

Research Article

A Study on Stock Graph Recognition Based on Wavelet Denoising and DTW Algorithm

Ji Wei Luo 

Shanxi University of Finance & Economics, No. 696, Wucheng Road, Taiyuan, Shanxi, China

Correspondence should be addressed to Ji Wei Luo; 201805040119@stu.sxufe.edu.cn

Received 7 November 2020; Revised 31 December 2020; Accepted 13 January 2021; Published 27 January 2021

Academic Editor: Adrian Neagu

Copyright © 2021 Ji Wei Luo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The “classical pattern” of stock price formation has long been widely used in the determination of future price trends of stocks, and the identification and analysis of classical price patterns have an important guiding role in investors’ decision-making and trading. The wavelet transform is a useful tool to remove some of the noise of time series because it has the characteristic of multiresolution. In this study, we propose a method for stock price pattern recognition based on the wavelet transform and dynamic time warp (DTW). A pattern recognition method with similar quantified results is developed to obtain accurate pattern recognition results. That is, using the wavelet transform to smooth the original price graph, and then using the DTW algorithm improved in this study to find the graph with the smallest distance from the target graph under the sliding window method, the identification and analysis of the target graph can be realized. In order to improve the recognition rate of the target graph, we preprocessed the raw price sequence using the moving average convergence and divergence (MACD) algorithm based on the control experiments set up in this study. The pattern recognition method used in this study will identify the price patterns of a certain time window as a whole, thus avoiding the problem of how to objectively select the important points that constitute a price pattern and the mathematical definition of different price patterns in the previous traditional methods.

1. Introduction

Technical analysis based on trading pattern curves is an important method for quantitative analysis of stocks, and in recent years, there have been a number of scholars who have proved the effectiveness of technical analysis from different perspectives as an effective means of extracting information from market prices [1–3]. With the development of financial engineering, computer technology, and mathematical algorithms, technical analysts using traditional trading model analysis methods now regularly use statistical software packages and other data processing techniques. The use of data analytics can minimize or even completely eliminate human subjective influences and turn trading into a scientific practice. Therefore, it is important to find a method that can quickly and efficiently identify trading pattern curves automatically to help investors grasp the volatility patterns of the stock market.

Existing stock pattern recognition methods can be divided into rule-based pattern recognition and template-based pattern recognition, both of these methods are based

on the extraction of important points and are subjective in nature. Therefore, this study attempts to use the overall graphical identification, that is, the overall identification of the price pattern of a certain time window the above-mentioned approach can avoid the problem of how to objectively select the important points that constitute the price pattern and the mathematical definition of different price patterns. Based on this, this study selects an arbitrary time window from the A-share market as the target graph, which avoids mathematical definition of price patterns and increases the universality of the method. The appropriate wavelet parameters are then selected by empirical analysis and the original price series is smoothed using the wavelet transform. An improved DTW algorithm is used to accurately identify the target pattern by finding the stock among the remaining stocks that is most similar to the target pattern. Scholars also recently used machine learning or other intelligent algorithms for financial time series graph identification, but the basic approach is consistent with the abovementioned two categories [4–6].

The main tasks of this study include the following: (1) the entropy weighting method is used to determine the weights of the two indicators of information extraction rate and information extraction error, and then to determine the composite indicators according to their weights, and finally to determine the number of optimal decomposition layers of the wavelet and the selection of the wavelet threshold function based on the value of the composite indicators. (2) The DTW algorithm is improved to overcome the problem that the algorithm itself is far away from similar shapes with certain phase differences. This study uses an improved DTW algorithm with a sliding time window algorithm to find the most similar time window to the target graph. (3) Attempt to compare the average rise and fall of the target graph to its similar graph sixty days later and determine that it is a realistic guide.

The overall structure of this article is as follows. Section 1 is an introduction that reviews the background of graphic recognition technology and the main work of this study. Section 2 presents work on graph recognition. Section 3 briefly introduces wavelet denoising and the theoretical foundations of MACD metrics and explains the principles of the wavelet threshold function. Section 4 describes technological improvements related to the characteristics of financial time series. Section 5 empirically determines the number of layers of the wavelet optimal decomposition as well as the threshold function and the graphs most similar to the target graphs found by the abovementioned methods. Section 6 compares the optimal graph with the forty-day rise and fall of the target graph. Finally, we conclude this study in Section 7. The specific research ideas are shown in Figure 1.

2. Related Work

The current research on automatic price pattern recognition has fallen into two main categories as rule-based pattern recognition and template-based pattern recognition.

Lo et al. [7] were the first to mathematically define ten price patterns such as head-and-shoulders and double-bottoms using the partial pole method. He proposed a systematic, automated method for technical pattern identification using nonparametric kernel regression and applies the method to a large number of U.S. stocks from 1962 to 1996 to evaluate its effectiveness for technical analysis. The abovementioned method creatively realizes the automatic recognition of price patterns by computer. Suh et al. [8] defined a Bezier curve as consisting of seven significant points and then identifies the pattern of the “cup with handle” by calculating the correlation coefficient between the time window and the Bezier curve. Savin et al. [9] used the pattern recognition algorithm of Andrew W with some modifications to determine whether “head-and-shoulders” (HS) price patterns have predictive power for future stock returns. Zapranis and Tsinaslanidis [10] represented the ideal circular base as an isosceles triangle formed by three key points and then used the two congruent circles on the outer ring of this triangle to control price fluctuations. The final two-dimensional coordinates of price and time were used for automatic pattern recognition.

Template-based pattern recognition is the network lattice method represented by Leigh et al. [11]. Leigh represented the flag as a weight-based grid and used it as a template to compare it with a time series window and determined whether the sequence matches the template based on the price percentile and a genetic algorithm. Wang and Chan [12] used the flag template method, which represents the top and bottom of the dome as a network lattice based on price fluctuations, as a template for identifying and examining buy signals. Royo et al. [13] defined flags from a new perspective and modified the grid weights.

Directly using an existing stock pattern curve as a template before using either of these methods will likely lead to overfitting and will be inefficient, also the accuracy and reliability of the analysis will be compromised [14, 15]. Therefore, when studying complex time series, scholars will first perform some preprocessing on the raw data. In order to improve the mining efficiency, scholars have proposed an approximate representation of the time series, which is a compressed representation of the time series, by retaining the main information of the time series and ignoring the tiny details. The current time series approximation methods have included discrete Fourier transform [16], discrete wavelet transform [17], piecewise aggregate approximation [18], symbolic aggregate approximation [19], and singular value decomposition [20]. Pattern recognition belongs to a type of time series mining. In conjunction with the above overview of price pattern recognition, time preprocessing before pattern recognition can be broadly divided into two categories. (1) Nonlinear curve description. Commonly used methods are nonparametric kernel regression methods and Gaussian kernel estimation methods. Dong used a Gaussian kernel estimation method to smooth the estimation process on the raw data [21]. (2) Approximate representation based on segmentation. For example, Keogh et al. [22] proposed a bottom-up approach for segmenting time series and extracting their features. Gong reduced the complexity of time series using perceived important points proposed by Chung [23, 24].

Wavelet transform has been an important way to preprocess time series, especially for nonstationary, nonlinear, and heteroskedastic signals [25]. Therefore, many scholars have used wavelet transforms to preprocess the raw data before building stock prediction models. The most used predictive model is the neural network algorithm [26–28]. The abovementioned prediction models show that the prediction accuracy of the original data is greatly improved after wavelet transform processing. In addition to neural network algorithms, there are other algorithms such as SVMs that are combined with wavelet transforms. The experimental results also showed that the hybrid model incorporating the wavelet transform is better than the prediction using a single algorithm [29–32]. DTW is a kind of elastic measurement for calculating time series similarity [33–35]. As DTW can extract morphological features of time series and handle time transformations and distortions, it can be used for pattern representation of stock time series. For example, Han developed a portfolio optimization model

