

An Effective Approach for Obtaining a Group Trading Strategy Portfolio Using Grouping Genetic Algorithm

CHUN-HAO CHEN¹, YU-HSIUAN CHEN¹, JERRY CHUN-WEI LIN², AND MU-EN WU³

¹Department of Computer Science and Information Engineering, Tamkang University, Taipei 25137, Taiwan

²Department of Computing, Mathematics and Physics, Western Norway University of Applied Sciences, 5063 Bergen, Norway

³Department of Information and Finance Management, National Taipei University of Technology, Taipei 10608, Taiwan

Corresponding author: Mu-En Wu (mnasia1@gmail.com)

This work was supported by the Ministry of Science and Technology, China, under Grant MOST 107-2218-E-001-009- and Grant MOST 106-2221-E-032 -049-MY2.

ABSTRACT To determine an appropriate trading time for buying or selling stocks is always a difficult task. The common way to deal with it is using trading strategies formed by technical or fundamental indicators. Lots of approaches have been presented on how to form trading strategies and how to set suitable parameters for those strategies. Furthermore, some approaches were also designed to optimize a trading strategy portfolio, which is a set of strategies where the return and risk of the portfolio can be maximized and minimized, respectively. To provide a more useful trading strategy portfolio, we first define a group trading strategy portfolio (GTSP). Then, an algorithm that utilizes the grouping genetic algorithm is designed for solving the GTSP optimization problem. In the chromosome representation, the grouping, strategy, and weight parts are employed to encode a possible GTSP. The fitness value of a chromosome is calculated by the group balance, weight balance, portfolio return, and risk to assess the quality of every possible solution. Genetic operators, including crossover, mutation, and inversion, are applied on the population to form a new offspring. Evolution is continued until the stop conditions are reached. Lastly, experiments were conducted on two real datasets with different trends to show that the advantages and the effectiveness of the presented approach.

INDEX TERMS Group trading strategy portfolio, grouping genetic algorithm, portfolio optimization, trading strategy, trading strategy portfolio.

I. INTRODUCTION

Portfolio selection is a challenging area of research and optimization problem due to the nature of the financial markets in which various assets can be chosen to form a portfolio, including stocks, funds, and futures [12], [19], [23], [31]. In financial markets, because they are uncertain and are easily influenced by government policies and global economics, one of the main goals is to obtain a portfolio that can maximize profits and avoid risks [24]. The Mean-Variance (M-V) model, which is the most well-known model, was presented in 1952 by Markowitz for obtaining an efficient frontier that contains a set of portfolios [29]. In addition, in the past decades, based on the M-V model, various models have also been presented. Because finding the efficient frontier is an optimization problem, lots of approaches based on evolutionary algorithms have been published to

obtain the weights of a portfolio from a set of assets [5], [27], [28], [33].

When a portfolio is suggested, the following issue is how to determine the trading times. In other words, when to buy and sell assets to form the portfolio is a problem that is always troublesome for investors. For deciding the trading times for a portfolio, trading strategies are a common way to address that. In general, technical or fundamental indicators are employed to generate trading strategies [10], [24]. For example, assume that the technical indicator, that is, the commodity channel index (CCI), is utilized to form a trading strategy. Then, the buy and sell signals could be generated as “When the CCI value of a stock crosses -100 from the bottom, then a buy signal is suggested” and “When the CCI value of a stock crosses 100 from the top, a sell signal is suggested”, respectively. To generate useful and actionable trading strategies is a

complicated task. Lots of algorithms have also been designed to address that in different ways [3], [7], [15], [26].

However, when a trading strategy portfolio is provided, the suggested trading strategies may not satisfy traders for a variety of reasons. In this situation, traders can more conveniently adjust their trading plans if other trading strategies could be provided to replace them. To deal with this, a group trading strategy portfolio (GTSP) problem is firstly defined in this paper. Because obtaining a GTSP is a grouping problem, an optimization approach based on the grouping genetic algorithm (GGA) is designed for obtaining a GTSP in which Brown *et al.* indicated the GGA is an enhanced approach of the genetic algorithm (GA) for solving grouping problems [2].

To find a GTSP, according to the selected technical indicators, the candidate trading strategies are first generated in the proposed approach. Then, a subset of the candidate trading strategies is selected using the three given ranking functions. Based on the subset of the trading strategies, the population starts to be initialized. In the encoding scheme, the grouping, weighting and trading strategy parts are utilized to represent a possible GTSP. The evaluation function, which is composed of the group balance, weight balance, portfolio return factor and risk factor, is presented to assess the quality of a possible solution. The genetic operators of crossover, mutation, and inversion are applied on the population to generate and search the different solutions. The evolution is continued until it reaches the termination condition. The experiments on the sideway trend and downtrend datasets were conducted to show the advantages of the proposed approach. The contributions of this paper are stated as follows.

- (1) First, the group trading strategy portfolio optimization (GTSPPO) problem is defined in order to provide more convenient ways for traders to establish trading plans.
- (2) Then, an approach is designed to solve the GTSPPO in order to optimize a GTSP using the GGA.
- (3) Experiments on the sideways trend and downtrend datasets showed that the proposed approach could increase profits and reduce risks compared to the benchmark trading strategy, which is the buy and hold strategy (BHS).
- (4) In addition, when the stop-loss and take-profit points (SLTP) are taken into consideration, we also verify that the proposed approach can reach better returns than that without the SLTP.
- (5) Finally, the derived GTSP can be utilized to provide various trading strategy portfolios for traders to adjust more effective trading plans.

The remaining part of this paper is organized as follows. Literature reviews are described in Section 2. The problem definition and framework of the proposed approach are described in Section 3. The main elements of the proposed approach are stated in Section 4. The proposed algorithm and an example are given in Section 5. Extensive experiments are shown in Section 6. At last, conclusions and future work are provided in Section 7.

II. LITERATURE REVIEWS

Related studies are described in this section. In this section, the trading strategy optimization and trading strategies for stock portfolios are reviewed in Sections II.A and II.B, respectively.

A. TRADING STRATEGY OPTIMIZATION

Fu *et al.* presented two GA-based approaches for finding the appropriate parameters for technical indicators and portfolio weighting [15]. For parameter optimization, seven technical indicators were employed to form trading strategies. Then, a two-dimensional array, which was the encoding scheme, was utilized to represent those parameters for the trading strategies. The fitness is evaluated by the return of the trading strategies in a chromosome. In addition, to deal with finding trading rules from numerical indicators, Chang-Chien *et al.* presented an algorithm for constructing an associative classifier using the GA [6]. For obtaining the relationship from numerical data in stock trading problems, three relations were considered to express the associated classification rules. At last, the phenotype encoding structure was designed to represent those rules. The algorithm, called the GA-ACR algorithm, was employed to find the rules for constructing the classifier with high accuracy. According to the event-based concept of directional changes, Kampouridis *et al.* proposed an approach to find trading strategies for foreign exchange markets using GA [25]. In their approach, the profitability was used to evaluate a chromosome. As a result, new and profitable trading strategies could be generated. Utilizing news sentiment, Feuerriegel *et al.* designed a supervised and reinforcement learning algorithm for deriving trading strategies to provide investors a more useful decision supports [16]. Besides, Wen *et al.* proposed a framework in accordance with the cointegration theory and complex network. In that paper, the static and dynamic features of pairs trading strategy were examined, and the results indicated investors should carefully chose pairs for trading because some trading pairs were unsteady [32]. To construct profitable currency portfolios, Chandrinos *et al.* focused on designing trading strategy using the Donchian Channel [4]. They first developed the modified renko bars. Subsequently, parameters of the trading strategy were tuned based on the given training datasets and compared with the three optimization algorithms.

B. TRADING STRATEGIES FOR STOCK PORTFOLIOS

Chou *et al.* proposed an approach to construct a rule-based dynamic stock trading system [10]. In their approach, technical indicators were first employed to form trading strategies. Then, using the designed quantum-inspired Tabu search algorithm, an optimal combination of trading strategies could be found. At last, the algorithm adopted a sliding window to reduce over-fitting and to realize the dynamic system. Furthermore, using maximizing the profit gained and the rate of successful transactions as two objective functions, they also presented a multi-objective quantum-inspired Tabu

search algorithm for identifying the optimal combination of trading rules [11]. Leu and Chiu [26] proposed an algorithm to obtain a stock portfolio trading strategy by utilizing the GA and fuzzy time series. First, it employed the GA to find a stock portfolio. The evaluation was designed in accordance with the fuzzy time series. In addition, to avoid huge risk, the periodic checking and stop-loss point checking were utilized to get the trading signals to sell or buy assets for a stock portfolio. Chang and M. Lee designed formulated investment strategies for stocks based on the Markov decision process and the GA [3]. The Markov decision process was utilized to obtain trading signals. Then, the GA was used to find the stock portfolio and capital allocation for assets. In addition, Berutich *et al.* proposed a genetic programming based approach to learn trading rules that can be used to manage a stock portfolio [1]. To overcome the over-fitting problem, in their approach, the new random sampling method was utilized for increasing the robustness of the strategies. Chen *et al.* designed a combination genetic algorithm to optimize an investment strategy portfolio [7]. It first generated security-rule pairs using the ten given technical indicators and ten stocks, where each indicator had sell and buy rules. Then, the selected security-rule pairs were utilized to form a possible investment strategy portfolio that was encoded into an individual. The fitness value of an investment strategy portfolio using chromosomes was evaluated by its return.

