Discovering Financial Investment Strategy through Wavelet-based SOM Networks

Sheng-Tun Li[1], Shu-Ching Kuo[1,2]
[1]Department of Industrial and Information Management, National Cheng Kung University, Taiwan, R.O.C.
Email:stli@mail.ncku.edu.tw
[2]Department of Information Management,Diwan College of Management,Taiwan, R.O.C.
Email: shu-ching@mail.dwu.edu.tw

Abstract

Recently, the recession of the global economy drives the coming of a new era of low interest rates, which resulted in the stock market as an alternative investment channel for investors. The diversity and complication of domain knowledge existing in the stock market make it very difficult for investors to make an effective investment decision. In the literature, it has been shown that intelligent decision support models can alleviate the burden of investors. In this study, we are in an attempt to develop such model by applying the over-whelming self-organization map (SOM) network. In contrast to most similar approaches toward the problem of stock prediction, we investigate the feature extraction issue in dealing with stock time series using discrete wavelet transform. We design a two-level SOM network and perform a novel trajectory analysis so that both short-term and long-term prediction can be provided. In addition, we not only endeavor ourselves to improving the accuracy of uncovering trading signals, but also to maximize the profits of trading. The experimental results demonstrate that the proposed decision model can help investment make the most profitable decisions.

KEY WORDS: knowledge discovery, discrete wavelet transform, self-organizing map network, decision support systems, financial investment.

1. Introduction

Recently, due to the recession of the global economy, US Federal Bank reduced the interest rate more than ten times. In the following, the fixed deposit interest rate of Central Bank of Taiwan (CBT) dropped drastically below 1%. With the coming of low interest-rate era, using fixed deposits as an investment tool was not feasible anymore and this thus drives investors to look for other investment channels, in which the stock market is perhaps with the lowest barrier. However, the diversity and complication of domain knowledge existing in the stock market make it very difficult for investors to make correct investment decisions promptly, since transaction speed has become much faster nowadays. To alleviate the burden of investors, there are a number of decision support systems which can gather real-time pricing information for supporting decision-making in financial investment have been proposed [8,22].

Financial investment is a knowledge-intensive industry. In the past years, with the electronic transaction technology advances, vast amount of transaction data has been collected and the emergence of knowledge discovery technology sheds light toward building up a financial investment decision support system [5,14,15,17]. Data of financial market, in particular the stock market, are essentially time series which bring more challenges than the traditional discrete data for uncovering the hidden knowledge [1]. Approaches toward predicting stock time series can be classified into two categories: linear statistic mode and non-linear artificial-intelligence (AI) model. The former model include the well-known regression, AR, MA, ARIMA, GARCH, and ARCH-M Since the fluctuations in the stock market are highly non-linear in nature, the effectiveness of the aforementioned methods is limited. On the other hand, the AI model including neural network, fuzzy system, and genetic algorithm has been shown successful in dealing with the prediction of stock market [3,11,13,21]. In addition, hybrid systems incorporating efficient feature-extraction methods can enhance the prediction performance. With machine learning, artificial neural network models the nonlinear characteristics of time series and allows appropriate learning, expression, and presenting for decision-making purposes. For example, the over-whelming self-organizing map (SOM) network has been recognized as an outstanding model in knowledge discovery in databases (KDD) or data mining due to its salient capability in clustering and visualization. The SOM network has been widely applied in diverse application domains [6,19,20].

Data from the stock market are essentially time-series, which bring more challenges than the traditional discrete data for uncovering the hidden knowledge. Significant features or noise could be embedded in the time series and thus need particular treatment. In this study, we tackle this issue by using the emerging discrete wavelet transform (DWT) technology. Differing from current work in the study of prediction of stock movement which mostly emphasizes on the improving the prediction accuracy, we focus on the most important issue in stock market, i.e., maximizing the profits of trading.

