Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 34 (2008) 935–951
Expert Systems with Applications
www.elsevier.com/locate/eswa

Knowledge discovery in financial investment for forecasting and trading strategy through wavelet-based SOM networks

Sheng-Tun Li [a,b,*], Shu-Ching Kuo [b,c]

[a] Institute of Information Management, National Cheng Kung University, No. 1, Ta-Hsueh Road, Tainan 701, Taiwan, ROC
[b] Department of Industrial and Information Management, National Cheng Kung University, Taiwan, ROC
[c] Department of Information Management, Diwan College of Management, Taiwan, ROC

Abstract

The stock market has been a popular financial investment channel in the recent era of low interest rates. How to maximize profits is always the main concern for investors; and different investors have different preferences about the holding periods of their investments. In this study, in contrast to other related studies, we propose a hybrid approach on the basis of the knowledge discovery methodology by integrating K-chart technical analysis for feature representation of stock price movements, discrete wavelet transform for feature extraction to overcome the multi-resolution obstacle, and a novel two-level self-organizing map network for the underlying forecasting model. In particular, a visual trajectory analysis is conducted to reveal the relationship of movements between primary bull and bear markets and help determine appropriate trading strategies for short-term investors and trend followers. The forecasting accuracy and trading profitability of the proposed decision model is validated by performing experiments using the Taiwan Weighted Stock Index (TAIEX) from 1991 to 2002 as the target dataset. The resultant intelligent investment model can help investors, fund managers and investment decision-makers of national stabilization funds make profitable decisions.

Keywords: Knowledge discovery; Self-organizing map network; Wavelet transform; Financial investment; Trajectory analysis

1. Introduction

Due to fears of inflation, the US Federal Bank has raised the interest rate more than ten times since 2004, but the fixed deposit interest rate of the Central Bank of Taiwan (CBT) dropped drastically below 2%. Using fixed deposits as an investment tool is thus not feasible anymore, and this has driven investors to look for other investment channels, of which the stock market is perhaps the most popular. However, the highly non-linear, dynamic complicated domain knowledge inherent in the stock market makes it very difficult for investors to make the right investment decisions promptly (Abhyankar, Copeland, & Wong, 1997; Hiemstra & Jones, 1994). It is thus of great necessity to develop a decision support system which can gather real-time pricing information, alleviate one burden of investors and help them maximize profits.

Financial investment is a knowledge-intensive industry. In recent years, with the advances in electronic transactions, vast amounts of data have been collected. In this context, the emergence of knowledge discovery technology enables the building up a financial investment decision support system (Boginski, Butenkob, & Pardalos, 2006; Bose & Mahapatra, 2001; Chen, Han, & Yu, 1996; Enke & Thawornwong, 2005; Irma, Stelios, & Steven, 2002; Pyle, 1999; Wang & Weigend, 2004). In the recent literature, various techniques of knowledge discovery have been employed in the stock market, and they can be divided into technical analysis and forecasting of stock time series data. The former uses the charting heuristics of technical analysis to identify the bull flag by template match and to establish

* Corresponding author. Address: Institute of Information Management, National Cheng Kung University, No. 1, Ta-Hsueh Road, Tainan, 701 Taiwan, ROC. Tel.: +886 6 2757575x53126; fax: +886 6 2362162.
E-mail address: stli@mail.ncku.edu.tw (S.-T. Li).

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.10.039

trading rules. Some related works include Bo, Linyan, and Mweene (2005); Leigh, Modani, Purvis, and Hightower (2004); Leigh, Modani, Purvis, and Roberts (2002); Leigh, Paz, and Purvis (2002). The latter focuses on the forecast of stock prices, which often employs a statistical or artificial intelligence (AI) approach to facilitate the trading strategy-making. The statistical approach, such as (Evstigneev, Hens, & Schenk-Hopp, 2006) and (Boginski, Butenko, & Pardalos, 2005), is built upon on the assumption that the stock time series data is stationary, normally distributed and has no outliers. Alternatively, a variety of AI technologies have been shown successful to construct the decision models such as decision tree (Wang & Chan, 2006; Wu, Lin, & Lin, 2005), case based reasoning (CBR) (Chun & Park, 2005; Oh & Kim, 2007), neural networks (Armano, Marchesi, & Murru, 2005; Chen, Leung, & Daouk, 2003; Thawornwong & Enke, 2004; Vellido, Lisboa, & Meehan, 1999), fuzzy system (Wang, 2003), and hybrid ones (Chun & Kim, 2004; Fang & Xu, 2003; Kim & Han, 2000; Zhang, 2003).

The aforementioned works mostly concentrate on improving the forecasting accuracy or the trading strategy from stock time series data. Yet, there are still several important issues worthy of attention. Firstly, in order to effectively predict future stock prices, various methods of technical analysis are usually integrated into the underlying forecasting model. However, most of them only provide basic information about the price movement and trend in markets. It has been evidenced that a visualized analysis can help investors catch future stock price movement patterns more promptly (Lee & Jo, 1999). Secondly, the stock time series can be noise prone or have some hidden, significant features. Moreover, the analytic result of stock data could be sensitive to multi-resolution scales, as in most time series data (Li & Shue, 2004). Finally, present forecasting models are mainly concerned with the prediction accuracy, which tends to lead to so-called black-box models, and investors or managers cannot benefit from the knowledge discovery in the analytic process.

In this paper, we propose a hybrid forecasting model which addresses these three important issues. Technically speaking, candlestick charts (K-charts) (Nison, 2001), appealing visualization aids to give investors the day-to-day sentiment, are utilized. We then appeal to discrete wavelet transform (DWT) which is allowed to decompose a time series into subsequences in different resolution scales and to extract the hidden significant, temporal features. To facilitate the knowledge discovery in the forecasting process, we perform a novel trajectory analysis using a modified self-organization map (SOM) (Kohonen, 1997) neural network to give insights into the relationship of movements between the so-called Primary Bull and Primary Bear Markets, helping determine appropriate trading strategies. We conduct a series of experiments on the proposed model using the Taiwan Weighted Stock Index (TAIEX) from 1991 to 2002 as the target dataset to validate its forecasting accuracy and trading profits. The resultant intelligent investment decision support model can help investors, fund managers and investment decision-makers of national stabilization funds make profitable decisions. In addition, financial experts can benefit from the ability of verifying or refining the tacit investment knowledge offered by the newly discovered knowledge.

The remaining sections of this paper are organized as follows. Section 2 provides the literature review. Section 3 introduces the essential concepts of K-charts, DWT and the SOM network. Section 4 describes the underlying research methodology and mining process. Section 5 presents the experimentation design and results. Section 6 discusses the forecasting and trading profits for different trading strategies. Finally, concluding remarks and future work are given in Section 7.

2. Literature review

In a stock market, how to forecast stock prices accurately and find the right time to trade are obviously of great interest to investors. To achieve this objective, various approaches have been employed, which can be separated into two categories: statistical and AI. The statistics school includes autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GARCH) volatility (Franses & Ghijsels, 1999), and smooth transition autoregressive (STAR) (Sarantis, 2001). These models rely on the assumption of linearity among variables and normal distribution. However, the assumption of linearity and normal distribution may not hold even though it has been shown successful in dealing with stock price movement in past decades. On the other hand, with greater in the stock market and the increasing need for more efficient forecasting models, the AI school, operating without the limitation of such an assumption, has been confirmed to outperform the conventional statistical methods experimentally (Enke & Thawornwong, 2005; Hansen & Nelson, 2002; Ture & Kurt, 2006; Zhang, 2003).

The AI approaches, mainly including neural network, fuzzy system, and genetic algorithm, have achieved impressive results in dealing with stock market prediction (Armano et al., 2005; Chen et al., 2003; Chun & Kim, 2004; Kim & Han, 2000; Shen & Loh, 2004; Thawornwong & Enke, 2004; Vellido et al., 1999; Wang, 2002). This superiority is attributed to several important advantages: nonlinearity, robustness, stability and adaptability (Lee & Jo, 1999; Wang & Chan, 2006). This study adopts the SOM network, a well-known unsupervised neural network, as the underlying forecasting model. There are numerous works on applying SOM in financial markets in the literature. Gafiychuk, Datsko, and Izmaylova (2004) utilized the SOM network to investigate deep relationships inherent to the financial market and the GMDH algorithm to analyze the clusters obtained. Chen and He (2003) proposed a systematic and automatic approach which used SOM to identify charts behavior and a normalized equity cure to reveal trading signals, respectively. They concluded that SOM can help one foresee the movement of a stock index in the near future and improve profitability. Chi, Peng, Wu, and Yu (2003) integrated SOM and fuzzy neural network techniques to establish the relationship of changes between the stock technical indicators and stock index in order to understand how the trend of stock index changes. The resultant network is demonstrated to achieve higher prediction accuracy than traditional networks. By taking advantage of the ability of SOM in pattern recognition and the knowledge of human technical traders, Resta (2000) tackled the issue of forecasting over S&P 500 high-frequency (ex. 30–60 min) data in addition to the regular daily data.

In addition to forecasting models, effective data pre-processing is definitely helpful for enhancing the forecasting quality (Pyle, 1999). For time series forecasting problems, multi-resolution analysis has been well recognized to extract features and build the model (Beltra'n et al., 2006; Murtagh, Starck, & Renaud, 2004; Starck & Murtagh, 2002). Especially, wavelet transforms (WT) provide a useful decomposition of time series in order to reveal vague temporal structures. It has been widely used for image compression, noise removal, object detection, and large-scale structure analysis in a variety of applications (Kandaswamy, Kumar, Ramanathan, Jayaraman, & Malmurugan, 2004; Starck & Murtagh, 2002; Subasi, 2005a; Subasi, 2005b). Petrosian, Prokhorov, Homan, Dashei, and Wunsch (2000) showed that the ability of specifically designed and trained recurrent neural networks (RNN) combined with wavelet preprocessing, was more accurate in predicting EEG signals. Murtagh et al. (2004) investigated the applicability of Haar wavelet transform in S&P 500 stock data and showed that the multi-resolution approach outperforms the traditional single resolution approach in modeling and prediction. Recently, the marriage of the WT's feature extraction and SOM's pattern recognition shows a very promising scheme for mining time series data. For example, the automatic segmentation of natural and synthetic diphthongs in the presence of additive noise (Torres, Gurlekian, Rufiner, & Torres, 2006), dynamic muscle fatigue detection (Moshou, Hostens, Papaioannou, & Ramon, 2005), air pollution management (Li & Shue, 2004), financial time-series visualization and interpretation (Moshou & Ramon, 2004); and astronomical image segmentation (Núñez & Llacer, 2003). Nevertheless, comparatively little work is reported about the topic of stock market forecasting, which is the major issue of our study. Furthermore, we propose a new trajectory analysis on a modified SOM so that more informative patterns about stock movements can be revealed.

