PATTERN EXTRACTING ENGINE USING GENETIC ALGORITHMS

PATAYA DANGPRASERT, SOMSONG CHIRAPHADHANAKUL
Graduate School, Assumption University
Bangkok, Thailand.

Abstract: *Humans assess situations in terms of the pattern of circumstances that constitutes the situation, and they act in accordance with what that pattern might represent or what it might portend. Patterns which well-captured relation of important characteristics of observed experiences are good bases for forecasting future value. Likewise in classification and cluster forming, the well-defined patterns that extracted from obtainable data direct classifying process and cluster forming to dependable results. In which case, pattern is one major factor that has impact on the efficacy of forecasting, classification, and clustering methodologies.*

This paper reviews works of Ph.D. dissertations at Assumption University of which using GAs in pattern searching for forecasting, classification and cluster forming process. One of these researches, Genetic Forecasting algorithm [1], corroborated that the pattern extracting engine of which using GAs provided the better results than that using neural networking. Application of this work both to forecast commercial bank deposit task and to classify business for financial distress task prove the acclamation. Also there is an effort to improve the cluster forming results by introducing messy genetic algorithms (mGAs) in the place of simple GAs for pattern searching [2].

Keywords*: Genetic algorithms, pattern extracting, forecasting, classification,*

INTRODUCTION

Situations in this world generally cannot be assessed in terms of isolated facts or even in terms of a body of isolated facts. Rather, we find that we need to describe situations in terms of patterns of interrelated facts. In addition, humans assess situations in terms of the pattern of circumstances that constitutes the situation, and they act in accordance with what that pattern might represent or what it might portend. Information in pattern format seems to help human to adapt well to circumstances. Patterns which well-captured relation of important characteristics of observed experiences are good bases for forecasting future value. Likewise in classification and cluster forming, the well-defined patterns that extracted from obtainable data direct classifying process and cluster forming to dependable results. In which case, pattern is one major factor that has impact on the efficacy of forecasting, classification, and clustering methodologies.

The interrelation may be implicit in the sense that we know that all those facts pertain to the same object or situation. In other cases, a pattern may be meaningful only because of explicit relationships among the various features of the pattern [3]. Sequentially, the significance of a pattern can be grasped only if all aspects are available simultaneously for consideration. Thus various techniques have been put to use for capturing or extracting them.

Techniques that in use to extract those meaningful aspects for good pattern include conventional techniques using statistical approaches such as discrimination, multiple regression, correlation analysis, time series, and curve fitting. In the last decade, artificial intelligence concepts include knowledge engineering and expert systems, fuzzy logic, neural networks, and genetic algorithms are introduced to do the task.

This paper reviews works of Ph.D. dissertations at Assumption University of which using GAs in pattern searching for forecasting, classification and cluster forming process. One of these researches, Genetic Forecasting algorithm [1], corroborated that the pattern extracting engine of which using GAs provided the better results than that using neural networking. Application of this work both to forecast the bank deposit task and to classify business for financial distress task prove the acclamation. Also there is an effort to improve the cluster forming results by introducing messy genetic algorithms (mGAs) in the place of simple GAs for pattern searching [2].

PATTERN AND STRUCTURE IN DATA

A search for pattern in data is a search for structure of information that those data generate them. The data may be qualitative, quantitative or both. They may be numerical, pictorial, textural, linguistic, or any combination. More reliable forms of data analysis have evolved with modern technology. At present, statistical pattern recognition is a major technique in the search. The type of search performed depends not only on the data and the models, but upon the structure expect to find. The structure here is the way pattern can be recognized so that relationships between the variables in the process can be identified.

Representations of the pattern depend on the data, the method of search, and the model used. In terms of information, the data contain it, the search recognizes it, and the pattern represents it. Various techniques have been put to use for capturing or extracting pattern. Those techniques that in use to extract those meaningful aspects for good pattern include conventional techniques using statistical approaches such as discrimination, multiple regression, correlation analysis, time series, and curve fitting. In the last decade, artificial intelligence concepts include knowledge engineering and expert systems, fuzzy logic, neural networks, and genetic algorithms are introduced to do the task.

Recently, works of Ph.D. dissertations at Assumption University use GAs in pattern searching for forecasting, classification and cluster forming process. The following sections: Genetic forecasting algorithm with financial applications and Data mining algorithm by mGAs clustering technique are the brief of these valuable researches.

