Applied Soft Computing 36 (2015) 534–551

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Review article

Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review

Yong Hu[a,*], Kang Liu[b], Xiangzhou Zhang[a,e], Lijun Su[b], E.W.T. Ngai[c], Mei Liu[d]

[a] Big Data Decision Institute, Jinan University, Guangzhou, PR China
[b] School of Business, Guangdong University of Foreign Studies, Higher Education Mega Center, Guangzhou 510006, PR China
[c] Department of Management and Marketing, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China
[d] Department of Internal Medicine, Division of Medical Informatics, University of Kansas Medical Center, Kansas City, KS 66160, USA
[e] School of Business, Sun Yat-sen University, China

ARTICLE INFO

Article history:
Received 19 August 2014
Received in revised form 7 June 2015
Accepted 12 July 2015
Available online 21 July 2015

Keywords:
Literature review
Evolutionary computation
Algorithmic trading
Stock trading rule
Rule discovery
Classification framework

ABSTRACT

Despite the wide application of evolutionary computation (EC) techniques to rule discovery in stock algorithmic trading (AT), a comprehensive literature review on this topic is unavailable. Therefore, this paper aims to provide the first systematic literature review on the state-of-the-art application of EC techniques for rule discovery in stock AT. Out of 650 articles published before 2013 (inclusive), 51 relevant articles from 24 journals were confirmed. These papers were reviewed and grouped into three analytical method categories (fundamental analysis, technical analysis, and blending analysis) and three EC technique categories (evolutionary algorithm, swarm intelligence, and hybrid EC techniques). A significant bias toward the applications of genetic algorithm-based (GA) and genetic programming-based (GP) techniques in technical trading rule discovery is observed. Other EC techniques and fundamental analysis lack sufficient study. Furthermore, we summarize the information on the evaluation scheme of selected papers and particularly analyze the researches which compare their models with buy and hold strategy (B&H). We observe an interesting phenomenon where most of the existing techniques perform effectively in the downtrend and poorly in the uptrend, and considering the distribution of research in the classification framework, we suggest that this phenomenon can be attributed to the inclination of factor selections and problem in transaction cost selections. We also observe the significant influence of the transaction cost change on the margins of excess return. Other influenced factors are also presented in detail. The absence of ways for market trend prediction and the selection of transaction cost are two major limitations of the studies reviewed. In addition, the combination of trading rule discovery techniques and portfolio selection is a major research gap. Our review reveals the research focus and gaps in applying EC techniques for rule discovery in stock AT and suggests a roadmap for future research.

Contents

* Corresponding author. Tel.: +1 973 596 6597; fax: +1 973 596 5777.
E-mail addresses: henryhu200211@163.com (Y. Hu), researcher_lk@foxmail.com (K. Liu), zhxzhou86@foxmail.com (X. Zhang), onnowa@163.com (L. Su), eric.ngai@polyu.edu.hk (E.W.T. Ngai), meiliu@kumc.edu (M. Liu).

http://dx.doi.org/10.1016/j.asoc.2015.07.008
1568-4946/

1. Introduction

Stock investment has attracted a lot of attention. In 2012, the value of share trading-electronic order book trade in the United States stock markets reached USD 23,226,924 million [1]. However, the computerization of stock trading from order book to exchange has been generating large amounts of real-time data [2,3]. At the same time, government, institutions, social media, and listed companies have been releasing an ocean of data on the operating performance of listed companies, such as news, financial statements, and macroeconomic information [2–4]. Discovering useful knowledge from these substantial high-dimensional financial data [5–12] and catching investment opportunities faster than other investors in such a noisy and dynamic market environment [2,10,11,13–15] are significant challenges faced by investors constantly.

Algorithmic trading (AT), an important automatic analysis and trading decision approach for equity investment, gained prominence in the early 1990s and accounted for at least 50% of the total US equity trading volume in 2012 [16]. Generally, AT refers to the use of sophisticated computer algorithms to automatically make certain trading decisions in the trading cycle, including pre-trade analysis (data analysis), trading signal generation (buying and selling recommendations), and trade execution (order management) [17–20]. One of the advantages of AT is the effectiveness and efficiency of machine learning techniques in financial big data analysis [2,21,22]. However, some AT learning models are considered as “black boxes” [23–28] because they involve difficulty in providing easy-to-understand explanations on the interactions between the model inputs and the outputs. Trading with black boxes makes investors uncomfortable and elicits mistrust in the model [2,29,150]. To address this issue, an increasing number of researchers have investigated rule discovery techniques for finding explicit trading rules that can provide explicit knowledge to guide trading.

Rule discovery is an important aspect of data mining because it can generate a set of symbolic rules that describe the relationship among variables in a natural way, and rules can be better understood by the human mind than any other data mining model [30,31]. Numerous studies have demonstrated the necessity of rule discovery in AT, such as the following: (1) increasing investors' approval of the system by improving the comprehensibility of the system decision logic. Thus, investors can justify system decisions using their domain knowledge, and potential investment risks can be reduced [32,33,151]. (2) facilitating the discovery of new knowledge and integration of new and old knowledge [33,34]. (3) reducing errors derived from noise, feature subset selection, or inaccurate parameter settings [35].

Evolutionary computation (EC) has been widely employed in rule discovery. EC is generally defined as a computing tool to solve realistic problems by simulating the evolutionary mechanisms of nature [36–41]. It is mainly based on a population, uses probabilistic transition rules, and directly applies the objectives from the user as “fitness” [36,37,40–42]. Applying EC techniques for rule discovery in stock AT is becoming popular because EC is able to find a sufficiently good solution for a wide range of problems within a relatively short time [14,42]. Previous research suggested that applying EC-based models to trading rule discovery could yield promising results [32,43–45]. Moreover, EC can be employed to optimize the individual trading rule [14,26,44,46,47] and the parameters of the underlying rule discovery algorithm [44,48,49]. This study focuses on the distinctive competence of EC in rule discovery. Over the past few years, several literature reviews have been conducted on stock prediction models. Atsalakis and Valavanis [50] reviewed the application of neural and neural-fuzzy techniques in stock prediction. Guresen et al. [51] presented a comparative survey of different neural network models in NASDAQ Stock exchange index prediction. Bahrammirzaee [52] provided a comparative analysis among artificial neural networks, expert system and hybrid intelligent systems in credit evaluation, portfolio management and financial prediction and planning. However, at present, a systematic and comprehensive review of the EC techniques applied to explicit trading rule discovery remains lacking. Therefore, the objectives of this paper are as follows: (1) develop a classification framework for the application of EC techniques for rule discovery in stock AT; (2) provide a systematic and comprehensive review; and (3) determine research gaps and propose suggestions for directions of future research.

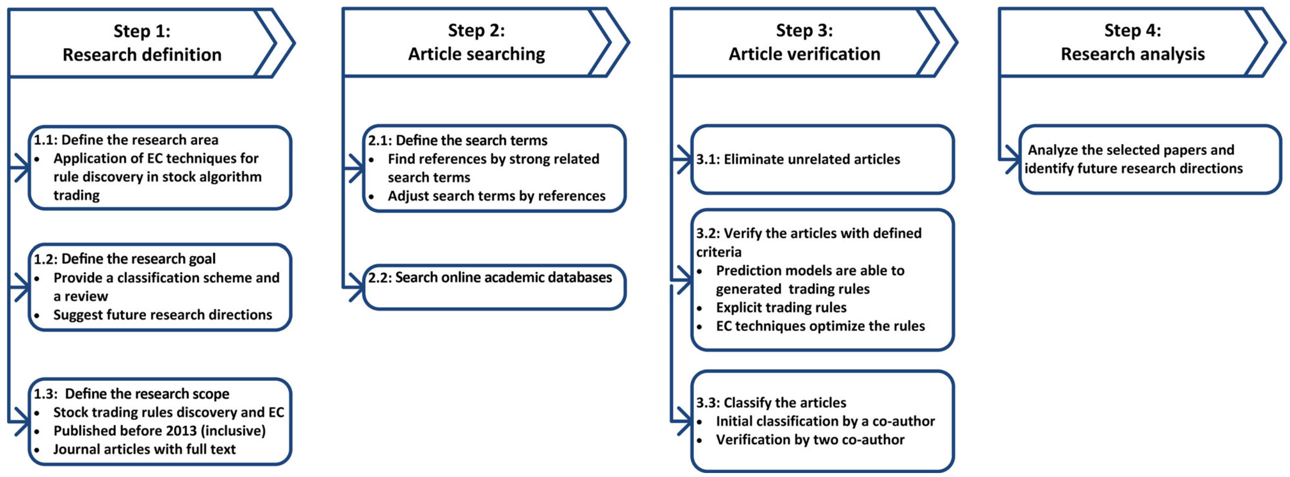

Fig. 1. Research framework.

The contribution of our paper is that it provides readers a clear picture of the application of EC techniques to rule discovery in stock AT by developing a classification framework for it. Furthermore, by summarizing the details of the comparative test in these studies, we have explicitly presented the most important factors, which are influencing the performance of different proposed models, precisely analyzed the major limitations in existing research, and provided a specific roadmap for future research. Thus, this paper assists readers in conducting further studies and in evaluating their models more adequately.

The remaining sections of this paper are organized as follows: Section 2 presents the research methodology of our paper. Section 3 introduces our proposed classification framework for the application of EC techniques for rule discovery in stock algorithmic trading. Section 4 analyzes the reviewed articles from different perspectives. Section 5 discusses the result and presents research gaps. Finally, Section 6 concludes and provides suggestions for future research.

2. Research methodology

Our research can be divided into four main phases: research definition, article searching, article verification, and research analysis, as shown in Fig. 1.

Step 1: Research definition

Step 1.1: Define the research area. The research area is the application of EC techniques for rule discovery in stock algorithmic trading, as shown in Fig. 2.
Step 1.2: Define the research goal. The research goal is to provide a classification framework and a systematic and comprehensive review of EC techniques for rule discovery applied to stock algorithmic trading, as well as to determine research gaps and suggest directions for future research.
Step 1.3: Define the research scope. The research scope is the articles on the application of EC techniques for rule discovery in stock algorithmic trading published in academic journals with their full text available before 2013 (inclusive).

Step 2: Article searching

Step 2.1: Define the search terms. First, representative articles were found as references by using search terms that were strongly related to the research area (see Fig. 2). Then, the search terms were adjusted based on keywords and information found in

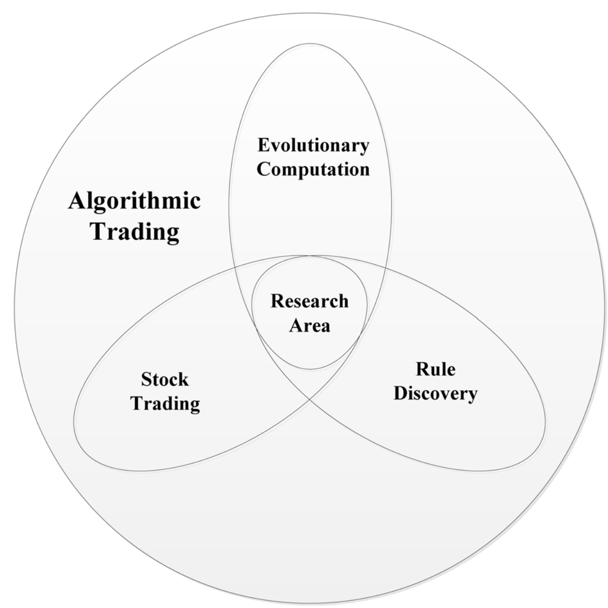

Fig. 2. Research area.

articles and authoritative site. Finally, we determined our search terms as "evolutionary computation," "stock," and "trading rule".
Step 2.2: Search the articles in online databases. More than 550 articles were found in the following online journal databases:
- ACM
- IEEE Transactions
- MIT Press journals
- Science Direct
- Springer-Link Journals
- Taylor & Francis Online

Step 2.3: Search the articles via other ways. To collect as much related studies as possible, we utilized the reference information of the papers under our research scope (verified in Step 3). By adding the papers we collected in our previous searches, a total of 650 papers were found in Step 2.