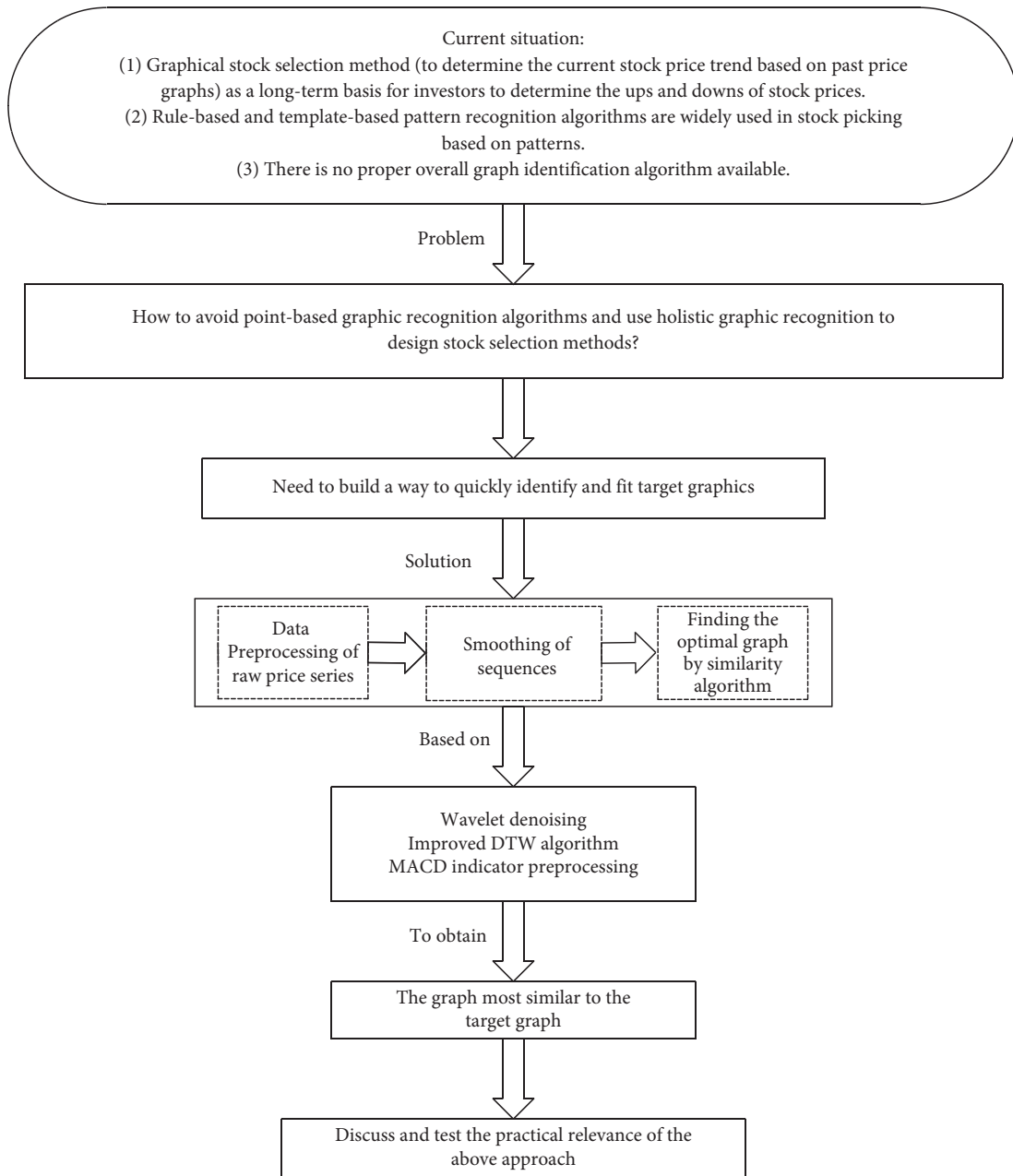


FIGURE 1: Research idea.

with strict constraints to obtain a pattern representation of stock time series based on DTW [35].

According to the abovementioned research, at present for the price pattern automatic identification has achieved some results, there is still room for improvement: (1) in the study of automatic price pattern recognition, the time series preprocessing method is relatively homogeneous. Due to the needs of pattern recognition, the automatic identification of price patterns uses a single method. (2) The extraction of important points in the perception of important points method is highly subjective. (3) In identifying patterns, most studies have used the sliding window method, which allows for the existence of other points between the extreme points of the pattern, but does not consider the relationship

between the extreme points of the pattern and other points, which is prone to misjudgment. For this reason, the study will combine the advantages of DTW and wavelet transform to establish a time series expression based on holistic graphs.

3. Theoretical Foundation

3.1. *Wavelet Denoising Principle.* Set a noise-containing sequence $F(t)$, which can be represented by the following model:

$$Ft = st + et. \tag{1}$$

In the above equation, $s(t)$ is the true sequence and $e(t)$ is the white noise.

Suppose that the noise $e(t)$ is white noise with zero mean and variance σ^2 . Set the wavelet function to φ_a , $W_{e,a,t}$ to be the wavelet transform of $e(t)$, which is obtained from the continuous wavelet transform theory.

$$W_{e,a,t} = \int_{-\infty}^{\infty} e(u)\varphi_a(t-u)du. \quad (2)$$

It is known from the hypothesis that

$$E\{e(u)e(v)\} = \sigma^2\delta(u,v). \quad (3)$$

Substituting the formula in (2) into (3) gives the following reasoning:

$$E\{|w_e(a,t)|^2\} = \frac{\sigma^2\|\varphi\|^2}{a}. \quad (4)$$

The purpose of sequence $f(t)$ denoising is to suppress the noisy part of the sequence $e(t)$ to obtain the true sequence $s(t)$; it follows from the abovementioned reasoning that the noise wavelet transformed $w_e(a,t)$ is inversely proportional to the scale a . Therefore, it is possible to denoise the sequence using the different characteristics exhibited at each scale after the noise and sequence wavelet transform.

3.2. Wavelet Denoising Principle. The method was proposed by Donoho and Johnstone [36]. The basic principle is that the orthogonal wavelet decomposition has the ability of time-frequency local decomposition. When performing sequence processing, the wavelet component exhibits a large amplitude, in contrast to the uniform variation of the noise in the high-frequency part. The larger amplitude wavelet coefficients are overwhelmingly useful sequences; the smaller amplitude coefficients are generally noise after wavelet decomposition. That is, it can be assumed that the wavelet transformation coefficient of the useful sequence is greater than the wavelet transformation coefficient of the noise. Then, a suitable threshold is selected according to certain rules, wavelet coefficients greater than the threshold are retained while wavelet coefficients smaller than the threshold are processed accordingly, and finally the useful sequence is reduced according to the processed wavelet coefficients. The key to the threshold denoising method is to pick an appropriate threshold value. According to the relationship between wavelet coefficients and the threshold, the threshold function can be divided into hard threshold denoising, soft threshold denoising, and zero-threshold denoising method. In addition, there are other threshold denoising methods, such as suppressing detail denoising (suppressing certain layers to obtain detail coefficients based on actual denoising requirements). These four threshold denoising methods are compared in detail in the experimental section of this study.

Suppose that the time series equation is as follows:

$$Ft = st + et. \quad (5)$$

The result of doing the wavelet transform on both sides of the above equation simultaneously is as follows:

$$W_{e,a,t} = \int_{-\infty}^{\infty} e(u)\varphi_a(t-u)du. \quad (6)$$

The wavelet transform of the actual sequence is equal to the sum of multiple wavelet transforms. The orthogonal wavelet transform of the original time series is able to remove the $f(t)$ correlation to the greatest extent possible, retaining a few and relatively large amplitudes on the wavelet coefficients to a large extent, while the noise $e(t)$ will be distributed across all time axes at all scales with relatively small amplitude after wavelet transformations. According to the abovementioned principle, the wavelet coefficients of the noise are minimized at each scale of the wavelet transform, and then the processed wavelet coefficients are used for sequence reconstruction to suppress the noise. The specific steps of the abovementioned method are shown in Figure 2.

3.3. MACD Fundamentals. In order to get better experimental results, this study uses the MACD algorithm to process the raw data to form a new MACD curve and then uses the DTW algorithm to compare the distance between the MACD curves. The MACD algorithm is a trend-following momentum indicator that shows the relationship between two moving averages of a security's price. This preprocessing method can be used to eliminate price disturbances and enhance the trend of the original price graphs, ultimately resulting in a better fit. The mathematical principle of the MACD algorithm is as follows:

$$\text{EMA}(t) = \frac{n-1}{n+1} \times \text{EMA}(t-1) + \frac{2}{n+1} \times \text{price}. \quad (7)$$

$\text{EMA}(t)$ is a moving average of t days, $\text{EMA}(t-1)$ is yesterday's moving average, n is the period, and price is the original price series. t -day maps t different weights. From the above equation, the weights increase over time, so $\text{EMA}(t)$ is a better reflection of the future trend of the price chart line. DIF (the difference between the short-term average price of a stock and its long-term average price) is defined as follows:

$$\text{DIF} = \text{EMA}(t_1) - \text{EMA}(t_2). \quad (8)$$

$\text{EMA}(t_1)$ is a fast-moving average and $\text{EMA}(t_2)$ is a slow-moving average; the difference between the two can reflect the future upward or downward trend of the price chart line. DEA (difference exponential average) is defined as follows:

$$\text{DEA} = \frac{2}{n+1} \times \text{DIF} + \frac{n-1}{n+1} \times \text{DEA}'. \quad (9)$$