3. Research methodologies

In this section, we briefly review the basic concepts about the underlying technologies used in the study.

3.1. K-chart patterns

Technical analysis is commonly used among traders, and the candlestick chart (K-chart) is a useful visualization tool which can unmask the fluctuation of stock prices and help develop a trading strategy based on patterns (Lee & Jo, 1999; Nison, 2001). Since K-charts were developed by Japanese rice traders in the 17th century, there has been a lot of effort to observe the future behavior of a market by using them. The candlestick is composed of upper and lower wick lines and a rectangle. The lines indicate the highest and lowest traded prices of a stock. The latter, so-called "real body," denotes the difference between the opening and closing prices of a stock. If the real body of a candlestick shows that the opening price is higher than the closing price, the candlestick is called a "black candlestick," otherwise "white candlestick." The white candlestick implies a rising signal of a stock price whereas the black candlestick implies a falling signal, as illustrated in Fig. 1. Investors thus can interpret the day-to-day sentiment from simply looking at the change in body color of the K-chart. In addition, there are multiple forms of K-chart patterns. For example, Shooting Star, composed of a small body with a long upper wick and a small or no lower wick, implicates a bearish pattern during an uptrend. On the other hand, Hammer pattern, constituted of a long lower wick with or without attachment to a small body, indicates a bullish pattern during a downtrend. Therefore the candlestick shapes provide important clues by which investors can predict future stock price movements. Thanks to its excellent visualization capability, we adopt K-charts as the feature representation of stock price movements.

3.2. Self-organizing map neural network

Since its initial introduction by Teuvo Kohonen in 1986, the SOM network has been recognized as one of the most successful unsupervised neural network models. Two key features significantly contribute to its success, namely

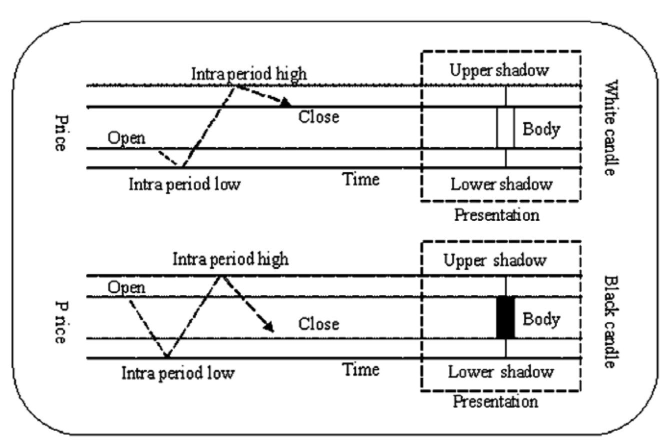

Fig. 1. K-chart terms and interpretation.

topology-preserving projection from high-dimensional space and clear visualization. In order to effectively facilitate data mining and data analysis, a variety of SOM models have been proposed, as witnessed by several thousand research papers (Chen & He, 2003; Deboeck, 1998; Deboeck & Kohonen, 1998; Flexer, 2001; Gafiychuk et al., 2004; Oja & Kaski, 1999; Penn, 2005). To support the short- and long-term strategic trading, we apply the two-level SOM model in this paper by following our previous work (Li & Shue, 2004). The two-level SOM network is composed of one input layer and two successive Kohonen feature maps in which neurons on the two contiguous layers are fully connected. The input to the network is a sequence of m patterns: $P_1, P_2, \ldots, P_t, \ldots, P_m$. Each input vector $P_t \in R^n$ is with n dimensionality representing the sliding window for period t. Each neuron i on the first Kohonen map is represented by an n-dimensional weight or reference vector $\mathbf{w}_i = [w_1, \ldots, w_n]^{\mathrm{T}}$, $i = 1, 2, \ldots, s$, where n is equal to the number of neurons in the input layer and s is the number of neurons on the first Kohonen feature map layer. When an input vector P_t is presented to the network, the connected relations $d(t,i)$ between each weight vector $\mathbf{w}_i$ and the input vector P_t is calculated according to the dissimilarity measure, such as the Euclidean distance function:

$$d(t,i) = \|P_t - \mathbf{w}_i(k)\| \tag{1}$$

Then the best-matching unit (BMU or winning unit) b in the map is the unit which has the smallest distance $d(t,i)$ as follows:

$$d(t,b) = \min_i \{d(t,i)\} \tag{2}$$

After the BMU is identified, the update of the synaptic weight vector associated with it and the neurons within a pre-defined neighborhood of the winning neuron are given by

$$\mathbf{w}_i(k+1) = \begin{cases} \mathbf{w}_i(k) + \eta h_{bi}(k)[P_t(k) - \mathbf{w}_i(k)], & \text{if } i \in N_c(k) \\ \mathbf{w}_i(k), & \text{if } i \notin N_c(k) \end{cases} \tag{3}$$

where $k = 0, 1, 2, 3, \ldots$ is the time lag, and η is a positive learning rate. $h_{bi}(t)$ is the neighborhood function which has the form $h_{ji}(t) = h(\|r_j - r_i\|,\ t)$, where $r_j,\ r_i \in R^2$ are the location vectors of neurons j and i, respectively. The neighborhood function should satisfy the condition that when $\|r_j - r_i\|$ increases, h_{ji} gradually decreases to zero. This leads to local relaxation or smoothing effects on the weight vectors of neurons in the neighborhood of the BMU. As a result, similar input vectors are grouped into a single neuron in the map when first-level SOM learning is accomplished, and the weight vector on the first-level Kohonen feature map is denoted as $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s]^{\mathrm{T}}$.

The second-level SOM is then trained by feeding the neurons with the weight vector $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s]^{\mathrm{T}}$ in the first Kohonen feature map into the second Kohonen feature map as inputs, and using the similar learning algorithm as in the first-level SOM. Table 1 summarizes the learning algorithm of the proposed two-level SOM network.

Table 1
Learning algorithm of the two-level SOM network

[Step 1.1] Initialize random weights of neurons in the first Kohonen layer and learning rate, and define the neighborhood function around a winner
[Step 1.2] Seek dissimilarity measure for a given input vector according to Eq. (1)
[Step 1.3] Select BMU b using Eq. (2)
[Step 1.4] Update the weights of the neighborhood of BMU using Eq. (3)
[Step 1.5] Repeat 1.2–1.4 until the learning of the first-level SOM finished
[Step 2.1] Initialize weights of neurons in the second Kohonen layer using ones in the first trained map. Define the learning rate and the winner neighborhood function
[Step 2.2] Repeat Steps 1.2–1.4 until training the second-level SOM is completed

Taking the topology of two-level SOMs into consideration, the first Kohonen feature map is chosen to be much larger than the second one, in such a way that the former plays the role of proto-clustering (or roughly clustering) and the latter performs clustering (or precisely clustering). Similar works on the two-level SOM network have been proposed in Li and Shue (2004); Lee, Suh, Kim, and Lee (2004); Li (2002); Vesanto and Alhoniemi (2000); Kiviluoto and Bergius (1998); Martı'n-del-Brı'o and Medrano (1995) whose objectives are in an attempt to effectively reduce the complexity of the reconstruction task and noise, to give a more accurate description and more easy interpretation of the data under investigation, and to reduce the computational cost. In addition, given a series of input patterns, one may record the trajectory of BMUs on each Kohonen map which will help in the decision-making of short- and long-term trading strategies. With these considerations, the two-level SOM network is used as the kernel model for this study.

3.3. Wavelet transform

In signal processing, Fourier transform has been widely used in identifying the frequency information and extracting features of a signal, however, it fails in dealing with non-stationary or transitory characteristics. For remedying such a shortcoming, wavelet transform (WT) has been proposed to process non-stationary signals and to investigate the temporal variation with a different scale (Li, Li, Zhu, & Ogihara, 2002; Percival & Walden, 2000). The continuous WT (CWT) is defined as the convolution of a time series $x(t)$ with a wavelet function $\psi(t)$ (Goswami & Chan, 1999):

$$\mathrm{CWT}_x^{\psi}(b,a) = \Phi_x^{\psi}(b,a) = \frac{1}{\sqrt{|a|}} \int x(t) \cdot \psi^*\left(\frac{t-b}{a}\right) \mathrm{d}t \tag{4}$$

where a is a scale parameter, b is the translation parameter and $*$ is the complex conjugate of $\psi(t)$. Let $a = 1/2^s$ and $b = k/2^s$, where s and k belong to the integer set Z. The

CWT of $x(t)$ is a number at $(k/2^s, 1/2^s)$ on the time-scale plane. It represents the correlation between $x(t)$ and $\psi^*(t)$ at that time-scale point. A discrete version of Eq. (4) is thus obtained as follows:

$$\mathrm{DWT}_x^{\psi}(k,s) = \Phi_x^{\psi}\left(\frac{k}{2^s},\frac{1}{2^s}\right) = \int_{-\infty}^{\infty} x(t)\cdot\psi^*\left(\frac{t-k/2^s}{1/2^s}\right)\mathrm{d}t \tag{5}$$

In this sense, Eq. (5) is the discrete wavelet transform (DWT) of $x(t)$, which separates the signal into components at various scales corresponding to successive frequencies.