GENETIC FORECASTING ALGORITHM WITH FINANCIAL APPLICATIONS

A forecasting method using GAs was introduced in [1] to overcome the traditional forecasting techniques that entail some shortcomings and problems. Such problems including in multiple regression that more than one independent variables may have high correlation among themselves. Multicollinearity occurs when these variables interfere with each other and force variables with high correlation to be left out. This may create a marginal error since the missing components may retain significant factors in the forecasting process. In time series methods, a large data set is required to extrapolate the past behavior but the causal relationships among factors are not exposed. Thus, some important indicators may be left out. With GAs, search process can avoid the differentiation and can escape from the local minimum [4].

The genetic forecasting algorithm combines the powerful search of GAs and their capability to learn patterns relationship of past data in order to forecast future values. It is also capable of revealing the causal relationships among components. Conceptually, the algorithm is composed of two loops, the genetic forecasting loop and the pattern learning loop as in Fig. 1. Investigation is performed by computer simulations to forecast commercial banks deposit in Thailand and to predict bankruptcy for companies.

GENETIC FORECASTING ALGORITHM

Fig. 1 illustrates a simplified idea of the proposed genetic forecasting algorithm. Conceptually, the algorithm is composed of two loops: the genetic forecasting loop, and the pattern learning loop.

The genetic forecasting loop, Fig. 2, aims to minimize an error between the actual value and the forecasting value as well as to minimize the patterns error. The pattern learning loop, Fig. 1, serves as patterns relationship learning and adds more predictive power to the algorithm.

The genetic forecasting algorithm uses the following mathematical model to represent the pattern of the problem domain :

$$y(k+1) = b_0 + b_1x_1(k) + b_2x_2(k) + ... + b_nx_n(k) + \omega e(k), \quad (1)$$

where y is the output, x_1,....x_n are components of the inputs, e is an error, and k is a step.

At any instant step k, the input components $x_1(k)$, $x_2(k)$,...., and error, e(k), are used to calculate an output of the next step, $\hat{y}(k+1)$, according to the following equation:

$$\hat{y}(k+1) = \hat{b}_0+\hat{b}_1x_1(k)+\hat{b}_2x_2(k) + ... + \hat{b}_nx_n(k)+ \hat{\omega}e(k), \qquad (2)$$

where $\hat{b}_0$, $\hat{b}_1$,…. and $\hat{\omega}$, are parameters estimated by using the input data at step k–1 and the output data y at step k. The parameters are optimally selected by genetic forecasting loop from a set of chromosomes in the population.

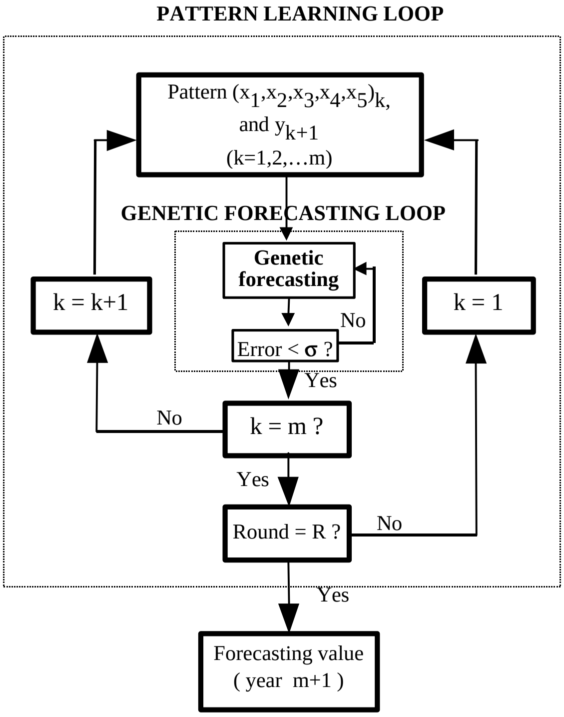

FIG. 1 : GENETIC FORECASTING ALGORITHM.