III. PROBLEM DEFINITION AND FRAMEWORK OF THE PROPOSED APPROACH

To state the problem to be solved in this paper, the definitions are described in this section as follows.

Definition 1 (A Trading Strategy (TS)): A TS is composed of a buy signal and a sell signal, where each signal is formed according to technical indicators. For example, a trading strategy could be “If the CCI value of a stock crosses -100 from the bottom, then a buy signal is suggested. If the CCI value of a stock crosses 100 from the top, a sell signal is suggested”.

Definition 2 (A Trading Strategy Portfolio (TSP)): A TSP has h TSs and is expressed as $TSP = \{TS_1, TS_2, \dots, TS_h\}$, where h is larger than 2. For instance, a TSP could be $\{TS_1$: “If the CCI value of a stock crosses -100 from the bottom, then a buy signal is suggested, and if the CCI value of a stock crosses 100 from the top, then a sell signal is suggested”, TS_2 : “If the 5-day moving average value crosses the 20-day moving average from the bottom, then a buy signal is suggested, and if the 20-day moving average value crosses the 5-day moving average from the top, then a sell signal is suggested” $\}$. Note that the return and risk values of a TSP are a common way to evaluate the quality of a TSP.

Definition 3 (Trading Strategy Group (TSG)): A TSG contains a set of trading strategies $TSG = \{TS_1, TS_2, \dots, TS_n\}$. The TSs in the same group have properties that are similar. For instance, the properties could be the ability to avoid risk.

Definition 4 (Grouping Problem): The purpose of the grouping problem is to divide n instances into a specific

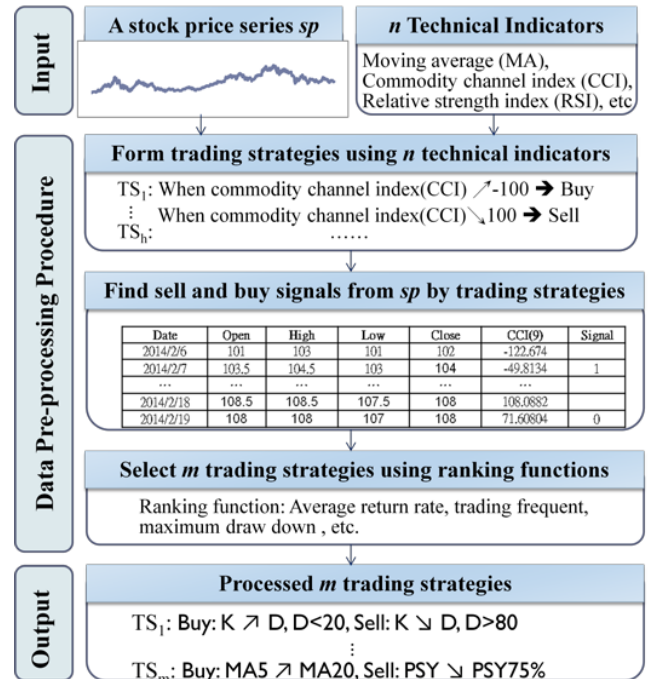


FIGURE 1. Data preprocessing procedure to form trading strategies.

number of groups, where each instance can only belong to a group. According to [13] and [14], given a set of instances $U = \{u_1, u_2, \dots, u_n\}$, the definition of grouping problem can be represented as: $\cup U_i = U$, and $U_i \cap U_j = \emptyset, i \neq j$. In general, the grouping problem has a cost function, which is formed from the hard constraints of each problem and used to evaluate the quality of combinations of elements generated from groups.

Definition 5 (Group Trading Strategy Portfolio (GTSP)): A GTSP consists of K TSGs and is represented as $GTSP = \{TSG_1, TSG_2, \dots, TSG_K\}$. Using a GTSP, lots of trading strategy portfolios can be formed, which is represented as $|GTS_1| \times |GTS_2| \times \dots \times |GTS_K|$. For instance, assume that a GTSP has three groups that contain 3, 4, and 3 TSs. Thus, $36 (= 3 \times 4 \times 3)$ TSPs can be derived in total.

Definition 6 (Group Trading Strategy Portfolio Optimization (GTSPo)) Problem: The main goal of the GTSPo problem is to optimize a GTSP in accordance with objective and subjective criteria given by users using evolutionary algorithms. For instance, the criteria could be minimizing the risk while maximizing the profit based on the allocated capitals of groups. Based on **Definitions 4** and **5**, we can map the GTSPo problem to a grouping problem.

To solve the GTSPo problem, the GGA is utilized in this paper to optimize a GTSP. In the following, before stating the proposed framework, the data preprocessing procedure to generate TSs is described in Fig. 1.

In Fig. 1, through the given stock price series and technical indicators, it shows that three steps are used in the data preprocessing procedure to generate the m processed TSs. First, according to the selected n technical indicators,

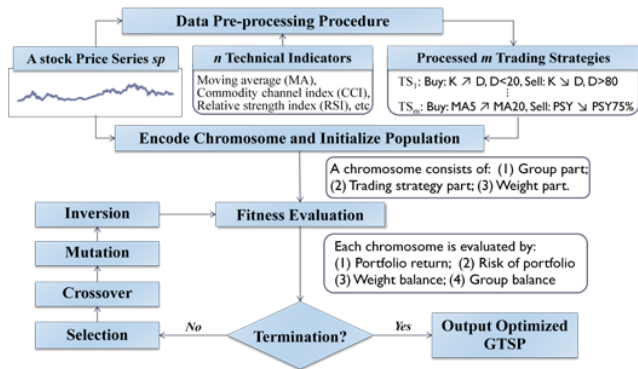


FIGURE 2. Framework of the proposed approach.

a combination of them is utilized to generate all possible TSs. Second, trading signals, including sell and buy signals, are derived from the stock price series for every TS. In the third step, m TSs are kept by using the three ranking functions, including the average return rate, the trading frequency, and the maximum draw down. By combining the data preprocessing procedure, the framework of the proposed algorithm is shown in Fig. 2.

In Fig. 2, first, the data preprocessing procedure is executed to derive m TSs, as described in Fig. 1. Then, an initial population is formed based on the predefined population size and the derived m TSs, where every chromosome in the population is a GTSP and is represented by grouping, trading strategy and weighting parts. The four factors, including the portfolio return, the risk of the portfolio, the weight balance and the group balance, are utilized to form the fitness function for the chromosome quality evaluation. According to fitness values of solutions, the selection operation is used to produce the next population. The genetic operators of crossover, mutation and inversion are executed to find new solutions. The evolution is repeated until the stop conditions are reached. Finally, the best GTSP that has the highest fitness value will be delivered to users for making trading plans.

IV. COMPONENTS OF PROPOSED APPROACH

In this section, the three main components of the proposed algorithm, including the encoding schema, the fitness function and reproduction, and the genetic operations, are stated in Sections VI.A to VI.C.

A. ENCODING SCHEMA

To effectively solve a task using the GA, it is important to design a data structure to represent a solution, which is known as the encoding schema. The purpose of this paper is to design an algorithm to obtain a GTSP that could reach both objective and subjective criteria. As stated in Section III, a GTSP is a set of groups where every group has TSs and the allocated capital. Thus, three parts, namely, the grouping, the trading strategy, and the weight parts, are utilized to represent a GTSP. The encoding schema is illustrated in Fig. 3.

In Figure 3, the grouping part indicates the number of groups to which TSs can belong. The trading strategy part

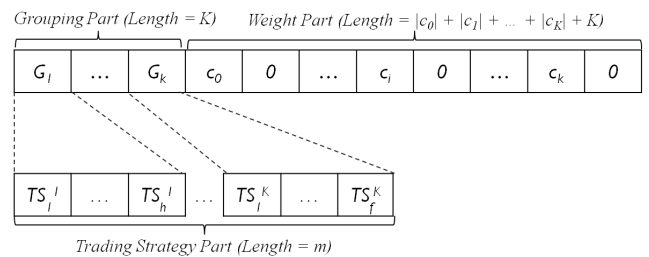


FIGURE 3. Encoding schema for a GTSP.