It is noted that DWT corresponds to multi-resolution approximation expressions, which permits the analysis of a signal in many frequency bands or at many scales. In practice, multi-resolution analysis is carried out using two 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. The number of repeats of this decomposition procedure will be dependent on the length of data. The down sampling procedure keeps the scaling parameter constant (1/2) throughout successive wavelet transforms, so that it benefits from simple computer implementation.

During the last ten years, DWT has been thoroughly developed and applied in various fields because it offers some unique characteristics that are the best local approximation to the signal, and represents the signal components with the greatest signal to noise ratio (namely feature extraction), and is thus useful in energy preservation, removing noise and trend detection (Mörchen, 2003). Moreover, its significant contribution to data mining has been confirmed recently, as the work Li et al. (2002) surveyed. For example, it can support different activities in the phases of data understanding, data preparation, modeling, and evaluation. In this study, DWT is utilized to extract non-stationary characteristic features from the K-chart stock data.

4. Research methodology and mining procedure

Since its emergence, knowledge discovery in database, KDD (or data mining), has shed new light on financial investment decision supports (Apte, Liu, Pednault, & Smyth, 2002). Conceptually, KDD is the process of discovering useful knowledge from large amounts of data stored in databases, data warehouses, or other information repositories. Performing KDD is an iterative process consisting of a number of stages with different sub-tasks, each of which is supported by various complementary techniques, including statistics, artificial intelligence, and domain-dependent ones, i.e. technical analysis (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Feelders, Daniels, & Holsheimer, 2000; Mitra, Pal, & Mitra, 2002; Zhou, 2003). In this section, we discuss the mining procedure and major techniques employed in the study.

4.1. Data preparation

The data under investigation was collected from the Taiwan Weighted Stock Index, published in the Taiwan Economic Journal Data Base. This data set covers 3,306 trading days ranging from January 3, 1991 to December 31, 2002, and includes the daily openings, highs, lows, and closings, which are the basic elements of K-charts. The set is partitioned into a training set for training the SOM model and a test set for validating whether the predictions are valid outside of the training samples. Both contain respectively the first 2814 days and the following 492 days of the period; as much as 85% of the data is used for training and 15% for testing.

The data preprocessing includes the process of sliding windows, and normalization. In this study, the objective is to forecast p_t, the stock price at time t from in a sliding window which covers the w most recent stock prices, $p_{t-1}, p_{t-2}, \ldots, p_{t-w}$. The system parameter w is the number of previous values that influence the prediction. Let variables p_t^o, p_t^h, p_t^l, and p_t^c be the opening, high, low, and closing prices at time t, respectively. Therefore, the sliding window frame contains $w+1$ of period t's opening price and, is denoted as $P_t^o = [p_t^o, p_{t-1}^o, \ldots, p_{t-w}^o]$, while the day high, low, and closing prices are expressed as $P_t^h = [p_t^h, p_{t-1}^h, \ldots, p_{t-w}^h]$, $P_t^l = [p_t^l, p_{t-1}^l, \ldots, p_{t-w}^l]$, and $P_t^c = [p_t^c, p_{t-1}^c, \ldots, p_{t-w}^c]$, respectively. Thus, the sliding window for the whole period t is represented as: $P_t = [P_t^o, P_t^h, P_t^l, P_t^c]$. In the case study, w was suggested to be 31 days by the senior chartist consulted. After the appropriate windowing, the 2814-day series in the training set is divided into 2783 (=2814 − 32 + 1) patterns and 461 (=492 − 32 + 1) patterns are constructed for the test set, and each pattern is with 128 ($32 * 4$) dimensionalities.

With regard to price differences within a sliding window, the well-known statistical z-score normalization is performed to scale all prices in a window to values with a unit of standard deviation and the normalized mean of zero. A stock price p is normalized to p' using the following formula:

$$p' = \frac{p-\mu_W}{\sigma_W} \tag{6}$$

where μ_W and σ_W are the mean and standard deviation of prices.

4.2. Data transformation

When performing DWT to preprocessed time series data, one has to decide an appropriate mother wavelet to be used and the number of levels of decomposition. There are a wide variety of wavelets proposed in the literature. Each has its application domain depending on different considerations such as the resolution capability, efficiency, computational cost, etc. In this study, the computationally efficient Haar wavelet is chosen since the window length, 32, is a power of two. The Haar wavelet is the simplest wavelet (Stankovic & Falkowski, 2003) and is defined as follows:

$$\psi_H(t) = \begin{cases} 1 & 0 \leqslant t \leqslant 1/2 \\ -1 & 1/2 < t \leqslant 1 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The appropriate number of levels of wavelet decomposition can be decided based on the nature of a time series, such as dominant frequency components (Mörchen, 2003), on an entropy criterion (Coifman & Wickerhauser, 1992), or on the application's characteristics (Li & Shue, 2004). Generally speaking, the number of levels that the time series can be decomposed into is determined by the length of the time series, i.e. five levels of decomposition in this study. In this study, we are interested in investigating the impact of weekly and biweekly trends, therefore the coefficients are partitioned into two levels with blocks of sizes 8, 8, and 16 (resolution days), i.e. *cA2*, *cD2*, and *cD1*, where *cA* and *cD* are the approximation and detail coefficients, respectively. *cA* represents the high-scale, low-frequency components of the time series, whereas *cD* is the low-scale, high-frequency components.

These approximation and detail records are reconstructed from the wavelet coefficients. The first high-pass filter provides the detail *D1*, and the first low-pass filter is approximation *A1*, which is obtained by superimposing detail *D2* on approximation *A2*. Finally, the original stock price is obtained by superimposing detail *D1* on approximation *A1*. Fig. 2 shows two different levels of approximation (identified by *A1–A2*) and detail (identified by *D1–D2*) of the window's opening, high, low, and closing prices (P_t). Wavelet transform acts like a mathematical microscope, zooming into small scales to reveal compactly spaced events in time, and zooming out into large scales to exhibit the global waveform patterns (Subasi, 2005a). The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the window's opening, highest, lowest, and closing prices in time and frequency.

4.3. Knowledge discovery process

The major task in the knowledge discovery process is training the two-level DWI-transformed SOM, namely DSOM, network. The first-level is a larger two dimensional, 10 * 10 SOM network, trained with the DWT coefficients extracted from the data transformation stage, and the resultant map forms the proto-clusters. The second-level is a smaller two-dimensional, $d_1 * d_2$, SOM and is trained with the proto-clusters in the first-level map. The parameters d_1 and d_2 are decided by the agglomeration coefficient of the well-known agglomerative hierarchical cluster (AHC) analysis. AHC groups data into a hierarchical structure according to the proximity of data. It begins with assigning each datum in a singleton cluster and is followed by subsequent one-by-one agglomeration stages of merging the most similar clusters until a stopping criterion is met, i.e. an appropriate or optimal clustering is achieved (Jain, Murty, & Flynn, 1999). For determining the optimal clustering in AHC, Everitt (1993) suggests a commonly used heuristic approach based on the analysis of the

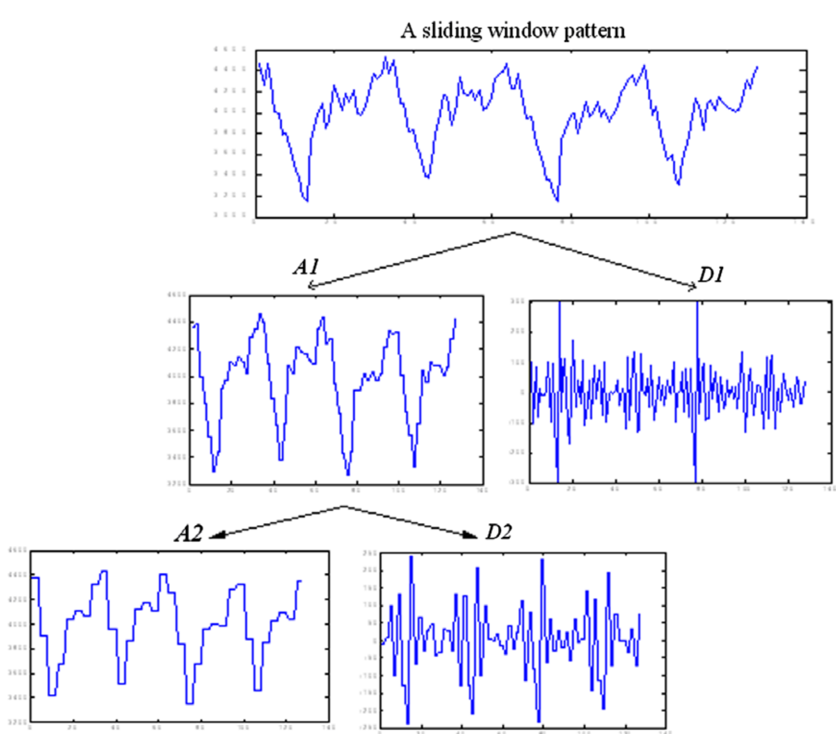

Fig. 2. The multi-resolution decomposition for a given P_t.

agglomeration schedule, which records the agglomeration process at each stage. One of the important elements in the schedule is the so-called agglomeration coefficient, the sum of the within-group variance of the two clusters merged at each successive stage. Such a coefficient indicates the degree of homogeneity of the clusters being merged at some stage. The smaller the value is, the more homogeneous the clusters are. Therefore, the optimal clustering result appears at the stage when there is the most significant increase in the change of the agglomeration coefficient at the next stage. To allow for relative comparisons, the percent change from one coefficient to the next was examined. Large percentage change values indicated the merging of two heterogeneous clusters and possible stopping points for the final cluster solution. The agglomeration coefficients and percentage change for clusters ranging from 10 down to 1 are illustrated in Table 2. The largest change arises from three to two, however, three clusters are too few to be subsequently analyzed. As a result, the second candidate, six clusters, was accepted as an optimal solution, and a topology of 2 * 3 was decided for the second level SOM.

