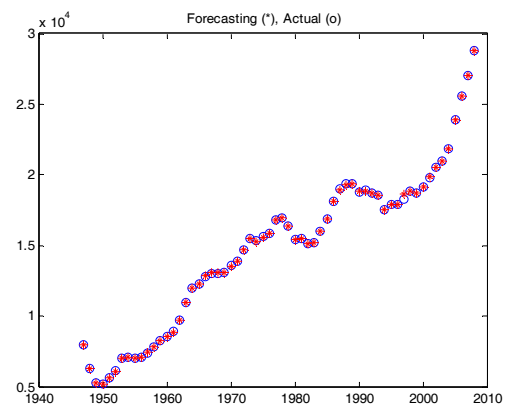


Figure 1. A comparison between actual and forecasted enrolment.

References

[1] Alina B. and Elena B., “A Hybrid Approach for Modeling Financial Time Series,” *the International Arab Journal of Information*

- Technology, vol. 9, no. 4, 2012. pp. 327-335, 2012.
- [2] Allahviranloo T., Hosseinzade F., and Adabitar M., "Fuzzy Efficiency Measure with Fuzzy Production Possibility Set," *Applications and Applied Mathematics: An International Journal*, vol. 2, no. 2, pp. 152-166, 2007.
- [3] Aydin I., Karakose M., and Akin E., "The Prediction Algorithm Based on Fuzzy Logic Using Time Series Data Mining Method," in *Proceedings of World Academy of Science, Engineering and Technology*, vol. 51, pp. 91-98, 2009.
- [4] Chen S. and Hsu C., "A New Method to Forecast Enrolments Using Fuzzy Time Series," *the International Journal of Applied Science and Engineering*, vol. 3, no. 2, pp. 234-244, 2004.
- [5] Chen S. and Chung N., "Forecasting Enrollments of Students by Using Fuzzy Time Series and Genetic Algorithms," *Information and Management Sciences*, vol. 17, no. 3, pp. 1-17, 2006.
- [6] Jilani T., Burney S., and Ardil C., "Multivariate High Order Fuzzy Time Series Forecasting for Car Road Accidents," in *Proceedings of World Academy of Science, Engineering and Technology*, Berlin, Germany, vol. 25, pp. 288-293, 2007.
- [7] Lee M., Efendi R., and Ismail Z., "Modified Weighted for Enrolment Forecasting Based on Fuzzy Time Series," *MATEMATEKA*, vol. 25, no. 1, pp. 67-78, 2009.
- [8] Mandal N., Choudhury J., Chaudhuri S., and Dilip D., "Soft Computing Approach in Prediction of a Time Series Data," *Journal of Theoretical and Applied Information Technology*, vol. 8, no. 3, pp. 1131-1141, 2008.
- [9] Memmedli M. and Ozdemir O., "An Application of Fuzzy Time Series to Improve ISE Forecasting," *WSEAS Transactions on Mathematics*, vol. 9, no. 1, pp. 12-21, 2010.
- [10] Şah M. and Degtiarev K., "Forecasting Enrolment Model Based on First-Order Fuzzy Time Series," in *Proceedings of World Academy of Science, Engineering and Technology*, vol. 1, pp. 375-378, 2005.
- [11] Savanandam S. and Deepa S., *Introduction to Genetic Algorithms*, Springer, Berlin, 2008.
- [12] Talal H., Jasim A., and Ibraheem K., "A Novel Algorithm to Forecast Enrollment Based on Fuzzy Time Series," *Applications and Applied Mathematics: An International Journal*, vol. 7, no. 1, pp. 385-397, 2012.
- [13] Timma K., "Forecasting Using Neural Networks and Genetic Algorithms," *Kanwal Rekhi School of Information Technology*, Indian Institute of Technology, Bombay, 2004.
- [14] Tofigh A., Mikaeilvand N., Narsis K., and Rasol M., Decomposition of Fully Fuzzy Linear Systems, *Applications and Applied Mathematics: An International Journal*, vol. 3, no. 1, pp. 77-88, 2008.
- [15] Zaharim A., Abdullah S., Ibrahim D., and Nopiah M., "Genetic Algorithm in Time Series Fatigue Analysis," *European Journal of Scientific Research*, vol. 28, no. 1, pp. 68-75, 2009.



Haneen Al-Wazan received her BSc degree in mathematics from the Department of Mathematics, College of Computers Sciences and Mathematics, Mosul University, Iraq 2005. Her MSc degree in numerical analysis from the Department of mathematics, College of Computers Sciences and Mathematics, Mosul University, Iraq 2008 and she is a PhD student in the research stage in time series analysis and forecasting using artificial intelligence techniques in the College of Computers Sciences and Mathematics, Mosul University, Iraq. Her research interests include numerical analysis, time series analysis and forecasting, fuzzy time series forecasting and forecasting using neural network. She is author of 7 research articles.



Kais Ibraheem received his MSc degrees in applied mathematics and PhD degrees in electrical engineering from Gadjah Mada University, Indonesia, in 1998 and 2003, respectively. He is currently a head of department of computer sciences faculty of education, Mosul University, Iraq. His research interests include artificial intelligence, wavelet, and numerical analysis.



Abdul-Ghafoor Salim received his BSc degree in mathematics from the Department of Mathematics, College of Sciences Mosul University, Iraq, 1984. His MSc degree in time series from the Department of Mathematics, College of Sciences Mosul University, Iraq, 1989. His PhD degree in time series from the Department of Mathematics, College of Sciences Mosul University, Iraq, 1996. His research interests include time series analysis, and forecasting. He is the author of over 30 research articles. He is the supervisor of 9 MSc thesis, and 5 PhD thesis. He participated at many conferences inside and outside the country.