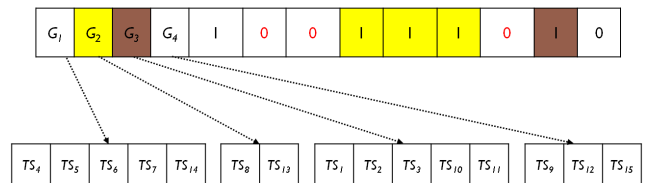


FIGURE 4. A possible chromosome.

expresses what TSs could be included in every group. Each c_j in the weight part is a ‘1’ string that represents the invested capital of the j -th groups, and c_0 represents the capital to be reserved. A symbol ‘0’ is employed to separate the two groups c_i and c_j . In accordance with the encoding schema, the initial population can be then generated randomly. Below, an example is given to demonstrate a chromosome. Assume that there are 15 trading strategies and that the number of groups is 4. A possible chromosome is shown in Fig. 4.

Figure 4 shows that there are four groups, G_1, G_2, G_3 and G_4 . Group G_1 contains the five trading strategies TS_4, TS_5, TS_6, TS_7 and TS_{14} . Since the number of ‘1’s in c_0 and the total number of ‘1’s in the weighting part are one and five, respectively, it means that 20% ($= 1 / 5$) capital is reserved. In the same way, we know that the number of ‘1’s in c_1 to c_4 is 0, 3, 1 and 0, respectively. Thus, the weights for groups G_1 to G_4 are 0%, 60%, 20% and 0%, respectively. In this GTSP, 10 ($= 2 \times 5$) trading strategy portfolios can be generated and provided to users.

B. FITNESS FUNCTION AND REPRODUCTION

To derive an optimized GTSP, various factors, such as objective and subjective criteria, should be taken into consideration for designing an effective fitness function. For evaluating the quality of a TSP, the maximum return and minimum risk are the two commonly known objective factors that should be considered. In addition, because every group has its own TSs that have similar properties and allocated weights, different factors that are the group balance and the weight balance should also be designed to access groups in a GTPS. As a result, the fitness function, which is shown in Formula (1), is used to evaluate a chromosome.

$$f(C_q) = PR(C_q) * Risk(C_q) * GB(C_q)^\alpha * WB(C_q), \quad (1)$$

where $PR(C_q)$ is the portfolio return, $Risk(C_q)$ is the risk of a GTSP, $GB(C_q)$ and $WB(C_q)$ are the group balance and weight balance, and α is a parameter used to show the influence of

the group balance. In the following, those four factors are described. The $PR(C_q)$ is defined in Formula (2).

$$PR(C_q) = \frac{\sum_{j=1}^{numTSP} return(TSP_j)}{numTSP}, \quad (2)$$

where $return(TSP_j)$ is the profit of a trading strategy portfolio TSP_j , and $numTSP$ is the number of TSPs that can be formed from C_q . The $return(TSP_j)$ is defined in Formula (3).

$$return(TSP_j) = \sum_{i=1}^K avgReturnRate(TS_i^j) * weight_i * Capital, \quad (3)$$

where $avgReturnRate(TS_i^j)$ is the average return rate of TS in group G_i , $weight_i$ is the weight of group G_i , and $Capital$ is the capital that a trader uses for trading. The higher the returns of TSPs generated from a chromosome are, the better the GTSP is. The $avgReturnRate(TS_i^j)$ is stated in Formula (4).

$$avgReturnRate(TS_i^j) = \frac{\sum_{h=1}^{frequency_i} returnRate(TS_{ih}^j)}{frequency_i}, \quad (4)$$

where $frequency_i$ is the number of trading times in the given training period, and $returnRate(TS_{ih}^j)$ is described in Formula (5).

$$returnRate(TS_{ih}^j) = \frac{sellPrice_h - buyPrice_h}{buyPrice_h}, \quad (5)$$

where the $sellPrice_h$ and $buyPrice_h$ are the sell and buy prices of the h -th trading using the i -th TS in TSP_j , respectively. Then, the risk of a chromosome is shown in Formula (6).

$$risk(C_q) = \frac{\sum_{j=1}^{numTSP} risk(TSP_j)}{numTSP}, \quad (6)$$

where $risk(TSP_j)$ is the risk of a TSP, $numTSP$ is the number of TSPs generated from C_q , and $risk(TSP_j)$ is defined in Formula (7).

$$risk(TSP_j) = \min(MDD(TS_1^j), \dots, MDD(TS_k^j)), \quad (7)$$

where $MDD(TS_i^j)$ is the maximum draw down of a TS in G_i of TSP_j , and K is the number of groups. It indicates that the risk of a TSP is calculated using the minimum and maximum draw down of TSs in the TSP. Note that the maximum draw down of TSs will be normalized from 0 to 1. The $MDD(TS_i^j)$ is given in Formula (8).

$$MDD(TS_i^j) = \min(returnRate(TS_{i1}^j), \dots, returnRate(TS_{ifrequent_i}^j)), \quad (8)$$

where $returnRate(TS_{ih}^j)$ is the same as that given in Formula (5). The $GB(C_q)$, which is the group balance of a chromosome, is defined in Formula (9).

$$GB(C_q) = \sum_{i=1}^K -\frac{|G_i|}{N} \log \frac{|G_i|}{N}, \quad (9)$$

Input: Two randomly selected chromosomes C_A and C_B .
Output: Two new chromosomes C_A^{new} and C_B^{new} .
Procedure: Crossover on weight part
1. Crossover () {
2. Generate two points s and t ;
3. WHILE C_A and C_B have equal number of 1 between s and t DO
4. $t = t - 1$;
5. END WHILE LOOP
6. FOR $i = s$ to t DO
7. $Temp = C_A[i]$;
8. $C_A[i] = C_B[i]$;
9. $C_B[i] = Temp$;
10. END i FOR LOOP
11. Output C_A^{new} and C_B^{new}
12. }

FIGURE 5. Crossover on the weight part.

where $|G_i|$ is the number of TSs in G_i , and N is the number of the given TSs. The purpose of the group balance is to ensure that the number of TSs is similar among groups in the chromosome C_q . The weight balance, $WB(C_q)$, is defined in Formula (10).

$$WB(C_q) = \sum_{i=1}^{K+1} -\frac{|c_i|}{T} \log \frac{|c_i|}{T}, \quad (10)$$

where $|c_i|$ is the length of the string c_i , and T is the summation of length of strings c_i , $0 \leq i \leq K$. The goal of the weighted balance is to reduce the probability that all capital is located in certain specific groups. Utilizing the designed fitness function, the fitness value of each chromosome can be evaluated. The reproduction can be executed in accordance with the fitness values to select the next population using different strategies. For instance, elitist selection or roulette wheel selection could be employed, where elitist selection is used to form the next population in the proposed algorithm.

C. GENETIC OPERATIONS

The genetic operations are introduced in this section, including crossover, mutation and inversion. For crossover operations, they are applied on the grouping and weighting parts. In the grouping part, it first picks two individuals randomly, where one is used as the base chromosome and the other one is the insertion chromosome. Then, parts of groups in the insertion chromosome are selected and put into the base chromosome. After insertion, groups in the base chromosome are removed if they are redundant. At last, groups are merged or split until the number of groups in a chromosome is correct. To execute the crossover operation on the weight part, the two-point crossover operator is employed to generate new chromosomes in which two points will first be determined and their sub sequences will be exchanged. Note that the appropriate arrangement should be made to ensure that the numbers of '1's and '0's in a chromosome are correct. Below, the procedure for the crossover on the weight part is shown in Fig. 5.

Figure 5 shows that two crossover points, s and t , are first generated (Line 2). The genes between the s and t of

TABLE 2. The portfolio returns of the ten chromosomes.

C_q	portfolioReturn(C_q)	C_q	portfolioReturn (C_q)
C_1	8549.419	C_6	10257.397
C_2	10775.223	C_7	7723.611
C_3	4039.327	C_8	7320.010
C_4	8088.963	C_9	10376.927
C_5	6612.954	C_{10}	8030.328

strings in the following sections. In other words, chromosome C_1 is represented as: $\{\{3, 4, 5, 10\}, \{0, 1, 6, 8, 9, 11, 12, 14\}, \{2, 7, 13\}, 0.29, 0.5, 0.14, 0.07\}$.