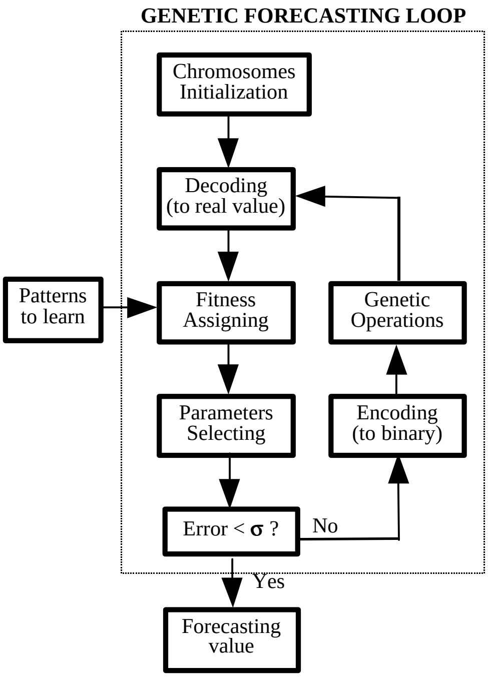

FIG. 2: GENETIC FORECASTING LOOP.

GENETIC FORECASTING LOOP

The genetic forecasting loop, the main forecasting function in the proposed algorithm, is illustrated in detail in Fig. 2. The aim of the genetic forecasting loop is to determine parameters b_0, b_1,... and ω, which are unknown. Genetic forecasting uses chromosomes to represent these parameters at step k. Each chromosome is composed of n+2 orders of parameters concatenating to form a string. A parameter is represented by ten bits. The first bit represents positive/negative sign, while the remaining nine bits represent the parameter value. To decode the nine binary bits to the real values, a decoding factor α is used. The range of parameter is limited by α. These decoded parameter values are parts of search spaces that GAs have to go around to find an optimum parameter. The genetic forecasting loop selects the fittest parameter set based on the selection mechanism (to be stated next) of which the fitness function is derived by taking into account the number of patterns in the pattern learning loop. The genetic forecasting loop is deemed satisfactory when the forecasting error is less than a predefined value σ.

PATTERN LEARNING LOOP

The pattern learning loop in the proposed algorithm enhances the forecast result. A data set, $(x_1,\dots x_n)_k$, and $y_{(k+1)}$, k=1,...m (m is a number of patterns to be learned), are sequentially presented to the genetic forecasting loop. To learn the patterns relationship between all x's at the step k, and y at the next step k+1, all patterns are recursively presented until R rounds which is predefined. At this point the optimal parameters are identified. The identified parameters and x's at step m, are then used to forecast the optimal output y at the step m+1, $\hat{y}_{opt}(m+1)$, by the following equation:

$$\hat{y}_{opt}(m+1) = \hat{b}_0+\hat{b}_1x_1(m)+\hat{b}_2x_2(m) +\dots+\hat{b}_nx_n(m) +\hat{\omega}e(m) \tag{3}$$

SELECTION MECHANISM

The selection mechanism evaluates the fitness function derived from two categories of errors: (i) the instant pattern k error, error[k], between the actual and the forecast values at any present pattern k:

$$\text{error}[k] = |\, y_{(Actual)} - \hat{y}_{(forecast)} \,|. \tag{4}$$

(ii) the total error of all patterns:

$$\text{error}[total] = \sum_{j=1}^{m} \text{error}[j] , \tag{5}$$

where m is the number of patterns for learning.

The selection mechanism selects optimal parameters at step k from a set of chromosomes in the population. It minimizes the following fitness function:

$$\text{fitness} = \text{error}[k] + \text{error}[total]. \tag{6}$$

The optimum identified parameters together with the inputs at step k are used to calculate the forecasting value at step k+1.

APPLICATION TO COMMERCIAL BANKS DEPOSIT FORECASTING

The genetic forecasting algorithm is introduced to forecast the commercial banks deposit. The simulations on computer uses five components of input data: $x_1,\dots.x_5$, i.e., gross domestic product, money supply (held by the public), interest rate, number of branches of commercial banks, and loans, respectively [5].

The investigation covers nineteen-year period of the input data from 1975 to 1993, and nineteen-year period of banks deposit data, from 1976 to 1994. As shown in Table 1, the data are divided into ten patterns of ten years, i.e., m=10, in each range for pattern learning, to forecast ten years of banks deposit, from 1986 to 1995. For example, the input data from 1975 to 1984 with banks deposit data from 1976 to 1985 are used for learning in order to forecast the banks deposit of 1986, the input data from 1976 to 1985 with banks deposit data from 1977 to 1986 are used for learning in order to forecast the banks deposit of 1987, etc.

