Nonlinear Modeling of Time Series Based on the Genetic Programming and Its Applications to Clustering of Feature in Stock Prices

呂，建軍

https://doi.org/10.15017/3000386

出版情報：経済論究. 123, pp.129-141, 2005-11-30. 九州大学大学院経済学会

Nonlinear Modeling of Time Series Based on the Genetic Programming and Its Applications to Clustering of Feature in Stock Prices

Jianjun Lu

1 Introduction

Many investment decisions are usually supported by the database system which maintains financial time series such as stock prices and exchange rates [6]. Especially, in the technical analysis of financial time series, traders recognize several characteristics as the features of time series which reveals in the segments of time series [20]. For example, traders expect a rise of stock price if some shapes of 'triangle' are found in the stock prices. Even though such kind of knowledge by expertise is valuable in investment, but these experiences are usually not shared by another investor. It is also doubtful whether the knowledge is able to be widely usable.

This paper deals with clustering of segments of stock prices by using the identification and modeling system for time series based on the Genetic Programming (GP) [17]. We apply the GP procedure in learning phase of the system where we improve the non-linear functional forms to approximate the models used to generate time series [2]-[4] [8]-[12] [21].

Conventional methods for clustering time series are categorized into two groups, and the first group contains various methods using the approximation system based on the basic functions, and the second group contains syntactical recognition methods [7]. For example, in the first group, the Fourie Transform and the Radial Basis Functions are usedto expand the time series into the components, and then we examine these components to find features of the time series [18] [19]. However, in these methods using the expansion into components we can not handledirect features in the time domain, and it is hard to find relations among temporal data.

The second method is resemble to the system for understanding natural language where the features of time series are described accordingto a kind of syntax so that we read and understand the meaning of the time series [22] [23]. However, usually the subsystem to recognize the features of the time series are not simple, and the comprehensive building of the whole system becomes a tough task. Contrast to these conventional methods, clustering system based on the GP procedure proposed in the paper is simple enough and we fully utilize the relations among temporal data. Moreover, the feature of the time series is described as a non-linear functional form, then the transient characteristics of the segments of time series is sufficiently retained.

The GP procedure has been successfully applied to the estimation of chaotic dynamics using the observed time series, and a direct-control method for chaotic dynamics is proposed based on the GP [5][8]-[12][21]. Moreover, the GP method has been widely used to emulate the agents' behavior in various markets such as the stock market. In this paper we extend the GP method to estimate and approximate nonlinear functional forms to fit the segments of time series, and its application to clustering of time series [1]-[4][21][24].

The GP method is effective in clustering phase as well as in learning phase. These non-linear functional forms are represented as tree structure (called individuals), and one tree corresponds to a model to generate time series. We have many individuals (pool) for the recognition whether a time series belongs to a certain cluster. The individuals are improved by using the GP procedurein learning phase so that the estimation becomes to be better. The scheme to maintain the pool of individuals is necessary for the GP procedure, but it also contributes to absorb the variation of the time series in clusters. Because, it is reasonable tounderstand that non-linear functional form to approximate the time series in a cluster is not a single form while the member of the cluster are regarded to be generated from the original time series by expanding or shrinking the time scale. Then, the variation of the individuals with relatively high capability in the pool can cope with clustering for various kinds of time series which are seemed to belong to the same cluster. The scheme is the same as the classifier systems used in the GA-based optimization where a set of strings of binary data are used repeatedly depending on the change of environment. If we usea single functional form to classify underlying time series,we may fail to approximate the time series.

As an application, we show clustering of artificially generated time series where clusters are obtained by expanding orshrinking by a transformation functions. Then, we apply the system to clustering of 8 kinds of segments of real stock prices which are typically found in the technical analysis.

In the following, in Section 2, we describe modeling of time series and the overview of the system. In Section 3, we explain the basics of the GP procedure. Section 4 describes the application toclustering of artificially generated time series, in Section 5 and Section 6, we show the result of clustering of segment of real stock prices.

