Stock Price Forecasting under Fuzzy Value Clustering Analysis

Jianfeng Li, Jun Zhai and Junfeng Guo

Management Science and Engineering, Transportation Management College, Dalian Maritime University
Dalian city, PR. China
Jianfeng_vvv@yahoo.com.cn

Abstract

A new approach is proposed according to values analysis to forecast the stock price through fuzzy clustering method. Firstly, some common hypothesis are brought forward in order to establish a frame of value analysis for stock forecasting, then, the algorithm of fuzzy clustering is introduced to judge the relative value of stocks as the rule of forecasting stock price under value analysis, and at last, a case choosing ten telecommunication companies is analyzed to prove the key points of view of this paper. The results of paper are useful for further research on value investment and wise decision-making in practice.

Keywords: *Stock forecasting; Fuzzy clustering method; Value mining*

1. Introduction

Stock investment is one of most popular forms in the financial market due to its high-expected profit, however, it‘s also followed by higher risk [1]. There are many factors affecting the performance of stock market, such as economic condition, operation of the enterprise and expectation of traders, therefore, mining the trend of stock and excavating the profitable object is a challenging task on account of stock’s high volatility and noisy environment [2].

Different methods have been proposed in numerous studies for wise decision-making, most of which are about stock price prediction. For example, Majhi, *et al.*, proposed a trigonometric functional link artificial neural network (FLANN) model for short (one-day) as well as long term (one month, two months) prediction of stock price of leading stock market indices: DJIA and S&P 500[3]; Lee proposed a prediction model based on a hybrid feature selection method and support vector machine (SVM) to predict the trend of stock market [4]; Slim applied stochastic neural networks in forecasting the volatility of index returns for TUNINDEX (Tunisian Stock Index), and found that the out-of sample neural network results were superior to traditional GARCH models [5]; Yudong and Lenan proposed an improved bacterial chemotaxis optimization (IBCO) which was integrated into BPN to develop an efficient forecasting model for prediction of various stock indices [6]; Atsalakis George and Valavanis Kimon proposed a hybrid model that linked two Adaptive Neuro-Fuzzy Inference System(ANFIS) controllers to forecast next day’s stock price trends of the Athens and the New York Stock Exchange (NYSE) [7];

In the current research, a large number of scholars mainly adopt some quantitative analysis methods, such as artificial neural network (ANN) and time serious analysis, to match the actual curve of stock price, and then, to mine the valuable stocks with ascending price for investment. However, there are some ignorable things to some extent, which are needed to emphasize for further research:

To begin with, interaction effect of similar stocks is overlooked. In many papers, only one stock is probed into by fitting the trend of this stock's price under many factors, such as economic policy and enterprise operation, but interaction between similar stocks is little paid attention to. Sometimes one stock's price fluctuates with others in the similar stock section, but in other times, it doesn't. That interaction phenomenon is very important for the stock investment.

In addition, to match the price curve is much difficult and not very indispensable. Fluctuation of stock price is very changeable and sometimes stochastic, which is affected by many factors, so it is difficult to match the price curve and predict the exact price next time "in reality", although some non-liner mathematic methods can be used to reflect true fluctuation relatively well on some historic data. What's more, it not very indispensable to forecast the exact price, because knowing the changing direction is enough to some extent for investment.

Last but not least, value analysis of stock is relatively few. The price of stock is always changeable with its real value, which is the most important on the stock investment. However, the current research about stock price prediction is largely based on the price scalar curve, but not on true value of the stock. Maybe it's the real reason that the value of stock ought to be judged with others, which is hard to calculate the precise number.

Therefore, according to the above reasons, a new kind of approach is proposed, which is based on fuzzy clustering method to evaluate the stock through mining the relative value. The rest of the paper is organized as follows. In Section 2, the frame of relative value mining is brought forward; in Section 3, the algorithm of fuzzy clustering is discussed; and then, in Section 4, the actual case is analyzed. At last, in Section 5, some valuable conclusions are presented finally.

2. The frame of value analysis for stock forecasting

A frame for stock investment is brought forward, which is in order to forecasting stock price by mining the stock's relative value, and then, to provider a good decision making on investment. Some common hypothesis is given below:

H1: price of stock is changeable with its true value or maybe expected value.