A weighted average was applied to the abovementioned DIFs, with the same weight settings as EMA. In the above equation, DEA' is the previous period DEA. HIST (histogram) is defined as follows:

$$\text{HIST} = 2 \times (\text{DIF} - \text{DEA}). \quad (10)$$

From the above formula, the HIST can better reflect the rise and fall of the original price series and eliminate the influence of price on the DTW algorithm. In this study, the closing price of 000009.sz from January 1, 2019, to July 24,

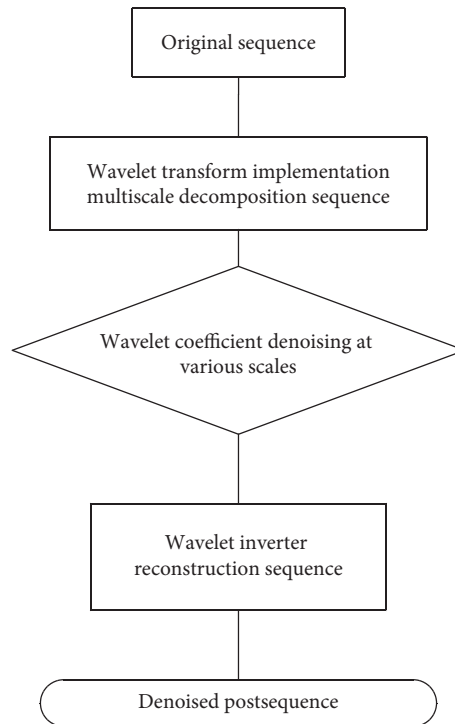


FIGURE 2: Basic steps of threshold denoising.

2020, is selected and the relevant graph line is plotted as shown in Figure 3.

It is obvious from the abovementioned chart, the original price sequence chart line (signal line in the chart) increase is basically consistent with the MACD column in the chart, so the MACD column line can reflect the increase of the original chart line. In addition to the above advantages, MACD removes price distractions when fitting graphs. Therefore, this study converts the standard graph line to an MACD line as shown in Figure 4.

3.4. The Process of Rolling Window with DTW. The article selected the raw price series with m data points as the target graphic. The remaining stock price series with n data points is then divided chronologically into subsegments with m data points. A total of (n/m) subsequence fragments can be obtained, and the subsequence fragments are labeled using $Window(i)$. The window (i) with the smallest distance from the target graph is obtained using the DTW algorithm. The rolling window process with DTW is shown in Figure 5.

4. Financial Time Series Wavelet Denoising Analysis and DTW Improvements

The time series can be smoothed using the abovementioned theory, but there is still the problem of parameter determination. According to the flow of financial time series processing by Lan and Ma [37], there are three important parameters that affect the final denoising effect: the choice of the wavelet base function, the number of decomposition layers, and the choice of the threshold function.

4.1. Characteristics of Financial Time Series. Stock price data, which is typical of financial data, commonly includes stock trading data and yield data calculated from trading data. However, both types of data exhibit an unsteady nature and the data are more volatile. In particular, financial time series are densely populated with singularities, so it is generally difficult to smooth out the financial time series. In addition to large fluctuations, there are small fluctuations in the data, which are highly stochastic and spread throughout the time window. Generally, large fluctuations contain more useful information and are therefore generally retained as useful information; a large number of small fluctuations are generally chosen to be removed due to their randomness and lack of value for analysis and prediction.

4.2. Selection of the Wavelet Function. At present, wavelet functions are more diverse and have different properties, which are suitable for different scenarios. The purpose of this study is to identify the graph of a stock in a given time window. Different wavelet functions have different denoising effects, so it is important to choose the right wavelet function. The following analysis of wavelet function characteristics is closely related to denoising.

Because there is often a correlation between wavelet function properties (it has been shown that there is no symmetric tightly-branched orthogonal wavelet base, except for the Harr wavelet), it is impossible to find a wavelet function that has all the optimal properties. The purpose of this study is to smooth the financial time series and different wavelet functions correspond to different denoising effects; it is crucial to choose the appropriate wavelet function. The

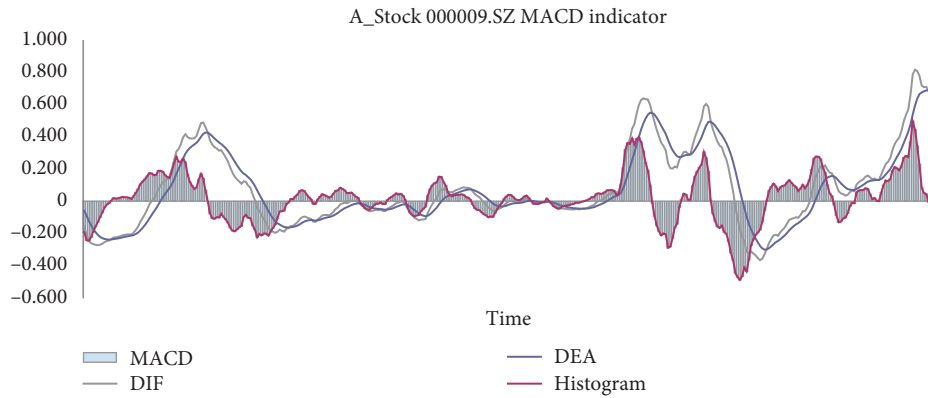


FIGURE 3: Indicator preprocessing.

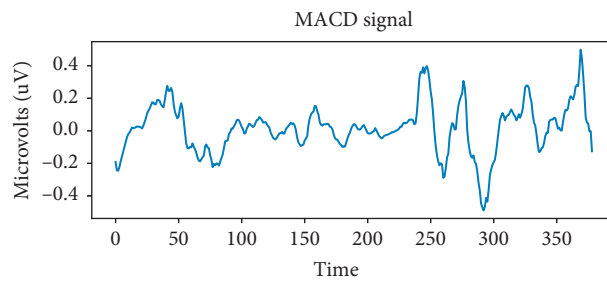


FIGURE 4: Target price series graphs are processed by MACD algorithm.

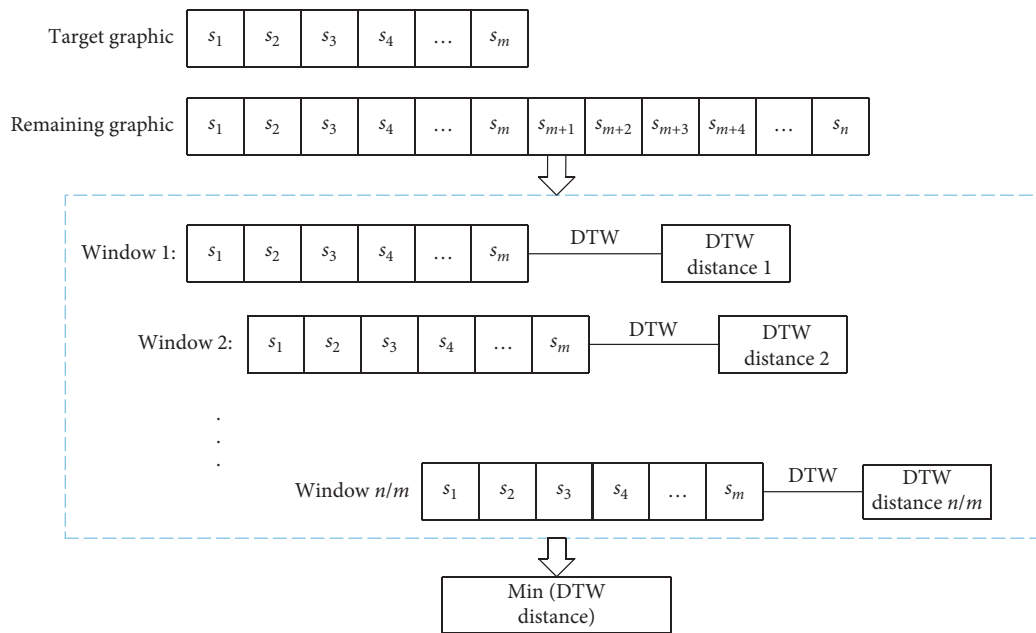


FIGURE 5: The process of rolling window with DTW.

following analysis focuses on a number of features that are closely related to the noise removal effect.

- (1) Orthogonality: the wavelet function with orthogonality can maintain the noncorrelation between the wavelet coefficients to improve the denoising performance of the wavelet function. More importantly, orthogonality is necessary to achieve fast wavelet

transforms, and thus orthogonality is necessary to achieve wavelet denoising.

- (2) Compact support: it means that the wavelet function is nonzero in the effective interval, so the smaller the support width of the wavelet or scale function, the better the local resolution of the wavelet and the finer the denoising effect.