2 Description of Feature of Time Series

2.1 Clustering of time series using features

In conventional studies, the clustering of time series is usually formalized based on the knowledge of expertise having sufficient experiences. For example, in the medical diagnosis expertise compare the underlying EEG (Electroencephalogram) or ECG (Electrocardiogram) with the dataset of times series which are accumulated corresponding to each disease [5][22]. It is also known in the technical analysis of stock trends, traders recognize the features of segments of stock prices so that they predict the rise or fall of price. Various patterns of stock trends are accumulated on the basis of specific features which are regarded to suggest us the rise/fall of stock price [20].

However, these human-based procedure are seemed to be costly and time consuming, then the automatization of the recognition process and the clustering of the time series may bring us the way of efficient investment. Moreover, it is also expected that the system can provide us the retrieval system of time series based on the feature of time series in relatively large database.

We must note following problems are needed to be resolved in clustering the time series based on the feature.

(1) Range of data

Depending on the time and situation in observing the time series, the amplitude of the time series are not necessarily restricted in a small range. Therefore, some kind of function is needed for adjusting the variability of range of time series.

(2) Expansion and shrinking of time scale

It may also happen that the length of continuation of observation of times series or the segments of time series embedded with features are not the same in any instance. We may classify a group of time series to the same cluster where each time series has a similar figure but it is deformed after the figure is expanded or shrunk along the time. The problem of expansion and shrinking of time scale is usually resolved by using the dynamic programming methods which are usually used in the recognition of voice. However, we may use another simpler method to treat the automatic clustering of conventional time series.

(3) Simplicity of feature description

It is also necessary to find simplified description of feature of time series so that we apply the various kinds of patterns of stock trends. Moreover, if the feature description is able to be a simple one, we can use the description also for the key in the retrieval of time series database.

The feature description and clustering method proposed in the paper based on the GP is expected to resolve above problems. For the first problem, the GP method improves the functional form for describing the relation among time series data, then difficulty caused by the difference of range occurred in each time series is almost always removed.

In term of the second problem, the functional form to describe the feature of time series is not a single form, but a pool of functional form is stored in the GP system. Then the variation and various patterns are realized in the same pool of individuals (functional forms) which are regarded to belong to the same cluster of time series.

For the simplified description of features (the third problem), we can control the complexity of the functional form in the GP method, and we have various level of feature description. If we use the shorter length of symbols to describe the feature, we have a simple form to detect the time series based on the overview. On the other hand, if we used longer length of symbols, then we can recognize the time series in more details.

2.2 Overview of system

In the following, we show the overview of the system for feature description and clustering method proposed in the paper based on the GP [7].

(1) Learning data

It is assumed that the time series data is stored and available in the system, each of which is divided into the same length. Moreover, it is assumed that the time series data used for learning process is available, each of which is accompanied with the cluster (category) to which the underlying time series is expected to be classified. In Figure 1 depicting the learning phase, it is assumed that such kinds of learning data are prepared for each clusters $i(i=1\sim n)$.

(2) Learning based on the GP

At first, the approximations of functional forms for each cluster are obtained to describe the generating models (features) of time series. These functional forms correspond to the individuals in pool in the GP procedure. The GP method is used for the approximation of functional forms. The functional form (individual) is not a single form, but is composed of a set of forms to approximate various generating models, while learning data includes variation of time series obtained from a single basic form by expanding or shrinking the time scale. Then, in the pool used for the GP method a number of individuals having relatively higher ability (called fitness) of approximation are retained in the system, and are used for clustering.

Since we use a set of learning data repeatedly to improve the fitness of individuals, the maximum value of the fitness of individuals in the pool does not increase monotonically, which is different form ordinary optimization processes or approximation using the GP.

Therefore, the iteration of the GP procedure is carried out until sufficient variation of functional form (individuals) for approximating the feature of time series is obtained.