4.4. Trajectory analysis

In order to explore more informative features embedded in the two-level DSOM network, we conduct a novel trajectory analysis. The first-level SOM network is trained by the sequences of wavelet-transformed sliding-window patterns. The successive BMUs on the level given the input patterns form a trajectory for recording the trend movement of the patterns, as shown in Fig. 3, where the X–Y plane denotes the coordinates of the 10 * 10 SOM map and coordinate Z represents the time's evolution, in which the trajectory coils up over time. It indicates that the trend moving around in a local region in the map represents fewer fluctuations in the stock market. On the other hand, when the movement is across homogeneous areas, a warning signal meaning significant market vibrations should be issued. Similarly, the series of BMUs forms another trajectory on the second-level 2 * 3 map, as shown in Fig. 4, which has much fewer fluctuations than the first-level SOM network. It can be observed that movement within the same or neighboring clusters implies the market belongs to the primary bull or the primary bear market. In summary, the trajectories on the first- and second-level maps provide detailed and abstractive views about the market, respectively. The trajectory analysis makes it easier to interpret stock movement process, uncover hidden financial knowledge, and facilitate the decision-making of trading strategies. A comprehensive investigation of these issues will be presented in Section 5.2.

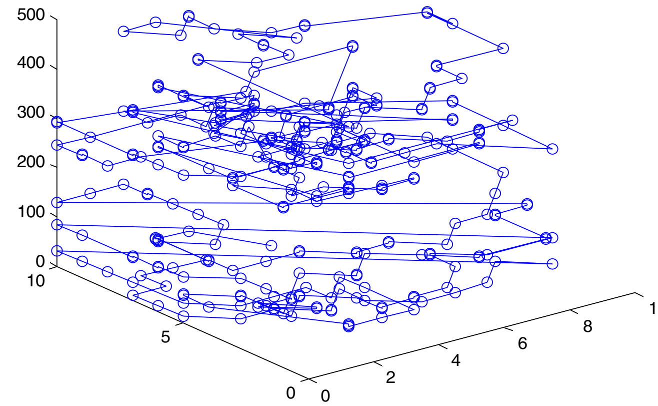

Fig. 3. The first-level SOM trajectory.

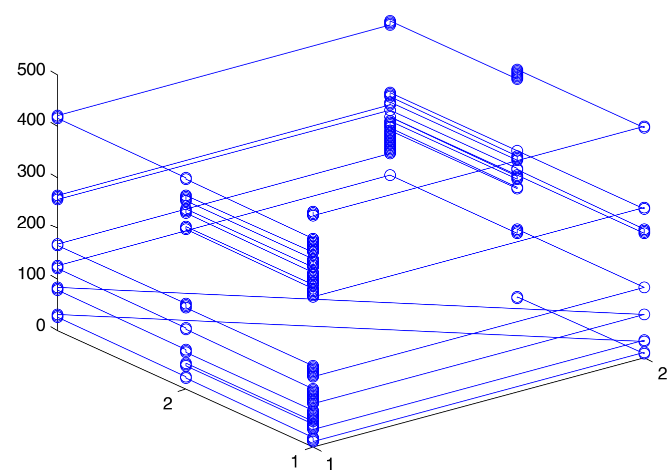

Fig. 4. The second-level SOM trajectory.

Table 2
The agglomeration coefficients in clusters

Number of clusters	Agglomeration coefficient	Percentage change in coefficient
10	776.18	10.67
9	858.97	12.66
8	967.71	14.46
7	1107.61	12.72
6	1248.46	16.22
5	1450.98	34.95
4	1958.04	29.06
3	2527.14	30.22
2	3290.89	102.41
1	6661.18	–

5. Experimental design and results

In this section, we discuss the experimental results and conduct performance evaluation in detail for the aspects of forecasting accuracy and profitability of trading strategies.

5.1. Generation of buying and selling signals

Different investors have different preferences about the holding periods of their investments. In this study, short-term traders are defined as those who care much about the vibrations of the underlying asset and get in and out of positions in a short time (one-step ahead), whereas trend followers are those who care more about sustained price movement direction rather than the fluctuations of the equity. Both preferences can be realized by the proposed two-level DSOM network.

5.1.1. Labeling one-step ahead signals

After training the first-level SOM network with 2783 patterns composed of the transformed openings, highs, lows and closings stock prices, each pattern is assigned to its BMU in the 10 * 10 SOM. Fig. 5 displays the distribution of DWT coefficient patterns, which confirms SOM's ability to cluster similar patterns. In order to reveal the trends of a real stock market, a wavelet reconstruction procedure is further performed and the resultant distribution is illustrated in Fig. 6, which provides a useful visualization tool for chartists. The chartist consulted in this study, who has more than 12 years experience in the analysis of stock investment, agrees that several interesting trends can be roughly identified which are worth further analysis, for example, *up-trends* (charts 78, 88–90, 98), *downtrends* (charts 1, 2, 11–12, 21–22), *rounding bottom* (charts 83–85, 95–97), *rounding top* (charts 8–10, 25–28). In addition, one-step ahead buying-and-selling trading signals, the important information for short-term investors, can be generated based on the rising or falling of the next day's stock price, defined as follows:

$$\text{One-step ahead Sig}(t) = \begin{cases} \text{Buy}, & \text{if } p^c(t) - p^c(t-1) > 0 \\ \text{Sell}, & \text{if Otherwise} \end{cases} \tag{8}$$

where $p^c(t)$ represents the closing price in day t. In this study, the length of the sliding window is 32 days, therefore, if the closing price of the 32nd day is higher than that of the thirty-first day, a buying signal is labeled as a red "o" otherwise a selling signal is labeled as a green "x" (see Fig. 6).

5.1.2. Labeling trend-following signals

By feeding the 10 * 10 trained weights of neurons in the first-level SOM map into the second-level SOM network, six clusters are obtained upon the completion of training, as discussed in Section 4.3. Fig. 7 illustrates the pattern distribution in the resultant map. With the help of the chartist, clusters one, two, and three show that the movement direction of the equity is downward, which suggests trend followers should take selling action over a long period of time. Other clusters, on the contrary, show the upward trend and investors should take buying action. In a sense, the market trend is divided into two categories: primary bull market (upward) and primary bear market (downward). By this, trend following signals can be labeled in the following way:

$$\text{Trend-following Sig}(t) = \begin{cases} \text{Sell}, & \text{if ClusterNumber } (t) < 4 \\ \text{Buy}, & \text{if ClusterNumber } (t) \geqslant 4 \end{cases} \tag{9}$$

We further project the clustering results in the second-level map onto the first-level map as shown in Fig. 8, i.e. charts 1–7, 11–16, 21–26, and 31–34 in the first-level map

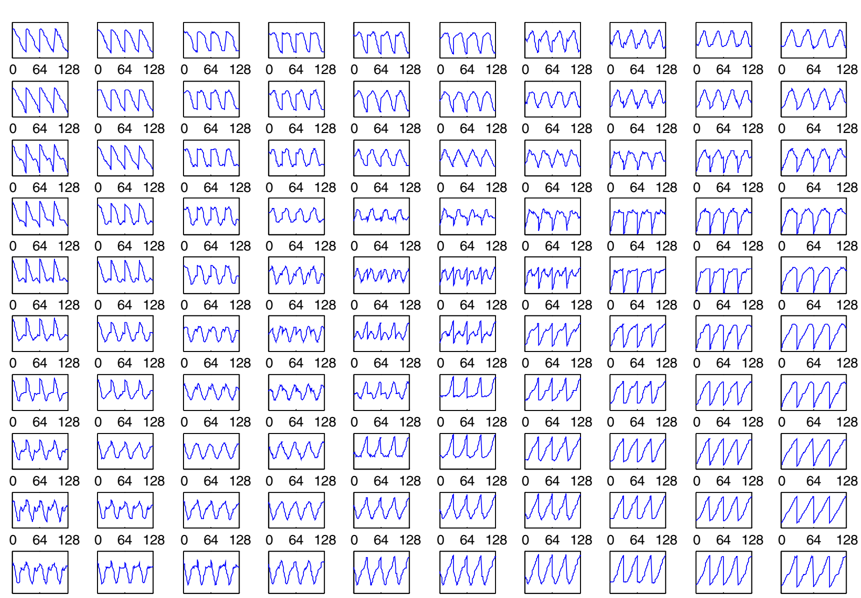

Fig. 5. The wavelet coefficients distribution in the first-level SOM.

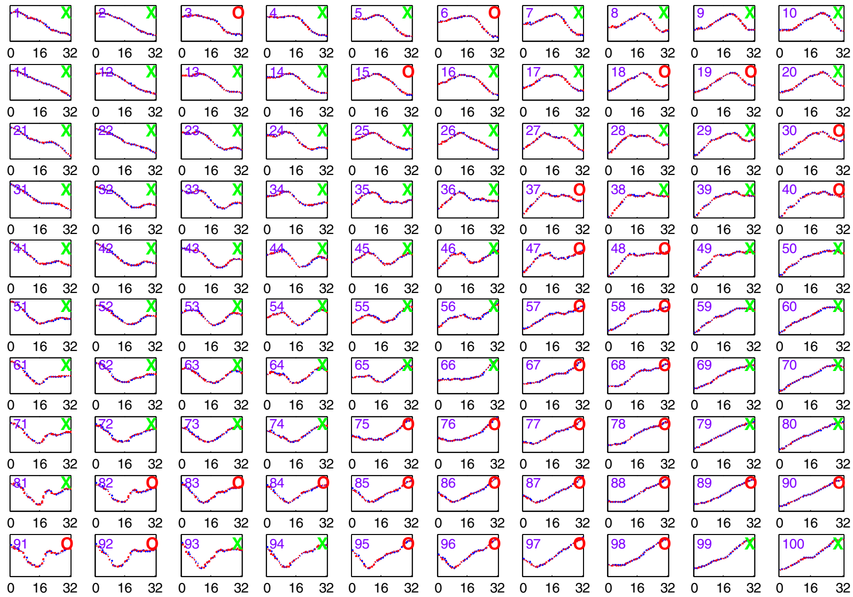

Fig. 6. The K-charts and one-step ahead signals in the first-level SOM.