The orders of model for study are set at n = 5, thus the number of parameters to be identified is 7. In estimating the model parameters, the initial parameters values are assigned in encoded chromosomes in binary values at random. A parameter is in the range of $-2^9\alpha$ to $+ 2^9\alpha$ where α is the decoding factor. Decoding factor is used to convert a binary-valued chromosome to a real-value parameter. Decoding factors for converting b_0, b_1,..,b_5, and ω are set to 20, 0.002, and 0.02, respectively. The GAs parameters for population size, crossover rate, mutation rate, number of crossing site, number of best individuals passed on to the next generation are set to 30, 1, 0.01, 1, and 2, respectively. The accepted absolute percentage error of forecasting σ is set at less than 0.01 %, and the number of rounds to learn patterns is set at R = 15. The instant pattern error and the total error of all patterns forming the fitness function are:

$$error[k] = | y_{(Actual)} - \hat{y}_{(forecast)} | / y_{(Actual)} \quad (7)$$

$$error[total] = \sum_{j=1}^{10} error[j] / 10. \quad (8)$$

The comparison of the forecast results of genetic forecasting algorithm and ANN are shown in Table 2 and Fig. 3, respectively. The architecture of ANN used in this paper consists of three layers, of which the input layer has five nodes, the hidden layer has seven nodes and the output layer has one node. The ANN is learned by cumulative delta technique. The findings indicate that the proposed algorithm provides closer results to the actual value than that of the ANN. In addition, the average error for ten-year forecasting of the proposed algorithm is also less than of ANN, 5.0723% and 8.2913%, respectively.

TABLE 1
DATA SERIES USED IN BANKS DEPOSIT FORECASTING

	Series estimating	For parameters	Series forecasting	For y
Forecast y at year	Patterns x at year	Patterns y at year	Estimated parameters year	Actual x year
86	75-84	76-85	85	85
87	76-85	77-86	86	86
88	77-86	78-87	87	87
89	78-87	79-88	88	88
90	79-88	80-89	89	89
91	80-89	81-90	90	90
92	81-90	82-91	91	91
93	82-91	83-92	92	92
94	83-92	84-93	93	93
95	84-93	85-94	94	94

TABLE 2
TEN YEAR FORECASTING RESULTS

Year	Actual	ANN		GAs	
	million baht	Value	|error| (%)	Value	|error| (%)
86	166,161.03	160,587.96	3.3540	159,204	4.187
87	194,154.16	171,727.34	11.5510	171,320	11.761
88	222,536.21	201,761.85	9.3353	216,940	2.515
89	267,907.18	235,297.41	12.1720	277,740	3.670
90	322,238.04	279,171.84	13.3647	324,961	0.845
91	370,007.70	351,413.78	5.0253	390,320	5.490
92	413,066.26	388,853.72	5.8617	415,388	0.562
93	476,546.35	433,124.71	9.1117	470,745	1.217
94	512,604.25	485,736.41	5.2414	578,776	12.909
95	572,766.93	527,544.91	7.8954	616,102	7.566
	Average	| % error |	8.2913		5.072

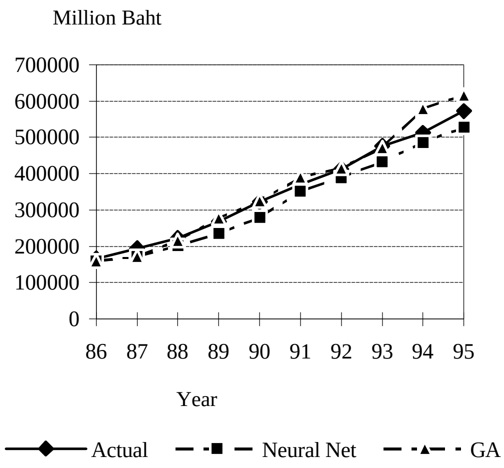

FIG. 3: ANN AND GAs BANKS DEPOSIT FORECASTING RESULTS.