The external stock price embodies its internal value. No mater how changeable the stock price is, it always fluctuates around its true value or maybe expected value. When the investor considers stock price is lower than its value, he would buy the stock, and otherwise, he would sell it. Sometimes that value is not accepted by the many persons, but expected only by the investor himself. Therefore, judging the stock value is very important on the decision-making for investment.

H2: value of stock is actually reflected on its company's financial index in large degree.

As a matter of fact, stock purchase means to own some part of that company, in that way, the value of stock actually is the judgment for the investor on the operation of company at present or in the future. Maybe someone argue there are also many factors, such as political atmosphere and economic environment, to affect investors to purchase the stock. However, in many scholars' opinion, those exterior factors in fact affect investors' approval on company operation, which is finally reflected on the expected change of financial index for that company.

H3: relative value of stock can be shown clearly by contrasting the similar stocks.

If value of stock can be reflected on its company's financial index (real or expectant), values of stocks are nearly equal to each other whose financial indexes are similar. In that way, although the absolute value of stock is difficult to calculated, the relative value of stock can be shown clearly and easily by contrasting the similar stocks, especial in the same industrial section. The big or small of stock values can be found out by comparing their financial index. Even in some special time, value of stock goes closely to others if its real or expectant indexes are changed.

According to the above hypothesis, an analytical frame of value analysis for stock forecasting is proposed, which is shown below (Figure 1), through that frame, the stock price is predicted roughly based on relative stock value, and it is enough for the investor to make a wise decision.

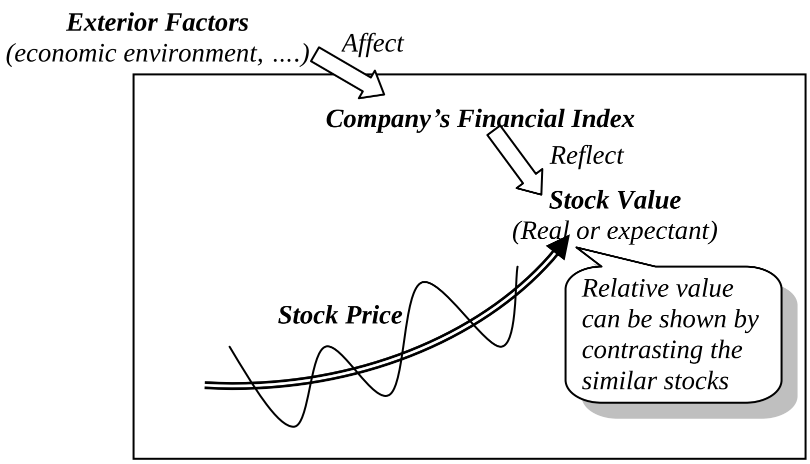

Figure 1. The Analytical Frame Under

In that analytical frame, there are four hinges:

- Exterior factors, such as political policy, economic environment and civil characteristics, affect the company's financial index.
- Company's financial index reflects the value of stock that is real at present or expectant in the future.
- Relative value of the stock can be shown by contrasting the company's financial indexes of the similar stocks.
- The price of stock always fluctuates with the value of stock, so knowing the value means predicting the price.

3. The algorithm of fuzzy clustering

Clustering is such a procedure that objects are distinguished or classified in accordance with their similarity. A formal mathematical definition of clustering is in the follow [8, 9]: let $X \in R^{n \times m}$ a set of data items representing a set of *n* points x_i in R^m. The goal is to partition X into K groups C_k such every data that belong to the same group are more "alike" than data in different groups. Each of the K groups is called a cluster. The result of the algorithm is an injective mapping $X \mapsto C$ of data items X_i to clusters C_k

As one of important clustering approaches, fuzzy clustering analysis which obtains the uncertainty degree of samples belonging to each class and expresses the intermediate property of their memberships, can trace back to the concept of fuzzy partition which is proposed by Ruspini [10, 11]. With this concept, some typical fuzzy clustering algorithms, such as methods based on the similarity and fuzzy relations [12], the transitive closure of fuzzy equivalent relation [13], the convex decomposition of data [14], or the dynamic programming and indistinguishable relation are developed one after the other by the scholars [15]. It has been applied in many aspects:

For example, in the educational aspect, students were allocated into some number of classes by means of fuzzy clustering algorithm based on each student's achievement of prerequisite subjects (Susanto, 2002) [16]; in the agricultural aspect, the soil samples were classified on base of the concentration of thirteen chemical elements through Gustafson-Kessel fuzzy clustering algorithm (Costel, 2007) [17], and in the documental aspect, similar documents were found through a kind of fuzzy clustering approach with the predefined fuzzy clusters being used to find extract feature vectors of the related documents (Ridvan, 2007) [18], and so on.

