Knowledge-Based Systems 115 (2017) 110–122

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Fuzzy cognitive maps in the modeling of granular time series

Wojciech Froelich [a,*], Witold Pedrycz [b,c,d]

[a] Institute of Computer Science, University of Silesia, ul. Bedzinska 39, Sosnowiec, Poland
[b] Computational Intelligence in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton T6R 2V4 AB Canada
[c] Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, 21589 Saudi Arabia
[d] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

ARTICLE INFO

Article history:
Received 30 June 2016
Revised 16 September 2016
Accepted 15 October 2016
Available online 19 October 2016

Keywords:
Fuzzy cognitive maps
Granular computing
Time series

ABSTRACT

In this study we propose a new approach to granular modeling of time series. In contrast to the existing fuzzy set-based models of time series, we engage information granules in time (granulation resulting in temporal segments). This method subsequently gives rise to information granules formed in the representation space of the series (in particular, the space of amplitude and space of changes of amplitude). Initially the time series is approximated as the sequence of granules forming a so-called granular time series (GTS). To develop a forecasting (prediction) model of the GTS, we cluster all information granules and regard the centers of the clusters obtained through fuzzy clustering as the concepts of the fuzzy cognitive map (FCM). We propose a matching mechanism to carry out description of the GTS and form the results as a vector of the concepts' activations. In this way the GTS is represented as the sequence of vectors of the concepts' activations, which is forecasted by the FCM. At the conceptual level, the forecasted granule is the FCM concept associated with the maximal degree of activation. At the numeric level, the predicted granule regarded as a fuzzy set is described in terms of its bounds and modal value. Experimental studies involving publicly available real-world data demonstrate the usefulness and satisfactory efficiency of the proposed approach.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Modeling time series is an important and challenging problem that has long been addressed by many researchers. Less well-known are the approaches that rely on the approximation of time series aiming to construct their higher-level representation.

Let us first observe that in several application domains the information on accurate numerical values of time series is less important; only knowledge of their approximation is required. For example, in the case of meteorological data analyzed in daily time scale (24-h intervals), it is important to predict minimal and maximal daily temperatures [1]. Also, the expected range of precipitation for every day is important because it influences water demand, which in turn affects the management of local water resources [2,3]. Apart from the ranges in which the daily temperatures and precipitation fall, it is beneficial to predict at least an approximated distribution of the amplitude or changes in amplitude of time series for the next day. Another example comes from the stock market. Information about the possible change of stock prices in the forthcoming period substantially influences the decisions of investors. They are usually not interested in the short-term random component of the time series, but rather in the range into which the stock prices are expected to fall in the following daily, monthly or even longer time periods. Also in this case, information on the distribution of data within the specified intervals is beneficial to investors. In both examples mentioned above, instead of forecasting numerical time series, there is an interest in modeling and realizing forecasting at the level of symbols or information granules.

When looking at the stated problem from a general perspective, we encounter the issues of specificity and generality. On the one hand, we have to deal with crisp, real-valued observations that are usually hard to predict. On the other hand, we generalize source data, making the representation less accurate but in turn increasing the likelihood of a good performance due to the operation at a higher level of abstraction.

The performance measures used for the numerical and granular forecasts are not directly comparable. However, when creating more abstract representations of time series, we attempt to get rid of the included random component that, apart from its statistical properties, is hard to predict. Moreover, we approximate time series, taking into account a specific criterion (which in our case is

* Corresponding author.
E-mail addresses: wojciech.froelich@us.edu.pl (W. Froelich), wpedrycz@ualberta.ca (W. Pedrycz).

http://dx.doi.org/10.1016/j.knosys.2016.10.017
0950-7051/© 2016 Elsevier B.V. All rights reserved.

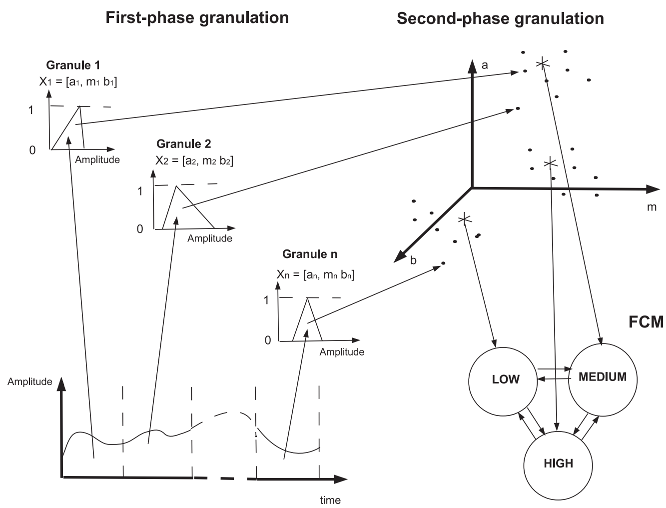

Fig. 1. General idea of the proposed approach.

the compromise between the generality and specificity of the representation). In this way we deal with the sequence of granules approximating numerical time series. The sequence of these granules is forecasted. The perfect forecast of this sequence is achieved when the actual and forecasted granules are equal. The performance of the forecasting is measured in terms of the overlap between the actual and forecasted granules. The more overlap recognized between these granules, the better the obtained forecast. For the measurement of the overlap between the granules, we propose a new function (to be explained later) expressed by formulas (28) and (29) for the amplitude and the change of amplitude or time series, respectively.