As is shown in Figure 1, in the Learning Phase the independent pools of individuals

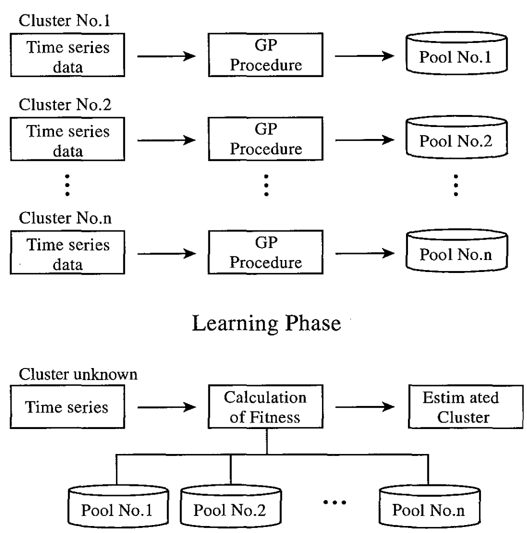

Figure 1: Overview of the systems

are organized for each cluster $i(i=1\sim n)$.

(3) Calculation of fitness of individuals

After applying the Learning Phase, we calculate the fitness ofindividuals in the Clustering Phase in Figure 2 for every individual stored in each pool $No.1\sim No.n$ by adopting (fitting) the observed data $x(t)$ of underlying time series whose cluster is not known and is needed to be determined by the system. Since the individual represented in the functional form is regarded as the prediction of time series at time t using data up to time $t-1$, we can calculate the prediction of $x(t)$ denoted as$\hat{x}(t)$. If the root mean square error (called *rmse*) between $x(t)$ and $\hat{x}(t)$ is small, then the fitness (ability) of a certain individual i realizing the underlying functional form is relatively high. Then, the fitness f_i of individual i is defined by the inverse of *rmse*.

In the Clustering Phase, we calculate the fitness f_i for every individual in every pool by fitting the observation $x(t)$ of time series with known cluster. Then, we estimate (determine) the cluster K of the time series by selecting the highest f_{max} among f_i where the individual i belongs to the K th pool.

3 Applying the GP to Nonlinear Function Approximation

3.1 Representation of equations

The prefix representation follows traditional representation by using the Lisp syntax [1]-[5][8]-[16][24]. Forexample, we have the next prefix representation.

$$\times - \times 3x(t-1)x(t-2)-x(t-3)4 \qquad (1)$$

For checking the validity of underlying parse tree, the so-called stack count (denoted as *StackCount* in the paper) is useful [2]-[4][8]-[12][16][24]. The *StackCount* is the number of arguments it places on minus the number of arguments it takes off from the stack.

3.2 Algorithm of the GP

Any two loci on the two parents genomes can serve as crossover points as long as the ongoing *StackCount* just before those points is the same. The crossover operation creates new offsprings by exchanging sub-trees between two parents.

(Step 1) Generate an initial population of random composition of possible function sand terminals for the problem at hand.

(Step 2) Execute each individual (evaluation of system equation) in population by applying the optimization of the constants included in the individual. Then, assign it a fitness S_i for individual i giving partial credit for getting close to the correct output.

(Step 3) Select a pair of individuals chosen with a probability p_i based on the fitness.

(Step 4) Then, create new individuals (offsprings) from the selected pair by genetically recombining randomly chosen parts of two existing individuals using the crossover operation applied at a randomly chosen crossover point.

(Step 5) If the result designation is obtained by the GP become larger than the prescribed value), then terminate the algorithm, otherwise go to Step 2.

4 Application to Clustering of time series

4.1 Cases for artificially generated time series

To test the capability of clustering method of the paper, we apply the system to the time series which are artificially generated by transforming the time scale for the original time series. At the beginning, we prepare 11 observations of time series, and they are regarded to be generated by using independent known generating function. Table 1 shows the overview of the time series.

The first time series is a chaotic time series generated by the logistic map, and is the target to test the ability of clustering proposed in the paper. The time series from denoted as No.2 through No.11 are selected from real world data, and they are assumed to be a kind of random effect to degrade clustering for No.1 time series. If the capability of clustering for the time series denoted as No.1 is not enough, then the time series belonging to No.1 cluster are partly misclassified to clusters from No.2 through No.11.