Step 3: Article verification

The classification of the articles found in Step 2 was carefully and repeatedly verified by the co-authors. Only articles related to

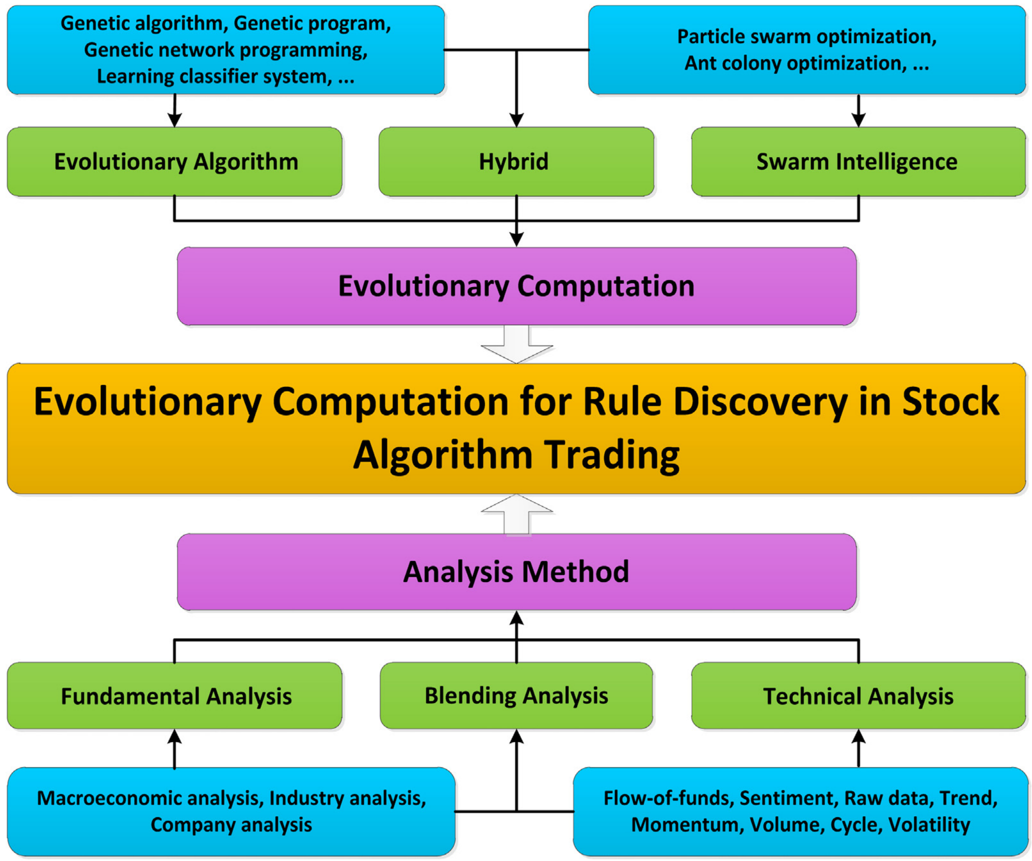

Fig. 3. Classification framework for EC techniques for rule discovery in stock algorithmic trading.

stock algorithmic trading and EC techniques for rule discovery were selected. The verification process was as follows:

Step 3.1: Eliminate unrelated articles. Papers that were obviously unrelated to our research area were eliminated by co-author respectively. Articles that did not refer to EC or stock investment were removed.
Step 3.2: Verify the articles with defined criteria. The remaining articles were verified independently. The articles with consistent verification results were selected, and the others were discussed by the co-authors to draw a final conclusion. The criteria for selecting articles were as follow: (1) the prediction models must be able to generate trading rules; (2) the trading rule discovered must be explicit knowledge which can directly reveal the relations between the market information and the prediction results; and (3) EC techniques must be applied to optimize the trading rules.
Step 3.3: Classify the articles. A co-author classified the selected articles. Then, the two other co-authors verified the results. The classifications assigned to the articles were approved by the co-authors or a discussion was held if any disagreements arose.

Through Steps 3.1–3.3, 51 papers were consequently selected from the 650 papers searched in Step 2.

Step 4: Research analysis

We analyzed the selected papers from different perspectives and identified some directions for future research. The details of the analysis and conclusion are presented in Sections 4 and 5.

3. Classification framework

To provide a systematic review of the available literatures on the application of EC techniques for rule discovery in stock AT, we propose a classification framework, which is shown in Fig. 3. The lists of abbreviations used in this paper are summarized in the Appendix.

3.1. Classification of analysis method

As Fig. 3 shows, all of the stock analysis approaches can be classified into fundamental, technical, or blending analysis. With the underlying analytical principle, completely different factors are utilized and different rules are obtained as a result. The category to which an analysis method belongs can be determined according to the input factors used in the rule discovery process. An introduction about the three aforementioned categories of analysis method is the following.

Fundamental analysis. This type of analysis method is based on the assumption that the internal value of each stock is determined by its potential profitability. Thus, based on the evaluation of the fundamental value of a stock, that is analyzing the factors that reflect and influence the profitability of a company to identify whether a stock is undervalued or overvalued, fundamental analysis could generate trading rules to determine which stock to buy/sell [7,35,53–56]. In practice, a number of well-known investors have been using fundamental analysis, the most influential of which is Benjamin Graham, hailed as the father of value investing. His investment philosophy includes the concept of investment diversification and buying obviously undervalued stocks with a high safety margin [57,58]. Another example is Warren Buffett, an adherent of Graham; some of his accomplishments can be attributed to fundamental analysis. Fundamental analysis is mainly based on three essential aspects, which are described as follows [7,56,59,60]:

i. Macroeconomic analysis, which analyzes the effect of the macroeconomic environment on the future profit of a company. The popular indicators include GDP, CPI, M1B, etc.
ii. Industry analysis, which estimates the value of the company based on industry status and prospect, such as by analyzing the billings (or revenues) of upper stream entities in an industry
iii. Company analysis, which analyzes the current operation status of a company to evaluate its internal value, mostly by examining the company financial reports.

Technical analysis. This analysis method considers the movement of stock price and volume as reflections of all the related information about the stock markets. In addition, previous market behavior patterns repeat in the future. Thus by analyzing the previous behavior patterns of price and volume, trading rules can be generated [35,53,54,59,61]. A common investment strategy based on technical analysis is trend following. Covel [29] described it in the following manner: "Instead of trying to predict a market direction, their strategy is to react to the market's movements whenever they occur". Thus, trend following judges the status of the current market promptly so that investment decisions can be made [6]. Unlike fundamental analysis, technical analysis lacks a consistent taxonomy. Therefore, by referencing Bodie et al. [7], CFA Institute [59], Colby [62], Fidelity Mutual Fund [63], Goldman Sachs [64], NASDAQ [65], Pring [6], Wikipedia [66], Yahoo Finance [67], and Market Technicians Association [68], we found that domains of technical analysis can be grouped into rational classification scheme as follows:

i. Sentiment, which mainly represents the behaviors of various market participants. The analysis of these indicators is grounded on hypothesis that different types of investors show different behaviors at the main market turning points. Indicators such as expert-public ratio, consulting services sentiment indicator, short interest ratio and put options to call options [6,59,62,68].
ii. Flow-of-funds, which is a type of indicator used to investigate the financial status of various investors to pre-evaluate their strength in terms of buying and selling stocks, then, corresponding strategies, such as short squeeze, can be adopted. The analyzable data for this strategy contains the capital of mutual funds or large investors, and the events like new and additional issues [6,59,68].
iii. Raw data, which include stock price series as well as price patterns such as K-line diagrams and bar charts. The former is commonly used for time series analysis or trend judgment combined with other indicators; the latter generally suggests price patterns can reflect the changes of markets sentiment which affect short movements of stocks [6,63–65,67,68].
iv. Trend, which is a type of price-based indicator for tracing the stock price trends. The corresponding strategy, called trend following strategy, proposes that economic and political events usually change market prices through changing market trends rather than by instantly returning to the most rational point. Thus, investors can gain profit by tracing the occurrent price trends. Common trend indicators include SMA and EMA [29,59,62,64,67].
v. Momentum, which is also a kind of price-based indicator but is used to evaluate the velocity of price change and judge whether a trend reversal in stock price is about to occur. The momentum indicators are analyzed based on the hypothesis that stock prices undergo a nominal cycle, and the cycle will manifest as price rebound and callback trend. Such indicators include RSI, MACD, and ROC, etc. [6,59,62,64,67].
vi. Volume. Volume-based indicators reflect the enthusiasm for investing of both buyers and sellers, which is also a basis for predicting stock price movements. Strategy applying volume indicators is grounded on the hypothesis that price movement is determined by the enthusiasm of buyers and sellers. Transactions are commonly suggested to expand during rising tendencies and shrink during downward tendencies. Popular indicators include volume, volume ratio, and OBV [6,64,66,67].
vii. Cycle. According to cycle theories, stock prices vary periodically; a long cycle may last more than 10 years and contains various short cycles that can be as short as a few days or weeks. Strategies that use cycle indicators aim to analyze the position of the current stock price in the cycle. For example, Elliott Wave and seasonal patterns are often used to examine the cyclical variations of stock prices [6,59,68].
viii. Volatility, which is commonly used to investigate the fluctuation range of stock prices. It can be used to evaluate the risk and identify the level of the support and resistance. Stock prices are generally recognized to fluctuate between the level of support and resistance, but continue to rise (fall) once they break through the level of resistance (support). Common volatility indicators include average true range and Bollinger band, etc. [6,62,66,152].

Blending analysis (fundamental and technical analysis). The analysis combining fundamental analysis and technical analysis will be regarded as the blending method.

3.2. Classification of EC techniques

Given the large search space and nonlinear nature of stock trading rule discovery, EC has been one of the most popular techniques in financial analysis. EC is based on population, utilizes probabilistic transition rules, and directly uses objectives defined from the user as "fitness" [36,37,40–42]. EC techniques possess the following advantages: (1) compared to the classical local optimization methods such as gradient-based and hill-climbing methods which are always based on deterministic rules, the population-based and stochastic characteristics of EC can greatly expand its search space and thus reduce the probability of trapping into local optima. (2) Compare to the inefficient random search, EC is directed toward better regions of the search space because it precisely uses fitness functions rather than function derivatives, which may be unavailable in some cases. Although the stochastic characteristic of EC cannot guarantee the best results, EC techniques are effective in generating nearly optimal results within a relatively short time. These advantages make EC more effective and robust in dealing with problems of discontinuity, non-differentiability, multiple modes, and noise [39–42,69,70]. Furthermore, its convenience in addressing a wide range of problems with different search space structures because of robustness of EC and the convenience of setting encoding of representation of solutions in EC techniques, EC suitable for many analytical tasks in investment.

The term "evolutionary computation" emerged in the early 1990s, which from the very beginning refers to the research areas of ES, EP and GA [37] (i.e., evolutionary algorithm, EA). Later, more algorithms such as GP and LCS have been included. Some researchers insist that EC only contains evolutionary algorithms which are based on the simulation of Darwinian evolution [39,70–75]. However, some suggest that EC should also contain swarm intelligence (SI), which is based on simulation of social behavior [36,38,41,42,76–79], because social behavior can be regarded as an evolutionary process. Thus, our review includes both EA and SI.

Table 1
Research on EC techniques for rule discovery in stock algorithmic trading.