Data from financial markets are essentially time-series, which bring more challenges than the traditional discrete data for uncovering the hidden knowledge. In this research, Taiwan Weighted Stock Index K-chart patterns as the target dataset, we tackle with these challenges by proposing an integrated solution on the basis of knowledge-discovery methodology which supports four important tasks of data mining: clustering, classification, forecasting and visualization. In particular, we adopt the approach of wavelet transform to extract features from the time-series data to overcome the impacts of smoothening resulting from clustering so that the important information regarding the trend changes can be kept. In order to provide a decision maker the functionality of visualization, we utilizes a two-level self-organization map that can transform high- dimensional, complicated, and nonlinear data onto low-dimensional ones with topology preservation. For clustering, the wavelet coefficients are used as the input vectors and the silhouette coefficient algorithm are applied to validate the clusters. Following, the trading signals are classified by performing pattern-match with K-Chart patterns in the trained SOM and the sliding-window data. Finally, the closing price in the next day is predicted based on the first 31 days pattern. This study is different from relevant research, we provide investors to judge Primary Bull Markets or Primary Bear Markets by trajectory analysis. The resulting intelligence investment decision support system can help fund managers and investment decision-makers of national stable funds make profitable decisions. In addition, financial experts can benefit from the ability of verifying or refining their tacit investment knowledge offered by the uncovered knowledge.

2. Underlying Technologies

It has been one of the greatest challenges to predict the stock market. Since stock prices vary dramatically, it is important to determine when to buy and sell in order to get high returns from stock investment. In technical analysis, we used candlestick to present daily market's open, high, low, and close prices and looked at the change in body color of the K-chart to interpret the day-to-day sentiment. Another issue is the selection of a proper technology in identifying homogeneous strategy using clustering, which involves tackling the problem of high dimensionality inherent in temporal data. In the machine learning literature, the SOM network has shown to be an effective clustering technology in isolating clusters in a high dimensional space [10]. In addition, wavelet transform could naturally play an important role in data mining since it provides presentations of data that make the mining process more efficient and accurate and it can be incorporated into the kernel of many data mining algorithms [16]. We will discuss these three major technologies in this section.

2.1 K-chart patterns

A K-chart analysis was used to elicit technical knowledge in this study. This charting technique has become very popular among traders. One of the major reasons is that K-chart is a useful tool to visualize the stock prices so that investors can catch up patterns promptly which can be used to predict future stock price movements [17]. As illustrated in figure 1, a candlestick is composed of a rectangle and two shadow lines. The former, so-called ''real body'', indicates the difference between the opening and closing prices of a stock. If the real body of a candlestick shows that the opening prices are higher than the closing prices, the candlestick is called a ''block candlestick'', otherwise it is named ''white candlestick.'' The white candlestick implies a rising signal of a stock price and the black candlestick implies a falling signal. Investors thus can interpret the day-to-day sentiment from simply looking at the change in body color of the K-chart. The stock price patterns represented by the candlestick shapes provide important clues to predict future stock price movements. So, we adopt K-charts as the feature representation of stock price movements.

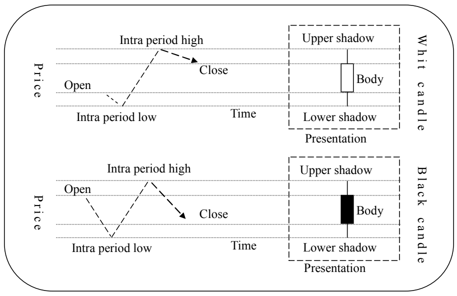

Figure 1. K-chart terms and interpretation

2.2. Self-organizing map neural network

The SOM neural network is one of the most popular unsupervised neural network models, which quantizes the data space and simultaneously performs a topology-preserving projection from the data space onto a regular two-dimensional grid [10]. The SOM network can be used for clustering, classification, and modeling. In particular, its outstanding visualization capability provides visual informative pictures of the data space. The versatile properties of the SOM network make it a valuable tool in data mining and knowledge discovery [9]. SOM have been successfully applied in various areas such as engineering, finance, bioinformatics, and many others. [19,20].

