Expert Systems with Applications 42 (2015) 4830–4839

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Description and prediction of time series: A general framework of Granular Computing

Rami Al-Hmouz [a,*], Witold Pedrycz [a,b,c], Abdullah Balamash [a]

[a] Electrical and Computer Engineering Department, King Abdulaziz University, Jeddah, Saudi Arabia
[b] Department of Electrical & Computer Engineering, University of Alberta, Edmonton, Alberta T6R 2V4, Canada
[c] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

ARTICLE INFO

Article history:
Available online 3 February 2015

Keywords:
Time series prediction and description
Granular Computing
Information granules

ABSTRACT

In this paper, we address problems of description and prediction of time series by developing architectures of granular time series. Granular time series are models of time series formed at the level of information granules expressed in the representation space and time. With regard to temporal granularity, time series is split into temporal windows leading in this way to the formation of temporal information granules. Information granules are also quantified and constructed over the space of amplitude and change of amplitude of the series collected over time windows. In the description of time series we involve clustering techniques and build information granules in the representation space (viz. the space of amplitude and change of amplitude) of the temporal data. Fuzzy relations forming the essence of the prediction model are optimized using particle swarm optimization. Experimental results are reported for a number of publicly available time series.

1. Introduction

Time series models have been devised by researchers to develop a well construct of accuracy and interpretability. A great deal of work has been explored in the aspect of prediction to achieve high accuracy. The representation of time series is still an ongoing challenge. In this regard, granular time series plays an important role of interpretation and representation of time series.

There have been a great number of studies on developing models of time series along with their numerous usages to prediction and control problems (Abonyi, Feil, Nemeth, & Arva, 2005; Aznarte & Benitez, 2010). We encounter a remarkable diversity of models of time series starting from some "standard" linear models and proceeding with a spectrum of advanced nonlinear models such as neural networks being augmented with significant approximation capabilities, fuzzy rule-based models (Chang & Liu, 2008), time-variant fuzzy time series (Liu, 2007; Song & Chissom, 1993), time invariant fuzzy time series (Hwang, Chen, & Lee, 1998; Liu, Wei, & Yang, 2009) and hybrid neuro-fuzzy architectures (Liu, Yeh, & Lee, 2012) quite commonly supported by design tools of evolutionary optimization (Cheng, Chen, Teoh, & Chiang, 2008). While the accuracy of these models has been the dominant development criterion in their design, with the ever-visible trend of designing user-centric models of data (the leitmotiv being so evidently visible in data mining), it becomes beneficial to construct models of time series that address this need and could be easily understood and interpreted. It is, however, apparent that humans perceive and organize knowledge at a higher level of abstraction than the one being supported by numeric models. To establish a coherent platform where time series models can be established at the higher level of generality, we resort ourselves to information granules and Granular Computing.

An abstract view of detailed numeric data (such as e.g., those encountered in time series) is realized with the aid of general entities – information granules. Information granules can be formalized and described as constructs of Granular Computing (Pedrycz, 2013; Zadeh, 1997, 2011) such as e.g., intervals, fuzzy sets, probabilities, rough sets and alike. Processing information granules and building models at their level give rise to a variety of classifiers (referred to as granular classifiers), predictors (granular predictors) and others. Some ideas of granular time series were discussed in the past, cf. Lu, Pedrycz, Liu, Yang, and Li (2014), Al-Hmouz, Pedrycz, Balamash, and Morfeq (2014). Notwithstanding the diversity of models studied there, we can look at the notion of information granularity in the context of time series form two points of view: (i) granulation of time, and (ii) granulation of representation space in which time series is described. With regard to the first point of view, one builds a collection of temporal windows so in

* Corresponding author.
E-mail addresses: ralhmouz@kau.edu.sa (R. Al-Hmouz), wpedrycz@ualberta.ca (W. Pedrycz), asbalamesh@kau.edu.sa (A. Balamash).

http://dx.doi.org/10.1016/j.eswa.2015.01.060
0957-4174/

this way we encounter a temporal granulation of the series. As far as the second facet of the above general taxonomy is concerned, we form information granules viewed as descriptors of the time series. Interestingly enough, in most of the existing studies, information granules are built in the representation space whereas the temporal space (time) is left in its numeric format. This may sound a bit counterintuitive. When interpreting time series we do not resort ourselves to individual discrete time moments but rather confine to some time intervals (temporal information granules).