- (3) Vanishing moment: since there are a certain number of singularities in the financial time series, it is necessary to apply a certain vanishing moment so that as many wavelet coefficients as possible are zero or generate as few nonzero wavelet coefficients as possible to increase the denoising effect. The larger the vanishing moment, the more wavelet coefficients are zero, so there is a positive correlation between the vanishing moment and the width of the support. That is why there is a compromise between the vanishing moment and the width of the support. The vanishing moment is proportional to the number of singularities present in the sequence, while the stock price sequence has more singularities, so the vanishing moment can be increased accordingly.
- (4) Symmetry: the wavelet function with symmetry will reduce the error after wavelet transformation, so this property is conducive to the recovery and reconstruction of the sequence after noise removal. However, this property has less influence on the noise removal effect than other properties. This study finds a number of commonly used wavelet functions to examine, the properties of which are shown in Table 1.

For the aforementioned wavelet function, SymN is the best choice for the following reasons: (1) it is orthogonal, enabling wavelet denoising. (2) There are more “spikes” in the stock price sequence, so it is necessary to ensure that the wavelet has a vanishing moment. SymN is the best choice because of the compromise between the vanishing moment and the width of the support. (3) SymN has better symmetry than DbN as well as Conif, which can reduce errors during sequence reconstruction.

4.3. Determination of the Number of Decomposition Layers

4.3.1. *Wavelet Noise Reduction Evaluation Index.* Since wavelet denoising is mostly used inside engineering, the requirements for denoising in engineering are different from the requirements for denoising in finance. In order to measure the effect of sequence denoising, the article defines two variables, information extraction rate (IER) and information extraction error (IEE), drawing on the wavelet threshold method selected by Liu [38] to measure the effect of sequence denoising, considering f as the original time series, \hat{f} as the denoised time series, and IEE and IER defined as follows:

$$\begin{aligned} \text{IER} &= \frac{\|\hat{f}\|_2}{\|f\|_2}, \\ \text{IEE} &= \|\hat{f} - f\|_2. \end{aligned} \quad (11)$$

From the above equation, the information extraction rate is positively related to the similarity of the denoised sequence to the original sequence. The larger the information extraction error, the smoother the denoised sequence is compared to the original sequence. The abovementioned

two indicators are negatively correlated, and if the information extraction rate is combined with the information extraction error to form a composite evaluation indicator, the process of increasing the number of decomposition layers will inevitably result in a maximum value. The graphical interpretation of this maximum value is the optimal ratio of the sequence in retaining detail information to its approximation information, which corresponds to the number of decomposition layers. The above demonstrates the feasibility of using a combination of information extraction rates and information extraction errors to form new composite evaluation indicators through theoretical analysis and also based on the numerical characteristics of the indicator changes.

4.3.2. Establishment of Composite Evaluation Indicators.

Different evaluation indicators have different scales and units, and in order to eliminate the incommensurability this brings, the study treats each evaluation indicator as non-quantitative. In order to preserve the difference between the information extraction rate and the information extraction error, this study chooses to use the averaging method in the no-quantization method. If there are m indicators in the composite evaluation, each with n units, the indicators are x_{ij} ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$), which represents the j th indicator value of the i -th unit, and y_{ij} represents the j th index value of the i -th unit after dimensionless homogenization. The formula is as follows:

$$y_{ij} = \frac{x_{ij}}{x_j}. \quad (12)$$

The mean value of each indicator after the mean method is 1 and its variance is as follows:

$$\text{var}(y_j) = E((y_j - 1)^2) = E\left(\frac{x_j - \bar{x}_j}{\bar{x}_j}\right)^2 = \frac{\text{var}(x_j)}{\bar{x}_j^2} = \left(\frac{\sigma_j}{\bar{x}_j}\right)^2. \quad (13)$$

The variance of each indicator after averaging is the square of the coefficient of variation \bar{x}_j of each indicator, which retains information on the degree of variation of each indicator. After averaging, the indicators become the same order of magnitude, and if they are to be combined linearly, weights need to be assigned to each indicator. Since the averaging method preserves the variability among the indicators, then weighting requires the elimination of the variability among the indicators, for which entropy weighting is chosen in this study. Entropy is an ideal metric for multiobjective decision-making and evaluation. It relies on the amount of decision-making information to improve the accuracy of decision-making, so it can objectively reflect the implicit information of the indicator data, improve the resolution of the indicators, avoid the selection bias caused by small differences in indicators, and reflect a full range of indicator information. There is an inverse relationship

TABLE 1: Characteristics of different types of wavelet functions.

	Compact support	Support width	Symmetry	Orthogonality	Vanishing moment
Harr	Yes	1	Yes	Yes	1
Mexican	No	inf	Yes	No	1
Meyer	No	inf	Yes	Yes	inf
DbN	Yes	$2N - 1$	Far	Yes	N
SymN	Yes	$2N - 1$	Near	Yes	N
CoifN	Yes	$6N - 1$	Near	Yes	$2N$
Since	No	inf	Yes	Yes	No
BionNr.Nd	Yes	$2Nr + 1, 2Nd + 1.$	Yes	No	Nr

between the entropy of an indicator and its weight. According to the concept of entropy, entropy Q is defined as follows:

$$Q_j = \frac{1 - H_j}{n - \sum_{j=1}^n H_j}. \quad (14)$$

In the above equation, H_j is defined as follows:

$$H_j = \frac{-\left(\sum_{i=1}^m f_{ij} \ln(f_{ij})\right)}{\ln(m)}. \quad (15)$$

In the above equation, f_{ij} is defined as follows:

$$f_{ij} = \frac{1 + y_{ij}}{\sum_{i=1}^m (1 + y_{ij})}. \quad (16)$$

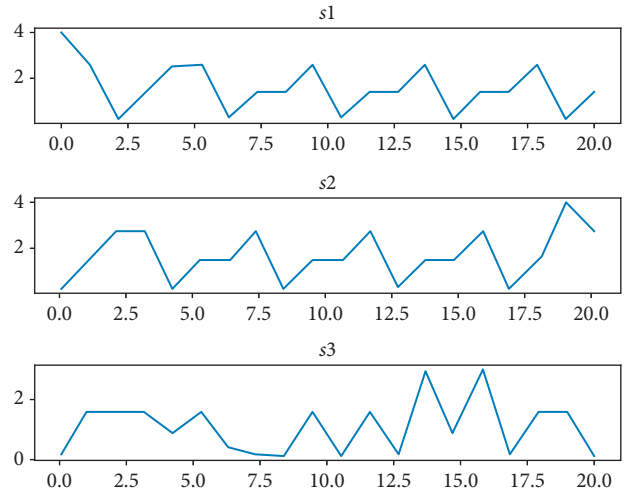
The above equation satisfies $\sum_{j=1}^n Q_j = 1$.

The use of entropy weighting allows for the combination of multiple metrics to determine weights, both to account for associations between multiple samples and to weaken the effect of outliers. The choice of entropy weighting and a measureless mean, i.e., a combination of differential and nondifferential, would theoretically be a good one.

According to the abovementioned method, wavelet decomposition is first performed on the original time series to calculate the IER_j and IEE_j for different decomposition scales J . Then, the dimensionless averaging process is performed to get the \overline{IER}_j and \overline{IEE}_j . Finally, the information extraction rate weights Q_{IER} and the information extraction error weights Q_{IEE} are calculated from the entropy weights. Then, the formula for the value of composite indicator S is as follows:

$$S_j = IER_j \times Q_{IER} + IEE_j \times Q_{IEE}, \quad j = 1, 2. \quad (17)$$

4.4. DTW Algorithm Improvement. The main objective of the study is to fit the graphs formed by stock prices in a particular time window, and the degree of similarity of the fit is achieved by the DTW algorithm. In order to improve the fit to the target graph, this study makes improvements to the traditional DTW algorithm. The article randomly constructs three sequences of length 20, where s_2 is derived from s_1 translation and s_3 has no relationship with either of the above. The three sequences are shown in Figure 6.

FIGURE 6: s_1 , s_2 , and s_3 graphs.

The distance between s_1 and s_2 is 5 and the distance between s_1 and s_3 is 4.12, calculated using the traditional DTW algorithm. Thus, it can be seen that the traditional DTW algorithm is not very effective for the distance of the postshift sequence. This study proposes an improved DTW algorithm.

- (1) Determine the largest common substring of the two sequences seq_1 and seq_2 , the length of which is given by a .
- (2) seq_1 and seq_2 are numerical sequences that set a maximum standard deviation of the offset tolerance in finding the longest public sequence. In other words, two values within this standard deviation are considered to be part of the public sequence as well.
- (3) Define the attenuation factor α as follows:

$$\alpha = 1 - \frac{a \times a}{\text{len}(sql_1) \times \text{len}(sql_2)}. \quad (18)$$

The length of the common subsequence of both sequences is negatively correlated with its attenuation coefficient, thus enabling improvement of the algorithm. The distance between s_1 and s_2 is 0.982 and the distance between s_1 and s_3 is 2.512, as recalculated by the improved algorithm. It follows that the improved DTW algorithm effectively reduces the distance between the periodic sequences after translation.