A general SOM network is composed of an input layer and a Kohonen layer. The input layer contains neurons for each element in the input vector. The Kohonen layer is formed by neurons which are located on a regular, usually two-dimensional grid and are fully connected with those at the input layer. Each neuron i in the map is represented by a n -dimensional weight or reference vector, where n is equal to the number of neurons in the input layer. The neurons on the map are connected to adjacent ones by a neighborhood relation dictating the topological structure of the neurons. When an input vector $x \in R^n$ is presented to the network, the neurons on the map compete with each other to be the winner (or the best-matching unit, BMU), which is the closest to the input vector in terms of some kind of similarity measure. Therefore, similar input vectors are grouped into a single neuron or neighboring ones in the map when learning is accomplished.

2.3 Wavelet Transform

Wavelet transform can be regarded as an extension of Fourier transform, which has been widely used in transferring signals to the frequency domain for extracting features from time series. Fourier transform can provide the frequency information of a signal, however, it cannot identify the time when the embedded frequency varies with time and thus cannot be applied to non-stationary signals. For remedying such a shortcoming, wavelet transform has been proposed to extract the time-frequency information simultaneously. In particular, the discrete wavelet transform (DWT) provides the capability to investigate the temporal variation with a different scale [19]. It decomposes signals into basis functions that are dilations and translations of a single prototype wavelet function:

$$f(x) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} C_n^m \psi_{m,n}(x)$$

where $\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n)$, are obtained by translates and dilates of the wavelet function $\psi(x)$. The DWT coefficients can be calculated by the inner products $(\psi_{m,n}(x), f(x))$ which are the estimation of signal components at $(2^{-m}n, 2^m)$ in the Time-Frequency place.

DWT corresponds to multiresolution approximation expressions which permits the analysis of a signal in many frequency bands or at many scales. In practice, mutiresolution analysis is carried out using 2 channel filter banks composed of a low-pass and a high-pass filter and each filter bank is then sampled at a half rate (1/2 down sampling) of the previous frequency. By repeating this procedure, it is possible to obtain wavelet transform of any order. The down sampling procedure keeps the scaling parameter constant (1/2) throughout successive wavelet transforms so that it benefits from simple computer implementation.

The study in Li et. al. [16] provides a comprehensive survey of wavelet application in data mining. In particular, the authors show how wavelet transform can be successfully applied to support the creative process of data mining in the phases of data understanding, data preparation, modeling, and evaluation. In this study, we apply DWT to extract features from the non-stationary characteristics of stock time-series data.

3. Research Methodology

Data mining is an iterative process which involves various steps. A reference model for data mining will greatly help the success of the problem under investigation. In this section, we present the process model used in mining Taiwan Weighted Stock Index by adapting the reference model proposed by Fayyad et. al [4]. Figure 2 shows the self-explanatory model and the associated steps, which will be discussed in detail in this section. The time-sequence data under mining is decomposed into experimental and control sets for comparing the performance evaluation. The former set is analyzed by the combination of DWT and SOM whereas the latter is by SOM alone.

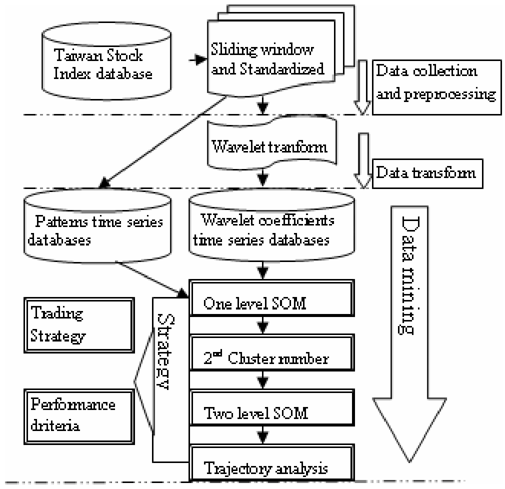

Figure 2. The proposed mining process model

3.1 Data collection and preprocessing

The data of the Taiwan Weighted Stock Index was gathered from the Taiwan Economic Journal Data Base. The dates ranged from January 3, 1991 to December 31, 2002, with a total of 3,306 exchange days. This research extracted information from the opening, high, low, and closing prices, and these numbers are the basic elements of what the K-chart consists of. To simulate a prospective use of the neural network, the data was divided into training and test sets. The training set was constructed from January 3, 1991 through December 31, 2000. The testing set was from January 2, 2001 to December 31, 2002, with 492 data sets.