Since the time series No.1 is generated by a deterministic function (chaotic dynamics), but for other time series we have no apparent functions for generating the time series, even though we can approximate the function which generated the time series with some range of fitting error. For these reasons, the time series from No.2 through No.11 are not the target to test the ability of the system, but we use them as the random factor to test the robustness of the system of clustering.

Table 1: Result of clustering of segments (p:%)

numbers	overview
No.1	logistic map
No.2	demand of goods 1
No.3	demand of goods 2
No.4	demand of goods 3
No.5	demand of coal
No.6	birth of lynx
No.7	sunspots
No.8	number of airplane passengers
No.9	demand of durable goods
No.10	demand of nondurable goods
No.11	number of house building

At the next step, we apply the transformation to these original time series by changing the time scale by expanding and shrinking the time scale. The function for the transformation is called as the warping function, and is depicted in Figure 3. By applying various warping function to the original time series in Table 1, we have a set of clusters of time series where a group of time series generated from a certain original time series in Table 1 by applying warping function should belong to the same cluster.

Definition of warping function

The warping function is defined as a curve obtained by overlayinga half cycle sinusoidal wave or a full cycle sinusoidal wave having the amplitude a on the straight line with 45 degree angle between two axis. Namely, the beginning and the ending points of the time scale for both original and transformed time series are the same, but intermediate sample points are shifted according to the transformation depicted in Figure 3. The sample points at time t_1 in the original time series is

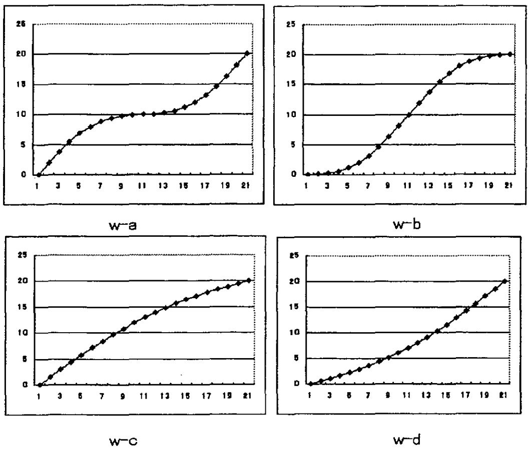

Figure 2: Definition of warping function

shifted to time t_2 in the transformed time series. The form of the warping functions are denoted as $w-a$, $w-b$, $w-c$, $w-d$ from the left top through the right bottom in Figure 3.

If we take 30 different values for the amplitude a, we have 30 time series for each clusters in Table 1. and these 30 time series should be theoretically classified into the same cluster.

Learning by the GP

Then, we apply the Learning Phase to the generated times series for each cluster. The condition for the learning by the GP is summarized as follows. Maximum length of individual: 20 Operators included in individuals: $,+,-,\times,abs$ Operands included in individuals: $x(t-1)$, $x(t-2),\cdots$, $x(t-10)$
Number of individuals in pool: 1000
Maximum generation of GP: 900

The way of learning is given as follows. Generated 30 time series are repeatedly used for learning, namely, a time series is used for 30 generation and the time series for learning is switched to another time series. A set of 1000 individuals is assignedas a pool to a cluster among 11 clusters, and the fitness of the individuals is improved by applying the GP procedure at 900 times. Finally we obtain a LCS composed of individuals possessing higher fitness for clustering the time series.

By the way, the maximum number of the GP generation is limitedto 900, while our purpose is to obtain a pool of individual having relatively higher fitness as a group rather than the optimization or the ultimate approximation of the problem. Since 30 time series are repeatedly used for learning for the Learning Phase including the original time series as well as transformed time series, the maximum fitness of the pool of the individuals dose not increase monotonically. An example of the maximum fitness along the GP generation is depicted in Figure 4.