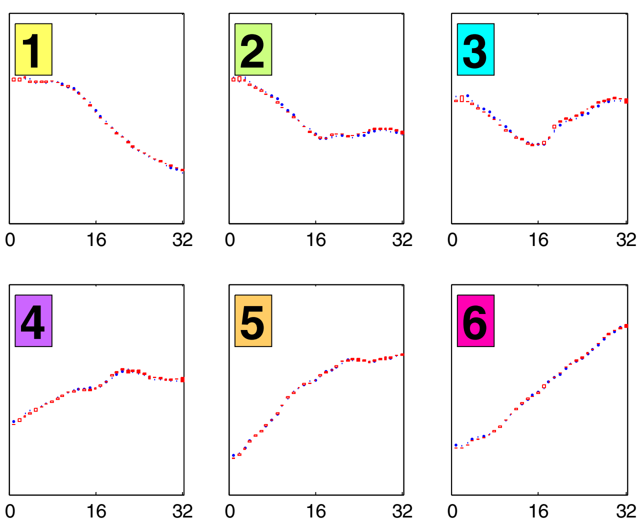

Fig. 7. The K-chart pattern distribution in the second SOM map.

are assigned to cluster one and so on. It is worth noting that the fluctuation of trading signals is much less sensitive to the market movement compared to the case of one-step ahead. The trend-following signals will thus help reduce the trading risk, cost, and tax associated with frequent trades.

5.2. Trading strategy making

After labeling trading signals, we proceed to validate the effectiveness of trajectory analysis using the test data set composed of 461 sliding-window patterns collected from the period of 2001/1/1 to 2002/12/31. The test patterns are matched with the patterns stored as BMU weights in the trained first-level and second-level SOM maps. Figs. 9 and 10 show the K-chart movement (blue color) and trajectory (red color) on both maps, and demonstrate how the trajectory reflects the trend of price movement. For example, the significant change in price at point A coincides with the dramatically variation in trajectory that was led from the shift from a primary bear market to a primary bull market. Furthermore, it is interesting to note that the fluctuation of the second-level SOM is much smaller than the first-level SOM. In the following, one-step ahead and trend-following buying-and-selling signals are labeled accordingly in Figs. 11 and 12, in which the signals agree with the market movement. These signals are very useful when making a trading strategy in the next stage.

Making a trading strategy concerns how to trade by using the aforementioned buy and sell signals. There are many possibilities for trading strategies according to the different considerations of investors. In this study, we consider three general strategies: a buying strategy, a selling strategy, and a buying and selling strategy.

(1) Buying strategy: Investors buy the stock and hold until the selling signal appears, and the short-term selling strategy is not considered.

Strategy-Buy(t)

$$= \begin{cases} \text{Buy}, & \text{if Sig}(t-1) = \text{Sell and Sig}(t) = \text{Buy} \\ \text{Hold}, & \text{if Sig}(t-1) = \text{Buy and Sig}(t) = \text{Buy} \\ \text{Sell}, & \text{if Sig}(t-1) = \text{Buy and Sig}(t) = \text{Sell} \end{cases} \quad (10)$$

(2) Selling strategy: Investors sell the stock until the buying signal appears and the short-term buying strategy is not taken into consideration.

Strategy-Sell(t)

$$= \begin{cases} \text{Buy}, & \text{if Sig}(t-1) = \text{Sell and Sig}(t) = \text{Buy} \\ \text{Hold}, & \text{if Sig}(t-1) = \text{Sell and Sig}(t) = \text{Sell} \\ \text{Sell}, & \text{if Sig}(t-1) = \text{Buy and Sig}(t) = \text{Sell} \end{cases} \quad (11)$$

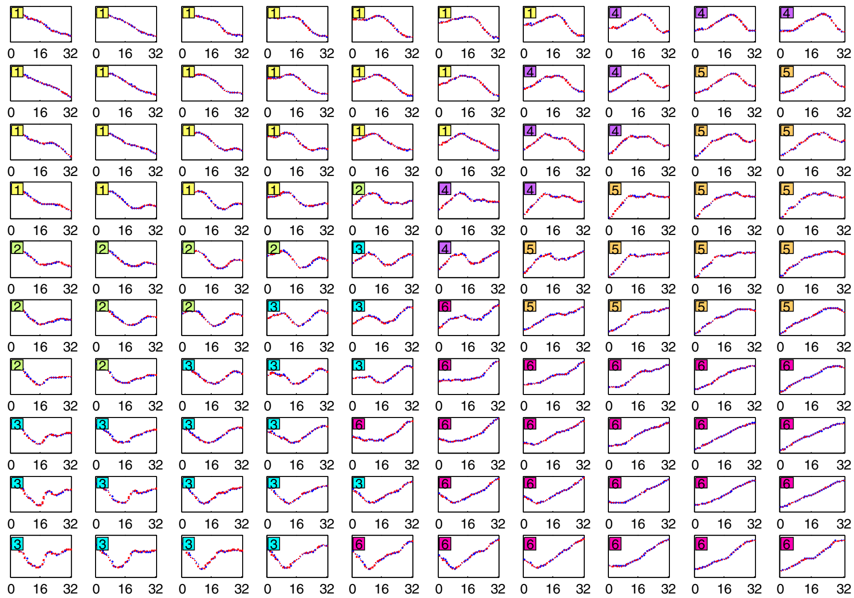

Fig. 8. Mapping trend-following signals onto the first-level SOM.

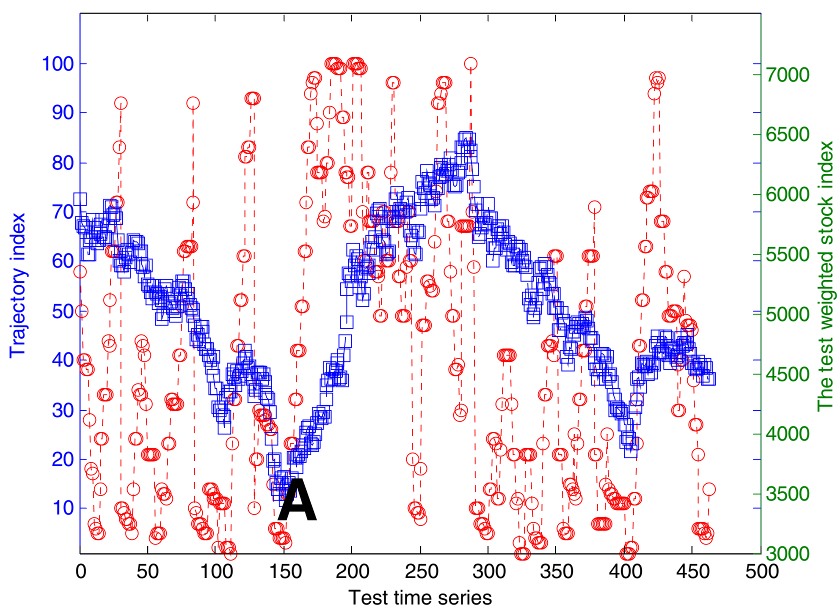

Fig. 9. Trajectory in the first-level SOM.

(3) Buying and selling strategy: It includes both short-term buying and selling trading strategies.

$$\text{Strategy-BS}(t) = \begin{cases} \text{Exit Short}, & \text{if } \mathrm{Sig}(t-1) = \text{Sell and } \mathrm{Sig}(t) = \text{Buy} \\ \text{Exit Long Position}, & \text{if } \mathrm{Sig}(t-1) = \text{Buy and } \mathrm{Sig}(t) = \text{Sell} \\ \text{Hold Long Position}, & \text{if } \mathrm{Sig}(t-1) = \text{Buy and } \mathrm{Sig}(t) = \text{Buy} \\ \text{Hold Short Position}, & \text{if } \mathrm{Sig}(t-1) = \text{Sell and } \mathrm{Sig}(t) = \text{Sell} \end{cases} \tag{12}$$

In general, an investor chooses an appropriate trading strategy on the basis of his/her individual plan, risk management and future needs for capital, however, the fundamental principle of buying at the lowest point and selling at the highest is always followed. We apply these trading strategies to the test data set in the first and second SOM and the results are depicted in Figs. 13 and 14. Both show that the timing of selling and buying adheres to the fundamental principle. Moreover, the trading times of the second-level SOM, 11, is much fewer than the first one, 62, and thus incurs lower trading costs.

Table 3 illustrates the trading process in the second-level SOM, i.e. trend-following, using the three strategies. The second, third, and fourth columns are the trading time, buying price and selling price, respectively. According to the selling strategy, the fifth column shows that it carried

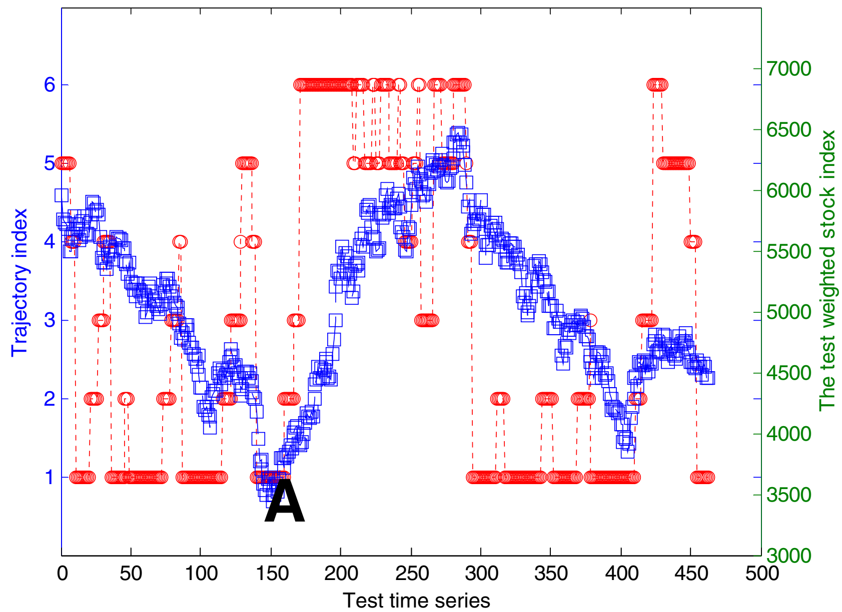

Fig. 10. Trajectory in the second-level SOM.