APPLICATION TO BANKRUPTCY PREDICTION

The investigation for predictive ability of the genetic forecasting algorithm is performed by applying the algorithm to analyze the financial data for bankruptcy. In this paper the same training data set as in Odom and Sharda method [6] is used : a total of 74 firms, 38of which went bankrupt and 36 were nonbankrupt. In addition the other set of 55 firms: 27 bankrupt firms and 28 nonbankrupt firms, are used as testing data. The five input components : $x_1,\ldots x_n$, are working capital to total assets, retained earnings to total assets, earning before interest and taxes to total assets, market value of equity to total debt, and sales to total assets, respectively [7]. Decoding factors for b_0, $\{b_1,\ldots,b_5\}$, and ω are set to 0.5, {0.125}, and 0, respectively. The GAs parameters for population size, crossover rate, mutation rate, number of crossing site, number of best individuals passed on to the next generation are set to 30, 1, 0.02, 1, and 2, respectively. Furthermore, according to Altman's study in Z-score analysis for classifying bankrupt/nonbankrupt firms, companies with Z-score below 1.81 are always bankrupt, whereas Z-score above 2.99 are nonbankrupt. In this investigation, the criteria for classifying bankrupt/nonbankrupt in Altman's Z-score analysis is employed. The accepted criteria for prediction in genetic forecasting loop, σ, is set at less than 120. This value tells how far the predictive output is allowed to be different from the decision values (above 2.99 and below 1.81).

In bankruptcy prediction, the training data set is recursively presented to the genetic forecasting algorithm until all patterns are correctly classified. The fitness function used in bankruptcy prediction is as follow:

$$\text{fitness} = \sum_{j=1}^{m} \text{misclassify}_j \,, \qquad (9)$$

where m is the number of training data.

In bankruptcy prediction, the fitness function in Eq. (9) is simplified from Eq. (6), so that it contains only error[total]. This is because the evaluation of an instant pattern k as indicated in the first term of Eq. (6) is not meaningful for learning in the predictive domain.

As the results of training with a set of 74 firms data, genetic forecasting algorithm as well as four neural networks : Athena model, perceptron model, back-propagation model [8], and Odom and Sharda model [6] can correctly classify all training data, both nonbankrupt firms and bankrupt firms, while discriminant analysis [7] correctly classifies all nonbankrupt firms but only correctly classifies 33 of the 38 bankrupt firms. Table 3 shows the results of prediction for a set of 55 firms test data.

TABLE 3
BANKRUPTCY PREDICTION FOR TESTING DATA

NONBANKRUPT FIRMS										
WC/TA	RE/TA	EBIT/T	MVE/T	S/TA						
0.2234	0.3931	0.1168	1.1371	1.7523						
0.1725	0.3238	0.1040	0.8847	0.5576						
0.2955	0.1959	0.2245	1.1606	1.8478						
0.5542	0.4316	0.1065	0.8375	1.6678						
0.2489	0.4014	0.1669	1.4609	7.1658						
0.3813	0.3194	0.2044	2.8513	0.9851						
0.4512	0.4114	0.1146	1.7185	1.5543						
0.1904	0.2011	0.1329	0.5586	1.6623						
0.5248	0.6437	0.2478	6.3501	1.2542						
0.4058	0.4497	0.1497	1.1076	1.7428						
0.2947	0.3724	0.1104	0.9410	1.3568						
0.4327	0.6494	0.2996	8.2982	1.2865						
0.1630	0.3555	0.0110	0.3730	2.8307		a			os	
0.5189	0.3627	0.1015	0.9764	0.7466						
0.4792	0.3495	0.1076	0.8105	1.7224						
0.0669	0.2904	0.0978	0.7659	4.3912						
0.3449	0.1270	-0.0083	0.1059	0.8611	d	a	p	b	os	ga
0.0272	0.0503	0.0184	0.1413	1.2008	d	a	p	b	os	ga
0.6302	0.3324	0.1524	1.1259	1.5579						
0.2170	0.2507	0.0826	0.3404	1.9889						
0.4078	0.1316	0.1095	0.3233	1.8226			p	b	os	ga
0.2864	0.2823	0.1856	2.7709	2.7730						
0.1841	0.3344	0.0857	2.1230	2.1686						
0.0732	0.3526	0.0587	0.2349	1.7432						
0.0106	0.0200	0.0226	1.8870	1.2274	d	a	p	b	os	ga
0.5398	0.1723	0.2019	34.503	1.1388						
0.3750	0.3326	0.1290	0.9487	1.2529						
0.2921	0.2390	0.0673	0.3402	0.7596						
			Misclassified		3	4	4	4	5	4