One of the approaches used in the approximation of raw data is information granulation. For example, concept learning from raw data via multi-granularity has been used in [4]. By using granulation, it is possible to form a higher-abstraction model representing time series [5]. The outcome of the granulation is the granular time series (GTS), which is subject to forecasting. The ease of human interpretation of the GTS is one of the benefits of granulation. For the forecasting of the GTS, a specialized granular model is required. This model is used to predict the sequence of granules instead of the numerical time series.

In Section 2 we provide a review of previous works related to the modeling and forecasting of granular time series. We also stress in that section the originality of our study with respect to the previous studies.

There are two main objectives of this study:

1. Approximation of numerical time series, aiming to represent them in the form of the GTS.
2. Construction of an effective model for the forecasting of the previously formed GTS.

The general idea behind the proposed approach is illustrated in Fig. 1.

To perform the approximation of time series, we partition the time domain of the series into intervals of the same length. For every interval, we construct an information granule in the form of a triangular fuzzy number. These information granules are developed with the use of the principle of justifiable granularity (PJG). The sequence of the granules is the granular time series. We refer to this process as the first phase of information granulation.

The PJG is a general method for the construction of data representation in the form of entities called granules. In comparison to clustering techniques, the PJG does not explicitly consider the distance between data points. Instead, the PJG optimizes the parameters of the granules, finding a trade-off between the two criteria of generality and specificity. For these contradicting criteria, the user specifies the related functions that depend on the dimensionality of data and the planned application. The goal of the optimization performed by the PJG is to form the most appropriate (justified) representation of data in terms of the two provided functions.

When it comes to the second objective, to form the forecasting model of the GTS, we cluster all previously obtained granules by regarding them as entities described in a three-dimensional space, namely by the use of their bounds (a,b) and modal values (m). The representatives of the clusters become the concepts used in the fuzzy cognitive map. We call this process the second-phase granulation, and the concepts of the FCM the second-phase granules. In addition, to interpret the GTS at a conceptual level, we order the second-phase granules with respect to their modal values and assign them linguistic terms, e.g., 'LOW', 'MEDIUM', 'HIGH' (as presented in Fig. 1).

To perform the forecasting of the GTS, we need to discover and exploit temporal relationships between the second-phase granules. To accomplish this task, we decided to apply the FCM model. FCMs are a specific type of artificial neural network (ANN). However, in the case of FCMs there is no need to specify input and output nodes, as there usually is for most ANNs. All nodes of the FCM play both roles. After initiating the states of the FCM's concepts, the reasoning formula is used enabling to forecast the future states of all concepts. In addition, contrary to most applications of ANNs, FCMs do not contain hidden nodes. All FCMs concepts are explicitly related to data. Thanks to this, the laborious task of selecting the number of hidden nodes and hidden layers of the ANN is no longer necessary. From a practical point of view, recent studies revealed excellent performances from FCMs for the time series forecasting task [6–9].

To form the forecasting model of the GTS, we assume the second-phase granules as the concepts of the FCM. The concepts play a role similar to that of the regressors in auto-regressive forecasting models. The connections (directed arcs) between the concepts are interpreted as the temporal dependencies recognized between the regressors. Note that the concepts are ordered according to the linguistic terms assigned to them – this means that the arc between two concepts specifies the dependency between the linguistic values of the time series observed at consecutive time stamps. In addition, every of the discovered arc is labeled by the weight, which is the strength of the considered relationship. The strengths of the relationships between concepts are crucial for the effectiveness of forecasting demonstrated by the FCM; therefore, we perform genetic optimization to adjust them.

To make the proposed model functional, the relationships of every first-phase granule to all second-phase granules (the concepts of the FCM) are described by the vector of numerical values. To calculate the elements of this vector, we propose a specialized function evaluating the degree of matching between granules. Consequently, for the entire GTS, we obtain the sequence of activation vectors of the FCM concepts. The FCM is then applied to the forecasting of the activation vectors. The result of this forecasting is twofold. At the conceptual level, it is the concept with the highest predicted activation grade. This concept represents the cluster in which the predicted granule is expected to fall. The predicted state of the FCM is degranulated to come up with the numeric results. As the second result of prediction, we obtain the first-phase granule that predicts the GTS in the following time interval.

The key contributions of this study can be summarized as follows:

- We propose a new approach to granular modeling of time series. This approach is based on two phases of information gran-

Download English Version:

https://daneshyari.com/en/article/4946494

Download Persian Version:

https://daneshyari.com/article/4946494

Daneshyari.com