Because of advance in mining the similarity, fuzzy clustering method is applied to contrast the financial indexes of the similar stocks here. The basic steps of the algorithm are shown below (Figure 2).

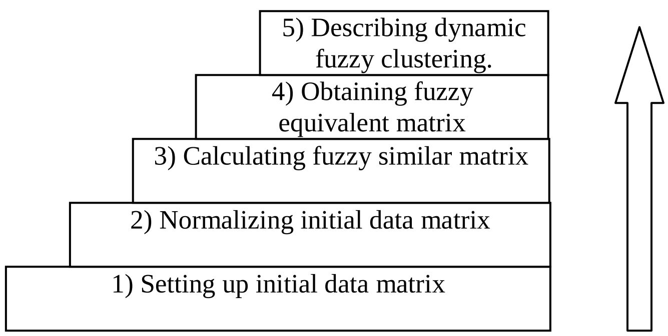

Figure 2. The Basic Steps of the Fuzzy

When dynamic fuzzy clustering graph is gained (step 5), value clustering of stocks is achieved too, i.e. the values of stock are similar in one cluster, what's more, their price are also similar according to the rule that the value determine the price. In that way, if a stock is attributed to a cluster, which price is away from the others' normal prices, its price will be near to that one. It can be called as individual effect (Ieff), which is just exterior expression of the difference of relative value when some financial indexes change (or expected). In addition, the whole stocks in one cluster have an overall trend, which can be called as overall effect (Oeff). The change of the stock price (Cprice) is under the impaction of those two effects both are positive or negative, which is shown in formula 1.

$$Cprice = Ieff + Oeff \qquad (1)$$

Thus, good chances of investment can be realized through above algorithm, especial when some companies' financial indexes change, which is fit for long-term investment. Even in the moment of whole stock market being week, some investing chances may be grasped, when Leff go beyond Oeff. What's more, some large loss also can be avoided through that algorithm.

4. The Case Analysis

Ten telecommunication companies are selected, which are the listed companies in the stock market of China. Fuzzy clustering method is applied to those companies according to nine general financial indexes (attributes). They are shown below:

Company:
X_1: Suzhou New Sea Union Telecom Technology Co.,Ltd;
X_2: Nanjing Panda Electronics Company Limited;
X_3: Hengtong Optic-electric Co.,LTD;
X_4: Beijing Electronic Zone Investment and Development Co.,Ltd;
X_5: Fiberhome Telecommunication Technologies Co.,Ltd;
X_6: Besttone Holding Co.,Ltd; X_7: Jiangsu Zhongtian Technology Co.,Ltd;
X_8: Shenzhen SDG Information Co.,Ltd;
X_9: Eastern Communication Co.,Ltd;
X_{10}: Kaile Science and Technology Co.,Ltd.Hubei

Financial indexes:
A_1:Net assets per share;
A_2:Retained earnings per share;
A_3:Accumulation fund per share;
A_4:Gross profit rate of sale;
A_5:Operating profit ratio;
A_6:Net profit ratio;
A_7:Weighted return on equity;
A_8:Diluted return on equity;
A_9:Current ratio