Analysis method classification	EC classification	EC technique	Proposed model	Optimization object	Input variables	Outperform B&H	Ref.
Fundamental analysis	Evolutionary algorithm	GA	GA	RPara.	PCF, D/BE, SG, TG, ROA, P/B		[85]
		GP	GP, Fuzzy system	RCons.	39 financial variables	Yes	[32]
Technical analysis	Evolutionary algorithm	GA	GA	RComb.	MA, SLMA, MAD, SLEMA, MTM, RSI, MACD, PSY, RCI		[86]
					EMA, HMA, ROC, RSI, MACD, TSI, OBV	Changeable	[87]
					200 rules	Yes	[43]
				RCons.	Index	Changeable (NTC)	[88]
				RPara.	Dimbeta		[89]
					Index, MA, VOP		[45]
			GA, ACR	RCons.	Price, MA, DMI, ADX, CCI, RSI, ROC, WMS %R, MACD, Stochastic (%K&%D), Vol		[90]
			GA, ANFIS, Subtractive clustering	RComb.	TAIEX, Difference of Hang Seng, NASDAQ		[91]
				RCons.	MA, AR, PSY, RSI, Stochastic (%K&%D), WMS %R, Vol		[24]
			GA, CBFDT, CBRWC, SRA	MF	MA, BIAS, RSI, Stochastic (%K&%D), MACD, WMS %R, PSY, Vol		[10]
			GA, FTS	MF	Index		[92]
			GA, Fuzzy system	RCons.	Price, MA, Price change, OBV	Changeable	[46]
					Price		[93]
			GA, Fuzzy system, GARCH	MF, RPara.	Price		[94]
			GA, Fuzzy system, SOM NN, SRA	MF, RComb.	Price		[95]
			GA, Hierarchical GA	RComb., RPara.	SMA, MACD, MFI, MTM, RSI, Stochastic (%K&%D), WMS %R	Changeable	[96]
			GA, Neurofuzzy system	MF	VAMA	Changeable	[56]
			GA, PVC, BMA, BOOST	RComb.	MAs, ROC, Stochastic (%K&%D), MACD, OBV, EMV, S&R	Changeable	[47]
			GA, RST, CPDA, MEPA	RComb.	MA, MTM, Stochastic (%K&%D), RSI, PSY, WMS %R, AR, VR, Vol	Changeable (NTC)	[27]
			GA, SAX	RCons.	Price	Changeable	[97]
			HCGA, Fuzzy system	MF	PPO	Yes	[26]
			Meta-GA	RPara., Parameters of GA	MA		[44]
			MOEASI-II	RPara.	MACD, RSI		[98]
			Phenotypic GA	RCons.	MA, ADX, DMI, Stochastic (%K&%D), RSI, MACD, ROC, WMS %R, CCI		[99]
		GA&GP	Hybrid of GA and GP	RCons.	MA, RSI, Stochastic	Changeable	[100]
			ITGAP	RCons.	Price, MA, EMA, ROC, RSI, Stochastic (%K), TD, VHF	Changeable	[101]
		GNP	GNP with control nodes, SL	RCons.	CC, ROD, G/D, RSI, ROC, RCI, Stochastics, MACD, VR	Yes (NTC)	[102]
			GNP with rule accumulation, SL	RCons.	CC, ROD, G/D, RSI, ROC, RCI, Stochastics, MACD, VR	Changeable	[103]
			GNP, SL	RCons.	CC, ROD, G/D, RSI, ROC, RCI, Stochastics, MACD, VR	Yes (NTC)	[104]
			TA-GNP, SL	RCons.	CC, ROD, G/D, RSI, ROC, RCI, Stochastics, MACD, VR	Yes (NTC)	[28]
		GP	GP	RCons.	MA, Vol		[105]
					Index, MA, lag, Norm, Slope, Volatility	Changeable	[106]
					Price, MA, lag (price, vol), Norm, RSI, ROC, Volatility, Vol, MA (vol)	Changeable	[61]
					Price	Changeable	[107]
					Price, MA, lag (price, vol), Norm, RSI, ROC, Volatility, Vol, MA (vol),	Changeable (NTC)	[108]

Table 1 (*Continued*)

Analysis method classification	EC classification	EC technique	Proposed model	Optimization object	Input variables	Outperform B&H	Ref.
					Price, MA, Norm	No	[14]
					Price, MA, lag	No	[109]
			GP, polynomial NN	RCons.	Index		[48]
			GP, SOM	RCons.	MA, +DI, −DI, CCI, BIAS, MACD, Stochastic (%K&%D), WMS %R, MTM, ROC, RSI, B/S MTM, B/S Will, PSY, VR		[110]
		LCS	XCS	RComb.	Price spread, Price spread ratio, Bias		[111]
	Swarm intelligence	PSO	PSO	RComb.	DMI, LIN, MA, PSR, MACD		[112]
			PSO, Fuzzy system, FLANN	MF, RCons.	SMA, EMA, ROC		[23]
			PSO, Fuzzy system, Random forests	RPara.	EMA, R/S line		[113]
			PSO, FTS, SVM	MF	One-day variation, two-days average variation		[114]
			PSO, Neurofuzzy system, LSE, Self-constructing clustering	RPara.	TAIEX, TAIFEX, DJIA, NASDAQ, SMA, BIAS		[115]
Blending analysis	Evolutionary algorithm	GA	GA	RPara.	Interbank rate, MA		[116]
					Interbank rate, MA		[117]
			GA, FDT, K-means, SRA	MF	CS, RS, P/E, TN, NWMV, MA, BIAS, RSI, Stochastic (%K&%D), MACD, PSY, WMS%R, Vol		[25]
			GA, Fuzzy system, Fuzzy clustering, Sugeno and Yasukawa method	MF	P/E/S, Price, MA, LRS, −DI, +DI, R2, MTM, DI, MACD, MA-MACD, RSI, Vol, ATR, Price channel		[118]

Note: The Optimization object "MF", "RComb.", "RCons.", "RPara." presented in row 5 mean membership functions, rule combination, rule construction and parameters of rule; "NTC" in row 7 indicated there is no information about transaction cost in paper.

Evolutionary algorithm (EA). EAs were proposed in the late 1950s and the early 1960s. Initially, the representative algorithms were EP, ES and GA. EA represents optimization algorithms that search the space by simulating the genetic evolutionary process of Darwin's theory including selection, mutation, recombination and reproduction. It also uses fitness functions as performance measures to drive the evolution towards better regions of the search space [39,71–75,80,81]. At present, the main sub-fields of EAs include EP, ES, GA, GP, LCS, etc.

Swarm intelligence (SI). SI, first proposed by Beni and Wang [82], primitively describes a paradigm about cellular robotic systems. Afterwards, Bonabeau et al. [83] expanded this definition to "algorithms or distributed problem solving devices inspired by the collective behavior of insect colonies and other animal societies." The latter definition is widely accepted [41,42,72,84]. At present, the main sub-fields of SI include PSO and ACO.

Hybrid EC techniques (evolutionary algorithm and swarm intelligence). Algorithms that combine EA and SI will be regarded as hybrid EC techniques.

Notably, some researchers also suggest that other techniques, such as simulated annealing (SA), be included into EC [36,76], however, following our investigation, this view seems to not have been widely accepted. Thus, these techniques are excluded from our review.

4. Classification and detail information of the reviewed articles

The 51 articles are classified based on the proposed classification framework, as shown in Table 1. This section provides a summary and analysis of the selected literatures from different perspectives.

4.1. Distribution of articles by analysis method

The distribution of articles classified by analysis methods and EC techniques is shown in Table 2. Technical analysis dominates this field, accounting for 45 out of the 51 articles. Only two out of the 51 articles use pure fundamental analysis, and four out of the 51 papers are devoted to Blending method. As Chavarnakul et al. [56] suggested, the popularity of technical analysis maybe due to the following reasons: the gains enjoyed by technical analysis have largely led to its wide adoption among investors and financial analysts. Many Wall Street financial advisors exploit technical analysis in practice [119], and most large investments are made within this technical environment [120]. In addition, major brokerage firms that comment on individual stocks and analysts that publish newsletters mostly adopt technical analysis [121]. Surveys also indicate that many investors make investment decisions using technical analysis tools [122]. However, according to the results of the selected papers in comparison with the B&H strategy, the profitability of the models for technical trading rule discovery seems unsteady.

Because of the popularity of technical analysis, popular indicators are technical indicators. According to Table 1, these technical indicators comprise MA (trend indicator, 29 out of the 51 articles), price or index (raw data, 21 out of the 51 articles), RSI (momentum indicator, 19 out of the 51 articles), stochastic (momentum indicator, 16 out of the 51 articles), MACD (momentum indicator, 16 out of the 51 articles), ROC (momentum indicator, 13 out of the 51 articles), Vol (volume indicator, 9 out of the 51 articles), WMS %R (momentum indicator, 8 out of the 51 articles), VR (volume indicator, 6 out of the 51 articles), MTM (momentum indicator, 5 out of the 51 articles), PSY (momentum indicator, 5 out of the 51 articles), and BIAS (momentum indicator, 5 out of the 51 articles). In general,

Table 2
Distribution of articles by analysis method and EC techniques.

Analysis method	EC classification	EC techniques	Number of papers		
Fundamental analysis			2		
	Evolutionary algorithm			2	
		GA			1
		GP			1
Technical analysis			45		
	Evolutionary algorithm			40	
		GA			24
		GA & GP			2
		GP			9
		GNP			4
		LCS			1
	Swarm intelligence			5	
		PSO			5
Blending analysis			4		
	Evolutionary algorithm			4	
		GA			4

Table 3
Distribution of articles by input variables.

Analysis type	Analysis factor classification	Number of papers	Ref.
Fundamental	Macroeconomic	2	[116,117]
	Company	4	[25,32,85,118]
Technical	Raw data	26	[14,28,43,45,46,48,61,88,90–95,97,101–104,106–109,111,115,118]
	Trend	35	[10,14,23–25,27,28,43–47,61,86,87,89,90,96,99–106,108–110,112,113,115–118]
	Momentum	33	[10,14,23–28,43,45–47,61,86,87,90,91,96,98–104,106,108,110–112,114,115,118]
	Volume	18	[10,24,25,27,28,43,47,56,61,87,90,102–105,108,110,118]
	Volatility	10	[43,45–47,61,101,106,108,113,118]

the popular classifications of factors shown in Table 3 are trend (35 out of the 51 papers), momentum (33 out of the 51 papers), raw data (26 out of the 51 papers) and volume (18 out of the 51 papers). Other types of factors, such as sentiment, flow-of-funds, and cycles, still lack research.

For fundamental factors, 4 out of the 6 articles that analyzed individual stocks used company analysis factors mainly from financial statements, whereas the other two articles that studied the index used the macroeconomic factor of interbank rate as input.

4.2. Distribution of articles by EC techniques

According to Table 2, currently the commonly used EC techniques in trading rule discovery are mainly GA and GP, among which GA is the most prominent accounting for 29 out of the 51 articles (one paper on fundamental trading rule discovery, 24 papers on technical trading rule discovery, and four papers on blending trading rule discovery), followed by GP with 10 out of the 51 articles (one fundamental and nine technical). Moreover, according to Table 1, 16 out of the 51 articles combined fuzzy logic techniques (such as fuzzy system and neurofuzzy system) with EC techniques, among which 12 out of the 16 articles combined it with GA. Research attaches importance to the ability of models based on fuzzy logic to generate linguistic rules. Relatively, other EC techniques, such as GNP, LCS, and PSO, have not received considerable attention because related research began in 2009.

This lack of attention may be due to the following reasons. The theoretical foundation of GA was proposed in 1975 [123], GP was proposed in 1992 [124], PSO was proposed in 1995 [125], and GNP was proposed in 2000 [126]. The longer the history of an algorithm, the easier it is for researchers to comprehend and discover. A firmer ground for research is also developed. The first article in this field by Allen and Karjalainen [14], which applied GP as an extension of genetic algorithms to find rules, was published in the *Journal of Financial Economics*, and this article attracted the attention of researchers to GP and GA. In addition, compared with other EC techniques, such as ES, the fine combinatorial optimization capability of GA and GP allows for the easier optimization of certain rules, such as those based on Boolean functions, which have been widely used by investors who focus on technical analysis.

The optimization objects of models based on different EC techniques are observably different. According to Table 1, GP and GNP are always used to reconstruct rules, i.e. they searching both the optimal structure and parameters of rules (RCons.). This fact can be attributed to the variable lengths of the strings that symbolize the decision tree representation of GP and to the graph structure representation of GNP. PSO and GA are often applied to optimize the parameters of the membership functions (MF) which the rules are based on and the parameters of rules (RPara.) or to combine rules with logical statements or the weights of the rules (RComb.). These optimization objects often combine more priori knowledge than the models based on GNP and GP. However, we cannot conclude which optimization object can gain better results based on the selected papers.

4.3. Evaluation scheme

Table 4 summarizes the primary measures used to evaluate model performance. It shows that most researchers test their

Table 4
Primary measures.