Simulation result

Table 2 summarizes the result of classification for the time series in No.1 cluster by using the system proposed in the paper. The

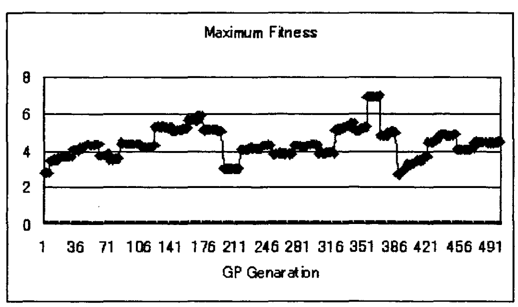

Figure 3: Example of plot of maximum fintess along GP generation

Table 2: Result of clustering (probability of true clustering%)

	w-a	w-b	w-c	w-d
No.1	100	90	100	100

value in Table 2 shows the rate of time series which are truly classified to cluster No.1 among the total 30 time series. For simplicity, the data for learning and testing are the same in the simulation study. In Table 2, the result for four warping function are summarized. As a result, the ability of the system for clustering artificially generated time series is good enough. About 96% of the time series of cluster No.1 is truly classified to cluster No.1.

5 Application to clustering of segments in stock trends

5.1 Segments in stock trend

Now we apply the method of the paper to clustering of segment sincluded in the stock trends (prices). In the technical analysis of stock price, the parts of the stock price (called as segments) are characterized with their features as almost standardized forms [20]. Figure 5 shows these basic 8 patterns of the segments as rough sketches. It is known that traders forecast the rise/fall of underlying stock by checking and recognizing the appearance of these 8 patterns of segments The notations and characteristics of these 8 segments are summarized as follows.

(a) downtrends descending parallel trend channel.

(b) uptrends ascending parallel trend channel.

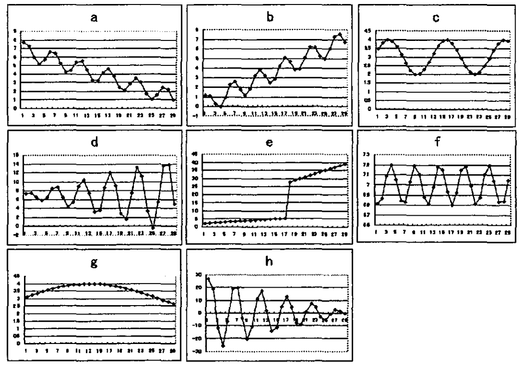

Figure 4: Overview of 8 patterns of segments in stock prices

(c) double top also called an "M" formation.

(d) broadning a rise and fall in expanding triangle.

(e) breakout there is a feature wave form in the former steps, and the stock prices keeping monotonously rising.

(f) rectangular a rise and fall in two balanced lines.

(g) rounding a big curve that is upward or downward and the peak value of stock prices appears at both ends.

(h) triangle a rise and fall in symmetrical triangle.

Clustering segments of stock price

Since the process of clustering of the time series describedin the preceding sections is the same for the clusters of segments of stock price, we concisely summarize only the result of simulation for clustering of segments. The condition for the simulation study is as follows:

Number of time series for each cluster: 30

Maximum length of individual: 20
Operators included in individuals:, $+$, $-$, $\times$, *abs*
Operands included in individuals:
$x(t-1)$, $x(t-2)$, $\cdots$, $x(t-10)$
Number of individuals in pool: 1000
Maximum generation of GP: 600

For simplicity, the data for learning and testing are the same in the simulation study.

Table 3 shows the result of clustering. In Table 3, the numbers in vertical column mean the original clusters to which the time series belong, and the numbers in horizontal row mean theestimated clusters obtained by the system which is determined by clustering process. The values in the table denote the rate of clustering. Theoretically, only the diagonal elements take the value of 100%, but due to the classification error, a part of time series are misclassified into another clusters.

However, the rate of proper clustering ranges from 70% through 97%, and the average value of proper clustering is about 85%. Then we can confirm the ability of clustering system.