The initial partition matrix $\{X_{ij} \mid i=1,\cdots,m;\ j=1,\cdots,n\}$, m=10, n=9, is constructed (Time: 2012, Jan to Sep), which is shown below:

$$X = \begin{pmatrix} 2.46 & 1.0795 & 0.2465 & 35.5 & 14.17 & 14.09 & 8.71 & 8.53 & 1.38 \\ 2.47 & 0.443 & 0.7137 & 10.75 & 3.78 & 3.96 & 4.15 & 4.11 & 1.2 \\ 11.55 & 5.5509 & 4.3617 & 20.68 & 4.61 & 4.79 & 11.53 & 11.23 & 1.06 \\ 3.46 & 1.1447 & 1.2778 & 63.88 & 47.12 & 35.58 & 6.39 & 6.47 & 1.77 \\ 10.68 & 2.6978 & 6.6023 & 24.38 & 4.34 & 6.46 & 8.76 & 7.41 & 1.83 \\ 4.447 & 0.3961 & 2.5415 & 30.67 & 11.07 & 8.7 & 5.47 & 5.36 & 2.89 \\ 6.17 & 2.0718 & 2.9026 & 21.27 & 8.47 & 6.95 & 6.71 & 6.69 & 1.83 \\ 3.3541 & 0.3613 & 1.9969 & 15.96 & 4.03 & 3.35 & 4.82 & 4.7 & 1.61 \\ 2.0273 & 0.292 & 0.6936 & 12.05 & 3.78 & 4.89 & 3.69 & 3.68 & 2.9 \\ 3.219 & 1.1766 & 0.7312 & 31.34 & 15.59 & 10.17 & 12.84 & 12.07 & 1.54 \end{pmatrix}$$

These initial data should be converted into the same dimension for comparison, and compressed into the range [0,1]. Formula 1 is mainly followed below (here, attribute 9 current ratio must be firstly transferred to the "absolute distance from 2", then normalized):

$$x_{ij}^{'} = \frac{x_{ij} - \min\limits_{1 \le i \le m}\left\{x_{ij}\right\}}{\max\limits_{1 \le i \le m}\left\{x_{ij}\right\} - \min\limits_{1 \le i \le m}\left\{x_{ij}\right\}} \; (j = 1, 2, \cdots, n) \qquad (2)$$

$$X^{'} = \begin{pmatrix} 0.0454 & 0.1497 & 0 & 0.4658 & 0.2397 & 0.3332 & 0.5486 & 0.5781 & 0.4156 \\ 0.0465 & 0.0287 & 0.0735 & 0 & 0 & 0.0189 & 0.0503 & 0.0513 & 0.1818 \\ 1 & 1 & 0.6475 & 0.1869 & 0.0192 & 0.0447 & 0.8568 & 0.8999 & 0 \\ 0.1505 & 0.1621 & 0.1623 & 1 & 1 & 1 & 0.2951 & 0.3325 & 0.9221 \\ 0.9086 & 0.4575 & 1 & 0.2565 & 0.0129 & 0.0965 & 0.5541 & 0.4446 & 1 \\ 0.2541 & 0.0198 & 0.3611 & 0.3749 & 0.1682 & 0.166 & 0.1945 & 0.2002 & 0.0649 \\ 0.435 & 0.3384 & 0.4179 & 0.198 & 0.1082 & 0.1117 & 0.3301 & 0.3588 & 1 \\ 0.1393 & 0.0132 & 0.2754 & 0.0981 & 0.0058 & 0 & 0.1235 & 0.1216 & 0.7143 \\ 0 & 0 & 0.0703 & 0.0245 & 0 & 0.0478 & 0 & 0 & 0.0519 \\ 0.1251 & 0.1682 & 0.0763 & 0.3875 & 0.2725 & 0.2116 & 1 & 1 & 0.6234 \end{pmatrix}$$

In order to cluster the companies, the similarity of vectors in the sample is needed. In this way, fuzzy similar matrix R can be gotten. The formula 3 (weighted *Minkowski metric)* is followed, where W= $\begin{pmatrix}0.2 & 0.2 & 0.15 & 0.05 & 0.05 & 0.1 & 0.1 & 0.1 & 0.05\end{pmatrix}$ and $p = 4$:

$$r_{ij} = \begin{cases} 1 & \text{i = j} \\ 1 - \left\{ \sum\limits_{k=1}^{m} \left| x_{ik}^{'} - x_{jk}^{'} \right|^{p} \times W \right\}^{\frac{1}{p}} & \text{i} \ne \text{j} \end{cases} \qquad (3)$$