Sliding window and ordinary normalization are two preprocesses before we use DWT and SOM algorithms analysis the Taiwan Weighted Stock Index trading signals. The sliding window technique is used to divide the time series into small windows, each moving window that is extracted from the time sequence data [2,7]. The length of the sliding window was set to thirty-two. The variables p_t^o, p_t^h, p_t^l, and p_t^c were set, respectively, as the opening price, the high price, the low price, and the closing price of period *t*. *W* represents the length of the sliding window sizes; therefore, the sliding window frame of period *t*'s opening price is $x_t^o = [p_t^o, p_{t-1}^o, p_{t-2}^o, \dots, p_{t-w+1}^o]$. With the same formula applied to the high price, the low price, and the closing price, $x_t^h = [p_t^h, p_{t-1}^h, p_{t-2}^h, \dots, p_{t-w+1}^h]$, $x_t^l = [p_t^l, p_{t-1}^l, p_{t-2}^l, \dots, p_{t-w+1}^l]$, and $x_t^c = [p_t^c, p_{t-1}^c, p_{t-2}^c, ..., p_{t-w+1}^c]$ are obtained, respectively. Thus, the sliding window for the whole period t is $x_t = [x_t^o, x_t^h, x_t^l, x_t^c]$. With the sliding window's movement on the time-axis, pieces of pattern can be formed, like the sequence database. Each of these patterns contain $[x_t^o, x_t^h, x_t^l, x_t^c]$ opening price, high price, low price, and closing price; and the dimensions for each pattern is 128 (32 ×4). In the meantime, the data is used ordinary normalization (minimum and maximum values in the [0, 1] range).

3.2 Mining process

This study uses quantitative methods and programmed methods for profitable decision-making. A K-chart analysis is used in order to elicit technical knowledge. The K-chart combines opening, highest, lowest, and closing prices, with utilizing sliding windows (32 days) as a segmentation algorithm. By proceeding with DWT processes, we extract features from the time-series data and data mining. For clustering, two-level SOM network are used as extraction of the whole cluster prototypes of the time series patterns. The first level given 10×10 clusters, the trading signals are classified by performing pattern-matching with K-Chart patterns in the trained SOM and the sliding-window data. Then, the closing price in the next day is predicted based on the patterns of the first 31 days. So far, first-stage daily prediction and profit testing is developed. Subsequently, the two-level SOM clusters patterns numbers are decided by the agglomeration coefficients. The second step uses the trained two-level SOM and conducts a corresponding one-level and two-level trajectory analysis, which develop trend predictions for the Primary Bull Markets and the Primary Bear Markets.

3.3 prediction model

In predicting the clustering patterns trained by the SOM neural network, we propose a novel trajectory analysis to record and interpret pattern series movement process. In addition, this study develops two decision support models: step prediction model and long-trend prediction model.

In the trajectory analysis of the SOM network, the test pattern series uses each clustering pattern to compare with the minimum mean square error, in order to search for the most similar pattern and to record the pattern serial numbers according to their sequence. Figure 3 is the trajectory analysis concept, where test pattern T_1 to test pattern T_n can contrast with the clustering patterns of the first level, which with the passing of time, also goes from pattern series number c_k to pattern number c_m ($k, m \in 1, 2, ..., 100$). The clustering pattern of the first level SOM can map to the second level SOM $c_1^{'}$ to $c_N^{'}$. Hence, when a test pattern is in progress, we can find the closest clustering pattern in the first level, reflect it onto the second level, and obtain its serial number. Therefore, the clustering patterns movement is identified, which implies knowledge when we record the pattern serial numbers according to time series moving.