Object	Measures	Number of papers
Profitability	Return/profit/asset value (Consider Transaction cost, compare with B&H or Index)	32 (23, 23)
	Sharpe ratio (Consider Transaction cost)	9 (7)
Accuracy	RMSE	9
	Accuracy	6
	MAPE	5
	MAE	4
	Hit rate	3
	R^2	3

Table 5
Detail of evaluation scheme of researches compared with B&H.

	Sample (granularity; span; training; testing)	Research targets	Transaction	Other influenced factors	Fitness	Compared EC techniques	Ref.
Steadily outperform B&H	Daily; 4 years; **Best**: 180 days; 20 days	10 stocks in first section of Tokyo market (includes different trend)	–	–	Profit	GNP-RL, GNP-Candlestick, GA	[28]
	Daily; 4 years; 3 years; 1 year	16 stocks in first section of Tokyo market (includes different trend)	–	–	Total return	GNP-Actor Critic, GNP-Candle Chart, GA	[104]
	Daily; 4 years; 3 years; 1 year	10 stocks in first section of Tokyo market (includes different trend)	–	–	Profit	Conventional GNP, GA	[102]
	Daily; 16, 20 years; 1, 6 years; 10 years	NOL and HIS (uptrend)	0.20%	–	R-Square	–	[26]
	Daily; 3 years; 1 year; 1 year	24 stocks of CAC40 (includes different trend)	0.25%	–	Return, percentage accuracy, risk, maximum loss	–	[43]
	Yearly; 11.5 years; **Best**: 11 years; 1 year	8 electronics stock in TSE	Needless	Change of fuzzy system	Return, risk	–	[32]
	Daily; 7 years; 6 months; 6 months	30 stocks of DJIA	0.02/Share, Mini: 14 USD	Run of experiments	Return	–	[87]
Changeable comparison result	Daily; 12 years; 7 years; 5 years	S&P 500,99 stocks of S&P 500; **Good**: Avg., uptrend of index and downtrend of stocks; **Bad**: Avg., uptrend of stocks	Considered	Run of experiments; enable strategies (long/short) in different trend; enable'Days to Buy/Sell'gene	Total return	–	[97]
	Daily; 9 years; 5 years; 1 year	S&P 500; **Good**: downtrend; **Bad**: uptrend	0, 0.5%; **Good**: 0	**Good**: Using VAMA(21-63) as inputs **Bad**: using VAMA(5) as input	RMSE	–	[56]
	Daily; 6 year; 10 months; 2 months	TAIEX; **Good**: Avg., uptrend, downtrend; **Bad**: violent	–	–	Percentage accuracy	GAs	[27]
	Daily; 13 years; 3 years; 1 year	STI; **Good**: downtrend; **Bad**: Avg., uptrend	0.50%	Risk level	Risk-adjusted profit	–	[100]
	Daily; 4 years; 3 years; 1 year	10 Iranian companies with the highest liquidity; **Good:** using return as measure, Avg., downtrend, steady; **Bad**: using return as measure, uptrend	0.52%	**Good**: using risk adjusted return as measure; **Changeable:** using return as measure	Excess risk adjusted return	–	[61]

	Daily; 8 years; - ; -	S&P 500, S&P auto, S&P banks	0, 0.5%	**Good:** using return as measure; **Bad:** using risk adjust return	Sharp ratio or excess return	–	[107]
	Daily; 18 years; **Best:** 10 years; 1 year	S&P 500 (includes different trend)	0.25%	–	**Good:** Sharp ratio; **Bad:** risk, return	–	[101]
	Daily; 16 years; 120 days; 20 days	Stocks of MSCI Europe index; **Good:** Avg. downtrend	Considered	–	Return, penalty function	–	[46]
	Daily; 28 years; 7 years; 5–20 years	Russell 1000 2000 3000; **Good:** Russell 2000	0, 0.25%; **Bad:** 0.25%	–	Excess return	–	[106]
	Daily; 7 years; 27 months; 9 months	S&P 500; **Good:** downtrend	0.25%	–	Daily returns, accuracy	–	[47]
	Daily; 4 years; 3 years; 1 year	16 stocks in first section of Tokyo market; **Good:** Avg., downtrend; **Bad:** uptrend	Refers to securities company	–	Total return	–	[103]
	Daily; 10 years; -; -	IBEX-35; **Good:** steady, violent, downtrend; **Bad:** uptrend	–	–	R-Square	–	[88]
	Daily; 2003 days; 1,6 years; 1 year	14 stock of 14 section in TSE 300; **Good:** downtrend, steady; **Bad:** uptrend	–	Run of experiments, parameters of model	Excess return	–	[108]
	Daily; 8 years; 1.5, 2 years; 1.5, 2 years	8 stock of 4 section in HK; **Good:** downtrend; **Bad:** uptrend	Considered	Change of variables	Profit	Traditional GA	[96]
Underperform B&H	Daily; 67 years; 7 years; 15–65 years	S&P 500	0.1%, 0.25%, 0.5%	–	Compounded return	–	[14]
	Daily; 68 years; 7 years; 15–60 years	S&P 500	0.25%	–	Daily excess return, Sharpe ratio	–	[109]

Note: indicates the factors are key factors which means they can determine the comparison result in the experiments (i.e. whether the model can outperform B&H); "good"/"bad" indicates the margins of outperform/underperform.

indicates the factors are influenced factors which have some influence to the profitability of model but the experiments did not show that they can determine the comparison result; "best" means the best setting proposed by researchers.

indicates although there is some change in that factors, however, the change did not generate any significant effect to the result or the effective cannot catch the attention of authors of researches.
The remaining grids have no color means the factors in those grids has no change in the comparison experiments.

Table 6
Detail of evaluation scheme of researches did not compare with B&H.

Sample (granularity; span)	Research target	Fitness	Compared EC techniques	Ref.
Daily; 11 years	23 Stock of DJAI	MSE		[93]
Daily; 22 years	DJIA	Percent profit, Sharpe ratio	NSGA-II, GA	[112]
Daily; 10 years, 1 year	TAIEX, DJIA	RMSE		[92]
Daily; 13 years, 9 years, 17 years	S&P 500, DJIA, BSE	RMSE	Hybrid of LLWNNand PSO	[23]
Daily; 3 years	Top 10 stock of S&P 500	Precision rate		[99]
Daily; 4 years	Top 10 stock of S&P 500	Confidence		[90]
Daily; 7 years	5 Stock of S&P 500, S&P 500	Accuracy		[10]
Seasonally; 20 years	S&P 500	Accumulated return		[85]
Daily; 4 years	DJAI	Regularized cross-validation error	Traditional GP	[48]
Daily; 365 data points	An automotive company	Difference or errors between desired and simulated outputs		[118]
Daily; 26 years	IGBM	Return rate		[117]
Daily; 2 years	IBM, DELL, British airlines, Ryanair airlines	MSE	Hybrid of HMM, ANN and GA	[95]
Daily; 14 years	TAIEX-FISI	R-Square, MSE		[110]
Intraday; 4 years, 5 years, 6 years	TX, MSCI			[111]
Daily; 7 years	TAIEX, NASDAQ	A cost function		[94]
Daily; 14 years	KGHM	A cost function		[113]
Daily; 1430 records	3 Electronic companies in Taiwan	Hit ratio		[25]
Daily; 6 years	TAIEX	RMSE		[115]
Daily; 13 years	S&P 500	Compounded return		[45]
Daily; 6 years	25 stock of Madrid Stock Exchange, IBEX-35, General Index	Accumulated return	Traditional GA	[44]
Daily; 2800 observations	A UBS mutual fund investing in emerging stock markets	Accumulated wealth		[89]
Daily; 25 years	IGBM	Excess return		[116]
Daily; 1 years	8 Stocks of NYSE	Prediction rates		[86]
–	–	–		[105]
Daily; –	TAIEX	RMSE		[24]
Daily; 7 years	TAIEX	RMSE		[91]
Daily; 15 years	TAIEX	Average difference		[114]
Daily; 6 years	DJIA	Profit, risk		[98]

models through the simulation of trading in stock markets (32 out of the 51 papers) and present the return/profit/asset value of their models. Most of them have also considered transaction cost (23 out of the 32 papers). Thus, profitability is an important evaluation indicator in this field, which is reasonable because profit is the goal of many investors. Most studies (23 out of the 32 papers) compared their models with the B&H strategy or market index, aiming to prove whether the model can defeat the market average level (i.e., if their model is meaningful for investors). In addition, the Sharpe ratio is the most popular indicator for measuring risk-adjusted return. The most frequently used indicators that measure prediction accuracy are RMSE, MAPE, Hit rate, MAE, and accuracy.

To better analyze the existing state-of-the-art research, we further summarize the detail of the research evaluation scheme. Specifically, in this part, we divide the papers into two classifications. Table 5 shows that 23 studies compared their models with B&H, the most popular trading strategy used for comparisons. And the analysis of different studies based on the comparisons can help us avoid the problem that different market situation in studies will influence the profitability of models. The evaluation schemes of the remaining studies are presented in Table 6.

In general, the difference between the research works in Tables 5 and 6 is in their fitness function. Most of the fitness functions in the studies presented in Table 5 are indicators of profitability, such as return, Sharp ratio, and excess return. In Table 6, they are RMSE, MSE, Hit rate, R-Square, and so on, which are used to evaluate the prediction accuracy of the models. In fact, the profitability of the model minimizing prediction error is easily influenced by the preset strategies on how to use the prediction to trade, introducing significant difficulty in evaluating their profitability.

In Table 5, the evaluation scheme detail of the 23 papers compared with that of the B&H strategy is presented. Results show that outperforming the B&H strategy steadily is not easy; only 7 out of 51 papers have achieved it in their experiments. However, 4 of the 7 did not present any information on transaction cost, which indicates the costs were probably not considered in experiment (one of them are based on yearly data and transaction cost is not so important for it). One of the seven used R-Square as fitness, which means that the profitability can be easily influenced by the preset strategies. Thus, only the results of the three seem reliable.

Most of the comparison results are changeable. Transaction cost and market trend are two key factors that influence the comparisons. Although researchers conduct their experiment mainly based on day trading rather than high-frequency trading, half the number of the studies which did not present any information on transaction cost steadily outperforming the B&H, and only 3 of 17 research works, which definitely considered the transaction cost, can do so. In addition, each time researchers consider different levels of transaction cost in their research, they can always be a key factor (2 out of 3) or at least an important factor (1 out of 3) for the margins of excess return.

The most frequent key factor that influences the margins of excess return is market trend. In the classification of changeable results, the market trend is a key factor of the margins in 11 of 14 research works. A more interesting result is that in 11 of 11 papers, downtrend is a condition under which a model can gain excess return, and none of the papers observed the bad performance because of downtrend. Conversely, only one paper showed that its model can gain better performance in the uptrend than in other market trends, and 8 of 11 papers reported that their models underperforms B&H in the uptrend. The results suggest that existing models perform much better in the downtrend than in the uptrend. Furthermore, [97] suggests that disabling short strategies in the uptrend or disabling long strategies in the downtrend can improve the return; however, this approach cannot change the comparison result in its experiments. Actually, even compare differences of other factors among the research works, including sample, research target, optimization object, proposed model, input variables, and transaction cost, we cannot obtain any exact conclusion regarding the phenomenon based on such a small number of studies.

Table 7
Distribution of articles by stock market regions.

Region	Number of papers
USA	23 (41.1%)
Taiwan	11 (19.6%)
Japan	4 (7.1%)
Spain	4 (7.1%)
Singapore	3 (5.4%)
Hong Kong	2 (3.6%)
Iran	1 (1.8%)
Asia	1 (1.8%)
Canada	1 (1.8%)
Europe	1 (1.8%)
France	1 (1.8%)
India	1 (1.8%)
Ireland	1 (1.8%)
Poland	1 (1.8%)
UK	1 (1.8%)

However, by combining the information presented in Tables 1–3 and 5, we suggest that the phenomenon may be related to the input variable and the problem of liquidity in different market trends, which are discussed in Section 5.