5. Experimental Demonstration and Analysis

5.1. Methodology. This study arbitrarily selects five stocks in the A-share market as the target graphics and determines the wavelet denoising parameters by the abovementioned entropy weighting method. An improved DTW algorithm is then used to find other stock time windows from the stock market that are the least distant from it. In order to achieve better experimental results, this study introduces MACD algorithm to convert the original price series into “MACD graph line” on the basis of wavelet denoising. Then, the distance between the calculated original price series is transformed into the distance between “MACD lines,” which improves the accuracy of the result.

5.2. Source of Experimental Data. All stock data in this study are from the Tushare website (<https://tushare.pro>) and include trading date, opening price, high price, low price, closing price before compounding, and uncompounded upside and downside. Stock data are daily data and the time frame studied is from the listing of each stock to July 24, 2020.

5.3. Wavelet Denoising Preprocessing

5.3.1. Stock Price Sequence Wavelet Denoising Process. (1) Choose the appropriate wavelet base function and decomposition layer N for the time series f .

(2) Threshold quantification of the decomposed high-frequency coefficients is performed, and for each layer from 1 to N , the appropriate threshold and the appropriate threshold function are selected to obtain the estimated wavelet coefficients.

(3) The original time series is reconstructed (performing wavelet inversion) based on the low-frequency coefficients (scale coefficients) of the N th layer after wavelet decomposition and the high-frequency coefficients (wavelet coefficients) of each layer after threshold quantization. The wavelet reconstruction is then performed using the reconstruction algorithm to obtain the denoised postdecomposition sequence.

Because the daily stock closing price sequence is essentially a one-dimensional discrete sequence, this focuses on the wavelet decomposition of the one-dimensional discrete sequence. Experiments show that when a white noise sequence is wavelet decomposed, it can be found that the decay rate of the high-frequency coefficient after wavelet decomposition is proportional to the number of layers of wavelet decomposition, and the variance of the same coefficient is inversely proportional to the number of layers of wavelet decomposition. Thus, for the general case, there is a positive relationship between the number of decomposed layers and the smoothness of the denoised sequence. Therefore, in this study, we have to find the most suitable wavelet

denoising parameters based on the characteristics of the stock price sequence.

5.3.2. Wavelet Denoising Parameter Investigation. This study selects five stocks in the China A-share market (ticker symbols: 300327.SZ, 300332.SZ, 300336.SZ, 601567.SH, and 601616.SH) as target patterns I–V, respectively. The results of denoising the abovementioned five target graphics using hard threshold, soft threshold, zero threshold, and inhibited detail factor threshold functions combined with the wavelet decomposition layer number method are shown in Tables 2–5.

From Table 5, it can be seen that the integrated index score of wavelet decomposition layers and threshold function shows a trend of increasing and then decreasing, and when the number of wavelet decomposition layers is 4 and the suppression detail factor threshold method is selected, the integrated evaluation index value is the largest, and this experimental result is consistent with the results of the previous mathematical derivation. The data of the abovementioned five stocks are obtained by random sampling, but they all show the maximum value on the same parameters as above, which fully demonstrates the generalizability of the method.

The abovementioned five graphics are compared with the original graphics after the abovementioned wavelet denoising method which is shown in Figure 7.

5.4. Finding Time Windows Similar to the Target Graph Using the DTW Algorithm. This study selects five stocks in the China A-share market (ticker symbols: 300327.SZ, 300332.SZ, 300336.SZ, 601567.SH, and 601616.SH) as target patterns I–V, respectively. The optimal DTW distances and similarities found using the abovementioned methods, respectively, are shown in Table 6. The time complexity of all the abovementioned algorithms is $O(n^2)$.

The above results show that there is a significant improvement in the final similarity of the original data smoothed using wavelet denoising compared to using the original data directly (the average similarity of the five graphs of the three methods is 0.3612, 0.5075, and 0.8936). This result verifies the need for wavelet denoising. The final similarity of using MACD data compared to direct use of raw data is also significantly improved. Therefore, the final graph recognition method used in this study is to first transform the original data into MACD graphs, then use wavelet denoising to smooth the MACD graphs, and finally use the improved DTW algorithm to achieve effective recognition of the target graphs. The most similar graphs of the five abovementioned SDs after wavelet preprocessing and direct search using the DTW algorithm are shown in Figure 8. The most similar graphs found using the DTW algorithm after the abovementioned five target graphs were processed by MACD and then denoised using wavelets, as shown in Figure 9.

TABLE 2: Suppression of detail factor.

	Layers	1	2	3	4	5	6	7
Figure 1	IER	1.000270	1.000198	1.000035	0.999953	0.999475	0.998558	0.998343
	IEE	0.212469	0.289914	0.327015	0.355167	0.370909	0.374698	0.378330
	<i>s</i>	12.59581	25.69571	49.61283	54.72754	51.94654	26.17834	14.00664
Figure 2	IER	1.000029	1.000167	0.999944	0.999775	0.999281	0.998238	0.997986
	IEE	0.236178	0.317151	0.360568	0.377266	0.389581	0.392551	0.396433
	<i>s</i>	15.35511	21.54187	47.11543	56.45679	50.98837	24.79653	15.51747
Figure 3	IER	0.999970	1.000110	0.999984	0.999735	0.999507	0.998720	0.998531
	IEE	0.283456	0.352092	0.392585	0.410939	0.420073	0.424726	0.427760
	<i>s</i>	15.09589	22.75856	50.73586	56.59787	48.42030	23.92601	18.50077
Figure 4	IER	1.000053	1.000136	1.000007	0.999898	0.999618	0.999604	0.999455
	IEE	0.202496	0.283091	0.328215	0.350692	0.363871	0.376480	0.375991
	<i>s</i>	22.47839	46.72721	48.40989	59.92729	47.47580	26.86061	21.57036
Figure 5	IER	1.000005	1.000287	1.000015	0.999749	0.999071	0.997442	0.997055
	IEE	0.139754	0.202370	0.246294	0.271735	0.287797	0.291735	0.297407
	<i>s</i>	19.90815	28.85956	52.12604	52.42208	47.60686	22.53001	16.47265

TABLE 3: Hard threshold.

	Layers	1	2	3	4	5	6	7
Figure 1	IER	1.000053	1.000136	1.000007	0.999998	0.999818	0.999804	0.999855
	IEE	0.202496	0.283091	0.428215	0.450692	0.463871	0.476480	0.485991
	<i>s</i>	13.53791	22.83471	39.61283	44.95648	38.71361	28.72435	16.53761
Figure 2	IER	1.000034	1.000047	0.999931	0.999923	0.999714	0.999738	0.998976
	IEE	0.228416	0.305499	0.343774	0.389739	0.395459	0.397032	0.423563
	<i>s</i>	14.19504	21.59701	37.90262	46.37774	40.63924	24.07936	15.33218
Figure 3	IER	0.999990	1.000009	0.999993	0.999935	0.999904	0.999718	0.999476
	IEE	0.349785	0.368139	0.405443	0.408998	0.420211	0.424952	0.436943
	<i>s</i>	15.20355	22.23524	40.21399	46.11387	38.44541	23.07473	18.85263
Figure 4	IER	1.000042	1.000036	0.999997	0.999947	0.999857	0.999844	0.999837
	IEE	0.208219	0.243726	0.304472	0.320692	0.348696	0.377230	0.381473
	<i>s</i>	22.30446	35.93901	37.28737	48.33225	39.09336	26.44241	24.57120
Figure 5	IER	1.000142	1.000007	1.000005	0.999994	0.999971	0.999522	0.999473
	IEE	0.239648	0.238462	0.241003	0.378103	0.380470	0.392062	0.394378
	<i>s</i>	18.44058	28.11498	32.09637	42.04788	37.49007	22.32213	17.15010

TABLE 4: Soft threshold.

