Expert Systems with Applications 39 (2012) 7458–7470

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

How many reference patterns can improve profitability for real-time trading in futures market?

Suk Jun Lee [a], Kyong Joo Oh [a,*], Tae Yoon Kim [b]

[a] Department of Information and Industrial Engineering, Yonsei University, 262, Seongsanno, Seodaemun-ku, Seoul 120-749, South Korea
[b] Department of Statistics, Keimyung University, 1000, Sindang-dong, Dalseo-ku, Daegu 704-701, South Korea

ARTICLE INFO

Keywords:
Rough set
Dynamic time warping
Real-time trading
Reference pattern
Random pattern sampling
Template cardinality

ABSTRACT

Investors in futures market used to employ trading system which depends on reference pattern (template) to detect real-time buy or sell signal from the market. Indeed they prepare in advance a number of reference patterns that market movement might follow, and then match the current market with one of reference patterns. One popular way to prepare templates is to fix a relatively small number of them which represent possible market movements efficiently. The underlying assumption of this approach is of course that the current market movement is close enough to one of the templates. However, there is always a calculated risk that the current market is close to none of them sufficiently. In this article we investigate the issue of appropriate number of templates (or template cardinality *I*) in terms of profitability. We will show that one may improve profitability by increasing *I* and that random pattern sampling plays a key role in such case. An empirical study is done on the Korean futures market.

1. Introduction

Developing a more accurate and more profitable investment strategy for the futures market equipped with real-time data has been a challenging issue among investors and professional analysts (Gay & Jung, 1999; Veld-Merkoulova, 2003; Wang & Chan, 2007). Challenge is that real-time data add versatility as well as unpredictability to the futures market (Dhar, Chou, & Provost, 2000) and thus it is hard to build consistently profitable strategies. Quite a few researchers use artificial intelligence expert systems to develop trading system based on a relatively small set of templates (or reference patterns) characterizing the stock market movements (Chang, Wang, & Yang, 2004; Choudhry & Garg, 2008; Chung, Fu, Ng, & Luk, 2004; Kim & Shin, 2007; Li & Kuo, 2008; Oh & Kim, 2002) but their performances usually turn out to be rather limited when they are applied to the futures market in practice. Major hindrance to their approach seems that frequently their templates go awry or fail to match the current market accurately enough. One obvious solution to this hindrance is to enhance template variety by increasing template cardinality *I* (or the number of reference patterns prepared). This study aims to implement and investigate such idea. In particular we employ random pattern sampling to enhance template variety with an increased *I*. Our study reveals relation of template cardinality and trading system profitability. In fact we will show that one may improve profitability by increasing *I*, but such improvement might become relatively stable for *I* too large.

For our study, we develop a procedure called *expanded real-time rule-based trading system* (*eRRTS*) which expands the set of templates of *RRTS* (*real-time rule-based trading system*) developed by Lee, Ahn, Oh, and Kim (2010). Notice that *RRTS* (and hence *eRRTS*) attaches trading rule to each reference pattern and then activates the corresponding trading rule when it finds a reference pattern close to the current movement. Lee et al. (2010) argues that a relatively small number of templates for *RRTS* is unavoidable since with enormous amount of real-time data it is usually hard to figure out the current pattern from the entire past patterns. A major technical reformation of *RRTS* to *eRRTS* is that it enhances template variety with an increased *I*. The rest of the paper is organized as follows. Section 2 briefly reviews the technical background of this research and Section 3 describes the *eRRTS* construction procedure. In Section 4 empirical studies are discussed to illustrate the *eRRTS* construction procedure and address our proposition. Concluding remarks are presented in Section 5.