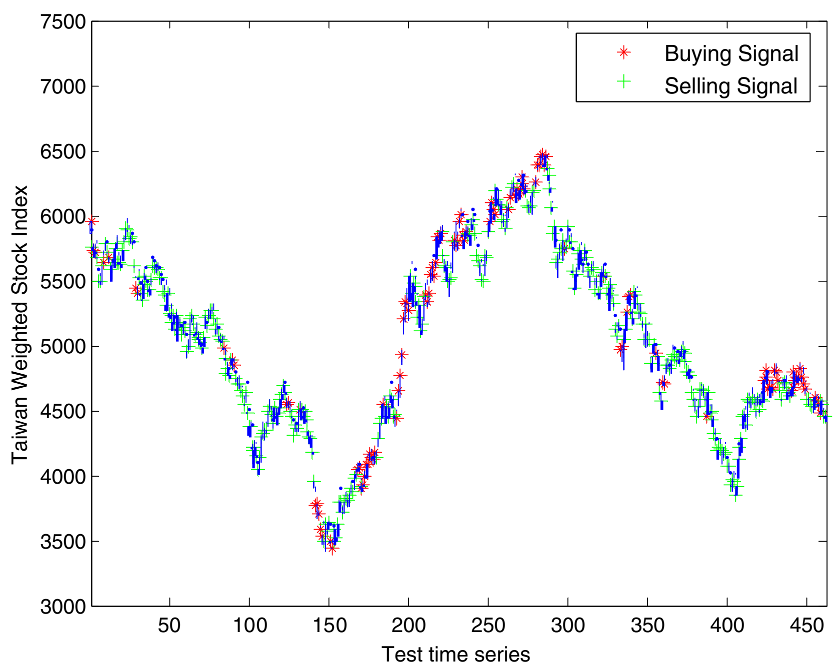

Fig. 11. Trading signals in the first-level SOM.

out five transactions, of which three were profitable and two times were not (in red color), and the total profit was 2366.7 (Taiwan Dollar, NTD). Similarly, columns six and seven show that buying and buying-and-selling strategies achieve total profits 1004.9 and 3371.7, respectively.

We also considered another popular trading strategy, i.e. buy-and-hold, in which the stock was bought at price 5950 on January 1, 2001 and then was held until December 31, 2002 and sold at price 4452.45. The profit is thus equal to −497.55. Therefore, from the viewpoint of total profit, the ranking from the least to the most for these four strategies is buy-and-hold strategy (−497.55), buying strategy (1004.9), selling strategy (2366.7), and buying and selling strategy (3371.7). The profit comparison is confirmed by Fig. 15. On the other hand, similar results are obtained for one-step ahead DSOM, as shown in Fig. 16. Both show that the buying and selling strategy achieves the best earning returns and the differences become more obvious with time. Therefore, the following performance evaluation to be discussed will be based on this strategy.

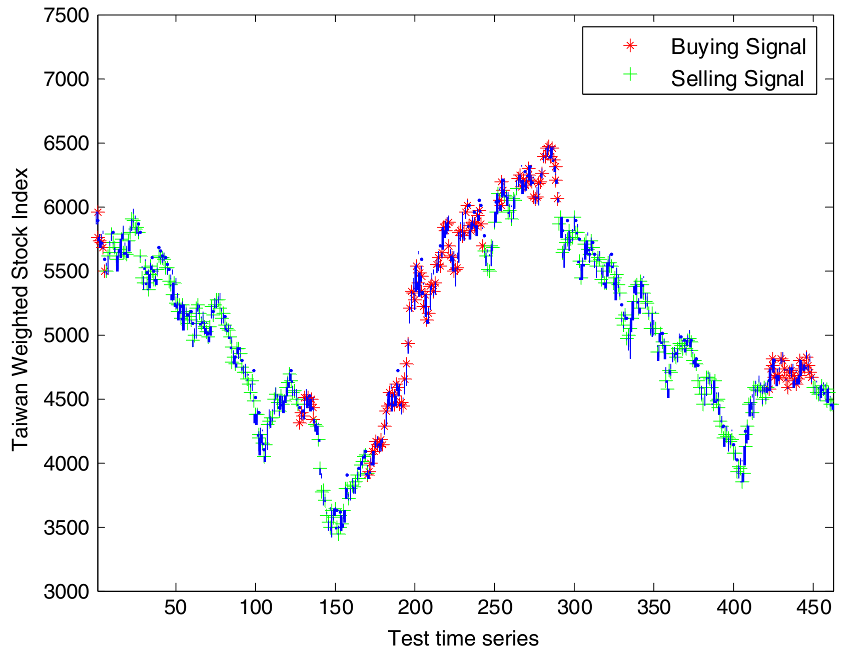

Fig. 12. Trading signals in the second-level SOM.

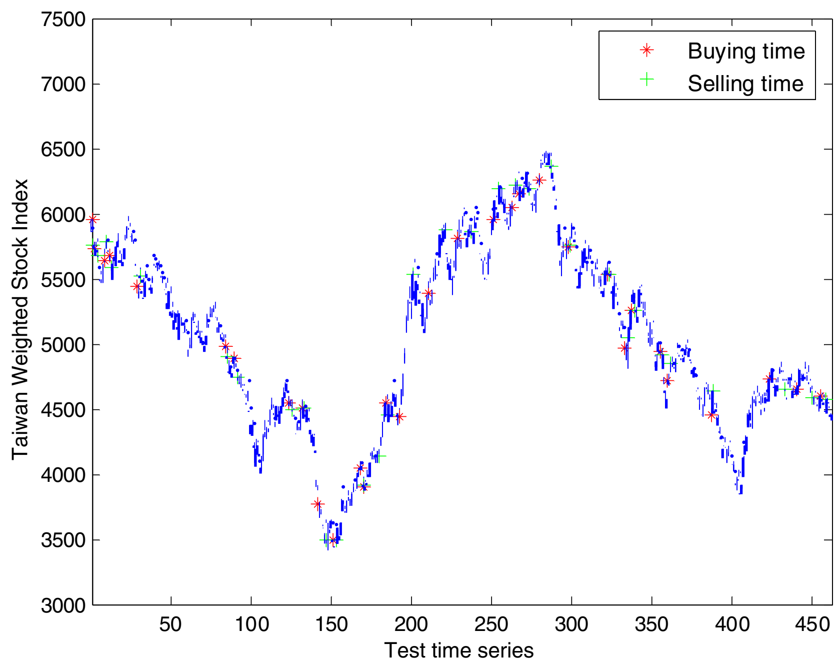

Fig. 13. The trading time in the first-level SOM.

6. Analyses of forecasting and trading strategies

In this section, we evaluate and analyze the forecasting ability and trading profitability of the proposed DSOM network. In order to compare the performance with the traditional model in detail, a variety of performance metrics are calculated which are defined as follows.

Signal correct rate (SCR): This value is the ratio of correct labeling of buying and selling signals, the greater, the better:

$$\mathrm{SCR} = \frac{C}{N} \tag{13}$$

where N is the total number of buying and selling signals labeled, and C is the number of buying and selling signals correctly labeled.

Gain rate (GR): This value is the probability of successful trades, the greater the better.

$$\mathrm{GR} = \frac{n_g}{n} \tag{14}$$

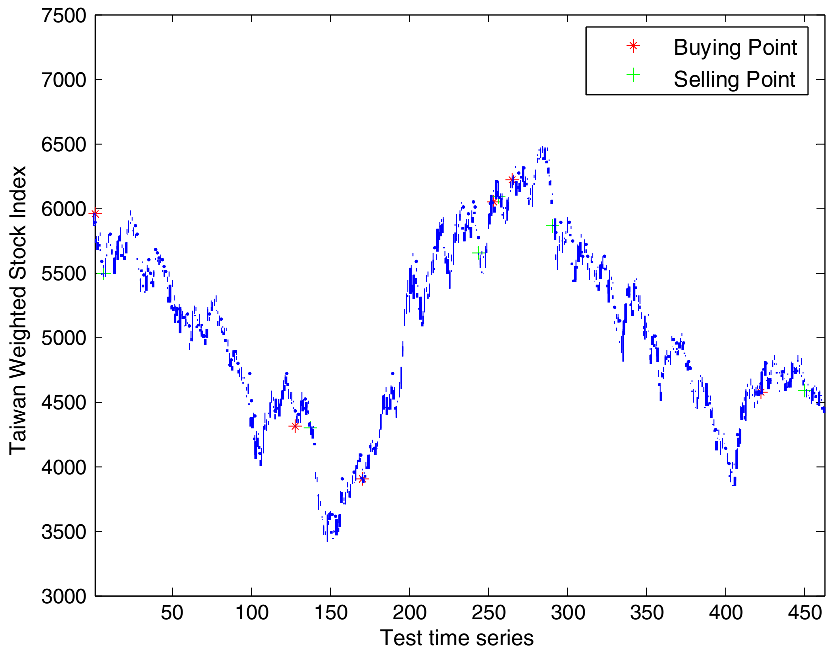

Fig. 14. The trading time in the second-level SOM.