Periods for training and testing are also interesting factors that influence the return, both the long training period (about 10 years) with a short testing (1 year) period [32,101] and the short training period (180 days) with a shorter testing period (20 days) [28], are suggested by researchers. The periods in the remaining research works vary, but mainly in the middle in terms of training period (3–6 years) and with a short testing period (1 year). In addition, some researchers reported the problem of instability of results when using EC; in their experiments, the excess returns in the best run are several times that of the average level [87,97,108]. Other factors that influence the profit reported by the researchers include the research targets, algorithm parameters, input variable, fitness function, risk level, and measure.

As the analysis above, many factors can influence the result; comparing the different kinds of EC techniques according to the results of the different studies is extremely difficult. Performing the comparisons in one study is more feasible. However, according to both Tables 5 and 6, studies on the comparisons among different kinds of EC techniques are lacking. Only 10 out of 51 papers implemented such comparisons, and a large number of these comparisons are among different versions of the same type of EC techniques, such as the comparisons among GNP, TA-GNP, GNP-RL, GNP-Candlestick, and GNP-SL.

4.4. Surveyed markets

The surveyed stock markets (Table 7) are mainly from the US (23 out of the 51 papers) and Taiwan (11 out of the 51 papers). The remaining articles are also primarily from mature markets such as Japan, Singapore and European countries. However, given that they are hotspots of the current world economy, emerging markets such as China and India lack studies.

4.5. Distribution of articles by EC techniques and publication year

Table 8 presents the distribution of articles according to EC techniques and year of publication, while Fig. 4 shows the overall research tendency. Obviously, the related studies increased significantly from 2009 to 2013. As a whole, they exhibit a rising trend.

4.6. Distribution of articles by journal

Table 9 shows that the selected articles are distributed across 24 journals. Among these are, mainstream academic journals about decision support systems and expert systems, including

Table 8
Distribution of articles by EC techniques and publication year.

Year	GA	GP	GA & GP	GNP	LCS	PSO	Total
1999		1					1
2000							
2001							
2002	1						1
2003		1					1
2004	1	2					3
2005	1	1					2
2006	1	1					2
2007	3	1					4
2008							
2009	7		1	2			10
2010	4	1	1	1			7
2011	2	2			1	1	6
2012	3					2	5
2013	6			1		2	9
Total	27	10	2	4	1	5	51

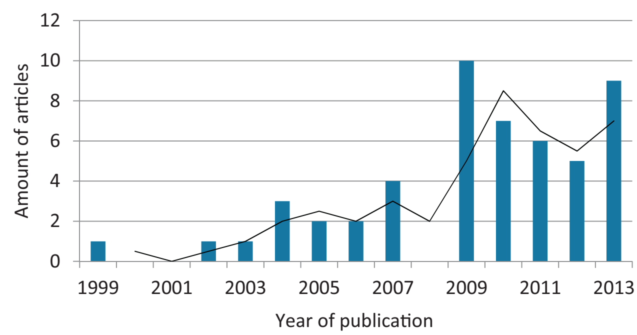

Fig. 4. Distribution of articles by year of publication.

"*Decision Support Systems*" (two papers), "*Information Sciences*" (three papers), "*Expert Systems with Applications*" (16 papers), and "*Knowledge-Based Systems*" (two papers) composing more than 45.8% (22 out of the 51) of the total articles published. Mainstream journals in operational research and soft computing (which mainly

Table 9
Distribution of articles by journal.

Journal title	Number of papers	Percentage
Expert Systems with Applications	16	31.37%
Applied Soft Computing	4	7.84%
Information Sciences	3	5.88%
Applied Financial Economics	2	3.92%
Computers & Operations Research	2	3.92%
Decision Support Systems	2	3.92%
European Journal of Operational Research	2	3.92%
IEEE Transactions on Evolutionary Computation	2	3.92%
Knowledge-Based Systems	2	3.92%
Procedia Computer Science	2	3.92%
Applied Economics Letters	1	1.96%
Applied Intelligence	1	1.96%
Applied Stochastic Models in Business and Industry	1	1.96%
Computers & Mathematics with Applications	1	1.96%
International Journal of Forecasting	1	1.96%
International Review of Economics & Finance	1	1.96%
Journal of Financial Economics	1	1.96%
Mathematical and Computer Modeling	1	1.96%
Economic Modeling	1	1.96%
Natural Computing	1	1.96%
Neurocomputing	1	1.96%
Physica A: Statistical Mechanics and its Applications	1	1.96%
Quantitative Finance	1	1.96%
Soft Computing	1	1.96%
Total	51	100.00%

include fuzzy logic, artificial neural network, and evolution computation) follow and include "*European Journal of Operational Research*" (two papers), "*IEEE Transactions on Evolutionary Computation*" (two papers), "*Computers & Operations Research*" (two papers), "*Applied Soft Computing*" (four papers), and "*Soft Computing*" (one paper) totaling 11 articles (22.9%). Mainstream academic finance journals come last and include "*Journal of Financial Economics*" (one paper), "*Quantitative Finance*" (one paper), and "*Applied Financial Economics*" (two papers), totaling only four papers. Overall, the two journals ("*Expert Systems with Applications*" and "*Applied Soft Computing*") seem to appreciate articles in this field better.

5. Discussion

5.1. Limitations in existing researches

5.1.1. Limitation in imbalanced performance

Based on the preceding analysis, this section provides discussions on the research gaps and future research directions in the reviewed area.

For further research, the imbalanced performance of trading rules in different market trends (i.e., good in downtrend and bad in uptrend as presented in Table 5) and the improvement of the evaluation scheme are two key problems that need to be addressed to further improve the trading rule discovery. Even though the discovery of effective rules in the downtrend market can also assist investors, two problems occur. First, predicting the future market trend is difficult for investors. Second is the question of whether the existing model can gain excess return over a long period because the markets are always in an uptrend in a long history. Moreover, the two studies that underperformed the B&H strategy are the only research works that tested their model with long history data [14,109]. Thus, the contribution of the studies, which only claim good performance mainly in the downtrend and have not any effective approach to predict the trend, is questionable.

5.1.2. Limitation in feature selection

According to the existing studies, the reason that causes the imbalance performance cannot be determined exactly, and more researchers should examine this problem. However, by analyzing the distribution of studies based on the classification framework, we observe the occurrence of the phenomenon in different research targets, techniques, and optimization objects. Thus, we think these factors do not cause the phenomenon. Thus, we wonder about the role of input variables in this phenomenon. As mentioned in Section 4.1, most of the studies are based on technical analysis, and a small number of technical indicators account for a major share of selected factors in the research works. Thus, we wonder if the inclination of factor selection causes the phenomenon of imbalanced comparison result. To verify this point we try our best to find some researches about stock trading rule discovery with data mining techniques base on unpopular factors. According to several studies which mainly involved fundamental analysis and compared their result with uptrend markets, we did observe those models can more easily outperform B&H in uptrend [13,34,35].

5.1.3. Limitation in transaction cost selection

Moreover, by further analyzing the difference between reality and experiments, as well as the results presented in Table 5, we suggest that the change of liquidity may be another factor that causes the imbalanced performance of models. Liquidity is the ability of a market to facilitate the purchase or sale of an asset without causing an observable change in the price level [127]. Buying or selling stock with lower liquidity increases or decreases the stock price more easily, and the result is a decrease in the expected return. Researchers often use transaction cost to avoid the influence of liquidity. Transaction cost includes the cost of commissions, trading fee, slippage, and liquidity; Narang [2] suggested that the cost of liquidity is the most important among these costs. Based on the analysis in Section 4.3, change of transaction cost significantly influences the profitability of a model. The influence is more severe for technical trading, to which most of the selected papers belong, because transaction frequency in technical trading is often much higher than that in fundamental trading. The problem is that the liquidity of a market is not constant; many researchers have suggested that liquidity is influenced by many market conditions such as drying up in the downtrend market [128–130]. Results of previous studies on liquidity suggest two things: one is that transaction costs in the downtrends should be much larger than those in the uptrends and the other is that the loss in the downtrend and uptrend with the same percentage is different [128–130]. However, none of the selected studies have considered these two problems in their experiments.

5.2. Future research directions for limitations

To solve the problem of imbalanced performance in different market trends, we suggest the following research directions for researchers: improving performance in the uptrend, predicting future market trends, and discovering other predictable classification standards to replace the classification of market trends.

5.2.1. Improving performance in uptrend market

Because of the poor performance of existing models in the uptrend, the performance improvement of models in the uptrend has significant research potential in terms of space and value. As mentioned, we suggest that the inclination of factor selection causes the phenomenon of poor performance in the uptrend. Thus, for further research, other analysis factors, such as technical indicators of sentiment, flow of funds, and cycles, as well as fundamental factors, can be considered; all of these factors lack sufficient research as the analysis in Section 4.1 indicates. Sentiment is a highly popular topic in financial analysis with machine learning [131,132]. Instead of using classical indicators, such as expert-public ratio, recent studies have focused on the sentiment hidden in the text message released by social networks [132,133]. However, researches about using effective techniques to discovery explicit knowledge from this kind of message are lacking. Flow of fund can be a reference to the liquidity of stocks, which are controlled by several large investors [6]. Cycles are a popular approach to evaluate the long-term movement from a technical perspective [6]. Fundamental analysis is also an extremely popular approach for investors [55]. Compared with technical analysis, which learns the market pattern empirically, fundamental analysis reveals the stock value through both empirical and theoretical ways, which can improve the comprehension of investors regarding stock price. Thus, even famous technical investors pay attention to fundamental factors [134]. In addition, although certain fundamental factors have been employed in practice [7], the majority of research is limited to using indicators that are mainly from financial statements. Other fundamental factors are hardly ever considered. However, Bodie et al. [7] illustrated that, for some companies, macroeconomic and industrial environments may play more important roles than company performance within the industry. In addition, given that insiders usually monopolize the operational information of companies, using more transparent data from the macroeconomy and industry environments may neutralize the informational disadvantage of the investors. In the big data era, researchers can obtain much more data, such as news on the Internet which has not been well exploited by investors yet [4,135,136]. The efficiency of markets also requires a constant search for effective factors [7,14,15,137].

5.2.2. *Predicting future market trends*

As we have discovered the advantages of different investment approaches (i.e., EC performs well in the downtrend and B&H performs well in the uptrend), predicting future market trends and changing the investment approach based on the prediction is also an alternative to solve the imbalanced performance. Trend, momentum, cycle, and raw data (chart data) in technical and fundamental analyses are popular for the long-term price trend prediction [13,61,138–140]. Data granularity may be among the problems faced by researchers. As shown in Tables 5 and 6, the most commonly used granularity is the daily. Daily data may be suitable for short-term movement analysis. However, evaluating the long-term trend based on recent movement over a few days is difficult; the judgments are always based on several weeks, months, and years of data [6], which means that the search space of the model based on daily data are extended greatly and selecting the suitable data granularity for the input and output (e.g., suitable for short-term prediction model) is highly important.

5.2.3. *Searching more predictable and influential classification standards*

Emphasizing the strengths of the model in other predictable classifications rather than the market trend is also a feasible approach. Table 5 shows that changing the research target can change the profitability of the models. This phenomenon has been widely observed by investors. It is related to the concept drift of effect of variables and the analysis approach among different kinds of stocks [153,154]. Because of the budget limitation and the selectivity of investment approach, accurately predicting the price movement of some predictable classification stock can also bring investors great help. And investing in the classification of stocks, which the investor is good at, is also very normal even for well-known investors such as Warren Buffett [141]. Researchers can consider studying how to identify the strength of models in predictable classifications, such as industry, growth or value, large-cap or small-cap firms, and mature market or emerging markets, which are not investigated sufficiently.

5.3. *Additional research directions*

Furthermore, using the information presented in Section 4, we provide additional suggestions to improve the evaluation scheme and application, which include considering liquidity and transaction cost in more precise and positive ways, and combining portfolio selection techniques and period for training and learning.