Table 3: Result of clustering of segments (p:%)

segment No.	p	segment No.	p
a	83	b	87
c	97	d	77
e	97	f	93
g	70	h	70

6 Clustering by using sliding window

6.1 Clustering of continuous stock prices (artificial time series)

In previous sections, we showed time series belonging to several clusters in artificially generated time series and segments of stock prices are classifier by using clustering algorithm proposed in the paper. However, in a long record of time series, it mayhappen the cased where the time point of the beginning andending of the segment are not known beforehand. Then, we examinethe capability of clustering scheme of the paper even in the cases where we must also find the time interval in which the features (segments) of the stock prices are included.

Similar to the ordinary method for the recognition of time interval in a time series, we use the sliding window method to find the beginning and ending of the segments. The method is based on overlapping the time window along the time so that the nonlinear modeling using the GP and the detection of the time interval are simultaneously realized. The algorithm is summarized as follows.

(1) Sliding windows

Given the whole time series as $x(t)$, $t=t_1, t_2, \cdots, t_M$, and the interval T where the segments listed in Figure 3 are usually included in the time interval T. We assumethat clustering system of the paper process the time series of the length T, and we call the length T as the length of window. Then, we define the fraction of window, for example, $I=T/5$.

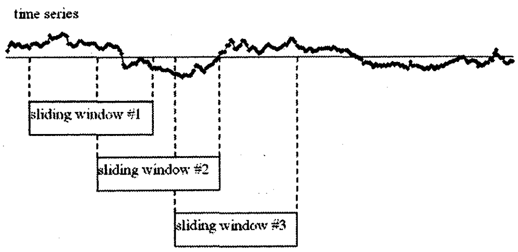

Figure 5: Overview of sliding windows

Then, we move the starting point T_s and the ending point T_e of the window such as $T_s=1+k\times I$, $T_e=T_s+T$ where k is integer. Then, the all of the set of windows (called as sliding windows) cover the whole time series by changing the starting point and ending point incrementally.

By using the scheme of sliding window, we can find the true interval in the time domain in which the segments of the time series are included by observing the maximum fitness of clustering system.

(2) Record of estimated clusters and maximum fitness

Then, we apply clustering method of the paper to the portion of time series which are taken out from the whole time series by using the sliding windows. According to clustering algorithm, we have a certain estimated cluster to which the portion of time series to be classified, and the maximum value of fitness obtainedby the system. Then, we take a record of these two values along the time (denoted as K_i for the cluster and F_i, $i=1, 2, \cdots, 5\times M$ for the fitness, respectively. It isexpected that the value F_i is large enough, the windowin the time period captures the true segment of time series, otherwise we estimate that the system fails to recognize true segments.

(3) Miscellaneous segments

Now, we assume that in the whole time series of stock price $x(t)$, we fined miscellaneous segments different from the segments listed in the Figure 5. These miscellaneous segments are assumed to be the segments found in ordinary stock price but having no dominant feature of characteristics observed in real stock prices. For the simulation study, we mix these miscellaneous segments as well as typical segments listed in Figure 5 so that we have more realistic situation forclustering the stock prices.

However, we must note that these miscellaneous segments are artificially generated, but quite different from typical segments. Otherwise, the clustering algorithm does notwork efficiently.

(4) Finding segment and threshold value of F_i

As we explain in (2), clustering system obtain the records of K_i and F_i by using the sliding windows and the pool of individuals assigned to each clusters. But, simultaneously, as we describe in (3), we mix the miscellaneous segments different from the segments listed in the Figure 5. Therefore, for the time period when the sliding window covers the miscellaneous segments, the maximum fitness F_i will be limited to be a small value. Moreover, if the sliding window is placed on the boundary between two segments, the maximum fitness F_i will be also small.