2. Technical background

In futures markets, agents purchase contracts to buy a specific amount of a commodity or financial instrument at a specific price for delivery at a specific time in the future. In the stock futures market, stock price movements are commodities and an investor can earn a profit by buying a contract to buy when he expects a bull market or by buying a contract to sell when he expects a bear

* Corresponding author. Tel.: +82 2 2123 5720; fax: +82 2 364 7807.
E-mail address: johanoh@yonsei.ac.kr (K.J. Oh).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2012.01.101

market. Thus the direction of stock price movements (or buy and sell signal) presents an opportunity to generate profits in both bull and bear markets. For detection of buy and sell signal, various types of trend patterns have been categorized and analyzed where trend pattern usually indicates a specific trend of stock price over a period of time (Kim, Lee, Lee, & Chae, 2002). In practice, a small number of trend patterns are usually utilized as templates. For instance, there are five different types of templates characterized by Levy (1971): ascending channel, descending channel, cup and handle, head and shoulders and triangle.

Recently Lee et al. (2010) developed *RRTS* which provides trading rules against stream of real-time data. In *RRTS*, 6 reference patterns are used as templates (i.e., short-term ascending, short-term descending, long-term ascending, long-term descending, flat top and flat bottom) where each of them has its own trading rule(s). Note that the trading rule(s) is designed to issue buy, sell or hold signal when some prescribed market conditions are satisfied. Technically *RRTS* heavily relies on rough set theory for attaching specific rule(s) to each reference pattern since the rough set theory provides an important mathematical method for generating rules from noisy and versatile data such as real-time data (Walczak & Massart, 1999). In fact *RRTS* utilizes the existing various technical oscillators used in financial markets for generating trading rules (refer to Appendix A). For instance, for a given reference pattern it picks out a few technical oscillators according to profitability and uses them for generating trading rules. *RRTS* is implemented by sorting out a proper reference pattern for the current market movement, which is done technically by measuring distance between the 6 reference patterns and the current market movement.

One drawback of this approach is that there is always a calculated risk that the current market is close enough to none of the given templates. To overcome such difficulty, we propose to increase template cardinality I so that templates with an increased I enhance representation of the underlying pattern population. The methodology for achieving this is random sampling of period of length α, i.e., select I periods of length α:

$$L_{RP}(I) = \{L_1, \ldots, L_I\} \quad \text{with} \quad \alpha I \leqslant N \tag{1}$$

from the entire training period:

$$L_{TP} = \cup_{s=1}^{J} L_s, \quad L_s \neq L_t \quad \text{for } s \neq t, \quad \ell(L_s) = \alpha \quad \text{for all } s \quad \text{and} \quad [\alpha J] = N. \tag{2}$$

Here we assume that the length of entire training period is N. L_{RP} forms a set of periods randomly selected from L_{TP} and movement of stock market on each L_i means one reference pattern. This procedure is termed as random pattern sampling here and below.

Besides random pattern sampling, there are two major technical tools for *eRRTS*, i.e., DTW (dynamic time warping) algorithm and rough sets theory. DTW is employed for finding proper reference pattern from I templates against the current market movement while rough sets theory for attaching specific rule to each template. The DTW algorithm, originally introduced by Bellman (1957), is one of many techniques that can be used to measure similarity between two time series (or two patterns) particularly when the two are not aligned properly on time axis (Bagnall, Ratanamahatana, Keogh, Lonardi, & Janacek, 2006; Keogh & Kasetty, 2003; Niennattrakul, Ruengronghirunya, & Ratanamahatana, 2010). It has been used in various fields (e.g., gesture recognition (Gavrila & Davis, 1995), robotics (Schmill, Oates, & Cohen, 1999), speech processing (Rabiner & Juang, 1993) and medicine (Caiani et al. 1998)). To describe DTW briefly, let us assume that two time series are given, $R = (r_1, r_2, \ldots, r_j, \ldots, r_p)$ and $T = (t_1, t_2, \ldots, t_j, \ldots, t_q)$ where p and q might be unequal. To align these two, we build a $p \times q$ matrix where the (ith,jth) element of the matrix is the distance $D(r_i, t_j)$ between the two points r_i and t_j using the Euclidean distance $D(r_i, t_j) = (r_i - t_j)^2$. Here each matrix element (i,j) corresponds to the alignment between the points r_i and t_j. Fig. 1 from Keogh and Pazzani (2001) provides an illustration of this. A warping path W is a contiguous set of matrix elements. The kth element of W is defined as $w_k = (i_k, j_k)$ and hence $W = W(w_1, \ldots, w_k, \ldots, w_K)$. Here several constraints on W are imposed, i.e., (i) $w_1 = (1,1)$and $w_K = (p,q)$, (ii) if $w_k = (i_k, j_k)$ and $w_{k-1} = (i_{k-1}, j_{k-1})$, then $0 \leqslant i_k - i_{k-1} \leqslant 1$ and $0 \leqslant j_k - j_{k-1} \leqslant 1$. These conditions respectively imply that: (i) W must start and finish at diagonally opposite corners of the matrix; (ii) the allowable steps in W are to be contiguous cells (including diagonally contiguous cells) and the points in W to be monotonically spaced in time. Thus, we have $\max(p,q) \leqslant K < p + q - 1$ for $W = w_1, \ldots, w_k, \ldots, w_K$.