Sub-steps 1.3: Repeating the previous two substeps, an initial population with ten chromosomes is generated as follows:

C_1 : $\{\{3, 4, 5, 10\}, \{0, 1, 6, 8, 9, 11, 12, 14\}, \{2, 7, 13\}, 0.29, 0.5, 0.14, 0.07\}$;

C_2 : $\{\{3, 5, 8, 11, 13\}, \{1, 4, 6, 7, 10, 12\}, \{0, 2, 9, 14\}, 0.05, 0.18, 0.11, 0.66\}$;

C_3 : $\{\{1, 3, 8, 11, 13\}, \{0, 4, 6, 7, 10, 12, 14\}, \{2, 5, 9\}, 0.62, 0.26, 0.07, 0.05\}$;

C_4 : $\{\{3, 8, 13, 14\}, \{1, 6, 7, 12\}, \{0, 2, 4, 5, 9, 10, 11\}, 0.29, 0.13, 0.04, 0.54\}$;

C_5 : $\{\{1, 6, 7, 8, 11, 12\}, \{2, 3, 10\}, \{0, 4, 5, 9, 13, 14\}, 0.44, 0.05, 0.29, 0.22\}$;

C_6 : $\{\{0, 1, 2, 5, 8, 9\}, \{3, 4, 10\}, \{6, 7, 11, 12, 13, 14\}, 0.08, 0.02, 0.65, 0.25\}$;

C_7 : $\{\{1, 4, 6\}, \{0, 3, 5, 8, 9, 10, 11, 12\}, \{2, 7, 13, 14\}, 0.33, 0.25, 0.23, 0.19\}$;

C_8 : $\{\{1, 2, 7, 9, 11, 12\}, \{4, 6, 10, 13, 14\}, \{0, 3, 5, 8\}, 0.21, 0.04, 0.67, 0.08\}$;

C_9 : $\{\{1, 5, 7, 8, 10, 12\}, \{0, 3, 4, 6\}, \{2, 9, 11, 13, 14\}, 0.11, 0.12, 0.4, 0.37\}$;

C_{10} : $\{\{0, 1, 3, 5, 6, 7, 12, 13\}, \{2, 9, 10, 14\}, \{4, 8, 11\}, 0.24, 0.17, 0.01, 0.58\}$.

STEP 2: The fitness value of each chromosome is calculated using the following sub-steps.

Sub-step 2.1: The trading strategy portfolios are first generated. Take chromosome C_1 as an example. According to its grouping part $[G_1: \{3, 4, 5, 10\}, G_2: \{0, 1, 6, 8, 9, 11, 12, 14\}, G_3: \{2, 7, 13\}]$, there are 96 ($= 4 \times 8 \times 3$) possible trading strategy portfolios. All of them are collected in a set $TSP = \{tsp_1, tsp_2, \dots, tsp_{96}\} = \{\{3, 0, 2\}, \{3, 0, 7\}, \{3, 0, 13\}, \dots, \{10, 14, 13\}\}$.

Sub-step 2.2: The portfolio return of each chromosome is calculated as follows.

Sub-step 2.2.1: The return of each trading strategy portfolio is then calculated. Take $tsp_1: \{3, 0, 2\}$ as an example. According to the weight part of chromosome C_1 , which is $[0.29, 0.5, 0.14, 0.07]$, we assume that investment capital is 100000, and the profits of tsp_1 are 24658 ($= [0.354 \times (100000 \times 0.5) + 0.320 \times (100000 \times 0.14) + 0.354 \times (100000 \times 0.07)]$).

Sub-step 2.2.2: The portfolio return of a chromosome can be calculated using Formula (2), and the results are shown in Table 2.

Sub-step 2.3: The risk in each chromosome is then calculated as follows.

TABLE 3. Normalized MDD for every strategy.

Strategy Name	MDD	Normalized MDD
TS_0 : Buy: PSY \nearrow 25% ; Sell: PSY \searrow 75%	0.320	0.742
TS_1 : Buy: MOM \nearrow 0 ; Sell: PSY \searrow 75%	0.424	0.945
TS_2 : Buy: +DI \nearrow -DI ; Sell: PSY \searrow 75%	0.354	0.809
TS_3 : Buy: CCI \nearrow -100 ; Sell: PSY \searrow 75%	0.354	0.809
TS_4 : Buy: D < 20, K \nearrow D ; Sell: PSY \searrow 75%	0	0.117
TS_5 : Buy: MA5 \nearrow MA 20 ; Sell: PSY \searrow 75%	0.294	0.691
TS_6 : Buy: BIAS \nearrow -4.5% ; Sell: PSY \searrow 75%	0.354	0.808
TS_7 : Buy: DIF \nearrow MACD \searrow EM, or DIF \nearrow 0 ; Sell: PSY \searrow 75%	0.452	1
TS_8 : Buy: D < 20, K \nearrow D ; Sell: +D \nearrow -DI	0	0.117
TS_9 : Buy: PSY \nearrow 25% ; Sell: +D \nearrow -DI	0.013	0.143
TS_{10} : Buy: WMS%R \searrow 80 ; Sell: PSY \searrow 75%	0.308	0.719
TS_{11} : Buy: D < 20, K \nearrow D ; Sell: MOM \searrow 0	0	0.117
TS_{12} : Buy: MOM \nearrow 0 ; Sell: +D \nearrow -DI	-0.033	0.053
TS_{13} : Buy: PSY \nearrow 25% ; Sell: MA5 \searrow MA 20	-0.018	0.082
TS_{14} : Buy: +DI \nearrow -DI ; Sell: MA5 \searrow MA 20	-0.060	0

TABLE 4. The risks of ten chromosomes.

C_q	risk(C_q)	C_q	risk(C_q)
C_1	0.093	C_6	0.136
C_2	0.165	C_7	0.129
C_3	0.163	C_8	0.162
C_4	0.132	C_9	0.155
C_5	0.100	C_{10}	0.200

TABLE 5. The group balances of all chromosomes.

C_q	GB(C_q)	C_q	GB(C_q)
C_1	1.019	C_6	1.113
C_2	1.178	C_7	1.019
C_3	1.089	C_8	1.178
C_4	1.125	C_9	1.178
C_5	1.113	C_{10}	1.019

Sub-step 2.3.1: The maximum draw down of every trading strategy is normalized to 0 to 1. After normalization, the maximum draw down of each trading strategy is shown in Table 3.

Sub-step 2.3.2: The risk of each tsp_j in the set TSP is calculated. Take $tsp_1: \{3, 0, 2\}$ as an example. The risk of tsp_1 is calculated as 0.742 ($= \min(0.809, 0.742, 0.809)$) by Formula (7).

Sub-step 2.3.3: The risk of a chromosome is set by Formula (6), and the results are shown in Table 4.

Sub-step 2.4: The group balance of each chromosome is calculated according to its grouping part. Take chromosome C_1 as an example. Since the grouping part of C_1 is $[G_1: \{3, 4, 5, 10\}, G_2: \{0, 1, 6, 8, 9, 11, 12, 14\}, G_3: \{2, 7, 13\}]$, the $GB(C_1)$ is calculated as 1.019 using Formula (9). The group balances of the other chromosomes are shown in Table 5.

Sub-step 2.5: The weight balance of each chromosome is then calculated according to its weight part. Take chromosome C_1 as an example. Since the weight part of C_1 is $[0.29, 0.5, 0.14, 0.07]$, the $WB(C_1)$ is calculated as 1.362 using Formula (10). The weight balances of the other chromosomes are shown in Table 6.

Sub-step 2.6: The fitness value of chromosome C_1 is 1124.463 ($= 8549.419 \times 0.093 \times 1.019^2 \times 1.362$) using Formula (1) and assuming that α is 2. In the same way, the fitness values of other chromosomes can be calculated and the results are given in Table 7.

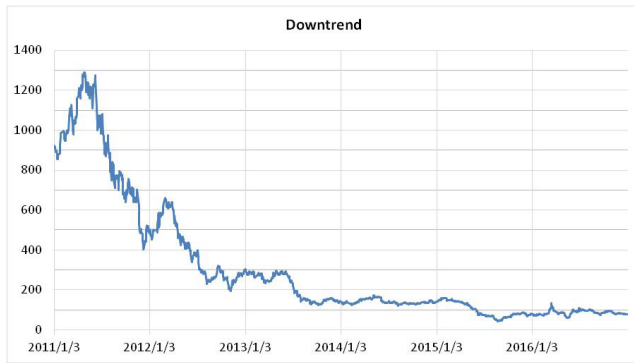


FIGURE 8. The downtrend of the dataset.