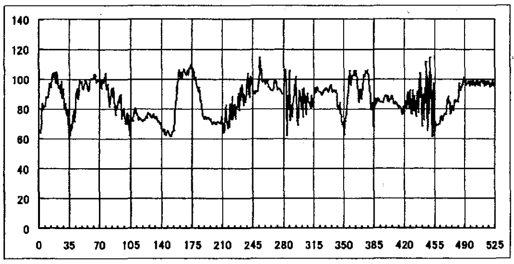

Figure 6: Example of generated time series

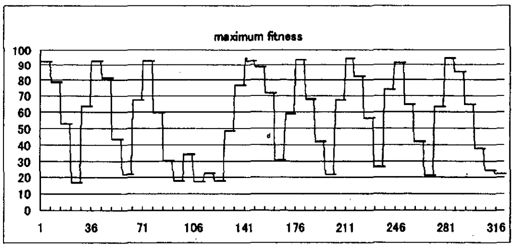

Figure 7: Example of maximum fitness

For these reasons, we define a threshold value F_s of F_i. We assume if the value of F_i is less than F_s, the sliding window is placed on the boundary between two segments, or the system recognize the miscellaneous segments different from the targets.

In the following, we summarize the simulation study forclustering the segments of stock price based on the artificially generated time series mentioned above. We select at random one of thesegments in Figure 5 and concatinate it with a miscellaneous segment. Then, we repeat the process until we obtain a sufficient length of time series to be examined. Figure 7 shows an example of a generated time series. Figure 8 shows the result of maximum fitness.

We show the result of clustering for the segments in Figure 5 as in Table 3 where the values mean the average rate of true classification for the segments. We define the false result of clustering as the cases where the cluster of the segment is not the same as the original segment, and the miscellaneoussegments are recognized as targeted segments. Moreover, the false estimation may happen when the target segments are missed to be detected by the system.

As is seen from Table 4, we have about 89% correct clustering by the system proposed in the paper.

6.2 Clustering of real stock price

Now, we apply the same algorithm for clustering for a long time series of stock price to clustering of real stock price by using the sliding windows as described in the-previous section. Since the conditions for the simulation is the same as in the previous section, we concisely summarize the result in

Table 4: Result of clustering

segmentNo.	p	segmentNo.	p
a	83	b	83
c	93	d	77
e	93	f	90
g	70	h	70

Table 5: Result of clustering

segmentNo.	p	segmentNo.	p
a	77	b	80
c	83	d	70
e	87	f	87
g	70	h	67

the following.

Different from the artificially generated time series, we have no evidence that the detected and estimated segments embedded in the time series are the true cluster. Therefore, we compare the result of clustering obtained by the system of the paper with the result of human subjects. The result is summarized in Table 5. The values in Table 5 mean the average rate of true clustering for each cluster. Namely, the false cases for clustering may happen for the misclassification, the wrong detection of segments and the lost of necessary recognition of segments.

As is seen from Table 5, the result is worse than the cases for the artificially generated time series in the previous section, but we have about 76% true clustering in the average, and the system is still seemed to be applicable for real applications.

7 Conclusion

This paper showed clustering of segments of stock prices by using modeling system for time series based on the GP. The pool of individuals improved by the GP for clustering a time series belongs to a certain cluster is useful to absorb the variation of the time series in clusters. The system was applied to clustering of artificially generated time series, and also to clustering of 8 kinds of segments of real stock prices. The problem to be solved is still remained in the configurati on of upper level of recognition system for longer time series in a syntactical manner, and the future works will be done by the authors.

Reference

[1] Chen.S.H. and C.H.Yeh.C.H., (2001), Evolving traders and the business school with genetic programming: A new architecture of the agent -based artificial stock market, Journal of Economic Dynamic and Control, vol.25, pp.363-393.

[2] Chen.X. and Tokinaga.S. (2002), Approximation of chaotic dynamics for input pricing at service facilities based on the GP and the control of chaos, IEICE Trans.Fundamentals, vol.E85-A, no.9, pp.2107-2117.

[3] Chen.X. and S.Tokinaga.S. (2003), Synthesis of multi-agent systems based on the co-evolutionary genetic Programming and its applications to the analysis of artificial markets, IEICE Trans.Fundamentals, vol.E86-A, no.10, pp.1038-1048.

[4] Chen.X. and S.Tokinaga.S. (2004), Multi-agent modeling of artificial stock markets by using the co-evolutionary GP approach, Journal of Operations Research Society of Japan, vol.47, no.3, pp.163-181.