There are quite a few number of warping paths that meet the above conditions. However, we are only interested in W^*, which minimizes the so-called warping cost:

$$\mathrm{DTW}(R, T) = \min_W \left(\sqrt{\sum_{k=1}^{K} D(w_k)/K} \right). \tag{3}$$

Note that K is used for scaling since warping paths may have different lengths and that the path W^* can be found very efficiently by using dynamic programming:

$$D_C(i,j) = D(r_i, t_j) + \min\{D_C(i-1, j-1),\ D_C(i-1, j),\ D_C(i, j-1)\}, \tag{4}$$

where $D_C(i,j)$ is the cumulative distance of the adjacent elements. Fig. 2 depicts how to implement DTW via cumulative distance.

In the rough set theory, knowledge about objects is explained in the form of an information table (Pawlak, 1997). The rows and columns of the information table correspond to objects and attributes, respectively, and attribute values fill the cells of the table. The primary concept of a rough set is a collection of rows that has the same value for one or more attributes, which creates the indiscernibility relationship regarding the finite set of objects (referred to as the universe). Any complete set of indiscernible objects is referred to as an elementary set, which forms the basis for the universe. Every subset of the universe can be expressed either precisely or

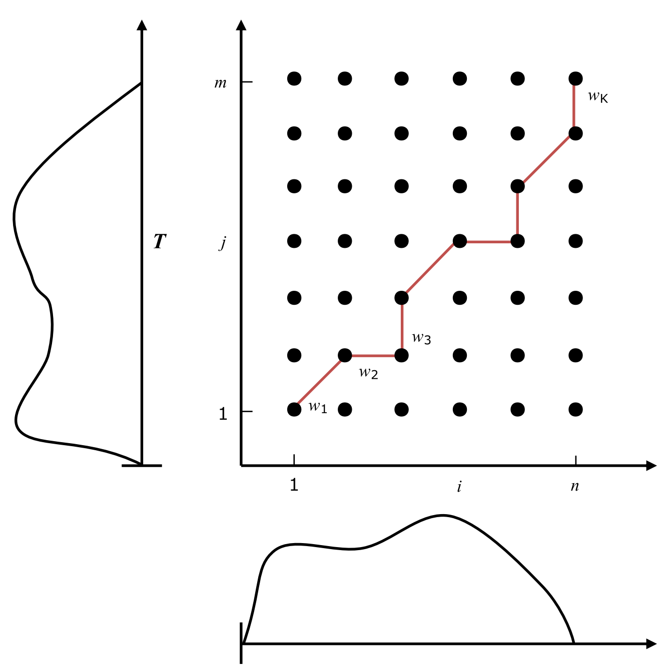

Fig. 1. An example of a warping path (Keogh & Pazzani, 2001).