$$r = \begin{pmatrix} 1 & 0.64 & 0.26 & 0.54 & 0.28 & 0.71 & 0.64 & 0.67 & 0.61 & 0.71 \\ 0.64 & 1 & 0.18 & 0.32 & 0.29 & 0.77 & 0.58 & 0.74 & 0.94 & 0.36 \\ 0.26 & 0.18 & 1 & 0.2 & 0.48 & 0.27 & 0.4 & 0.21 & 0.14 & 0.3 \\ 0.54 & 0.32 & 0.2 & 1 & 0.28 & 0.43 & 0.42 & 0.35 & 0.32 & 0.43 \\ 0.28 & 0.29 & 0.48 & 0.28 & 1 & 0.42 & 0.59 & 0.41 & 0.25 & 0.33 \\ 0.71 & 0.77 & 0.27 & 0.43 & 0.42 & 1 & 0.55 & 0.69 & 0.77 & 0.45 \\ 0.64 & 0.58 & 0.4 & 0.42 & 0.59 & 0.55 & 1 & 0.74 & 0.52 & 0.55 \\ 0.67 & 0.74 & 0.21 & 0.35 & 0.41 & 0.69 & 0.74 & 1 & 0.68 & 0.41 \\ 0.61 & 0.94 & 0.14 & 0.32 & 0.25 & 0.77 & 0.52 & 0.68 & 1 & 0.33 \\ 0.71 & 0.36 & 0.3 & 0.43 & 0.33 & 0.45 & 0.55 & 0.41 & 0.33 & 1 \end{pmatrix}$$

According to the fuzzy clustering approach, the matrix must have three qualities: “reflexivity”, “symmetry” and “transitivity”, which isn’t all owed by fuzzy similar matrix, however, through transitive closure, Fuzzy similar matrix R can be transformed into fuzzy equivalent matrix, which not only have “reflexivity” and “symmetry”, but also is provided with “transitivity”.

The transitive closure $t(R)$ is gained by means of the “square method” that fuzzy similar matrix is squared gradually.

$$R \to R^2 \to R^4 \to \cdots \to R^{2^i} \to \cdots \tag{4}$$

$$\left(R^2 = R \circ R = \bigvee_{k=1}^{n} (r_{ik} \wedge r_{kj}) \right)$$

When $R^k \circ R^k = R^k$, it means that R^k is provided with "transitivity", and R^k is just requisite transitive closure, and which is also the "optimal" fuzzy equivalent matrix R^*.

$$r = \begin{pmatrix}
1 & 0.64 & 0.26 & 0.54 & 0.28 & 0.71 & 0.64 & 0.67 & 0.61 & 0.71 \\
0.64 & 1 & 0.18 & 0.32 & 0.29 & 0.77 & 0.58 & 0.74 & 0.94 & 0.36 \\
0.26 & 0.18 & 1 & 0.2 & 0.48 & 0.27 & 0.4 & 0.21 & 0.14 & 0.3 \\
0.54 & 0.32 & 0.2 & 1 & 0.28 & 0.43 & 0.42 & 0.35 & 0.32 & 0.43 \\
0.28 & 0.29 & 0.48 & 0.28 & 1 & 0.42 & 0.59 & 0.41 & 0.25 & 0.33 \\
0.71 & 0.77 & 0.27 & 0.43 & 0.42 & 1 & 0.55 & 0.69 & 0.77 & 0.45 \\
0.64 & 0.58 & 0.4 & 0.42 & 0.59 & 0.55 & 1 & 0.74 & 0.52 & 0.55 \\
0.67 & 0.74 & 0.21 & 0.35 & 0.41 & 0.69 & 0.74 & 1 & 0.68 & 0.41 \\
0.61 & 0.94 & 0.14 & 0.32 & 0.25 & 0.77 & 0.52 & 0.68 & 1 & 0.33 \\
0.71 & 0.36 & 0.3 & 0.43 & 0.33 & 0.45 & 0.55 & 0.41 & 0.33 & 1
\end{pmatrix}$$

Setting the threshold to λ, which changes from the big to the small, the dynamic fuzzy clustering can result from the above optimal fuzzy equivalent matrix R^*.

$$\text{Supposing } \underset{\sim}{C} \in P(U), \text{ for } \forall \lambda \in [0,1],\ (\underset{\sim}{C})_\lambda = C_i \triangleq \{x \mid \underset{\sim}{C}(x) \geq \lambda\}, \tag{5}$$

The last, with the different threshold λ being set on, the dynamic fuzzy clustering for the telecommunication companies based on finance condition is shown according to the formula 5. According to the above analysis, the dynamic fuzzy clustering map can be gained, which is shown in Fig3.