[5] J.J.Fernandez.J.J, Farry.K.A, and Cheatham.J.B (1996), Waveform recognition using Genetic Programming: The myoelectric signal recognition problem in (eds J.Koza) Genetic Programming.

[6] Freitas.A.A (2002), Data Mining and Knowledge Discovery with Evolutionary Algorithms, Springer-Verlag, 2002.

[7] K.S.Fu.K.S (1974), Syntactical Method in Pattern Recognition, Academic Press.

[8] Ikeda.Y. (2002), Estimation of the chaotic ordinary differential equations by co-evolutionary genetic programming, IEICE Trans. Fundamentals, vol.E85-A, no.4, pp.424-433.

[9] Ikeda.Y and Tokinaga.S (2000), Approximation of chaotic dynamics by using smaller number of data based upon the genetic programming, IEICE Trans. Fundamentals, vol.E83-A, no.8, pp. 1599-1607.

[10] Ikeda.Y and Tokinaga.S (2001), Controlling the chaotic dynamics by using approximated system equations obtained by the genetic programming, IEICE Trans. Fundamentals, vol.E84-A, no.9, pp.2118-2127, 2001.

[11] Ikeda.Y. and Tokinaga.S. (2003),Synthesis of multi-agent systems based on the co-evolutionary genetic Programming by considering social learning and its applications to the analysis of artificial stock markets, Journal of Japan Society for Management Information, vol.12, no.3, pp.17-35.

[12] Ikeda.Y. and Tokinaga.S. (2004), Chaoticity and fractality analysis of an artificial stock market by the multi-agent systems based on the co-evolutionary Genetic Programming, IEICE Trans.Fundamentals, vol.E87-A, no.9, pp.2387-2394.

[13] J.Koza.J.R (1990), Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems, Report No.STAN-CS-90-1314, Dept.of.Computer Science Stanford University.

[14] Koza.J.R (1991), Evalution and subsumption using genetic programming, Proc of the First European Conference on Artificial Life, MIT Press.

[15] Koza.J.R (1992), Genetic Programming, MIT Press.

[16] Keith.M.J and Martin.M.C (1994), Genetic programming in C++: Implementation issues, in (ed) K.E.Kinnerar, Jr., Advance in Genetic Programming MIT Press.

[17] Lu.J and Tokinaga.S (2004), An Aggregated Approximation for Modeling of Time Series based on the Genetic Programming and its Application to Clustering,Special Interest Group of Operations Research Society of Japan "Modeling under Uncertainty".

[18] Tokinaga.S (1984), On the management of time-series data by feature description based upon generative grammar-By fitting ARMA model to EEG signals, Trans.IEICE, vol.J67-D, vol.10, pp.1099-1106.

[19] Tokinaga.S (1986), Automatic EEG classification besed on syntactical pattern recognition method-feature extraction by adaptive ARMA model fitting, Trans.IEICE, vol.E69, no.10, pp. 1125-1132.

[20] Tokinaga.S and Ishida.Y (1995), "An intelligent digital signal processing systems for stock trends based upon transioent wave detection by using Gabor representation and knowledge representation of waveform, IEICE Trans.Fundamentals, vol.J78-A, vol.2, pp.169-177.

[21] Tokinaga.S and Tominaga.M, (2003), An improvement method for the workflow management systems based on the reallocation of flows using the Genetic Programming and its applications, Journal of Operations Research Society of Japan, vol.46, no.3, pp.286-305.

[22] Trahanias.P and E.Skordalakis.E (1991), Syntactic pattern recognition of the ECG", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.12, no.7, pp.648-657.

[23] Tsai.W.H and Fu.K.S (1980), "Attributed grammar- a tool forcombining syntactical and statistical approaches topattern resognition", IEEE Transaction on Systems, Man and Cybernetics, vol.SMC-10, no.12, pp.873-885.

[24] Yababe.M and Tokinaga.S, (2002), Applying the Genetic Programming to modeling of diffusion processes by using the CNNand its applications to the synchronization, IEICE Trans.Fundamentals, vol.J85-A, no.5, pp.548-559.