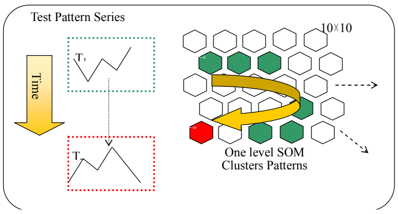

Figure 3. Trajectory analysis concept

To support investors in decision-making, step prediction model and long-trend prediction model are developed. The step prediction model mainly predicts the rising or falling of next day's price with time series moving. We feed these testing patterns using a 31 trading day window by the first level SOM pattern matching. The signal of buying or selling is defined as follows:

$$Sig(t) = \begin{cases} Buy, if & P(t+1)-P(t)>0 \\ Sell, if & Otherwise \end{cases}$$

On the other hand, the long-trend prediction model predicts the rising or falling signal according to trajectory analysis of price in Taiwan stock market. We divide the market trend in two parts: primary bull market and primary bear market. By this, we can observe the long-term trend of price in Taiwan stock market. The rule is as follows:

$$Trajectory\_Strategy(t) = \begin{cases} Buy, & if \quad PatternNumber\ (t) < 4 \\ Sell, & if \quad PatternNumber\ (t) \geq 4 \end{cases}$$

4. Experimental Results

This study uses short-term trend to produce trading signal of buying or selling. Predicting the trend pattern is more effective than predicting actual closing price because of predicting price can not provide a continuous trend prediction. However predicting the trend pattern can provide a complete support investment strategy. In the following, we only discuss the experimental results from the experimental set alone due to space limitation.

4.1 Step prediction evaluation

We constructed a trained 10×10 SOM from the experimental set, which shows the distribution of time series data as Figure 4 and demonstrates the characteristics of birds of a feather flock together. Since DWT characteristic coefficients can not completely represent the implied meaning of original data, a reconstruction result from DWT characteristic coefficients is shown in Figure 5, which clearly confirms that the K-chart patterns are cluster homogeneously. Then, we try to label the buying or selling signal based on the centroid of each cluster. If the closing price of the thirty-second day is higher than that of the thirty-first day, a buying signal is labeled as red "○" otherwise green "×" for selling signal. The resulting curve of trading strategy is shown in Figure 6, by which the investor can take appropriate actions when a trend pattern occurs.