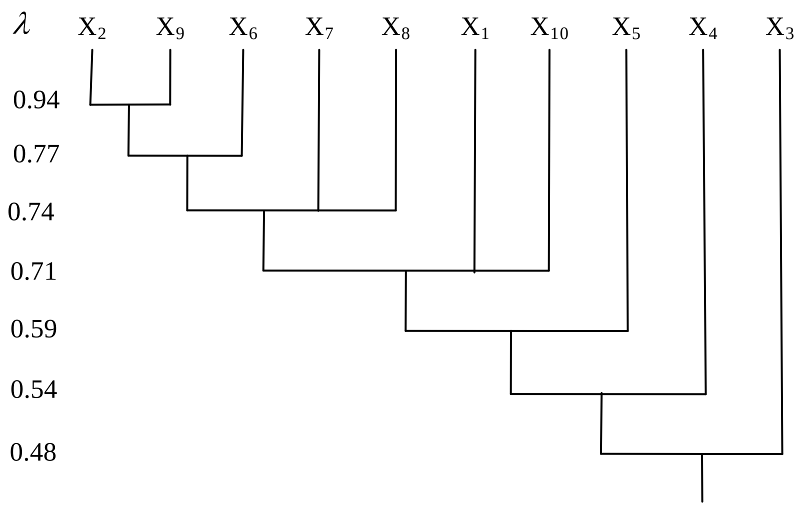

Figure 3. Dynamic Fuzzy Clustering Map

Here, the price data in September 3, 2012 is analyzed, which is just near to the time that the financial reports of companies are shown publicly, most companies' prices fluctuation is similar with the dynamic fuzzy clustering map. For example, at first, when $\lambda = 0.94$, the companies X_2 and X_9 are clustered firstly according to the financial indexes, and their stock prices(￥5.62 and ￥4.82) are also the lowest and similar to each other, in addition, when $\lambda = 0.74$, companies X_6, X_7 and X_8 are clustered into that class again, their prices are very near(￥8.19, ￥8.5 and ￥7.09), and then, when $\lambda = 0.71$, companies X_1 and X_{10} are clustered and stock price of company X_1 (￥7.00) are similar to the prices of companies X_6, X_7 and X_8, but stock price of company X_{10} (￥12.4) is not. In this way, companies X_5, X_4 and X_3 are clustered step by step, whose prices are ￥22.5, ￥6.57 and ￥20.18.

Here, it can be seen that stock prices of these companies are clustered into each other, which is fit for the dynamic fuzzy clustering map. Those companies except companies X_{10} and X_4 follow that rule, and the rate of accuracy is 80%. It is proven that the point in this paper is basically right.

What's more, although some companies such as X_{10} and X_4 are not totally fit for that discipline, but there are some changes to explain them very well. The data of stock prices of September 3, 2012 and February 22, 2013 are chosen. Because the overall effects (Oeff) are the almost unchanged in that two time points, the stock prices are mainly affected by their relative values, *i.e.*, individual effect (Ieff). For the company X_{10}, its stock price decrease from ￥12.4 to ￥9.99(Figure 4_A), which is near to the prices of other companies in the cluster when $\lambda = 0.71$; In addition, for the company X_4, because its total amount of stock is the "largest", it may affect its stock price, which is normally very low, but its price is going up from ￥6.57 to ￥8.85(Figure 4_B). In that way, the risk can be avoided and the chance can be grasped.

(A)

(B)

Figure 4. The Change of Stock Prices on Company X_{10} and X_4

5. Conclusion

This paper proposes a new approach based on fuzzy clustering method to evaluate the stock through mining the relative value. The frame of relative value mining is established, the algorithm of fuzzy clustering is discussed, and at last, a case is analyzed to prove the key points. The results of paper are useful for further research on value investment and wise decision-making in practice.