BANKRUPT FIRMS										
WC/TA	RE/TA	EBIT/T	MVE/T	S/TA						
0.0471	0.1506	-0.0150	0.1039	0.6253						
0.2770	-0.0417	0.0904	0.5245	1.9380						
0.4958	0.2199	0.0219	0.1267	3.0305						
0.1070	0.0787	0.0433	0.1083	1.2051						
0.1936	0.0778	-0.1830	0.6531	2.4263						
0.1611	0.0954	0.0307	0.2113	1.4529						
0.3732	0.3484	-0.0139	0.3483	1.8223	d			b		
0.2653	0.2683	0.0235	0.5118	1.8350	d			b		
-0.1599	-0.5018	-0.0889	0.1748	2.1608						
0.1123	0.2288	0.0100	0.1884	2.7186						
0.3696	0.2917	0.0621	0.5554	1.7326	d	a	p	b	os	ga
0.2702	0.1402	0.1668	0.2717	2.1121	d	a	p		os	ga
0.1144	-0.0194	0.0074	0.2940	1.5734						
0.4044	-0.1878	0.0768	0.2846	1.3489						
0.2787	0.1767	0.0305	0.1797	5.3003						
-0.0357	-0.9814	-0.0031	0.3291	2.1088						
-0.0179	-0.2902	0.0984	2.2848	2.1803						
0.4067	0.2972	0.0454	0.5001	2.0631	d	a	p	b		
0.2260	0.1620	0.0965	0.2737	1.9199	d					
0.0780	-0.2451	0.0627	0.0453	0.1451						
0.3422	0.2865	0.0778	0.5300	1.5564	d	a	p	b	os	ga
0.3440	0.1725	0.1386	0.2775	2.0030	d	a	p		os	ga
0.1756	0.1233	0.1046	0.7468	1.6774	d					
0.1186	0.1849	-0.0718	0.2117	0.1376						
0.3617	0.1312	0.0413	0.3706	2.1890						
0.1162	0.3026	0.0863	0.9220	0.9513	d	a	p	b	os	ga
0.2323	0.1095	0.1054	0.4661	0.9193	d					
			Misclassified		11	6	6	6	5	5
			Total		14	10	10	10	10	9

d : Misclassified by Discriminant Analysis
a : Misclassified by Athena Model
p : Misclassified by Perceptron Model
b : Misclassified by Back-propagation Model
os : Misclassified by Odom and Sharda Model
ga : Misclassified by Genetic Forecasting Algorithm

This indicates that genetic forecasting algorithm has better ability to classify bankrupt/ nonbankrupt firms from financial data. Genetic forecasting algorithm misclassifies 9 firms (4 nonbankrupt and 5 bankrupt) of 55 test data while 4 neural network methods misclassify 10 firms and discriminant analysis misclassifies 14 firms.

DATA MINING ALGORITHM BY MESSY GAs CLUSTERING TECHNIQUE

There is a technical report of [9] using data mining algorithms by applying simple GAs to evolve the interested clusters. Later they cited the linkage problem of simple GAs (sGAs) and replaced the engine with messy genetic algorithms (mGAs). The technique was covered in the Data Mining Algorithm by Messy GAs Clustering Technique (DMAmGAs) [2]. The mGAs with variable length chromosome and cut and splice operators has been reported to be more efficient than the former one. The operation of mGAs divide into a primordial phase and a juxtapositional phase. The primordial phase uses local search template and threshold selection operator to prepare a good building block population. The juxtaposional phase using cut and splice operator searches for an optimal solution. The mGAs form tight building blocks and combines them to an optimal solution in a way that respects to the schema theorem. The way mGAs represent their chromosome makes good building block less susceptible to be destroyed than that of sGAs.

The conceptual approach of the DMAmGAs is illustrated in Fig. 3. It composes of the data pruning interface process and the cluster forming process. The DMAmGAs acquires data from data warehouse by mean of data pruning interface. This process filters the information from database. It is the cluster forming process that mGAs is used as pattern extracting engine. The process extracts pattern from data in the warehouse and forms the cluster one by one until all interesting clusters are discovered.