Table 3
The long-term trading strategies on the second-level SOM in DSOM (Unit: NTD)

Trading number	Trading time	Buy price	Sell price	Selling strategy	Buying strategy	Buying and selling strategy
1	2001/1/2	5950.0	0.0	0.0	−450.4	−450.4
2	2001/1/10	0.0	5499.5	1189.2	0.0	1189.2
3	2001/7/12	4310.3	0.0	0.0	−8.2	−8.2
4	2001/7/26	0.0	4302.2	398.7	0.0	398.7
5	2001/9/13	3903.5	0.0	0.0	1752.6	1752.6
6	2001/12/31	0.0	5656.1	−392.0	0.0	−392.0
7	2002/1/14	6048.2	0.0	0.0	40.2	40.2
8	2002/1/18	0.0	6088.3	−130.8	0.0	−130.8
9	2002/1/30	6219.2	0.0	0.0	−351.3	−351.3
10	2002/3/19	0.0	5867.8	1301.7	0.0	1301.7
11	2002/9/23	4566.1	0.0	0.0	22.0	22.0
12	2002/11/1	0.0	4588.1	0.0	0.0	0.0
End time	2002/12/31	4452.45				
Total profit		−497.55		2366.7	1004.9	3371.7

where n is total trade times, n_g is total gain times. Generally speaking, the gain rate of one-step ahead trading is higher than with trend-following trading.

Total profit (TP): This value represents the profitability of the total trades. A profit is made when the value is greater than 0. The higher the value is, the better.

$$TP = G - L \tag{15}$$

where G is the total profit which is not written off with losses and is called the gross gain. L is the total loss which is not written off with gains, and is called gross loss.

Average profit (AP): The average profit obtained per trade, the higher the better.

$$AP = \frac{TP}{n} = \frac{G - L}{n} \tag{16}$$

Average gain (AG): The average profit per successful trade, the greater the better.

$$AG = \frac{G}{n_g} \tag{17}$$

Average ratio of gain to loss ($A_{G/L}$): This value is the ratio of average gains to average losses

$$A_{G/L} = \frac{AG}{AL} = \frac{n_l}{n_g} \tag{18}$$

where n_l is total loss times. When the ratio is greater than 1, there are more profitable trades than losing ones. When the ratio is less than 1, there are more losing trades than profitable ones.

Fig. 15. Total profits comparison of the trend-following approach.

Fig. 16. Total profits comparison of the one-step ahead approach.

Profit Factor (PF): This value is the ratio of the profit generated by profitable trades to the losses generated by losing trades.

$$\mathrm{PF} = \frac{G}{L} \tag{19}$$

In general, a two or higher PF is acceptable, and a higher PF value indicates less risk.

Pessimistic Ratio of Return (PRR): A performance metric proposed by Vince (1990), which is a statistical adjustment of the wins to losses ratio for estimating the worst return from trading results.

$$\mathrm{PRR} = \frac{(n_g - \sqrt{n_g}) \cdot G/n_g}{(n_l - \sqrt{n_l}) \cdot G/n_l} \tag{20}$$

6.1. Analysis of prediction accuracy

The traditional measure of prediction performance mostly uses mean squared error (MSE) or mean absolute error (MAE) on the actual and predicted closing prices, however Pesaran and Timmermann (1995) suggested that this measure may not be strongly related to trading profits. Instead, measuring the correct rate of buying and selling signals and gain rate is more effective, since it reflects the ratio of profit obtained. Thus, we adopt the signal correct rate and gain rate as the criteria of evaluating the accuracy for one-step ahead and trend-following prediction using the buying and selling trading strategies on the traditional SOM and DWT-transformed SOM models (DSOM). In addition, a statistical test is conducted to confirm the significance of the accuracy rate. The summary of evaluations is listed in Table 4.

As shown in Table 4, the signal correct rates for all strategies and models are above 50%, in which DSOM with the one-step ahead trading strategy achieves the highest rate (70.02%). The proportion test (Z-test) is then used to test the significant differences between correct rates and 0.5 at the 95% confidence level. The null hypothesis of the signal correct rate was calculated by conducting a one-sided test of H_0: $p \leqslant 0.5$ against H_1: $p > 0.5$. The results indicate that all strategies are significantly different and marked with symbol a. Therefore, one-step ahead and trend-following trading signals and are both capable of producing desirable accuracies in accordance with the work in Enke and Thawornwong (2005). Similarly, the gain rates are all over 50%, and DSOM with the one-step ahead trading obtains the highest rate (64.52%). Furthermore, the gain rate of the one-step ahead trading strategy is shown as reliable with respect to the Z-value, however, the one for the trend-following strategy is negative due to too few trading times (11 or 12 times).

6.2. Analysis of profits

The evaluation result of profit performance using various metrics is illustrated in Table 5. It demonstrates that the basic capability of earning profits for DSOM and traditional SOM models is witnessed by the fact of positive total and average profits. In particular, DSOM performs better than SOM in either one-step ahead (4940.95 NTD vs. 4431.35) or trend-following trading strategy (3230.95 v.s. 3021.55) in terms of total profit. Similarly, DSOM outperforms SOM in one-step ahead (79.69 vs. 63.31) and trend-following (306.51 vs. 251.80) trading strategies as well regarding the average profit. This confirms the contribution of DWT to the forecasting model.

Next, we consider the performance comparison of one-step ahead and trend-following trading strategies for different forecasting models. Symbol a marked in each row in Table 5 indicates which one is better than the other with respect to the metric in the row. It is worth noting that the trend-following case is always better than the one-step ahead case for all indexes except the total profit. Although the one-step ahead approach gives more total profit, it also has far more trades. When transaction costs and taxes are taken into consideration, this will result in higher costs for one-step ahead trading. In the long term, it is possible that the transaction costs might cancel out the profits earned. Thus, the trend-following trading approach is viewed as a better way to reduce transaction costs and increase profits.

Table 4
The summary of prediction accuracies

Trading strategy	One-step ahead trading strategy		Trend-following trading strategy	
	SOM	DSOM	SOM	DSOM
Signal correct rate (%)	64.83	70.02	65.70	59.33
SCR (Z-value)	6.3750[a]	8.6060[a]	6.7491[a]	4.0107[a]
Gain times	44	40	7	6
Loss times	26	22	5	5
Total trading times	70	62	12	11
Gain rate (%)	62.86	64.52	58.33	55.00
Gain rate (Z-value)	2.1519[a]	2.4297[a]	1.3939	0.3320

[a] Statistically significant at 5% confidence level.

Table 5
The summary of profit performance (Unit: NTD)

Trading strategy	SOM		DSOM	
	One-step ahead strategy	Trend-following strategy	One-step ahead strategy	Trend-following strategy
Total times	70	62	12	11
Total profit	4431.35[a]	3021.55	4940.95[a]	3230.95
Avg. profit	63.31	251.80[a]	79.69	306.51[a]
Avg. gain	187.38	551.21[a]	159.06	777.40[a]
Avg. G/L	1.28	3.29[a]	1.26	2.91[a]
Max. gain	804.80	1598.90[a]	1098.20	1752.60[a]
Max. loss	−473.30	−349.30[a]	−549.90	−450.40[a]
Gross loss	−3813.40	−836.90[a]	−2769.20	−1332.80[a]
Profit factor	2.16	4.61[a]	2.30	3.49[a]
PRR	2.28	5.19[a]	2.46	3.75[a]

[a] The better metric achieved.

7. Conclusions and future work

With the coming of the inflation era, investment in stocks has been of great importance for investors. However, the diversity and complication of domain knowledge existing in the stock market make it very difficult for investors to make the right decisions promptly. In this study, we apply knowledge discovery in the TAIEX stock database to develop an intelligent decision support model to provide investors with systematic trading strategies. The knowledge discovery process is mainly composed of feature representation by visual K-chart technical analysis, feature extraction by discrete wavelet transform, and modeling by two-level SOM networks. More importantly, a visual trajectory analysis is conducted to reveal the relationship of movements between primary bull and bear markets and help determine appropriate trading strategies. We take the different needs of short-term investors and trend followers in trading strategies into consideration, and compare the proposed model to the traditional SOM model in terms of forecasting accuracy and profitability.

According to analysis of the experimental results, both SOM models are capable of producing satisfactory forecasting accuracies and gain rates, either with the one-step ahead and trend-following trading strategies. In particular, DSOM outperforms SOM with respect to the signal correct rate and gain rate for the one-step ahead strategy, which is due to the contribution of DWT. The comparison of total trading times coincides with the nature of trend followers who prefer buy a position in the stock and hold it until the trend direction is changed. As far as the paramount objective of financial investment, maximizing profits and minimizing losses, is concerned, DSOM surpasses SOM in a variety of performance metrics. An interesting pattern was observed in that the total profit earned by the trend-following strategy is not as good as by the one-step ahead strategy. Nevertheless, the trend-following one is a more reliable investment approach when considering lower risk, transaction cost, and tax. With the findings of this study, it is expected that the proposed decision model can helps investors, and fund managers make the most profitable decisions.

One of the major limitations of the study is the data collection. The current simulated data set is TAIEX, which is not a tradable specific equity, therefore, the decision outcomes from the proposed model can only be a reference model for investors. To be more practical, however, the model can be applicable to tradable futures by considering the additional costs, such as transactional cost and tax, and other factors or indexes impacting the stock market, for example, volume, exchange rate, wholesale price index, and consumer confidence index. Another interesting future work is to enhance the decision model, by utilizing a rule extraction mechanism, with explanatory ability, so that financial experts can benefit from the ability of verifying or refining the tacit investment knowledge offered by the facts revealed about stock market movements.

Acknowledgements

The authors thank Mr. Men-Hsieu Ho for his help in the preliminary experimentation. This work was supported in part by NSC92-2416-H-006-044, Taiwan, ROC.

References

Abhyankar, A., Copeland, L. S., & Wong, W. (1997). Uncovering nonlinear structure in real-time stock-market indexes: the S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *Journal of Business and Economic Statistics, 15*, 1–14.

Apte, C. B., Liu, B., Pednault, E., & Smyth, P. (2002). Business applications of data mining. *Communications of the ACM, 45*(8), 49–53.

Armano, G., Marchesi, M., & Murru, A. (2005). A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences, 170*(1), 3–33.

Beltra'n, N. H., Duarte-Mermoud, M. A., Bustos, M. A., Salah, S. A., Loyola, E. A., Penã-Neira, A. I., et al. (2006). Feature extraction and classification of Chilean wines. *Journal of Food Engineering, 75*, 1–10.

Boginski, V., Butenkob, S., & Pardalos, P. M. (2006). Mining market data: a network approach. *Computers and Operations Research, 33*, 3171–3184.