	Layers	1	2	3	4	5	6	7
Figure 1	IER	1.000012	1.000008	1.000002	0.999962	0.999634	0.999623	0.998252
	IEE	0.123851	0.174044	0.237015	0.266467	0.270435	0.279749	0.301757
	<i>s</i>	12.78834	15.68995	29.00644	34.85393	21.85329	16.43920	8.92942
Figure 2	IER	1.000023	1.000001	0.999990	0.999823	0.999347	0.999324	0.998235
	IEE	0.140276	0.203249	0.257859	0.267006	0.270619	0.285135	0.282012
	<i>s</i>	7.03805	11.86719	27.61023	36.45679	20.15955	14.77345	8.99417
Figure 3	IER	0.999997	0.999985	0.999874	0.999748	0.999636	0.999524	0.998648
	IEE	0.175627	0.229226	0.240494	0.280939	0.305462	0.322143	0.330210
	<i>s</i>	5.80707	12.82072	30.02364	26.60017	28.52938	13.37045	7.23444
Figure 4	IER	1.000023	1.000015	0.999945	0.999837	0.999713	0.999638	0.999564
	IEE	0.100542	0.188980	0.229553	0.253115	0.265422	0.270101	0.274818
	<i>s</i>	12.80665	16.55634	25.57838	38.92002	36.11341	25.36742	19.54325
Figure 5	IER	1.000221	1.000104	0.999952	0.999424	0.999329	0.997894	0.993652
	IEE	0.139109	0.144395	0.150898	0.188750	0.197533	0.208805	0.228945
	<i>s</i>	9.70329	28.81310	32.34880	32.80551	27.30382	12.84004	10.47287

TABLE 5: Zero threshold.

	Layers	1	2	3	4	5	6	7
Figure 1	IER	0.999245	0.998251	0.996324	0.995873	0.995344	0.995106	0.994857
	IEE	0.317301	0.385397	0.422673	0.457069	0.473794	0.477979	0.505537
	s	8.14986	10.92051	15.50857	20.84403	16.76388	14.76127	9.90114
Figure 2	IER	0.999356	0.999235	0.998544	0.997775	0.996281	0.995238	0.994567
	IEE	0.323665	0.413166	0.462867	0.474033	0.480247	0.497387	0.525427
	s	5.64122	11.83068	17.15623	23.06459	18.91895	14.16127	6.12647
Figure 3	IER	0.999932	0.999865	0.998364	0.997459	0.996964	0.995487	0.994786
	IEE	0.398156	0.461597	0.487581	0.506123	0.522824	0.535530	0.621169
	s	5.36012	12.17355	16.32691	24.24204	18.10343	13.75792	9.55168
Figure 4	IER	0.999926	0.998426	0.997452	0.997265	0.996324	0.995687	0.994568
	IEE	0.291294	0.396925	0.431996	0.467571	0.477762	0.497790	0.56190
	s	2.63584	11.61053	18.77789	23.98553	17.69008	10.80362	5.23897
Figure 5	IER	0.999326	0.998564	0.997456	0.997254	0.996325	0.995268	0.993125
	IEE	0.237483	0.298172	0.321009	0.368405	0.396861	0.417914	0.451320
	s	7.95445	10.85890	17.66482	22.88596	19.68822	13.82083	8.41443

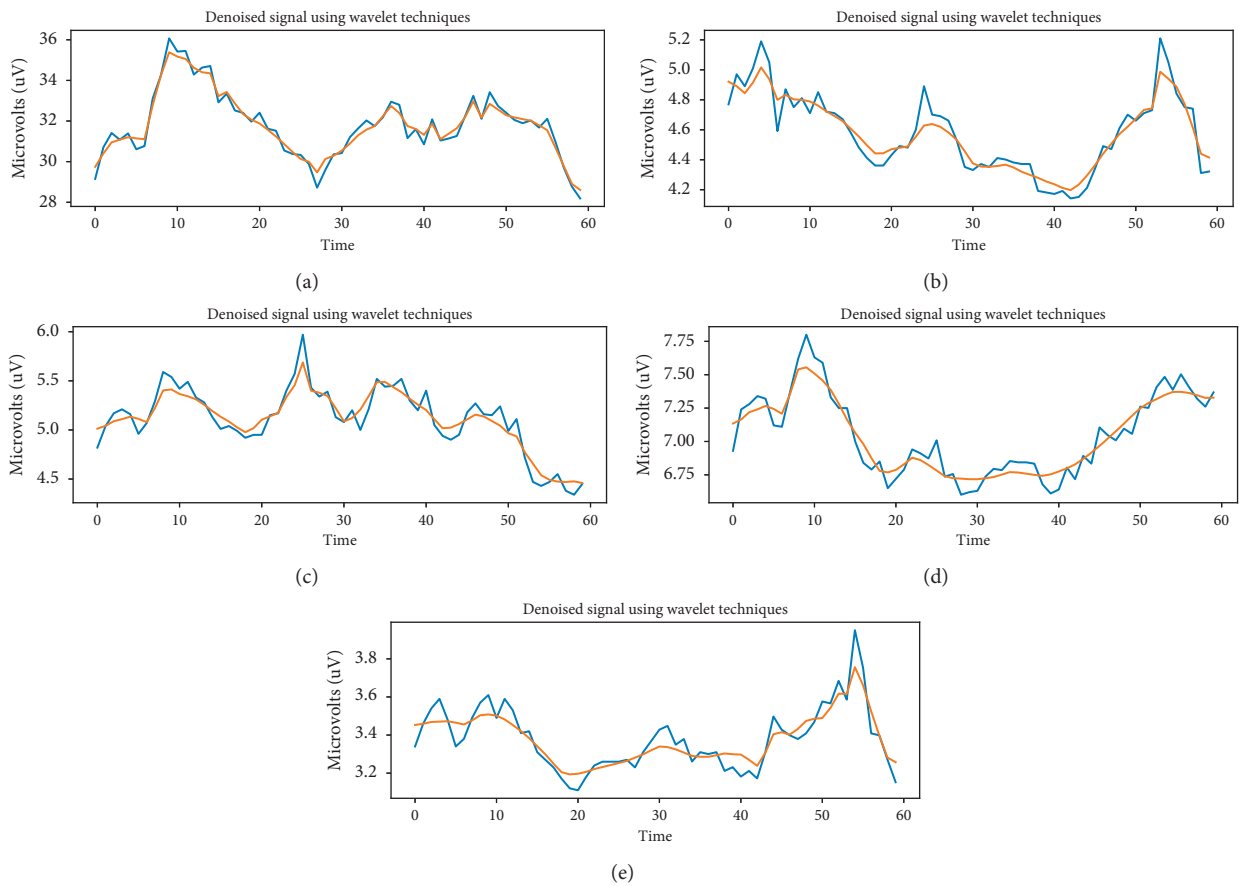


FIGURE 7: Comparison of the smoothing process with the original graphics.

TABLE 6: Comparison of various methods for optimal DTW distance analysis.

	Zero-step ahead	One-step ahead	Two-step ahead
	Preprocessing without wavelet denoising	Preprocessing with wavelet denoising	Preprocessing with MACD
Target graph I	2.236 (0.3090)	5.02 (0.1661)	0.1197 (0.8930)
Target graph II	2.637 (0.2749)	0.7637 (0.5669)	0.1682 (0.8560)
Target graph III	3.592 (0.2177)	0.9091 (0.5238)	0.1171 (0.8951)
Target graph IV	1.946 (0.3394)	0.6371 (0.6108)	0.1172 (0.8950)
Target graph V	4.897 (0.1695)	0.4918 (0.6703)	0.076 (0.9293)

TABLE 7: Comparison of 40-day average upside and downside analysis of various methods.

	Zero-step ahead Preprocessing without wavelet denoising	One-step ahead Preprocessing with wavelet denoising	Two-step ahead Preprocessing with MACD
Optimal graphics I	0.2354	0.0235	0.1546
Optimal graphics II	0.5773	-0.0107	0.3705
Optimal graphics III	0.1739	-0.0086	0.0668
Optimal graphics IV	-0.0094	0.3866	-0.004
Optimal graphics V	-0.0052	0.1394	-0.0678