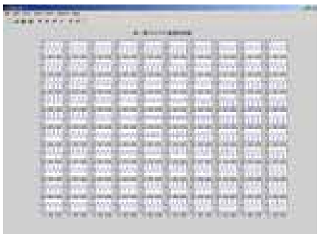

Figure 4. DWT and the first level SOM cluster feature map

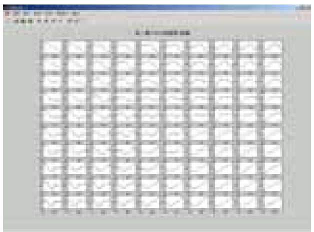

Figure 5. Clustering K-chart reconstructed by DWT

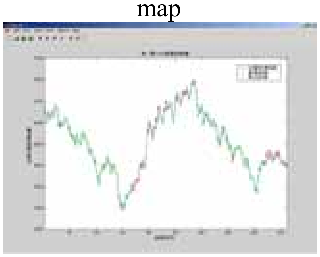

Figure 6. The trading strategy in the first level SOM with DWT

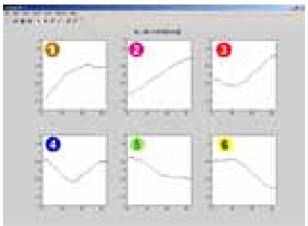

Figure 7. The clustering patterns in the second SOM with DWT

4.2 Long-trend prediction evaluation

To support the long-trend trading strategy, 2 two-level SOM network is constructed. The inputs to the second-level of the SOM are cluster centroids in the first-level SOM. In this study, the topology of the second-level SOM is 1*6, i.e. six clusters, is determined by using agglomeration coefficients of hierarchical cluster analysis in multivariate statistics. Figure 7 shows the long-trend patterns after training. By incorporating the long-trend pattern numbers into the first-level SOM, one obtains the movement trajectory as displayed in Figure 8, which demonstrating an interesting phenomenon that long-trend pattern numbers rotate in a clockwise direction. On the other hand, when taking Taiwan Weighted Stock Index and the pattern numbers defined by the second-level SOM pattern together into consideration (see Figure 9), two important trading trends can be identified by consulting a senior stock analyst: primary bull market and primary bear market. Pattern numbers (clusters) one, two, and three are all primary bull market, which suggest the investor to take buying strategy over a long period of time. On the contrary, the other clusters belong to the primary bear market and the investor should take selling action. Another interesting implication by observing Figure 9 is if the stock movement walks within the clusters in the same primary bull market or the primary bear market, it indicates that the market's vibration is quite limited. However, if the movement is walking from one market to the market, a warning signal of significant market vibration should be issued to the investor. Similar the trading approach in the step prediction, we may label the buying and selling signals based on the long-trend trajectory analysis. Figure 10 shows the trading signals with the two-level SOM network.

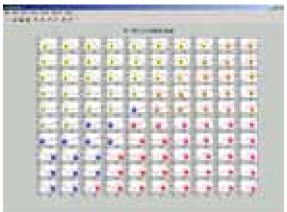

Figure 8. The stock movement trajectory analysis

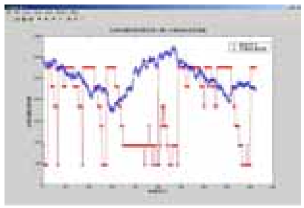

Figure 9. Stock Index and trajectory analysis with the two-level SOM network.

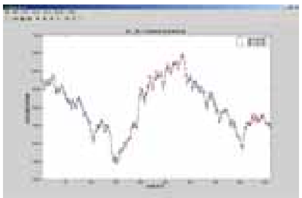

Figure 10. The trading strategy suggested by the two-level SOM network.

4.3 Performance evaluation

In order to compare the trading performance of the proposed DWT-based two-level SOM network, we calculate various performance indexes as shown in Table 1 by applying the suggested buying and selling strategies in Figures 6 and 10. Table 1 shows that both networks achieve positive total profits and confirms the basic prediction capability provided by the SOM-based network. It is worth noting that, in terms of the total profit, gross gain, gross loss, average loss, average G/L and maximal gain, the DWT-SOM network outperforms over the traditional SOM network in step prediction strategy. On the other hand, the DWT-SOM network obtains better performance than the SOM network in long-trend prediction strategy from the viewpoint of total profit, gross gain, average gain, average loss, average G/L, and maximal gain. In a word, the superiority of DWT-based SOM network is validated.

Table 1.Trading strategy comparison

Trading strategy	step prediction strategy		long-trend prediction strategy	
	SOM	DWT-SOM	SOM	DWT-SOM
Total profit	4431.35	4940.95	3021.55	3230.95
Avg. profit	63.31	79.69	251.80	179.50
Gross gain	8244.75	6362.50	3858.45	4686.85

Gross loss	-3813.4	-2769.20	-836.90	-1455.90
Avg. gain	187.38	159.06	551.21	669.55
Avg. loss	-146.67	-125.87	-167.38	-132.35
Avg. G/L	1.28	1.26	3.29	5.06
Max. gain	804.80	1098.20	1598.90	1958.90
Max. loss	-473.30	-549.90	-349.30	-555.20

5. Conclusions

Financial investment is a knowledge-intensive industry. Data of financial markets is essentially time-series which brings more challenges than the traditional discrete data for uncovering the hidden knowledge. In this study we explore the knowledge discovery in Taifex Index in Taiwan Futures Exchange in order to identify the important movement trends of stocks. In particular, we adopt the approach of wavelet transform to extract features from the time-series data to overcome the impacts of smoothening resulting from clustering so that the important information regarding the trend changes can be kept. With the outstanding abilities of clustering and visualization provided by the DWT-SOM network, a thorough trajectory analysis on the K-chart patterns provides satisfied accuracy of classification of trading signals and the profits of trading as well. In addition, the resulting visualization decision tool can help financial experts verify or refine their tacit investment knowledge offered by the uncovered knowledge.

