Physica A 469 (2017) 60–70

Contents lists available at ScienceDirect

Physica A

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

Transfer entropy coefficient: Quantifying level of information flow between financial time series

Yue Teng, Pengjian Shang *

Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing 100044, PR China

HIGHLIGHTS

- A new coefficient is proposed with the objective of quantifying the level of information flow between financial time series.
- We find that transfer entropy coefficient has superiority over transfer entropy.
- We know the measure of transfer entropy at different scales, which is corresponding to the relevant content of financial markets.
- We find that the direction of the information flow between two series may change with the increasing of the transfer entropy coefficient.
- We review that the change of transfer entropy coefficient is very complex.

ARTICLE INFO

Article history:
Received 28 June 2016
Received in revised form 14 October 2016
Available online 16 November 2016

Keywords:
Transfer entropy coefficient
Multiscale
Financial market
Information flow

ABSTRACT

In this paper, a new coefficient is proposed with the objective of quantifying the level of information flow between financial time series. This transfer entropy coefficient, which provides an assessment on the multiscale information flow between measurements, is defined in terms of the transfer entropy method and the multiscale method. The implementation of this transfer entropy coefficient is illustrated with simulated time series and financial time series. Examples taken from simulated and financial data demonstrate that the dynamic mechanism of a complex system cannot be detected solely on the basis of transfer entropy of single scale.

1. Introduction

Time series analysis has become an indispensable part of the financial market research, and it is one of the important methods of financial quantitative analysis [1–4]. Many research results of financial markets are based on time series analysis, and the importance of financial time series analysis methods have been widely recognized in the world. A measure that has been used in a variety of fields, and which is both dynamic and non-symmetric, is transfer entropy, developed by Schreiber et al. [5] and based on the concept of Shannon Entropy, first developed in the information theory by Shannon [6]. Transfer entropy shares some of the desired properties of mutual information but takes the dynamics of information transport into account. Transfer entropy has been widely applied in plenty of fields, such as the study of cellular automata of computer science, the study of the neural cortex of the brain, the study of social networks, statistics, and dynamical systems [7–14]. The assessment of information transfer in the global economic network helps to understand the current environment and the outlook of an economy [15–18]. Considering the applications of transfer entropy to finance [19–23], Marschinski and Kantz [24] analyzed the information flow between the S&P500 index of the New York Stock Exchange (USA) and the DAX

* Corresponding author.
E-mail address: 12271106@bjtu.edu.cn (Y. Teng).

http://dx.doi.org/10.1016/j.physa.2016.11.061

index of the Frankfurt Stock Exchange (Germany) and detected a nonlinear information transfer between both indices at the one minute scale, and introduced a measure called effective transfer entropy, which subtracts some of the effects of noise or of a highly volatile time series from transfer entropy. This concept is now amply used, especially in the study of the cerebral cortex, and is also used in financial markets [25]. Kwon and Yang [26] calculated the transfer entropy between 135 NYSE stock markets and identified the leading companies by the directivity of the information transmission. In a separate paper [20], they analyzed the information flow between 25 stock markets in the world. Their results indicate that the biggest source of information flow is the United States.

As is defined that the transfer entropy is proposed based on information theory, and it measures the information transfer between the two systems based on the past record values and the current observation value. The multiscale transfer entropy method is proposed based on transfer entropy, and it is realization of multiscale for transfer entropy. In this work, we established a model of transfer entropy coefficient and hope to analyze the function of the scale factor τ on transfer entropy of the two series.

Unlike the DCCA cross-correlation coefficient and multiscale cross-sample entropy, multiscale transfer entropy does not increase or decrease all the time along with the increasing of scale factor [27–34], its change is very complex and it is hard to describe the trend. Information from transfer entropy of one scale is limited, which cannot accurately depict the dynamic mechanism of a complex system. Information flow of the system changes with the time scale factor, and it is defective to calculate transfer entropy of single scale as the index of information flow. We must be careful in the interpretation of the direction of the information flow.

The remaining of this paper is organized as follows: Section 2 presents the transfer entropy, effective transfer entropy and the definition of transfer entropy coefficient. Section 3 presents the results with simulated modeling. Section 4 is devoted to provide the detailed results with the empirical study. Finally, it ends with a conclusion.

2. Methodology

2.1. Transfer entropy

Assume that Y is a discrete variable with probability distribution $p(y)$, where y labels the different values (or states) that Y can take. Then the Shannon entropy

$$H_Y(k) = -\sum_i p(y) \times \log_2 p(y) \tag{1}$$

gives the average number of bits needed to optimally encode independent draws from the distribution of Y. In the following statements log indicates the base 2 logarithm and the summation runs over all different values of Y. Shannon's formula is a measure for uncertainty. The more bits are needed to achieve optimal encoding of the process, the higher is its uncertainty. We can get the largest amount of uncertainty if all values of Y are equally likely, i.e., if Y is uniformly distributed and any realization of Y with the same probability can be generated by a random draw. The relationship between uncertainty and information follows from drawing on the Kullback entropy, which can be used to define the excess number of bits needed for encoding when improperly assuming a probability distribution $q(y)$ of Y different from $p(y)$:

$$K_Y = \sum_y p(y) \times \log \frac{p(y)}{q(y)}. \tag{2}$$

When it comes to the bivariate background, let us consider two discrete variables, X and Y, with marginal probability distributions $p_X(x)$ and $p_Y(y)$ and joint probability $p_{XY}(x, y)$. The mutual information of the two processes is given by reducing uncertainty compared to the circumstance where both processes are independent, i.e. where the joint distribution is given by the product of the marginal distributions, $p_{XY}(x, y) = p(x)p(y)$. The corresponding Kullback entropy, known as the formula for mutual information, given by [35]

$$M_{XY} = -\sum_{x,y} p(x, y) \times \log \frac{p(x, y)}{p(x)p(y)}, \tag{3}$$

where the summation runs over the distinct values x and y. Any form of statistical dependencies between different variables can be detected by mutual information. However, it is a symmetry measure and therefore any evidence related to the dynamics of information exchange is not available. Let us concentrate on a time series context. Here, dynamical structure can be introduced when we consider transition probabilities. Let X be a stationary Markov process of order k, then it holds for the probability to observe X at time $t+1$ in state x is conditional on the k previous observations that $p(x_{t+1}|x_t, \ldots, x_{t-k+1}) = p(x_{t+1}|x_t, \ldots, x_{t-k})$. The average number of bits needed to encode one more time series observation if the previous values are known is given by

$$h_X(k) = -\sum_x p(x_{t+1}, x_t^{(k)}) \times \log p(x_{t+1}|x_t^{(k)}), \tag{4}$$