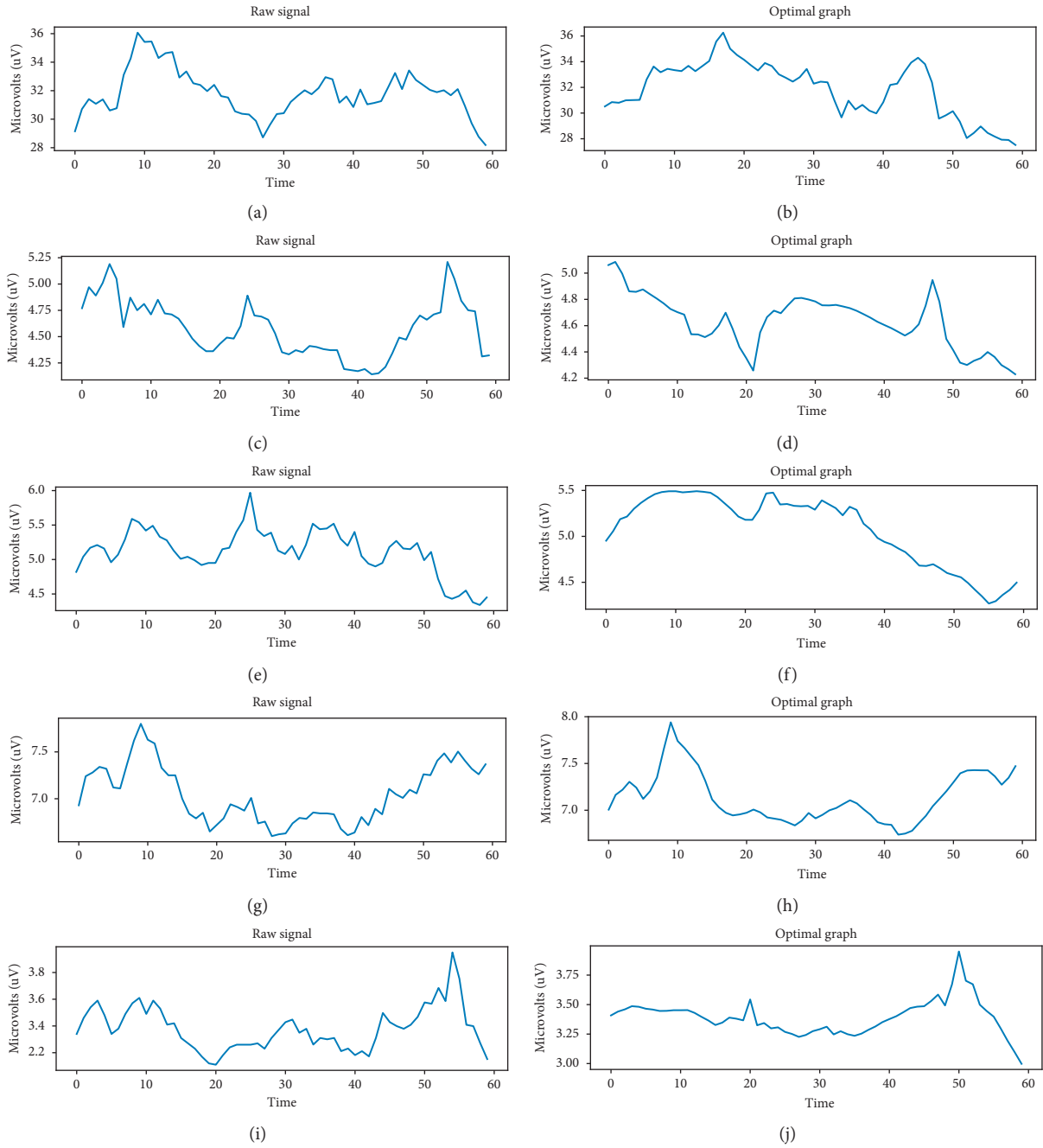


FIGURE 8: The graph most similar to the five target graphs.

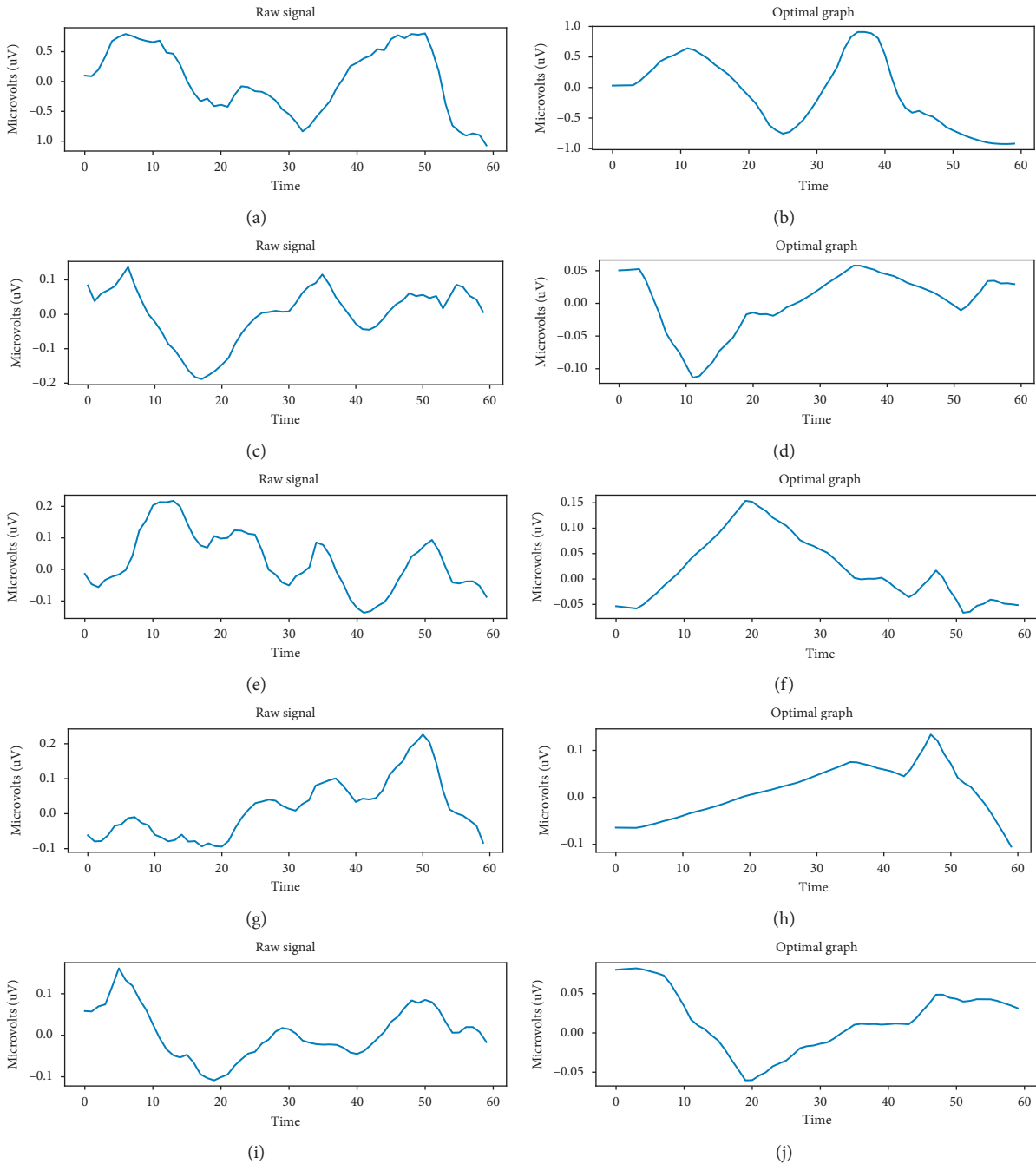


FIGURE 9: Graphs most similar to the five target MACD graphs.

6. Empirical Analysis of China’s A-Share Market

In order to find out the pattern of the stock market in the short-term and to compare the effectiveness of the above-mentioned two methods, the article calculates the average upside and downside of the current time window of the underlying graph and the average upside and downside of the 40-day time window of the optimal graph after the two treatments. The forty-day average increase and decrease in the underlying and optimal graphs are shown in Table 7:

As can be seen from Table 7, the average rise and fall of the optimal graph obtained using the MACD preprocessing

method compared to the direct use of the raw data is closer to the average rise and fall of the base graph over the next forty days. This result confirms the lagging nature of the MACD indicator and the effectiveness of the graphical stock selection, fully demonstrating the practicality of the method in practice.

7. Conclusions and Discussion

In this study, the wavelet transform is used to denoise the original stock price series, the SymN4 wavelet function is selected according to the characteristics of the stock time

series, and the comprehensive scores of different wavelet decomposition layers and wavelet threshold functions are calculated based on the entropy weight method; the results show that the 4-layer decomposition layers and the suppression detail factor threshold method have the highest scores, so the wavelet denoising parameters finally selected in this study are SymN4 wavelet function, 4-layer decomposition layers, and suppression detail factor threshold method. The improved DTW algorithm was then used to calculate the distance between the two sequences to select the optimal graph that is most similar to the target graph. To overcome the effect of price and other factors on the DTW algorithm, the raw sequence data were standardized using the MACD algorithm and then the optimal graph most similar to the target graph was selected using the DTW algorithm. The comparison experimental results show that the similarity between the final graph and the target graph after the MACD algorithm processing is greater than 85%, which fully demonstrates the effectiveness of the graph recognition method. Finally, by comparing the target graphs with the optimal graphs, it is found that the optimal graphs obtained by MACD preprocessing method are closer to the forty-day average increase/decrease of the underlying graphs (the difference between them is within 0.2), indicating that the method has strong practical guidance significance.

The graph recognition methods in this study can be nested in stock analysis software, and stock investors can select existing classical patterns or K-line patterns to find stocks that are similar to them for the purpose of graphical stock selection. This provides a practical technical basis for graphical stock selection techniques.