TABLE 8. The trading rules generated using the ten technical indicators.

BR#	Description	SR#	Description
BR ₁	Buy: MA5 \nearrow MA20	SR ₁	Sell: MA5 \searrow MA20
BR ₂	Buy: RSI \nearrow 30	SR ₂	Sell: RSI \searrow 70
BR ₃	Buy: WMS%R \searrow 80	SR ₃	Sell: WMS%R \nearrow 20
BR ₄	Buy: MOM \nearrow 0	SR ₄	Sell: MOM \searrow 0
BR ₅	Buy: PSY \nearrow 25%	SR ₅	Sell: PSY \searrow 75%
BR ₆	Buy: CCI \nearrow -100	SR ₆	Sell: CCI \searrow 100
BR ₇	Buy: D < 20, K \nearrow D	SR ₇	Sell: D > 80, K \searrow D
BR ₈	Buy: DIF \nearrow	SR ₈	Sell: DIF \searrow
	MACD(DEM), or DIF \nearrow 0		MACD(DEM), or DIF \searrow 0
BR ₉	Buy: BIAS \nearrow -4.5%	SR ₉	Sell: BIAS \searrow 5%
BR ₁₀	Buy: +DI \nearrow -DI	SR ₁₀	Sell: +DI \searrow -DI

respectively. When using BHS, the returns of 2012, 2013, 2014, 2015 and 2016 are -40% , -53% , -1% , -45% and -2% , respectively.

In the proposed approach, ten technical indicators presented in [7] and [15] were used in the data preprocessing procedure to form the trading strategies, and the parameter settings of the used indicators were set according to [22]. The ten technical indicators are the Moving Average (MA), the Relative Strength Index (RSI), the Williams%R (WMS%R), Momentum (MOM), Psychology (PSY), the Commodity Channel Index (CCI), the stochastic oscillator (KD), the Moving Average Convergence-Divergence (MACD), the Bias ratio (BIAS), and the Directional Movement Index (DMI). Trading rules are then generated using the ten technical indicators to identify buy or sell signals and are shown in Table 8.

Table 8 shows that there are twenty trading rules, ten for buying rules and ten for selling rules, that are used for finding trading signals. Then, a trading strategy can be formed using one buying rule and one selling rule. For example, a trading strategy could be “If MA5 rises above MA20, it generates a buying signal, and if MA5 falls below MA20, it generates a selling signal.” Because there are 10 buying and 10 selling rules, 100 different trading strategies can be formed. Hence, using the data preprocessing procedure with the ranking functions, m trading strategies can be selected to form a GTSP.

Experimental results on the two different market trends are shown and discussed in Sections VI.B and VI.C. Two sets of trading strategies are picked using the data preprocessing procedures, namely, TOP15 and TOP555, to generate the

TABLE 9. Comparison results of the derived GTSPs and BHS in terms of returns on the sideways trend dataset using the TOP15.

Training Period	Testing Period	GTSP without SLTP	GTSP with SLTP	BHS
2011	2012	-0.07	-0.06(-5% ~ 5%)	0.02
2012	2013	0.04	0.19(-15% ~ 15%)	0.34
2013	2014	0.10	0.24(-15% ~ 15%)	0.26
2014	2015	0.22	0.27 (-15% ~ 15%)	0.11
2015	2016	0.05	0.11 (-10% ~ 10%)	-0.20
2011 to 2012	2013	0.08	0.19(-15% ~ 15%)	0.34
2012 to 2013	2014	0.03	0.19(-15% ~ 15%)	0.26
2013 to 2014	2015	0.19	0.26 (-10% ~ 15%)	0.11
2014 to 2015	2016	0.08	0.15 (-15% ~ 15%)	-0.20
2011 to 2013	2014	0.27	0.30 (-15% ~ 15%)	0.26
2012 to 2014	2015	0.16	0.27 (-15% ~ 15%)	0.11
2013 to 2015	2016	0.08	0.16 (-5% ~ 10%)	-0.20
Average Returns:		0.103	0.188	0.100

GTSPs to verify the proposed approach. The top-15 trading strategies are selected from the 100 trading strategies to form the TOP15 set using the average return rate as a ranking function. To generate the TOP555 set, each of three ranking functions, which are the average return rate, trading frequency and maximum draw down, is utilized to pick five trading strategies. Then, the fifteen trading strategies are used to derive the GTSPs from the datasets with different training and testing intervals.

B. EVALUATION ON THE SIDWAYS TREND DATASET

In this section, experiments were conducted on the sideways trend dataset using the two trading strategies sets, TOP15 and TOP555. Tables 9 and 10 show the comparison results of the derived GTSPs using TOP15 and TOP555 on different training and testing periods with stop-loss and take-profit points (SLTP) and that using the BHS in terms of returns. Note that when the SLTP is considered in the proposed algorithm, it not only can reduce the transaction cost because the number of trading frequency is decreased but also provide the ability for the algorithm to avoid huge loss.

From Table 9, we can observe that the returns of the proposed approach with SLTP are better than that without SLTP. Then, when the training periods are more than two years, most of the derived GTSPs can achieve better returns than the BHS. In addition, when the testing period is 2016, the returns of the GTSPs trained using one, two and three years as training periods are 11%, 15%, and 16%, respectively. Because the return of the BHS is -20% , we can easily understand that the proposed approach is better than the BHS and can have positive returns on the sideways trend dataset. Comparing them in terms of average returns, we can also know that the proposed approach can have higher returns (18.8%) than the BHS (10%).

From Table 10, we can see that the returns of the proposed approach with SLTP are still better than that without SLTP. Comparing the proposed approach with the BHS, we can also observe that the proposed approach is effective because the returns of the proposed approach are almost positive. Especially in 2016, the returns of the derived GTSPs trained from one-, two- and three-year training datasets are

TABLE 10. Comparison results of the derived GTSPs and BHS in terms of returns on the sideways trend dataset using the TOP555.

Training Period	Testing Period	GTSP without SLTP	GTSP with SLTP	BHS
2011	2012	-0.07	-0.10(-5% ~ 5%)	0.02
2012	2013	0.04	0.20(-15% ~ 15%)	0.34
2013	2014	0.13	0.23(-15% ~ 15%)	0.26
2014	2015	0.19	0.27 (-10% ~ 15%)	0.11
2015	2016	0.08	0.13 (-10% ~ 10%)	-0.20
2011 to 2012	2013	0.06	0.19(-10% ~ 15%)	0.34
2012 to 2013	2014	0.04	0.16(-15% ~ 15%)	0.26
2013 to 2014	2015	0.12	0.22 (-10% ~ 10%)	0.11
2014 to 2015	2016	0.09	0.19 (-15% ~ 15%)	-0.20
2011 to 2013	2014	0.20	0.27 (-15% ~ 5%)	0.26
2012 to 2014	2015	0.13	0.15 (-5% ~ 15%)	0.11
2013 to 2015	2016	0.08	0.18 (-10% ~ 15%)	-0.20
Average Returns:		0.091	0.174	0.10

TABLE 11. Comparison results of the derived GTSPs and BHS in terms of returns on the downtrend dataset using the TOP15.

Training Period	Testing Period	GTSP without SLTP	GTSP with SLTP	BHS
2011	2012	-0.18	-0.18 (-5% ~ 5%)	-0.41
2012	2013	-0.06	-0.08 (-5% ~ 5%)	-0.54
2013	2014	-0.01	0.07 (-15% ~ 5%)	-0.01
2014	2015	-0.28	-0.16 (-10% ~ 15%)	-0.45
2015	2016	0.04	0.08 (-15% ~ 15%)	-0.02
2011 to 2012	2013	-0.08	0.00 (-15% ~ 15%)	-0.54
2012 to 2013	2014	-0.01	0.09 (-15% ~ 5%)	-0.01
2013 to 2014	2015	-0.45	-0.14 (-5% ~ 15%)	-0.45
2014 to 2015	2016	0.05	0.11 (-15% ~ 15%)	-0.02
2011 to 2013	2014	0.00	0.07 (-10% ~ 10%)	-0.01
2012 to 2014	2015	0.00	-0.05 (-10% ~ 15%)	-0.45
2013 to 2015	2016	0.04	0.05 (-10% ~ 10%)	-0.02
Average Returns:		-0.078	-0.016	-0.244

TABLE 12. Comparison results of the derived GTSPs and BHS in terms of returns on the downtrend dataset using the TOP555.