Acknowledgments

The authors would like to thank peer reviewers for commenting this article. This work is supported by the National Natural Science Foundation of China (70801007).

References

[1] S. -H. Liao, H. -h. Ho and H. -w. Lin, "Mining stock category association and cluster on Taiwan stock market", Expert Systems with Applications, vol. 35, **(2008)**, pp. 19-29.
[2] R. K. Lai, C.-Y. Fan, W. -H. Huang and P. -C. Chang, "Evolving and clustering fuzzy decision tree for financial time series data forecasting", Expert Systems with Applications, vol. 36, **(2009)**, pp. 3761-3773.
[3] R. Majhi, G. Panda and G. Sahoo, "Development and performance evaluation of FLANN based model for forecasting of stock markets", Expert Syst. Appl., vol. 36, **(2009)**, pp. 6800-6808.
[4] M. -C. Lee, "Using support vector machine with a hybrid feature selection method to the stock trend prediction", Expert Systems with Applications, vol. 36, no. 8, **(2009)**, pp. 10896-10904.
[5] C. Slim, "Forecasting the volatility of stock index returns: A stochastic neural network approach", Computational Science and its Applications, vol. 3, **(2004)**, pp. 935-944.
[6] L. Yang, C. W. Dawson, M. R. Brown and M. Gell, "Neural network and GA approaches for dwelling fire occurrence prediction", Knowl. Based Syst, vol. 19, **(2006)**, pp. 213-219.
[7] A. George, P. V. Kimon, "Forecasting stock market short-term trends using a neuro-fuzzy based methodology", Expert Syst. Appl., vol. 36, **(2009)**, pp. 10696-10707.
[8] K.- m. Koo and E. -y. Cha, "A Novel Container ISO-Code Recognition Method using Texture Clustering with a Spatial Structure Window", IJAST, vol. 41, **(2012)**, pp. 83-92.
[9] R. Sheikhpour and S. Jabbehdari, "A Two-Level Cluster based Routing Protocol for Wireless Sensor Networks", IJAST, vol. 45, **(2012)**, pp. 19-30.
[10] P. Shi, "An Efficient Approach for Clustering Web Access Patterns from Web Logs", IJAST, vol. 5, **(2009)**, pp. 1-14.
[11] E. H. Ruspini, "A new approach to clustering", Information and Control, vol. 15, no. 1, **(1969)**, pp. 22-32.
[12] S. Tamra, "Pattern classification based on fuzzy relations", IEEE SMC, vol. 1, no. 1, **(1971)**, pp. 217-242.
[13] L. Zkim, "Fuzzy relation compositions and pattern recognition", Inf. Sci., vol. 89, **(1996)**, pp. 107-130.
[14] R. -N. P. Singh and W. H. Bailey, "Fuzzy logic applications to multisensor multitargent correlation", IEEE Transa. Aero space Electron. Systems, vol. 33, no. 3, **(1997)**, pp.752-769.
[15] G. Xinbo, "Advances in theory and applications of fuzzy clustering", Chinese science bulletin, vol. 45, no. 11, **(2000)**, pp. 961-970.
[16] S. Susanto, I. Suharto and P. Sukapto, "Using fuzzy clustering algorithm for allocation of students", World transaction on engineering and technology education, vol. 1, **(2002)**, pp.245-248.
[17] C. Sarbu and K. Zehl, "Fuzzy divisive hierarchical clustering of soil data using gustafson-kessel algorithm", Chemometrics and Intelligent Laboratory Systems, vol. 86, **(2007)**, pp. 121-129.
[18] R. Sracoglu and N. Allahverdi, "A fuzzy clustering approach for finding similar documents using a novel similarity measure", Expert systems with applications, vol. 33, no. 3, **(2007)**, pp. 600-605.

Author

Jianfeng Li received Doctor's Degree from the Department of Management Science and Engineering of Harbin Institute of Technology (HIT), China, in January, 2006. The major fields are data mining, supply chain management, enterprise model and business process reengineering. Now, he works in Dalian Maritime University as associate professor, China. He has published quite a lot of articles in the journals, and has been presiding over some important scientific projects including the National Natural Science Foundation of China (NSFC).