5.3.1. *Transaction cost estimation and liquidity information utilization*

Previous research did not pay much attention to the selection of transaction cost. Six of 23 papers did not provide any information on transactions in their experiments. Two of 23 papers estimated the transaction cost based only on commissions and trading fee [87,103], which is improper in many cases because liquidity is a very important part of transaction cost [2]. Many of the remaining studies, such as [26,97], did not provide the sources of transaction cost. Based on the analysis in Section 4.3, transaction cost greatly influences the result, which means that further research should consider it more precisely in their experiments. In addition, researchers should also consider the dynamic nature of liquidity and its influence on transaction cost for a more precise evaluation and an improvement of the performance of rules and models. However, using transaction cost to eliminate the influence of their trading to liquidity is a passive approach; some large investors use their trading influence to evaluate the market conditions or change the trend of the stock deliberately [134]. This fact means that our influences can also be an advantage for us.

5.3.2. *Combining portfolio selection techniques*

Many of the selected papers did not consider the problem of portfolio optimization; the performances of the models were always average in the independent experiments of each of the research objects [61,103]. But this kind of evaluation scheme cannot take full advantage of budget. For dispersing the risk, taking full advantage of budget and managing conflict of trading signals, portfolio selection is a very important problem for investors [2]. In the framework on the process of AT proposed by Narang [2,17], portfolio selection is the main task of the second process (i.e., trading signal generation) following the process of stock prediction (i.e., pretrade analysis). Similarly, research on portfolio selection and stock trading rule discovery are always separate. A major problem is that directly combining the proposed techniques in the recent researches on both the trading processes is insufficient to obtain optimal strategies because the existing research on portfolio selection assumes that the portfolio will be kept for a long period [142–144]. However, research on trading rule discovery suggests that the direction of stock movement can change any time. Thus, for the combination of portfolio selection techniques and trading rule discovery techniques (especially the technical trading rule discovery), researchers need to consider correlation of stocks in different environments, relationships between trading rules (such as priority and sequence), transaction cost, and selection of fitness function.

5.3.3. *Optimization of period for training and learning*

The periods for training and testing are factors that have been observed to have large influences on the experiment result. The influences can be attributed to the effectiveness of the stock market and the noise of the stock market data [2,10,11,13–15]. Better periods for training and testing can improve the evaluation of the model performance. However, significant differences exist between the proposed optimal periods and the commonly used periods, and between the optimal periods proposed by different researchers. The relationship of stock classification, periods of training and testing, and profitability may be an interesting topic.

5.4. *Future roles of EC techniques*

Although we fail to observe the significant difference of results among using different techniques because of the variety of influenced factors and the lacking of researches of many techniques as well as comparison experiment among EC techniques. However, in implementing the proposed improvement, the selection of suitable EC techniques may be extremely important. For example, the PSO technique has become increasingly popular in recent years. In AT, PSO-based techniques are widely used for portfolio selection [142–144]. According to the comparative study between PSO and GA, the comparison of performance is optimization function-dependent; however, PSO outperforms GA in terms of operation efficiency in most cases [145–149]. The operation efficiency of techniques may be important for suggested improvements. For one thing, to improve the performance in the uptrend and predict future market trends, researchers may need to consider more variables than before, indicating the extension of search space. For another, because the recommendation of stocks (including information on direction, expected return, and risk) generated by rules is always difficult to predict, an online learning for portfolio selection may be necessary. Moreover, the tree structure representation of GP and the graph structure representation of GNP are suitable for managing the priority and sequence among rules involving different stocks. The EC techniques that combine reinforcement learning, such as GNP-SL, GNP-RL, and XCS, can be used to optimize the strategies considering the influence of investors' transaction behavior because collecting high-quality historical data samples may be very expensive. Moreover, the simplicity and advantage of GA in combination optimization can help researchers determine the best classification of stocks.

However, for further and more exact analysis of the advantage of different techniques in different situations, more comparative studies should be conducted. Finally, the problem of the unstable performance of models because of the stochastic characteristic of EC techniques should be considered in future research.

6. Conclusion

The application of EC techniques for rule discovery in stock algorithmic trading is an emerging research field in investment and is eliciting considerable research attention. To provide a systematic and comprehensive review, a classification framework for the related articles, and suggestions for future research, we reviewed 51 articles published in academic journals from the perspectives of evolutionary computation and investment theories. Although our research cannot claim to be exhaustive, we believe that this review can facilitate knowledge accumulation and further studies in this research area. The implications of this research include the following:

(1) Existing research bias toward the application of GA-based and GP-based techniques in technical trading rule discovery occurs, and existing research is not enough to clarify the advantage of different kinds of EC techniques, optimization object, and input variable in trading rule discovery. More comparative tests were needed.
(2) Market trend and transaction cost are the important factors that influence the performance of models in the selected papers. The corresponding imbalanced performance in different market trends and the improvement of evaluation scheme are two key problems that need to be considered in further enhancing the process of trading rule discovery.
(3) We suggest that the unity of factor selection and the problem of transaction cost selection are two important preconditions of the phenomenon of imbalanced performance in different market trends. However, further research is necessary to ascertain the preconditions. In addition, the absence of ways for market trend prediction and the problem of transaction cost selection are two major limitations of the studies that we have reviewed.
(4) To solve the problem of imbalanced performance in different market trends, we suggest three research directions for researchers: improving performance in the uptrend, predicting future market trends, and discovering other predictable classification standards to replace the classification of market trends.
(5) For the further improvement of the evaluation scheme and its application in this field, research can also consider liquidity and transaction cost in more precise and positive ways, as well as consider the optimization of portfolio selection and periods for training and learning.
(6) As future applications of EC techniques for rule discovery in stock AT may face more complex environments, the advantages of different EC techniques should gain more research attention to ensure improvements in this field.

This study has some limitations. First, most of the selected articles were extracted based on the descriptor of "evolutionary computation," "stock," and "trading rule" that mention the application of EC techniques in stock trading rule discovery. Second, non-English publications were excluded and related research may have been published in other languages.

Acknowledgments

This research was partly supported by the National Natural Science Foundation of China (71271061, 70801020), Science and Technology Planning Project of Guangdong Province, China (2010B010600034, 2012B091100192), and The Ministry of Education Innovation Team Development Plan, Guangdong Natural Science Foundation Research Team, and Business Intelligence Key Team of Guangdong University of Foreign Studies (IRT1224, S2013030015737, TD1202). The authors are grateful for the constructive comments of the referees on an earlier version of this paper.

Appendix A. Appendix

Modeling techniques

ANFIS	Adaptive network-based fuzzy inference system
ACR	Association rule
ACO	Ant colony optimization
BOOST	Boosting method
BMA	Bayesian Model Averaging
CBFDT	Case based fuzzy decision tree
CBRWC	Case base reasoning weighted-clustering
CPDA	Cumulative probability distribution approach
DT	Decision tree
EDDIE	Evolutionary dynamic data investment evaluator
EFuNN	Evolving fuzzy neural network
EP	Evolutionary programming
ES	Evolution strategy
FDT	Fuzzy decision tree
FLANN	Functional link artificial neural network
FTS	Fuzzy time series
GA	Genetic algorithm
GARCH	Generalized autoregressive conditional heteroskedasticity
GFS	Genetic fuzzy systems
GNP	Genetic network programming
GP	Genetic programming
HCGA	Hierarchical coevolutionary genetic algorithm
ITGAP	Incremental training genetic algorithm programming
LCS	Learning classifier system
LLWNN	Local Linear Wavelet Neural Network
LSE	Least square estimates
MEPA	Minimize entropy principle approach
MOEASI-II	Multi-Objective Evolutionary Algorithm with Super Individual II
NN	Neural network
Polynomial NN	Polynomial neural network
PSO	Particle swarm optimization
PVC	Plurality Voting Committee
RL	Reinforcement learning
RST	Rough sets theory
SA	Simulated annealing
SAX	Symbolic aggregate approximation
SL	Sarsa learning
SRA	Step-wise regression analysis
SOM	Self-organizing map
TA-GNP	Time adapting genetic network programming
XCS	Extended classifier system

Optimization objects

FS	Feature selection
MF	Membership function
RPara.	Rule parameter
RComb.	Rule combination
RCons.	Rule construction

Fitness functions

MAPE	Mean absolute percentage error
MSE	Mean squared error
RMSE	Root mean square error
MAE	Mean absolute error
R^2	Square of the Pearson product-moment correlation value

Factors

ADX	Average directional index	Trend
AR	Accumulative ratio	Momentum
ATR	Average true range	Volatility
B/S MTM	Buying/selling momentum indicator	Momentum
B/S Will	Buying/selling willingness indicator	Momentum
BIAS	Day bias	Momentum
CC	Candlestick chart	Raw data
CCI	Commodity channel index	Momentum
CPI	Consumer price index	Macroeconomic
CS	Capital stock	Company
DI	Demand index	Momentum
DMI	Directional movement index	Trend
D/BE	Debt over book equity	Company
EMA	Exponential moving average	Trend
EMV	Ease of movement value	Momentum
EPS	Earnings per share	Company
GDP	Gross domestic product	Macroeconomic
G/D	Golden/dead cross	Trend
HMA	Hull moving average	Trend
LIN	Linear regression	Trend
LRS	Linear regression slop	Trend
MA	Moving average	Trend
MA-MACD	Moving average MACD	Momentum
MACD	Moving average, convergence-divergence	Momentum
MAD	MA deviation rate	Trend
MFI	Money Flow Index	Momentum
MTM	Momentum	Momentum
Norm	Absolute value of the difference between two real value	Technical
NWMV	Net worth and market value ratio	Company
OBV	On-balance volume	Volume
PCF	Price-cash flow ratio	Company
PPO	Percentage price oscillator	Momentum
PSR	Parabolic stop and reverse	Trend
PSY	Psychology line	Momentum
P/B	Price-book value ratio	Company
P/E	Price per earning	Company
P/E/S	Price per earning per share	Company
RCI	Rank correlation index	Momentum
ROA	Return on assets	Company
ROC	Rate of change	Momentum
ROD	Rate of deviation from moving average	Trend
RS	Revenue situation	Company
RSI	Relative strength index	Momentum
SG	Sales growth	Company
SLEMA	EMA over shorter period versus EMA over longer period	Trend
Slope	The rate of change in price level over the past *n* days	Momentum
SMA	Simple moving average	Trend
SLMA	SMA versus LMA	Trend
S&R	Support/resistance indicator	Volatility
TD	Typical deviation	Volatility
TG	Turnover growth	Company
TN	Turnover number	Company
TSI	True Strength Index	Momentum
VAMA	Volume adjusted moving average	Volume
VHF	Vertical-horizontal filter	Volatility
Vol	Volume	Volume
Volatility	The variance in daily returns over the past n days	Volatility
VOP	Variance of price	Momentum
VR	Volume ratio	Volume
WMS %R	William's percent range	Momentum
PCF	Price-cash flow ratio	Company

Index & stock

BSE	Bombay stock exchange
CAC40	Cotation Assistée en Continu 40
DJIA	Dow Jones Industrial Average
HSI	Hang Seng Index
IBEX-35	Iberia Index
IGBM	General Index of the Madrid Stock Exchange
KGHM	The largest copper mining company in Europe
MSCI	Morgan Stanley Capital International Taiwan Index Futures
NOL	Neptune Orient Lines
STI	Straits Times Index
S&P500	Standard & Poor's 500 Index
TAIEX	Taiwan Weighted Stock Index
TAIEX-FISI	Finance and Insurance Sub-Index of TAIEX
TSE	Taiwan Stock Exchange Index
TX	Taiwan Stock Index Futures, traded at the Taiwan Futures Exchange