Reference

[1] Apte C. B., Liu B., Pednault E. and Smyth P., Business applications of data mining, *Communications of the ACM*, 2002,45(8) 49-53.

[2] Babcock B., Datar M., Motwani R., Sampling from a Moving Window over Streaming Data, *Proc. 13th SIAM-ACM Symp. on Discrete Algorithms*, 2002, 633-634.

[3] Cheng C., Low B., Chan P. and Motwani J., Improving the performance of neural networks in classification using fuzzy linear regression, *Expert Systems with Applications*, 2001, 20(2), 201-206.

[4] Fayyad U. M., Shapiro G. P., and Smyth P., From data Mining to Knowledge Discovery: An Overview, *Advances in Knowledge Discovery and Data Mining*, MA AAAI/ MIT Press, 1996, 1-36.

[5] Fayyad U. M., Piatetsky-Shapiro G. & Smyth P., The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM,*, 1996, 39(11), 27-34.

[6] Flexer A., On the ues of self-organizing maps for clustering and visualization, *Intelligent Data Analysis*, 2001,vol(5), 373-384.

[7] Golab L. and Ozsu M. T., Processing Sliding Window Multi-Joins in Continuous Queries over Data Streams, *University of Waterloo Technical Report*, February 2003. Available at http://db.uwaterloo.ca/ ~ddbms/publications/stream/multijoins.pdf.

[8] Jan M. F., Andrew S., and Whinston B., A Web-Based Financial Trading System, *Computer IEEE,* 1999, 32(4), 64-70.

[9] Kiang M. Y., Extending the Kohonen self-organizing map networks for clustering analysis, *Computational Statistics & Data Analysis* ,2001,38(2), 161-180.

[10] Kohonen T., Self-Organizing Maps, Springer-Verlag Berlin, Heideberg.

[11] Kuo R. J., Chen C. H. and Hwang Y. C., An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network, *Fuzzy Sets and Systems*, *118*, 2001, 21-45.

[12] Lee K. H. and Jo G. S., Expert system for predicting stock market timing using a candlestick chart, *Expert Systems with Applications*, 1997, vol(16), 357-364.

[13] Leigh W., Purvis R. and Ragusa J. M., Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support, *Decision Support Systems*, 2002, 32(4), 361-377.

[14] Leigh W., Modani N., Purvis R., and Roberts T., Stock market trading rule discovery using technical charting heuristics, *Expert Systems with Applications*, *23*, 2002, 155-159.

[15] Li S. L., A web-aware interoperable data mining system. *Expert Systems with Applications*, 2002, 22(2), 135-146.

[16] Li T., Li Q., Zhu S., & Ogihara M., A survey on wavelet applications in data mining. *ACM SIGKDD Explorations*, 2002, 4(2), 49-68.

[17] Liao S. H., Knowledge management technologies and applications-literature review from 1995-2002, *Expert Systems with Applications,* 2003, 25(2), 155-164.

[18] Mallat S., A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Pattern Analysis and Machine Intelligence,* 1989, 11(7), 674-693.

[19] Vesanto J., Alhoniemi E., Clustering of the self-organizing map, *IEEE Transactions on Neural Networks*, 2000, 11(3), 586-600.

[20] Vesanto J., Using SOM in Data Mining. *Licentiate's thesis at the Helsinki University of Technology*, Finland, April 2000.

[21] Zhou Z. H., Three perspectives of data mining. *Artificial Intelligence*, 2003, 143(1), 139-146.

[22] Zopounidis C., Doumpos C. M., and Matsatsinis N. F., On the use of Knowledge-based Decision Support Systems in Financial Management: A Survey, *Decision Support Systems*, 1997, 20(3), 259-277.