DMAmGAs is tested by computer simulation using the higher education planning for finding suitable location for new campus as problem domain. One of the main objectives is to expand the academic institutes to the rural areas. The actual data of a university database in Thailand and geographical database are used to provide training data. The major factors in use for considerations are (i) the number of prospected students in the area, (ii) the distance from the student's residences to the proposed campus, and (iii) the faculties to be announced. The DMAmGAs extracts the interesting clusters that suggest the centroid province for new campus and also names the faculties to be announced. The new campus must have at least three faculties with the minimum of 300 students.

The structure of chromosome is represented such that it compiles with the database structure of interest and the structure of objective function. It represents the location of the proposed campus and the geographical range of which this campus can serve.

The simulation results show that with mGAs as pattern extracting engine, it is faster to extract each clusters than that using sGAs in term of number of generations as seen in table 4. Both methodologies give the same result in term of suggested location, name and number of provinces that the proposed campus can serve and number of faculties to be announced on each campus.

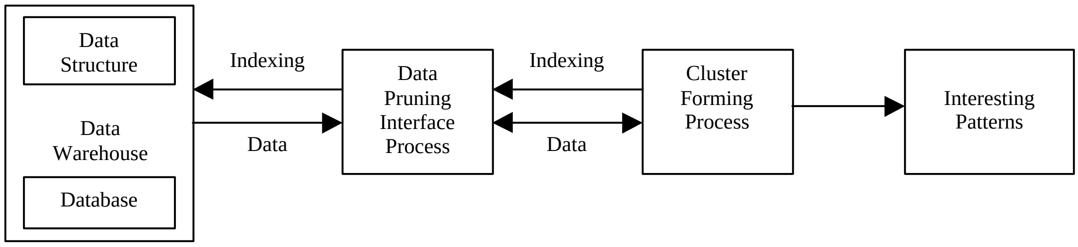

FIG. 3. THE CONCEPTUAL APPROACH OF DMAmGAs.

TABLE 4
THE NUMBER OF GENERATION USE FOR CONVERGENCE.

mGAs (Juxtapositional phase)	sGAs
2	9
3	6
5	9
4	12
2	8

CONCLUSIONS

The paper reviews the use of genetic algorithm as pattern extracting engine. Both sGAs and mGAs have been reported to be effective for the purpose. The forecasting algorithm with sGAs corroborated that the pattern extracting engine of which using GAs provided the better results than that using neural networking. Application of this work both to forecast the bank deposit task and to classify business for financial distress task prove the acclamation. Also there is an effort to improve the cluster forming results by introducing messy genetic algorithms (mGAs) in the place of simple GAs for pattern searching [2]. The results with lesser number of generations using in DMAmGAs suggests that the pre-process of chromosome in primordial phase, the variable length and the position independent representation of chromosome in mGAs render solution to the linkage problem of sGAs.

In mGAs the position independent chromosome make the building block harder to be destroyed by cut and splice operator. While the fix position chromosome with crossover operator in sGAs, the good building block is more susceptible to the destruction.

Patterns which well-captured relation of important characteristics of observed experiences are good bases for forecasting future value. Likewise in classification and cluster forming, the well-defined patterns that extracted from obtainable data direct classifying process and cluster forming to dependable results.

REFERENCES

[1] S. Chiraphadhanakul, *Genetic Forecasting Algorithm*, a doctoral dissertation, Assumption University, 1997.

[2] C. Intrasai, and V. Avatchanakorn, *Data Mining Algorithm by mGAs Clustering Technique*, A technical report, Assumption University, 1999.

[3] Y. Pao, *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, 1989.

[4] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.

[5] S. Chiraphadhanakul, *A Study on Factors effecting Changes in Commercial Banks Deposit in Thailand*, Thesis of Master of Accountancy, Department of Banking and Finance, Chulalongkorn University, Thailand, 1986.

[6] M. D. Odom, and R. Sharda, A Neural Network Model for Bankruptcy Prediction, *Neural Networks In Finance and Investing*, Heinemann Asia, 1994, pp.177-185.

[7] E. I. Altman, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, vol.23, September 1968, pp.589-609.

[8] E. Rahimian, S. Singh, T. Thammachote, and R. Virmani, Bankruptcy Prediction by Neural Network, *Neural Networks in Finance and Investing*, Heinemann Asia, 1994, pp.159-176.

[9] C. Intrasai, and V. Avatchanakorn, *Genetic Data Mining Algorithm with Academic Planning Application*, in *Proceeding 1998 IASTED International Conference on Applied Modeling and Simulation*, Alberta, Canada, paper no. 286-129.