Training Period	Testing Period	GTSP without SLTP	GTSP with SLTP	BHS
2011	2012	-0.21	-0.21 (-5% ~ 5%)	-0.41
2012	2013	-0.06	-0.08 (-10% ~ 5%)	-0.54
2013	2014	-0.02	0.05 (-15% ~ 5%)	-0.01
2014	2015	-0.27	-0.16 (-5% ~ 15%)	-0.45
2015	2016	0.04	0.10 (-5% ~ 15%)	-0.02
2011 to 2012	2013	-0.07	0.00 (-15% ~ 15%)	-0.54
2012 to 2013	2014	-0.01	0.04 (-15% ~ 5%)	-0.01
2013 to 2014	2015	-0.53	-0.09 (-5% ~ 15%)	-0.45
2014 to 2015	2016	0.06	0.14 (-15% ~ 15%)	-0.02
2011 to 2013	2014	0.00	0.04 (-10% ~ 10%)	-0.01
2012 to 2014	2015	0.00	-0.06 (-5% ~ 10%)	-0.45
2013 to 2015	2016	0.06	0.06 (-10% ~ 10%)	-0.02
Average Returns:		-0.084	-0.014	-0.244

13%, 19% and 18%, respectively, which are better than the BHS of -20%.

C. EVALUATION ON THE DOWNTREND DATASET

Experiments were then conducted on the downtrend dataset. In the following, the TOP15 and TOP555 are used to verify the effectiveness of the proposed approach. The comparison results of the proposed approach using the TOP15 and TOP555 on different training and testing periods and the BHS in terms of returns are given in Tables 11 and 12.

From Table 11, we can observe that irrespective of whether one-, two- or three-year training datasets are used for training,

TABLE 13. The optimized GTSPs on the sideways trend dataset.

Sideways Trend Dataset (BHS = -0.20)					
TOP15	Proposed Approach Without SLTP	Best GTSP: {14, 52, 89, 98, 96, 60}, {69, 41, 78, 87}, {5, 7, 43, 32, 23}, 0.14, 0.24, 0.37, 0.25			
		MAXROI	AVGROI	MINROI	
			0.18	0.10	0.04
	Proposed Approach With SLTP	Best GTSP: {32, 87, 96, 55, 41}, {60, 92, 1, 46, 65}, {44, 78, 38, 14, 43}, 0.13, 0.29, 0.29, 0.29			
MAXROI		AVGROI	MINROI		
		0.22	0.15	0.00	
TOP555	Proposed Approach Without SLTP	Best GTSP: {41, 60, 14, 32, 70}, {78, 87, 7, 73, 67, 66, 65}, {5, 69, 96}, 0.14, 0.27, 0.19, 0.4			
		MAXROI	AVGROI	MINROI	
			0.19	0.09	0.03
	Proposed Approach With SLTP	Best GTSP: {41, 60, 14, 32, 70}, {78, 87, 7, 73, 67, 66, 65}, {5, 69, 96}, 0.14, 0.37, 0.25, 0.24			
MAXROI		AVGROI	MINROI		
		0.24	0.19	0.15	

TABLE 14. The optimized GTSPs on the downtrend dataset.

Downtrend Dataset (BHS = -0.02)					
TOP15	Proposed Approach Without SLTP	Best GTSP: {38, 65, 8, 37, 26}, {28, 29, 44, 19, 82, 99}, {73, 46, 17, 71}, 0.13, 0.27, 0.26, 0.34			
		MAXROI	AVGROI	MINROI	
			0.09	0.06	0.02
	Proposed Approach With SLTP	Best GTSP: {28, 38, 17, 64, 20}, {73, 65, 77, 9, 44}, {82, 37, 74, 81, 26}, 0.13, 0.27, 0.27, 0.33			
MAXROI		AVGROI	MINROI		
		0.23	0.11	0.00	
TOP555	Proposed Approach Without SLTP	Best GTSP: {73, 46}, {28, 38, 15, 67, 69, 32, 23}, {65, 11, 12, 71, 14, 5}, 0.25, 0.62, 0.0, 0.13			
		MAXROI	AVGROI	MINROI	
			0.16	0.08	0.04
	Proposed Approach With SLTP	Best GTSP: {99, 90, 8, 16, 35}, {82, 73, 37, 28}, {74, 23, 60, 14, 88, 66}, 0.23, 0.12, 0.63, 0.02			
MAXROI		AVGROI	MINROI		
		0.22	0.17	0.12	

the proposed approach is better than the BHS in terms of returns. However, in the downtrend dataset, the proposed approach with SLTP can still reach better returns than that without the SLTP in most cases. Comparing the proposed approach with the BHS in terms of average returns, because their returns are -0.016 and -0.244, we can clearly know that the proposed approach is more effective at avoiding huge losses than the BHS. In addition, since the returns of the GTSPs trained from the two-year training periods are almost positive, we can also conclude that using the two-year training period to obtain GTSPs for making trading plans is appropriate when the market trend is a downtrend.

From Table 12, comparing the derived GTSPs with the BHS in terms of average returns, which are -0.014 and -0.244, it shows that the proposed approach has better returns than the BHS and also a higher ability to avoid losses. According to the results shown in Tables 11 and 12, we can say that the proposed approach is effective on the downtrend dataset, which means that the derived GTSPs have the ability to get positive returns with higher probability, or at least lose less money in bear markets.