Boginski, V., Butenko, S., & Pardalos, P. M. (2005). Statistical analysis of financial networks. *Computational Statistics and Data Analysis, 48*(2), 431–443.

Bo, L., Linyan, S., & Mweene, R. (2005). Empirical study of trading rule discovery in China stock market. *Expert Systems with Applications, 28*, 531–535.

Bose, I., & Mahapatra, R. K. (2001). Business data mining: a machine learning perspective. *Information and Management, 39*(3), 211–225.

Chen, M. S., Han, J., & Yu, P. S. (1996). Data mining: an overview from database perspective. *IEEE Transactions on Knowledge and Data Engineering, 8*(6), 866–883.

Chen, S. H., & He, H. (2003). Searching financial patterns with self-organizing maps. *Computational Intelligence in Economics and Finance*, 203–216.

Chen, A. S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan stock index. *Computers and Operations Research, 30*, 901–923.

Chi, S. C., Peng, W. L., Wu, P. T. & Yu, M. W. (2003). The study on the relationship among technical indicators and the development of stock index prediction system. *Fuzzy Information Processing Society, 22nd International Conference of the North American* (pp. 291–296).

Chun, S. H., & Kim, S. H. (2004). Data mining for financial prediction and trading: application to single and multiple markets. *Expert Systems with Applications, 26*, 131–139.

Chun, S. H., & Park, Y. J. (2005). Dynamic adaptive ensemble case-based reasoning: application to stock market prediction. *Expert Systems with Applications, 28*, 435–443.

Coifman, R. R., & Wickerhauser, M. V. (1992). Entropy-based algorithms for best-basis selection. *IEEE Transactions on Information Theory, 38*, 713–718.

Deboeck, G. J. (1998). Financial applications of self-organizing maps. *Neural Network World, 8*, 213–241.

Deboeck, G. J., & Kohonen, T. (1998). *Visual explorations in finance with self-organizing maps*. London: Springer-Finance.

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications, 29*, 927–940.

Everitt, B. (1993). *Cluster analysis* (3rd ed.). New York: Halsted.

Evstigneev, I., Hens, T., & Schenk-Hopp, K. R. (2006). Evolutionary stable stock markets. *Economic Theory, 27*, 449–468.

Fang, Y., & Xu, D. (2003). The predictability of asset returns: an approach combining technical analysis and time series forecasts. *International Journal of Forecasting, 19*, 369–385.

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

Feelders, A., Daniels, H., & Holsheimer, M. (2000). Methodological and practical aspects of data mining. *Information and Management, 37*, 271–281.

Flexer, A. (2001). On the use of self-organizing maps for clustering and visualization. *Intelligent Data Analysis*(5), 373–384.

Franses, P. H., & Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting, 15*(1), 1–9.

Gafiychuk, V. V., Datsko, B. Yo., & Izmaylova, J. (2004). Analysis of data clusters obtained by self-organizing methods. *Physica A, 341*, 547–555.

Goswami, J. C., & Chan, A. K. (1999). *Fundamentals of wavelets: theory, algorithms, and applications*. Wiley Publishers, pp. 149–152.

Hansen, J. V., & Nelson, R. D. (2002). Data mining of time series using stacked generalizers. *Neurocomputing, 43*(1–4), 173–184.

Hiemstra, C., & Jones, D. (1994). Testing for linear and nonlinear granger causality in the stock price–volume relation. *Journal of Finance, XLIX*, 1639–1664.

Irma, B. F., Stelios, H. Z., & Steven, W. (2002). Knowledge discovery techniques for predicting county investment risk. *Computers and Industrial Engineering, 43*, 787–800.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys, 31*(3), 264–323.

Kandaswamy, A., Kumar, C. S., Ramanathan, R. P., Jayaraman, S., & Malmurugan, N. (2004). Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine, 34*(6), 523–537.

Kim, K. J., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications, 19*(2), 125–132.

Kiviluoto, K. & Bergius, P. (1998). Two-level self-organizing maps for analysis of financial statements. In *Proceedings of the 1998 IEEE International Joint Conference on Neural Networks (IJCNN'98), 1*, 189–192.

Kohonen, T. (1997). *Self-organizing maps*. Berlin, Heidelberg: Springer-Verlag.

Lee, K. H., & Jo, G. S. (1999). Expert system for predicting stock market timing using a candlestick chart. *Expert Systems with Applications, 16*, 357–364.

Lee, S., Suh, Y., Kim, J., & Lee, K. A. (2004). Cross-national market segmentation of online game industry using SOM. *Expert Systems with Applications, 27*, 559–570.

Leigh, W., Modani, N., Purvis, R., & Hightower, R. (2004). A computational implementation of stock charting: abrupt volume increase as signal for movement in New York stock exchange composite index. *Decision Support Systems, 37*, 515–530.

Leigh, W., Modani, N., Purvis, R., & Roberts, T. (2002). Stock market trading rule discovery using technical charting heuristics. *Expert Systems with Applications, 23*, 155–159.

Leigh, W., Paz, N., & Purvis, R. (2002). Market timing: a test of a charting heuristic. *Economics Letters, 77*, 55–63.

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

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

Li, S. T., & Shue, L. Y. (2004). Data mining to aid policy making in air pollution management. *Expert Systems with Applications, 27*(3), 331–340.

Martı'n-del-Brı'o, B., & Medrano, N. (1995). Feature map architectures for pattern recognition: techniques for automatic region selection. In D. W. Pearson, N. C. Steele, & R. F. Albrecht (Eds.), *Artificial neural nets and genetic algorithms* (pp. 124–127).

Mitra, S., Pal, S. K., & Mitra, P. (2002). Data mining in soft computing framework: a survey. *IEEE Transactions on Neural Networks, 13*(1), 3–14.

Mörchen, F. (2003). Time series feature extraction for data mining using DWT and DFT. Department of Mathematics and Computer Science Philips-University Marburg, Technical Report, 33.

Moshou, D., Hostens, I., Papaioannou, G., & Ramon, H. (2005). Dynamic muscle fatigue detection using self-organizing maps. *Applied Soft Computing, 5*, 391–398.

Moshou, D., & Ramon, H. (2004). Financial applications of wavelets and self-organizing maps. *Computational Intelligence in Economics and Finance*, 234–249.

Murtagh, F., Starck, J. L., & Renaud, O. (2004). On neuro-wavelet modeling. *Decision Support Systems, 37*, 475–484.

Nison, S. (2001). *Japanese candlestick charting* (2nd ed.). Prentice Hall Press.

Núñez, J., & Llacer, J. (2003). Astronomical image segmentation by self-organizing neural networks and wavelets. *Neural Networks, 16*(3–4), 411–417.

Oh, K. J., & Kim, T. Y. (2007). Financial market monitoring by case-based reasoning. *Expert Systems with Applications, 32*(3), 789–800.

Oja, E., & Kaski, S. (1999). *Kohonen maps*. Amsterdam, Holland: Elsevier.

Penn, B. S. (2005). Using self-organizing maps to visualize high-dimensional data. *Computers and Geosciences, 31*, 531–544.

Percival, D. B., & Walden, A. T. (2000). *Wavelet methods for time series analysis*. Cambridge University Press.

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: robustness and economic significance. *Journal of Finance, 50*, 1201–1227.

Petrosian, A., Prokhorov, D., Homan, R., Dashei, R., & Wunsch, D. (2000). Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG. *Neurocomputing, 30*, 201–218.

Pyle, D. (1999). *Data preparation for data mining*. Morgan Kaufman.

Resta, M. (2000). ATA: the artificial technical analyst building intra-day market strategies. *Knowledge-Based Intelligent Engineering Systems and Allied Technologies, Fourth International Conference, 2*, 729–732.

Sarantis, N. (2001). Nonlinearities, cyclical behavior and predictability in stock markets: international evidence. *International Journal of Forecasting, 17*(3), 459–482.

Shen, L., & Loh, H. T. (2004). Applying rough sets to market timing decisions. *Decision Support Systems, 37*, 583–597.

Stankovic, R. S., & Falkowski, B. J. (2003). The Haar wavelet transform: its status and achievements. *Computers and Electrical Engineering, 29*, 25–44.

Starck, J. L., & Murtagh, F. (2002). *Astronomical image and data analysis*. Berlin: Springer-Verlag.

Subasi, A. (2005a). Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Systems with Applications, 28*, 701–711.

Subasi, A. (2005b). Epileptic seizure detection using dynamic wavelet network. *Expert Systems with Applications, 29*, 343–355.

Thawornwong, S., & Enke, D. (2004). The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing, 56*, 205–232.

Torres, H. M., Gurlekian, J. A., Rufiner, H. L., & Torres, M. E. (2006). Self-organizing map clustering based on continuous multi-resolution entropy. *Physica A, 361*, 337–354.

Ture, M., & Kurt, I. (2006). Comparison of four different time series methods to forecast hepatitis A virus infection. *Expert Systems with Applications, 31*, 41–46.

Vellido, A., Lisboa, P. J. G., & Meehan, K. (1999). Segmentation of the on-line shopping market using neural networks. *Expert Systems with Applications, 17*, 303–314.

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

Vince, R. (1990). *Portfolio money management formulas: mathematical trading methods for the futures, options, and stock markets*. John Wiley and Sons, p. 1.

Wang, Y. F. (2002). Predicting stock price using fuzzy grey prediction system. *Expert Systems with Applications, 22*(1), 33–39.

Wang, Y. F. (2003). Mining stock price using fuzzy rough set system. *Expert Systems with Applications, 24*, 13–23.

Wang, J. L., & Chan, S. H. (2006). Stock market trading rule discovery using two-layer bias decision tree. *Expert Systems with Applications, 30*, 605–611.

Wang, H., & Weigend, A. S. (2004). Data mining for financial decision making. *Decision Support Systems, 37*, 457–460.

Wu, M. C., Lin, S. Y., & Lin, C. H. (2005). An effective application of decision tree to stock trading. *Expert Systems with Applications, 31*(2), 1–5.

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159–175.

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