References

[1] World Federation of Exchanges 2012. Annual Report & Statistics, 2013.
[2] R.K. Narang, Inside the Black Box: the simple truth about quantitative trading, in: Wiley Finance, 1st ed., John Wiley & Sons, Hoboken, NJ, 2009, pp. 256.
[3] R. Pardo, The Evaluation and Optimization of Trading Strategies, 2nd ed., John Wiley & Sons, Hoboken, NJ, 2008, pp. 334.
[4] J. Malvey, A. Shorwty, L. Akoner, A First Perspective: The Transformational Influence of "Big Data" on the 21st Century Global Financial System, The Bank of New York Mellon Corporation, 2013.
[5] Y. Zhang, L. Wu, Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network, Expert Syst. Appl. 36 (5) (2009) 8849–8854.
[6] M.J. Pring, Technical Analysis Explained: The Successful Investor's Guide to Spotting Investment Trends and Turning Points, 4th ed., McGraw-Hill, New York, 2002.
[7] Z. Bodie, A. Kane, A. Marcus, Investments, 6th ed., McGraw-Hill, New York, 2005.
[8] R.W. Colby, T.A. Meyers, The Encyclopedia of Technical Market Indicators, 1st ed., Dow Jones-Irwin, USA, 1988, pp. 581.
[9] S.B. Achelis, Technical Analysis from A to Z, 2nd ed., McGraw-Hill, New York, 2013.
[10] P. Chang, C. Fan, J. Lin, Trend discovery in financial time series data using a case based fuzzy decision tree, Expert Syst. Appl. 38 (5) (2011) 6070–6080.
[11] A. Rodríguez-González, et al., CAST: using neural networks to improve trading systems based on technical analysis by means of the RSI financial indicator, Expert Syst. Appl. 38 (9) (2011) 11489–11500.
[12] M.A. Boyacioglu, D. Avci, An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange, Expert Syst. Appl. 37 (12) (2010) 7908–7912.
[13] N. Ren, M. Zargham, S. Rahimi, A decision tree-based classification approach to rule extraction for security analysis, Int. J. Inf. Technol. Decis. Mak. 5 (1) (2006) 227–240.
[14] F. Allen, R. Karjalainen, Using genetic algorithms to find technical trading rules, J. Financ. Econ. 51 (2) (1999) 245–271.
[15] B.G. Malkiel, E.F. Fama, Efficient capital markets: a review of theory and empirical work, J. Financ. 25 (2) (1970) 383–417.
[16] The New York Times, Times Topics: High-Frequency Trading, 2012.
[17] P. Treleaven, M. Galas, V. Lalchand, Algorithmic trading review, CACM 56 (11) (2013) 76–85.
[18] T. Hendershott, C.M. Jones, A.J. Menkveld, Does algorithmic trading improve liquidity? J. Financ. 66 (1) (2011) 1–33.
[19] I. Aldridge, High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems, 2nd ed., John Wiley & Sons, Hoboken, NJ, 2013, pp. 306.
[20] G. Nuti, et al., Algorithmic trading, Computer 44 (11) (2011) 61–69.
[21] P. Tan, Introduction to Data Mining, 1st ed., Pearson Education, New Jersey, 2007, pp. 792.
[22] E. Ngai, et al., The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature, Decis. Support Syst. 50 (3) (2011) 559–569.
[23] S. Chakravarty, P.K. Dash, A PSO based integrated functional link net and interval type-2 fuzzy logic system for predicting stock market indices, Appl. Soft Comput. 12 (2) (2012) 931–941.
[24] L. Wei, A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX, Econ. Model. 33 (2013) 893–899.
[25] R.K. Lai, et al., Evolving and clustering fuzzy decision tree for financial time series data forecasting, Expert Syst. Appl. 36 (2, Part 2) (2009) 3761–3773.
[26] H. Haoming, M. Pasquier, Q. Chai, Financial market trading system with a hierarchical coevolutionary fuzzy predictive mode, IEEE Trans. Evol. Comput. 13 (1) (2009) 56–70.
[27] C. Cheng, T. Chen, L. Wei, A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting, Inf. Sci. 180 (9) (2010) 1610–1629.
[28] Y. Chen, S. Mabu, K. Hirasawa, A model of portfolio optimization using time adapting genetic network programming, Comput. Oper. Res. 37 (10) (2010) 1697–1707.
[29] M. Covel, Trend following: how great traders make millions in up or down markets, in: Financial Times Prentice Hall Books, 1st ed., Financial Times Prentice Hall, New Jersey, 2004, pp. 336.
[30] B. Liu, H.A. Abbass, B. McKay, Classification Rule Discovery With Ant Colony Optimization, IEEE Computer Society, 2003.
[31] J. Li, On optimal rule discovery, IEEE Trans. Knowl. Data Eng. 18 (4) (2006) 460–471.

[32] P. Lin, J. Chen, FuzzyTree crossover for multi-valued stock valuation, Inf. Sci. 177 (5) (2007) 1193–1203.
[33] H. Lu, R. Setiono, H. Liu, Effective data mining using neural networks, IEEE Trans. Knowl. Data Eng. 8 (6) (1996) 957–961.
[34] G.H. John, P. Miller, R. Kerber, Stock selection using rule induction, IEEE Expert 11 (5) (1996) 52–58.
[35] M. Lam, Neural network techniques for financial performance prediction: integrating fundamental and technical analysis, Decis. Support Syst. 37 (4) (2004) 567–581.
[36] Wikipedia, Evolutionary Computation, http://en.wikipedia.org/wiki/Evolutionary_computation, 2014-11-20.
[37] D.B. Fogel, Evolutionary Computation: Toward a New Philosophy of Machine Intelligence, 3rd ed., John Wiley & Sons, Hoboken, NJ, 2006, pp. 384.
[38] M.T. Jones, Artificial Intelligence: A Systems Approach Computer Science Series, 1st ed., Jones & Bartlett Learning, Sudbury, MA, 2008, pp. 498.
[39] T. Back, D.B. Fogel, Z. Michalewicz, Handbook of Evolutionary Computation, 1st ed., IOP Publishing Ltd., Bristol, UK, 1997, pp. 988.
[40] D.B. Fogel, What is evolutionary computation? IEEE Spectr. 37 (2) (2000) 26–28.
[41] A.P. Engelbrecht, Fundamentals of Computational Swarm Intelligence, 1st ed., John Wiley & Sons, New York, 2005, pp. 672.
[42] J.F. Kennedy, J. Kennedy, R.C. Eberhart, Swarm Intelligence, 1st ed., Morgan Kaufmann, San Francisco, 2001, pp. 512.
[43] J. Korczak, P. Roger, Stock timing using genetic algorithms, Appl. Stoch. Models Bus. Ind. 18 (2) (2002) 121–134.
[44] L. Nunez-Letamendia, Fitting the control parameters of a genetic algorithm: an application to technical trading systems design, Eur. J. Oper. Res. 179 (3) (2007) 847–868.
[45] K. Mehta, S. Bhattacharyya, Adequacy of training data for evolutionary mining of trading rules, Decis. Support Syst. 37 (4) (2004) 461–474.
[46] A. Ghandar, et al., Computational intelligence for evolving trading rules, IEEE Trans. Evol. Comput. 13 (1) (2009) 71–86.
[47] M. Kaucic, Investment using evolutionary learning methods and technical rules, Eur. J. Oper. Res. 207 (3) (2010) 1717–1727.
[48] L.M. de Menezes, N.Y. Nikolaev, Forecasting with genetically programmed polynomial neural networks, Int. J. Forecast. 22 (2) (2006) 249–265.
[49] Z. Tan, C. Quek, P.Y.K. Cheng, Stock trading with cycles: a financial application of ANFIS and reinforcement learning, Expert Syst. Appl. 38 (5) (2011) 4741–4755.
[50] G.S. Atsalakis, K.P. Valavanis, Surveying stock market forecasting techniques: Part II. Soft computing methods, Expert Syst. Appl. 36 (3) (2009) 5932–5941.
[51] E. Guresen, G. Kayakutlu, T.U. Daim, Using artificial neural network models in stock market index prediction, Expert Syst. Appl. 38 (8) (2011) 10389–10397.
[52] A. Bahrammirzaee, A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems, Neural Comput. Appl. 19 (8) (2010) 1165–1195.
[53] E.F. Fama, Random walks in stock market prices, Financ. Anal. J. 51 (1) (1995) 75–80.
[54] M. Fasanghari, G.A. Montazer, Design and implementation of fuzzy expert system for Tehran Stock Exchange portfolio recommendation, Expert Syst. Appl. 37 (9) (2010) 6138–6147.
[55] W.F. Sharpe, G.J. Alexander, J.V. Bailey, Investments, 6th ed., Prentice Hall, New Jersey, 1999, pp. 962.
[56] T. Chavarnakul, D. Enke, A hybrid stock trading system for intelligent technical analysis-based equivolume charting, Neurocomputing 72 (16–18) (2009) 3517–3528.
[57] B. Graham, D.L.F. Dodd, S. Cottle, Security Analysis, 1st ed., McGraw-Hill, New York, 1934.
[58] B. Graham, The Intelligent Investor, 1st ed., Harper & Brothers, USA, 1949.
[59] CFA Institute, Ethical and professional standards and quantitative methods, in: CFA Curriculum 2013 Level 1, 1st ed., Pearson, London, 2013.
[60] Wikipedia, Fundamental Analysis, 2013.
[61] A. Esfahanipour, S. Mousavi, A genetic programming model to generate risk-adjusted technical trading rules in stock markets, Expert Syst. Appl. 38 (7) (2011) 8438–8445.
[62] R.W. Colby, The Encyclopedia of Technical Market Indicators, 2nd ed., McGraw-Hill, New York, 2002.
[63] Fidelity Mutual Fund, Fidelity Mutual Fund & Investment Management, 2013.
[64] Goldman Sachs, http://www.goldmansachs.com/investor-relations/stock-chart/index.html, 2013-11-20.
[65] NASDAQ, NASDAQ Stock Market – Stock Quotes – Stock Exchange News, 2013.
[66] Wikipedia, Technical Analysis, http://en.wikipedia.org/wiki/Technical_analysis, 2013-10-08.
[67] Yahoo Finance, Yahoo Finance Charts User Guide, 2013.
[68] Market Technicians Association, The MTA Knowledge Base, 2014.
[69] T. Back, U. Hammel, H. Schwefel, Evolutionary computation: comments on the history and current state, IEEE Trans. Evol. Comput. 1 (1) (1997) 3–17.
[70] Z. Michalewicz, M. Schoenauer, Evolutionary algorithms for constrained parameter optimization problems, Evol. Comput. 4 (1) (1996) 1–32.
[71] A.E. Eiben, J.E. Smith, Introduction to evolutionary computing, in: Natural Computing Series, 1st ed., Springer, Berlin, 2003.
[72] A.P. Engelbrecht, Computational Intelligence: An Introduction, 2nd ed., John Wiley & Sons, New York, 2007, pp. 628.
[73] K.A. De Jong, Evolutionary Computation: A Unified Approach, 1st ed., MIT Press, Cambridge, MA, 2006, pp. 256.
[74] T. Bäck, H. Schwefel, An overview of evolutionary algorithms for parameter optimization, Evol. Comput. 1 (1) (1993) 1–23.
[75] Y. Shi, R. Eberhart, A modified particle swarm optimizer, in: Proceedings of the 1998 IEEE World Congress on Computational Intelligence. Evolutionary Computation, 1998.
[76] D. Simon, Evolutionary Optimization Algorithms: Biologically-Inspired and Population-Based Approaches to Computer Intelligence, 1st ed., John Wiley & Sons, Hoboken, NJ, 2013, pp. 772.
[77] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995.
[78] R.C. Eberhart, Y. Shi, Computational Intelligence: Concepts to Implementations, 1st ed., Morgan Kaufmann, Burlington, 2007, pp. 496.
[79] Y. Shi, Particle swarm optimization: developments, applications and resources, in: Proceedings of the 2001 Congress on Evolutionary Computation, 2001.
[80] N. Nedjah, L.D.M. Mourelle, Swarm intelligent systems, in: Studies in Computational Intelligence, 1st ed., Springer, Berlin, Heidelberg, 2006.
[81] Wikipedia, Evolutionary Algorithm http://en.wikipedia.org/wiki/Evolutionary_algorithm, 2013-11-20.
[82] G. Beni, J. Wang, Swarm intelligence in cellular robotic systems, in: P. Dario, G. Sandini, P. Aebischer, P. Dario, G. Sandini, P. Aebischer (Eds.), Robots and Biological Systems: Towards a New Bionics?, Springer, Berlin, Heidelberg, 1993, pp. 703–712.
[83] E. Bonabeau, M. Dorigo, G. Theraulaz, Swarm Intelligence: From Natural to Artificial Systems, 1st ed., Oxford University Press, New York, 1999.
[84] M. Dorigo, et al., Ant Colony Optimization and Swarm Intelligence Lecture Notes in Computer Science, 2006 ed., Springer, New York, 2006, pp. 526.
[85] I. Contreras, et al., Using a GPU-CPU architecture to speed up a GA-based real-time system for trading the stock market, Soft Comput. 16 (2) (2012) 203–215.
[86] M. Tanaka-Yamawaki, S. Tokuoka, Adaptive use of technical indicators for the prediction of intra-day stock prices, Phys. A: Stat. Mech. Appl. 383 (1) (2007) 125–133.
[87] A. Gorgulho, R. Neves, N. Horta, Applying a GA kernel on optimizing technical analysis rules for stock picking and portfolio composition, Expert Syst. Appl. 38 (11) (2011) 14072–14085.
[88] M. Matilla-García, Are trading rules based on genetic algorithms profitable? Appl. Econ. Lett. 13 (2) (2006) 123–126.
[89] S. Papadamou, G. Stephanides, Improving technical trading systems by using a new MATLAB-based genetic algorithm procedure, Math. Comput. Model. 46 (1–2) (2007) 189–197.
[90] Y. Chang Chien, Y. Chen, Mining associative classification rules with stock trading data – a GA-based method, Knowl. Based Syst. 23 (6) (2010) 605–614.
[91] L. Wei, C. Cheng, A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting, Appl. Soft Comput. 13 (2) (2013) 911–920.
[92] Q. Cai, et al., A novel stock forecasting model based on fuzzy time series and genetic algorithm, Proc. Comput. Sci. 18 (2013) 1155–1162.
[93] J.L. Aznarte, et al., Financial time series forecasting with a bio-inspired fuzzy model, Expert Syst. Appl. 39 (16) (2012) 12302–12309.
[94] J. Hung, A fuzzy GARCH model applied to stock market scenario using a genetic algorithm, Expert Syst. Appl. 36 (9) (2009) 11710–11717.
[95] E. Hadavandi, H. Shavandi, A. Ghanbari, Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting, Knowl. Based Syst. 23 (8) (2010) 800–808.
[96] T. Fu, C. Chung, F. Chung, Adopting genetic algorithms for technical analysis and portfolio management, Comput. Math. Appl. 66 (10) (2013) 1743–1757.
[97] A. Canelas, R. Neves, N. Horta, A SAX-GA approach to evolve investment strategies on financial markets based on pattern discovery techniques, Expert Syst. Appl. 40 (5) (2013) 1579–1590.
[98] D. Jose Bodas-Sagi, et al., A parallel evolutionary algorithm for technical market indicators optimization, Nat. Comput. 12 (2013) 195–207.
[99] Y. Chang Chien, Y. Chen, A phenotypic genetic algorithm for inductive logic programming, Expert Syst. Appl. 36 (3, Part 2) (2009) 6935–6944.
[100] S.C. Chiam, K.C. Tan, A. Al Mamun, Investigating technical trading strategy via an multi-objective evolutionary platform, Expert Syst. Appl. 36 (7) (2009) 10408–10423.
[101] M.E.F. Garcia, A. Enrique, R.Q. Garcia, Improving return using risk-return adjustment and incremental training in technical trading rules with GAPs, Appl. Intell. 33 (2) (2010) 93–106.
[102] Y. Chen, et al., A portfolio optimization model using genetic network programming with control nodes, Expert Syst. Appl. 36 (7) (2009) 10735–10745.
[103] S. Mabu, et al., Enhanced decision making mechanism of rule-based genetic network programming for creating stock trading signals, Expert Syst. Appl. 40 (16) (2013) 6311–6320.
[104] Y. Chen, et al., A genetic network programming with learning approach for enhanced stock trading model, Expert Syst. Appl. 36 (10) (2009) 12537–12546.
[105] E. Tsang, P. Yung, J. Li, EDDIE-automation, a decision support tool for financial forecasting, Decis. Support Syst. 37 (4) (2004) 559–565.
[106] J. How, M. Ling, P. Verhoeven, Does size matter? A genetic programming approach to technical trading, Quant. Financ. 10 (2) (2010) 131–140.
[107] C. Fyfe, J.P. Marney, H. Tarbert, Risk adjusted returns from technical trading: a genetic programming approach, Appl. Financ. Econ. 15 (15) (2005) 1073–1077.