APPENDIX I

TS#	Trading Strategy	TS#	Trading Strategy
TS ₁	Buy: MA5 ↗ MA20, Sell: MA5 ↘ MA20	TS ₃₁	Buy: PSY ↗ 25%, Sell: CCI ↘ 100
TS ₂	Buy: RSI ↗ 30, Sell: RSI ↘ 70	TS ₃₂	Buy: PSY ↗ 25%, Sell: D > 80, K ↘ D
TS ₃	Buy: WMS%R ↘ 80, Sell: WMS%R ↗ 20	TS ₃₃	Buy: PSY ↗ 25%, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0
TS ₄	Buy: MOM ↗ 0, Sell: MOM ↘ 0	TS ₃₄	Buy: PSY ↗ 25%, Sell: BIAS ↘ 5%
TS ₅	Buy: PSY ↗ 25%, Sell: PSY ↘ 75%	TS ₃₅	Buy: PSY ↗ 25%, Sell: +DI ↘ -DI
TS ₆	Buy: CCI ↗ -100, Sell: CCI ↘ 100	TS ₃₆	Buy: CCI ↗ -100, Sell: MA5 ↘ MA20
TS ₇	Buy: D < 20, K ↗ D, Sell: D > 80, K ↘ D	TS ₃₇	Buy: CCI ↗ -100, Sell: RSI ↘ 70
TS ₈	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0	TS ₃₈	Buy: CCI ↗ -100, Sell: WMS%R ↗ 20
TS ₉	Buy: BIAS ↗ -4.5%, Sell: BIAS ↘ 5%	TS ₃₉	Buy: CCI ↗ -100, Sell: MOM ↘ 0
TS ₁₀	Buy: +DI ↗ -DI, Sell: +DI ↘ -DI	TS ₄₀	Buy: CCI ↗ -100, Sell: PSY ↘ 75%
TS ₁₁	Buy: MA5 ↗ MA20(20), Sell: RSI ↘ 70	TS ₄₁	Buy: CCI ↗ -100, Sell: D > 80, K ↘ D
TS ₁₂	Buy: MA5 ↗ MA20(20), Sell: WMS%R ↗ 20	TS ₄₂	Buy: CCI ↗ -100, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0
TS ₁₃	Buy: MA5 ↗ MA20, Sell: MOM ↘ 0	TS ₄₃	Buy: CCI ↗ -100, Sell: BIAS ↘ 5%
TS ₁₄	Buy: MA5 ↗ MA20, Sell: PSY ↘ 75%	TS ₄₄	Buy: CCI ↗ -100, Sell: +DI ↘ -DI
TS ₁₅	Buy: MA5 ↗ MA20, Sell: CCI ↘ 100	TS ₄₅	Buy: D < 20, K ↗ D, Sell: MA5 ↘ MA20
TS ₁₆	Buy: MA5 ↗ MA20, Sell: D > 80, K ↘ D	TS ₄₆	Buy: D < 20, K ↗ D, Sell: RSI ↘ 70
TS ₁₇	Buy: MA5 ↗ MA20, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0	TS ₄₇	Buy: D < 20, K ↗ D, Sell: WMS%R ↗ 20
TS ₁₈	Buy: MA5 ↗ MA20, Sell: BIAS ↘ 5%	TS ₄₈	Buy: D < 20, K ↗ D, Sell: MOM ↘ 0
TS ₁₉	Buy: MA5 ↗ MA20, Sell: +DI ↘ -DI	TS ₄₉	Buy: D < 20, K ↗ D, Sell: PSY ↘ 75%
TS ₂₀	Buy: RSI ↗ 30, Sell: MA5 ↘ MA20	TS ₅₀	Buy: D < 20, K ↗ D, Sell: CCI ↘ 100
TS ₂₁	Buy: RSI ↗ 30, Sell: WMS%R ↗ 20	TS ₅₁	Buy: D < 20, K ↗ D, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0
TS ₂₂	Buy: RSI ↗ 30, Sell: MOM ↘ 0	TS ₅₂	Buy: D < 20, K ↗ D, Sell: BIAS ↘ 5%
TS ₂₃	Buy: RSI ↗ 30, Sell: PSY ↘ 75%	TS ₅₃	Buy: D < 20, K ↗ D, Sell: +DI ↘ -DI
TS ₂₄	Buy: RSI ↗ 30, Sell: CCI ↘ 100	TS ₅₄	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: MA5 ↘ MA20
TS ₂₅	Buy: RSI ↗ 30, Sell: D > 80, K ↘ D	TS ₅₅	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: RSI ↘ 70
TS ₂₆	Buy: RSI ↗ 30, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0	TS ₅₆	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: WMS%R ↗ 20
TS ₂₇	Buy: RSI ↗ 30, Sell: BIAS ↘ 5%	TS ₅₇	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: MOM ↘ 0
TS ₂₈	Buy: RSI ↗ 30, Sell: +DI ↘ -DI	TS ₅₈	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: PSY ↘ 75%
TS ₂₉	Buy: WMS%R ↘ 80, Sell: MA5 ↘ MA20	TS ₅₉	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: CCI ↘ 100
TS ₃₀	Buy: WMS%R ↘ 80, Sell: RSI ↘ 70	TS ₆₀	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: D > 80, K ↘ D
TS ₃₁	Buy: WMS%R ↘ 80, Sell: MOM ↘ 0	TS ₆₁	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: BIAS ↘ 5%
TS ₃₂	Buy: WMS%R ↘ 80, Sell: PSY ↘ 75%	TS ₆₂	Buy: DIF ↗ MACD(DEM), or DIF ↗ 0, Sell: +DI ↘ -DI
TS ₃₃	Buy: WMS%R ↘ 80, Sell: CCI ↘ 100	TS ₆₃	Buy: BIAS ↗ -4.5%, Sell: MA5 ↘ MA20
TS ₃₄	Buy: WMS%R ↘ 80, Sell: D > 80, K ↘ D	TS ₆₄	Buy: BIAS ↗ -4.5%, Sell: RSI ↘ 70
TS ₃₅	Buy: WMS%R ↘ 80, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0	TS ₆₅	Buy: BIAS ↗ -4.5%, Sell: WMS%R ↗ 20
TS ₃₆	Buy: WMS%R ↘ 80, Sell: BIAS ↘ 5%	TS ₆₆	Buy: BIAS ↗ -4.5%, Sell: MOM ↘ 0
TS ₃₇	Buy: WMS%R ↘ 80, Sell: +DI ↘ -DI	TS ₆₇	Buy: BIAS ↗ -4.5%, Sell: PSY ↘ 75%
TS ₃₈	Buy: MOM ↗ 0, Sell: MA5 ↘ MA20	TS ₆₈	Buy: BIAS ↗ -4.5%, Sell: CCI ↘ 100
TS ₃₉	Buy: MOM ↗ 0, Sell: RSI ↘ 70	TS ₆₉	Buy: BIAS ↗ -4.5%, Sell: D > 80, K ↘ D
TS ₄₀	Buy: MOM ↗ 0, Sell: WMS%R ↗ 20	TS ₇₀	Buy: BIAS ↗ -4.5%, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0
TS ₄₁	Buy: MOM ↗ 0, Sell: PSY ↘ 75%	TS ₇₁	Buy: +DI ↗ -DI, Sell: MA5 ↘ MA20
TS ₄₂	Buy: MOM ↗ 0, Sell: CCI ↘ 100	TS ₇₂	Buy: +DI ↗ -DI, Sell: RSI ↘ 70
TS ₄₃	Buy: MOM ↗ 0, Sell: D > 80, K ↘ D	TS ₇₃	Buy: +DI ↗ -DI, Sell: WMS%R ↗ 20
TS ₄₄	Buy: MOM ↗ 0, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0	TS ₇₄	Buy: +DI ↗ -DI, Sell: MOM ↘ 0
TS ₄₅	Buy: MOM ↗ 0, Sell: BIAS ↘ 5%	TS ₇₅	Buy: +DI ↗ -DI, Sell: PSY ↘ 75%
TS ₄₆	Buy: MOM ↗ 0, Sell: +DI ↘ -DI	TS ₇₆	Buy: +DI ↗ -DI, Sell: CCI ↘ 100
TS ₄₇	Buy: PSY ↗ 25%, Sell: MA5 ↘ MA20	TS ₇₇	Buy: +DI ↗ -DI, Sell: D > 80, K ↘ D
TS ₄₈	Buy: PSY ↗ 25%, Sell: RSI ↘ 70	TS ₇₈	Buy: +DI ↗ -DI, Sell: DIF ↘ MACD(DEM), or DIF ↘ 0
TS ₄₉	Buy: PSY ↗ 25%, Sell: WMS%R ↗ 20	TS ₇₉	Buy: +DI ↗ -DI, Sell: BIAS ↘ 5%
TS ₅₀	Buy: PSY ↗ 25%, Sell: MOM ↘ 0	TS ₈₀	

D. EVALUATION ON THE DERIVED GTSPs

At last, the optimized GTSPs by the proposed approach using 2014 to 2015 for training and 2016 as the testing period on the two datasets are illustrated in Tables 13 and 14. Since a GTSP can generate many TSPs, we use three measurements, namely, the MAXROI, AVGROI and MINROI, to indicate the quality of the optimized GTSPs. The MAXROI represents that the TSP generated from a GTSP can have the largest return. The MINROI, on the contrary, represents that the

return of the TSP generated from a GTSP is the smallest. The AVGROI is the average return of all TSPs that can be generated from a GTSP.

Three observations can be found from Tables 13 and 14. First, the returns of the optimized GTSPs are positive and the proposed approach with the SLTP is again better than that without SLTP in terms of returns. Second, comparing the returns of the optimized GTSPs with the BHS on the two datasets, we know that regardless of which TSP is utilized

for trading, their returns are better than the BHS. Third, comparing the proposed approach with TOP15 to that with TOP555, we also observe that using TOP555 as an input trading strategy set could get more stable returns than TOP15. For instance, the MAXROI, AVGROI and MINROI on the sideways trend dataset for TOP15 are 0.22, 0.15 and 0.00, respectively, and those for TOP555 are 0.24, 0.19 and 0.15, respectively. Note that the descriptions of the trading strategies are shown in Appendix I.

VII. CONCLUSIONS AND FUTURE WORK

In stock markets, a common way to find trading signals is using TSSs, which could be generated by technical or fundamental indicators. To provide a more useful TSP, we first defined the GTSP problem and then proposed an algorithm to obtain a GTSP using the GGA. To encode a GTSP into a chromosome, three parts, named the grouping, the trading strategy, and the weight parts, are used to represent a GTSP. Then, the fitness value of every chromosome is calculated according to four factors, which are the portfolio return, risk, weight balance, and group balance. The genetic operators are executed to generate various offspring. After evolution, by using the optimized GTSP, various TSPs can be provided to traders for making trading plans. Experiments on two real datasets, the sideways trend and the downtrend, were made to show the effectiveness of the proposed algorithm. As a result, three merits are described as follows: (1) The proposed approach with SLTP is better than that without SLTP in terms of returns; (2) on both datasets, the obtained GTSPs can have a better ability to reduce losses and increase returns than using the BHS; and (3) using TOP555 as the inputted trading strategy set, the proposed approach can derive a better GTSP than that using TOP15, especially in a bear market. In the future, several ways could be used to enhance the proposed algorithm, such as using the fuzzy grouping genetic algorithm to find a GTSP effectively and employing island-based algorithms to efficiently optimize a GTSP. Besides, based on the proposed approach, enhanced optimization approaches could also be designed to select appropriate set of trading strategies, e.g., TOP5, TOP10, or TOP20, and/or to decide a suitable number of trading strategies for deriving a GTSP.

APPENDIX

See Appendix I.

ACKNOWLEDGMENT

This paper was presented at The Multidisciplinary International Social Networks Conference, in 2017.