The graph recognition methods used in this study are all based on a consistent time window size, so it will result in finding the optimal graph with the same time window size as the target graph, and the DTW algorithm can calculate the similarity of graphs with different size time windows, in addition to the overall recognition of the target graph. Therefore, this study does not fully use all the advantages of the DTW algorithm, and in real-life applications, if the problem of matching different size time window graphics can be solved, it will have more far-reaching implications for stock investors. In future work, this study will try to use genetic algorithm to further find the optimal time window size.

Data Availability

The stock price data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

The author greatly appreciates the reviewers' suggestions, and Shanxi University of Finance and Economics. This

research was funded by National Natural Science Foundation of China, grant no. 61173144, and the Science and Technology Innovation Project of Higher Education Department of Shanxi Province, grant no. 2019L0493.

References

- [1] M. J. S. A de Souza, D. G. F. Ramos, M. G. Pena, V. A. Sobreiro, and H. Kimura, "Examination of the profitability of technical analysis based on moving average strategies in BRICS," *Financial Innovation*, vol. 4, p. 3, 2018.
- [2] A. Silva, R. Neves, and N. Horta, "A hybrid approach to portfolio composition based on fundamental and technical indicators," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2036–2048, 2015.
- [3] M. Kaucic, "Portfolio management using artificial trading systems based on technical analysis," *Genetic Algorithms in Application*, vol. 15, pp. 281–294, 2012.
- [4] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, and E. Salwana, "Deep learning for stock market prediction," *Entropy*, vol. 22, no. 8, p. 840, 2020.
- [5] S. Li and A. Lin, "Exploring the relationship among predictability, prediction accuracy and data frequency of financial time series," *Entropy*, vol. 22, no. 12, p. 1381, 2020.
- [6] D. Wu, X. Wang, J. Su, B. Tang, and S. Wu, "A labeling method for financial time series prediction based on trends," *Entropy*, vol. 22, no. 10, p. 1162, 2020.
- [7] A. W. Lo, H. Mamaysky, and J. Wang, "Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation," *The Journal of Finance*, vol. 55, no. 4, pp. 1705–1765, 2000.
- [8] S. C. Suh, D. Li, and J. Gao, "A novel chart pattern recognition approach: a case study on cup with handle," in *Artificial Neural Networks* St. Louis, MI, USA, 2004.
- [9] G. Savin, P. Weller, and J. Zvingelis, "The predictive power of "Head-and-Shoulders" price patterns in the US stock market," *Journal of Financial Econometrics*, vol. 5, no. 2, pp. 243–265, 2006.
- [10] A. Zapranis and P. E. Tsinaslanidis, "A novel, rule-based technical pattern identification mechanism: identifying and evaluating saucers and resistant levels in the US stock market," *Expert Systems with Applications*, vol. 39, no. 7, pp. 6301–6308, 2012.
- [11] W. Leigh, R. Purvis, and J. M. Ragusa, "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*, vol. 32, no. 4, pp. 361–377, 2002.
- [12] J.-L. Wang and S.-H. Chan, "Trading rule discovery in the US stock market: an empirical study," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5450–5455, 2009.
- [13] R. C. Royo, F. Guijarro, and K. Michniuk, "Stock market trading rule based on pattern recognition and technical analysis: forecasting the DJIA index with intraday data," *Expert Systems with Applications*, vol. 42, pp. 5963–5975, 2015.
- [14] X. X. He, *Research on Key Issues in Time Series Data Mining*, University of Science and Technology of China, Hefei, China, 2014.
- [15] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys*, vol. 45, no. 1, pp. 1–34, 2012.
- [16] A. Rakesh, C. Faloutsos, and A. N. Swami, "Efficient similarity search in sequence databases," in *Proceedings of the*

- International Conference on Foundations of Data Organization and Algorithms*, pp. 69–84, Chicago, IL, USA, October 1993.
- [17] K. Chan and W. Fu, “Efficient time series matching by wavelets,” in *Proceedings of the 15th IEEE International Conference on Data Engineering*, pp. 126–133, Sydney, Australia, March 1999.
- [18] C. Jason, “Useful clustering outcomes from meaningful time series clustering,” in *Proceedings of the Sixth Australasian Data Mining Conference*, pp. 101–109, Gold Coast, Australia, December 2007.
- [19] J. Lin, E. J. Keogh, S. Lonardi, and Y. C. Chiu, “A symbolic representation of time series, with implications for streaming algorithms,” in *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, pp. 2–11, San Diego, CA, USA, June 2003.
- [20] K. Flip, H. V. Jagadish, and C. Faloutsos, “Efficiently supporting ad hoc queries in large datasets of time sequences,” *SIGMOD Record*, vol. 26, pp. 289–300, 1997.
- [21] M. Dong and X. S. Zhou, “Exploring the fuzzy nature of technical patterns of U.S. stock market,” in *Fuzzy System and Knowledge Discovery*, vol. 1, pp. 324–328, Springer, Berlin, Germany, 2002.
- [22] E. Keogh, S. Chu, D. Hart, and M. Pazzani, “An online algorithm for segmenting time series,” in *Proceedings of the 2001 IEEE International Conference on Data Mining*, pp. 289–296, San Jose, CA, USA, November 2001.
- [23] X. Gong, Y.-W. Si, S. Fong, and R. P. Biuk-Aghai, “Financial time series pattern matching with extended UCR suite and support vector machine,” *Expert Systems with Applications*, vol. 55, pp. 284–296, 2016.
- [24] F. L. Chung, T. C. Fu, R. Luk, and V. Ng, “Flexible time series pattern matching based on perceptually important points,” in *Proceedings of the International Joint Conference on Artificial Intelligence Workshop on Learning from Temporal and Spatial Data*, pp. 1–7, San Francisco, CA, USA, 2001.
- [25] S. Gao and Q. S. Li, “Evaluation index of wavelet de-noising based on entropy weight,” *Geomatics and Spatial Information Technology*, vol. 42, pp. 1–4, 2019.
- [26] T.-J. Hsieh, H.-F. Hsiao, and W.-C. Yeh, “Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm,” *Applied Soft Computing*, vol. 11, no. 2, pp. 2510–2525, 2011.
- [27] Z. Huang, L. Peng, X. Kunwu, X. U. Jiawei, S. Sanzhi, and S. O. Science, “Prediction model of futures based on wavelet denoising and neural network,” *Journal of Changchun University of Science and Technology (Natural Science Edition)*, vol. 42, pp. 129–132, 2019.
- [28] S. Lahmiri, “Wavelet transform, neural networks and the prediction of s&p price index: a comparative study of backpropagation numerical algorithms,” *Business Intelligence Journal*, vol. 5, pp. 235–244, 2012.
- [29] S. Lahmiri, “Randomness in denoised stock returns: the case of Moroccan family business companies,” *Physics Letters A*, vol. 382, no. 8, pp. 554–560, 2018.
- [30] S. Yang and J. Liu, “Time-series forecasting based on high-order fuzzy cognitive maps and wavelet transform,” *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 6, pp. 3391–3402, 2018.
- [31] X. Liang, X. Liang, W. Xu, and X. Wang, “A hybrid model for stock price based on wavelet transform and support vector machines,” in *Proceedings of the 2015 12th International Conference on Service Systems and Service Management*, pp. 1–7, Guangzhou, China, June 2015.
- [32] H. Yu, L. J. Ming, R. Sumei, and Z. Shuping, “A hybrid model for financial time series forecasting-integration of EWT, ARIMA with the improved ABC optimized ELM,” *IEEE Access*, vol. 8, pp. 84501–84518, 2020.
- [33] A. Mueen, N. Chavoshi, N. Abu-El-Rub, H. Hamooni, A. Minnich, and J. MacCarthy, “Speeding up dynamic time warping distance for sparse time series data,” *Knowledge and Information Systems*, vol. 54, no. 1, pp. 237–263, 2017.
- [34] A. Sharabiani, H. Darabi, S. Harford et al., “Asymptotic dynamic time warping calculation with utilizing value repetition,” *Knowledge and Information Systems*, vol. 57, no. 2, pp. 359–388, 2018.
- [35] T. Han, Q. Peng, Z. Zhu, Y. Shen, H. Huang, and N. N. Abid, “A pattern representation of stock time series based on DTW,” *Physica A: Statistical Mechanics and Its Applications*, vol. 550, p. 124161, 2020.
- [36] D. L. Donoho and I. M. Johnstone, “Ideal spatial adaptation by wavelet shrinkage,” *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [37] Q. J. Lan and C. Q. Ma, “Wavelet transformation methods for financial time series denoising,” *Science and Technology Management Research*, vol. 6, pp. 117–120, 2004.
- [38] J. F. Liu, “The establishment of effective portfolio based on wavelet transformation method and machine learning algorithm,” Master thesis, Southwestern University Of Finance And Economics, Chengdu, China, 2016.