[108] J. Potvin, P. Soriano, M. Vallée, Generating trading rules on the stock markets with genetic programming, Comput. Oper. Res. 31 (7) (2004) 1033–1047.
[109] C.J. Neely, Risk-adjusted, ex ante, optimal technical trading rules in equity markets, Int. Rev. Econ. Financ. 12 (1) (2003) 69–87.
[110] C. Hsu, A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming, Expert Syst. Appl. 38 (11) (2011) 14026–14036.
[111] Y. Hsu, A. Chen, J. Chang, An inter-market arbitrage trading system based on extended classifier systems, Expert Syst. Appl. 38 (4) (2011) 3784–3792.
[112] A.C. Briza, P.C. Naval Jr., Stock trading system based on the multi-objective particle swarm optimization of technical indicators on end-of-day market data, Appl. Soft Comput. 11 (1) (2011) 1191–1201.
[113] P. Ładyżyński, P. Grzegorzewski, Particle swarm intelligence tunning of fuzzy geometric protoforms for price patterns recognition and stock trading, Expert Syst. Appl. 40 (7) (2013) 2391–2397.
[114] S. Chen, P. Kao, TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines, Inf. Sci. 247 (2013) 62–71.
[115] C. Liu, C. Yeh, S. Lee, Application of type-2 neuro-fuzzy modeling in stock price prediction, Appl. Soft Comput. 12 (4) (2012) 1348–1358.
[116] J. Straßburg, C. Gonzàlez-Martel, V. Alexandrov, Parallel genetic algorithms for stock market trading rules, Proc. Comput. Sci. 9 (2012) 1306–1313.
[117] F. Fernandez-Rodriguez, C. Gonzalez-Martel, S. Sosvilla-Rivero, Optimization of technical rules by genetic algorithms: evidence from the Madrid stock market, Appl. Financ. Econ. 15 (11) (2005) 773–775.
[118] M.H. Fazel Zarandi, et al., A type-2 fuzzy rule-based expert system model for stock price analysis, Expert Syst. Appl. 36 (1) (2009) 139–154.
[119] W. Leigh, et al., Stock market trading rule discovery using technical charting heuristics, Expert Syst. Appl. 23 (2) (2002) 155–159.
[120] S. Thawornwong, Development and Analysis of Intelligent Computation Based Stock Forecasting and Trading Systems, University of Missouri, Rolla, 2003.
[121] W. Brock, J. Lakonishok, B. LeBaron, Simple technical trading rules and the stochastic properties of stock returns, J. Financ. 47 (5) (1992) 1731–1764.
[122] F. Fernández Rodríguez, S. Sosvilla Rivero, J. Andrada Félix, Technical Analysis in the Madrid Stock Exchange, FEDEA, 1999.
[123] J.H. Holland, Adaptation in Natural and Artificial Systems: An Introductory Analysis With Applications to Biology, Control, and Artificial Intelligence, 1st ed., University of Michigan Press, Oxford, England, 1975, pp. 183.
[124] J.R. Koza, Genetic programming: on the programming of computers by means of natural selection, in: Complex Adaptive Systems, 1st ed., MIT Press, Cambridge, MA, 1992.
[125] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of IEEE International Conference on Neural Networks 1995, Perth, Australia, 1995.
[126] H. Katagiri, K. Hirasama, J. Hu, Genetic network programming – application to intelligent agents, in: 2000 IEEE International Conference on Systems, Man, and Cybernetics, 2000.
[127] Wikipedia, Market Liquidity, 2015.
[128] R. Næs, J.A. Skjeltorp, B.A. Ødegaard, Stock market liquidity and the business cycle, J. Financ. 66 (1) (2011) 139–176.
[129] T. Chordia, R. Roll, A. Subrahmanyam, Market liquidity and trading activity, J. Financ. 56 (2) (2001) 501–530.
[130] A. Hameed, W. Kang, S. Viswanathan, Stock market declines and liquidity, J. Financ. 65 (1) (2010) 257–293.
[131] C. Yang, R. Zhang, Dynamic asset pricing model with heterogeneous sentiments, Econ. Model. 33 (2013) 248–253.
[132] J. Smailović, et al., Stream-based active learning for sentiment analysis in the financial domain, Inf. Sci. 285 (2014) 181–203.
[133] D.D. Wu, Z. Lijuan, D.L. Olson, A decision support approach for online stock forum sentiment analysis, IEEE Trans. Syst. Man Cybern. 44 (8) (2014) 1077–1087.
[134] E. Lefevre, Reminiscences of a stock operator, in: A Marketplace Book, 1st ed., John Wiley & Sons, Hoboken, NJ, 2004.
[135] D. Turner, M. Schroeck, R. Shockley, Analytics: the real-world use of Big Data in financial services, IBM Glob. Bus. Serv. (2013) 27.
[136] M. Scholtus, D. van Dijk, B. Frijns, Speed, algorithmic trading, and market quality around macroeconomic news announcements, J. Bank. Financ. 38 (2014) 89–105.
[137] B.G. Malkiel, The efficient market hypothesis and its critics, J. Econ. Perspect. 17 (1) (2003) 59–82.
[138] C. Pei-Chann, F. Chin-Yuan, A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting, IEEE Trans. Syst. Man Cybern. C: Appl. Rev. 38 (6) (2008) 802–815.
[139] Y. Chen, C. Cheng, A soft-computing based rough sets classifier for classifying IPO returns in the financial markets, Appl. Soft Comput. 12 (1) (2012) 462–475.
[140] J. Cheng, H. Chen, Y. Lin, A hybrid forecast marketing timing model based on probabilistic neural network, rough set and C4.5, Expert Syst. Appl. 37 (3) (2010) 1814–1820.
[141] R. Lowenstein, Buffett: The Making of an American Capitalist, Random House, 2013.
[142] G. Deng, W. Lin, C. Lo, Markowitz-based portfolio selection with cardinality constraints using improved particle swarm optimization, Expert Syst. Appl. 39 (4) (2012) 4558–4566.
[143] H.R. Golmakani, M. Fazel, Constrained portfolio selection using particle swarm optimization, Expert Syst. Appl. 38 (7) (2011) 8327–8335.
[144] H. Zhu, et al., Particle swarm optimization (PSO) for the constrained portfolio optimization problem, Expert Syst. Appl. 38 (8) (2011) 10161–10169.
[145] J. Kennedy, W.M. Spears, Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimodal problem generator, in: Proceedings of IEEE International Conference on Evolutionary Computation, 1998, Alaska, USA, 1998.
[146] R. Hassan, et al., A comparison of particle swarm optimization and the genetic algorithm, in: 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Austin, Texas, 2005, pp. 18–21.
[147] J. Vesterstrom, R. Thomsen, A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems, in: Congress on Evolutionary Computation 2004 (CEC2004), 2004.
[148] E. Elbeltagi, T. Hegazy, D. Grierson, Comparison among five evolutionary-based optimization algorithms, Adv. Eng. Inf. 19 (1) (2005) 43–53.
[149] D. Karaboga, B. Akay, A comparative study of artificial bee colony algorithm, Appl. Math. Comput. 214 (1) (2009) 108–132.
[150] X. Zhang, Y. Hu, K. Xie, et al., An evolutionary trend reversion model for stock trading rule discovery[J], Knowledge-Based Systems 79 (2015) 27–35.
[151] Y. Hu, B. Feng, X. Zhang, et al., Stock trading rule discovery with an evolutionary trend following model[J], Expert Systems with Applications 42 (1) (2015) 212–222.
[152] X. Zhang, Y. Hu, K. Xie, et al., A causal feature selection algorithm for stock prediction modeling[J], Neurocomputing 142 (2014) 48–59.
[153] Y. Hu, K. Liu, X. Zhang, et al., Concept drift mining of portfolio selection factors in stock market[J], Electronic Commerce Research and Applications (2015).
[154] W. Chen, Z. Hao, R. Cai, et al., Multiple-cause discovery combined with structure learning for high-dimensional discrete data and application to stock prediction[J], Soft Computing. (2015) 1–14.