REFERENCES

- [1] J. M. Berutich, F. López, F. Luna, and D. Quintana, "Robust technical trading strategies using GP for algorithmic portfolio selection," *Expert Syst. Appl.*, vol. 46, pp. 307–315, Mar. 2016.
- [2] E. C. Brown and R. T. Sumichrast, "Evaluating performance advantages of grouping genetic algorithms," *Eng. Appl. Artif. Intell.*, vol. 18, no. 1, pp. 1–12, 2005.
- [3] Y.-H. Chang and M.-S. Lee, "Incorporating Markov decision process on genetic algorithms to formulate trading strategies for stock markets," *Appl. Soft Comput.*, vol. 52, no. 10, pp. 1143–1153, 2016.
- [4] S. K. Chandrinou and N. D. Lagaros, "Construction of currency portfolios by means of an optimized investment strategy," *Oper. Res. Perspect.*, vol. 5, pp. 32–44, 2018.
- [5] T.-J. Chang, S.-C. Yang, and K.-J. Chang, "Portfolio optimization problems in different risk measures using genetic algorithm," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10529–10537, Sep. 2009.
- [6] Y.-W. C. Chien and Y. L. Chen, "Mining associative classification rules with stock trading data—A GA-based method," *Knowl.-Based Syst.*, vol. 23, no. 6, pp. 605–614, 2010.
- [7] J.-S. Chen, J.-L. Hou, S.-M. Wu, and Y.-W. Chang-Chien, "Constructing investment strategy portfolios by combination genetic algorithms," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 3824–3828, 2009.
- [8] C.-H. Chen, C.-B. Lin, and C.-C. Chen, "Mining group stock portfolio by using grouping genetic algorithms," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, May 2015, pp. 738–743.
- [9] C.-H. Chen, C.-Y. Lu, T.-P. Hong, and J.-H. Su, "Using grouping genetic algorithm to mine diverse group stock portfolio," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 2016, pp. 4734–4738.
- [10] Y.-H. Chou, S.-Y. Kuo, C.-Y. Chen, and H.-C. Chao, "A rule-based dynamic decision-making stock trading system based on quantum-inspired tabu search algorithm," *IEEE Access*, vol. 2, pp. 883–896, 2014.
- [11] Y.-H. Chou, S.-Y. Kuo, and C. Kuo, "A dynamic stock trading system based on a multi-objective quantum-inspired tabu search algorithm," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2014, pp. 112–119.
- [12] Y.-H. Chou, S.-Y. Kuo, and Y.-T. Lo, "Portfolio optimization based on funds standardization and genetic algorithm," *IEEE Access*, vol. 5, pp. 21885–21900, 2017.
- [13] E. Falkenauer, "A new representation and operators for genetic algorithms applied to grouping problems," *Evol. Comput.*, vol. 2, pp. 123–144, Jun. 1994.
- [14] E. Falkenauer, "A hybrid grouping genetic algorithm for bin packing," *J. Heuristics*, vol. 2, no. 1, pp. 5–30, 1996.
- [15] T.-C. Fu, C.-P. Chung, and F.-L. Chung, "Adopting genetic algorithms for technical analysis and portfolio management," *Comput. Math. Appl.*, vol. 66, no. 10, pp. 1743–1757, 2013.
- [16] S. Feuerriegel and H. Prendinger, "News-based trading strategies," *Decis. Support Syst.*, vol. 90, pp. 65–74, Oct. 2016.
- [17] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA, USA: Addison-Wesley, 1989.
- [18] J. J. Grefenstette, "Optimization of control parameters for genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, no. 1, pp. 122–128, Jan. 1986.
- [19] F. Hassanzadeh, M. Collan, and M. Modarres, "A practical approach to R&D portfolio selection using the fuzzy pay-off method," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 4, pp. 615–622, Aug. 2012.
- [20] Hoklie and L. R. Zuhail, "Resolving multi objective stock portfolio optimization problem using genetic algorithm," in *Proc. Int. Conf. Comput. Autom. Eng.*, 2010, pp. 40–44.
- [21] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI, USA: Univ. of Michigan Press, 1975.
- [22] Y. Kim and D. Enke, "Developing a rule change trading system for the futures market using rough set analysis," *Expert Syst. Appl.*, vol. 59, pp. 165–173, Oct. 2016.
- [23] R. Kumar and S. Bhattacharya, "Cooperative search using agents for cardinality constrained portfolio selection problem," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1510–1518, Nov. 2012.
- [24] S.-Y. Kuo, C. Kuo, and Y.-H. Chou, "Dynamic stock trading system based on quantum-inspired tabu search algorithm," in *Proc. IEEE Congr. Evol. Comput.*, Jun. 2013, pp. 1029–1036.
- [25] M. Kampouridis and F. E. B. Otero, "Evolving trading strategies using directional changes," *Expert Syst. Appl.*, vol. 73, pp. 145–160, May 2017.
- [26] Y. Leu and T.-I. Chiu, "An effective stock portfolio trading strategy using genetic algorithms and weighted fuzzy time series," in *Proc. North-East Asia Symp. Nano, Inf. Technol. Rel.*, 2011, pp. 70–75.
- [27] Y.-J. Liu and W.-G. Zhang, "Fuzzy portfolio optimization model under real constraints," *Insurance, Math. Econ.*, vol. 53, no. 3, pp. 704–711, 2013.
- [28] K. Lwin, R. Qu, and G. Kendall, "A learning-guided multi-objective evolutionary algorithm for constrained portfolio optimization," *Appl. Soft Comput.*, vol. 24, pp. 757–772, Nov. 2014.
- [29] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [30] H. M. Markowitz, Ed. *Harry Markowitz: Selected Works*. Singapore: World Scientific, 2009.

- [31] M.-E. Wu, C.-H. Wang, and W.-H. Chung, "Using trading mechanisms to investigate large futures data and their implications to market trends," *Soft Comput.*, vol. 21, no. 11, pp. 2821–2834, Jun. 2017.
- [32] D. Wen, C. Ma, G.-J. Wang, and S. Wang, "Investigating the features of pairs trading strategy: A network perspective on the Chinese stock market," *Phys. A, Stat. Mech. Appl.*, vol. 505, pp. 903–918, Sep. 2018.
- [33] H. Yao, Z. Li, and D. Li, "Multi-period mean-variance portfolio selection with stochastic interest rate and uncontrollable liability," *Eur. J. Oper. Res.*, vol. 252, no. 3, pp. 837–851, 2016.



CHUN-HAO CHEN received the Ph.D. degree in computer science and information engineering from National Cheng Kung University, Taiwan, in 2008. He is currently a Professor with the Department of Computer Science and Information Engineering, Tamkang University, Taiwan. He joined the Department of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung, Taiwan, as a Postdoctoral Fellow, during 2009. From 2010 to 2013, he has served as an Assistant Professor with the Department of Computer Science and Information Engineering, Tamkang University, where he was an Associate Professor, from 2013 to 2017. He has published over 100 research papers in referred journals and in international conferences. His research interests include data mining, time series, machine learning, evolutionary algorithms, fuzzy theory, portfolio selection, trading strategy, business data analysis, and time series pattern discovery.



YU-HSUAN CHEN received the B.S. and M.S. degrees from the Department of Computer Science and Information Engineering, Tamkang University, Taiwan, in 2014 and 2018, respectively. She currently serves as an Engineer with the Department of Information Technology, EVA Air, Taiwan. Her research interests include machine learning, financial data analysis, and genetic algorithms.



JERRY CHUN-WEI LIN received the Ph.D. degree from the Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan. He is currently an Associate Professor with the Department of Computing, Mathematics, and Physics, Western Norway University of Computer Sciences, Bergen, Norway. He has published over 200 research papers in referred journals and in international conferences. His research interests include data mining, soft computing, artificial intelligence, social computing, multimedia and image processing, and privacy-preserving and security technologies. He is also the Project Leader of SPMF: An Open-Source Data Mining Library, which is a toolkit offering multiple types of data mining algorithms. He also serves as the Editor-in-Chief for the international journal of *Data Science and Pattern Recognition*.



MU-EN WU received the Ph.D. degree in computer science from National Tsing Hua University, Taiwan, in 2009. From 2014 to 2017, he served as an Assistant Professor with the Department of Mathematics, Soochow University. He was with the Institute of Information Science, Academia Sinica, Taipei, Taiwan, as a Postdoctoral Fellow, from 2009 to 2014. He is currently an Assistant Professor with the Department of Information and Finance Management, National Taipei University of Technology, Taiwan. His research interests include cryptography, information theory, prediction market, money management, and financial data analysis.

...