In this study, time series is considered as a finite sequence of temporal data and denoted as $x_1, x_2, x_k, \ldots, x_N$ with N being the length of the sequence. Subsequently, the successive changes in values of the time series are denoted by $\partial x_1, \partial x_2, \ldots, \partial x_{N-1}$ where the determination of these changes could be associated with some eventual smoothing process. Furthermore, the resulting space of amplitude in which the time series is reported is denoted by $\mathbf{X}$ while the space of successive changes of the amplitudes with the values $\partial x_1, \partial x_2, \ldots, \partial x_{N-1}$ is denoted by $\partial\mathbf{X}$.

A time series comes in numeric nature, forming granular constructs of numeric time series can offer a meaningful interpretation of temporal relationships. One can look at time series as time windows of days, months and year, another can quantify time series over the amplitude such as low, medium and high amplitude. Information granules can also be formed say in the duration of a month no change of amplitude.

The motivation of this study is to form linguistic descriptors of time series over time and space with the aid of granular time series that give rise of better interpretation, representation and visualization of time series. Various aspects of building information granules are studied (especially realized with the aid of fuzzy clustering) and then associations among the granules are optimized. The associations are structured into a form of relational predictive models of time series. The parameters of these models, viz. fuzzy relations are optimized to minimize a certain performance index (again expressed at the level of information granules).

It is worth stressing that in spite of interesting research covered in the existing literature (Chen & Tanuwijaya, 2011; Karnik & Mandel, 1999; Lu et al., 2014), a comprehensive study on the conceptual information granules-based time series as well as a comprehensive design setting including a formation of information granules has not been reported so far; this is one of the factors clearly determining the originality of the undertaken study.

The paper is structured as follows. We start with a granular description of time series realized in time domain (Section 2) and then discuss a granular description of time series in Section 3. Models of granular time series are introduced in Section 4. Experimental studies are covered in Section 5 whereas conclusions are provided in Section 6.

In the paper, experiments are reported throughout consecutive sections so once a certain approach or algorithmic steps have been elaborated on, the ensuing functioning becomes illustrated.

2. Granular description of time series in time domain

In general, time series can be described in various and quite diversified ways. The three commonly encountered categories of the models concern architectures built in the space of the time frequency, and those of hybrid nature. Here we consider a description of the series realized in the time domain. An intuitive description is formed by taking into account amplitudes of the series (temporal samples) as well as their changes reported in successive discrete time moments. Granularity of information is sought as an essential vehicle to capture the essence of the data, make the models more compact and transparent as well as interpretable. Following the two-way taxonomy, we discuss information granulation of time series (realized by means of information granules built on the space of amplitude and changes of amplitude) and then look at the granulation of time.

2.1. Information granulation of time series

Information granules utilized in the description of time series are formed with regard to the space in which the series are characterized as well as the time variable itself. Let us look into the essence of these constructs.

2.1.1. Formation of linguistic (granular) descriptors in the space of amplitude and changes of amplitude

A simple, intuitively appealing, and algorithmically convincing way to form information granules is completed through clustering, especially fuzzy clustering (Bezdek, 1981; Pedrycz & Gomide, 2007). Clusters are sought as a synonym of information granules. The original data are clustered and the results of clustering (usually prototypes and partition matrices) are used to characterize information granules.

Two main approaches in which fuzzy clustering is exploited, are considered here:

(i) Clustering is realized for one-dimensional data, viz. a collection of amplitudes of the series $\{x_k\}$ and the changes reported for this series $\{\partial x_k\}$. These changes can be obtained after applying some smoothing procedure. The smoothing mechanism is a commonly used k-point moving average as reported in the literature (Contreras, Espínola, Nogales, & Conejo, 2003; Raudys, Lenciauskas, & Malcius, 2013) and taking on the following form

$$\partial x_l = \frac{\sum_{i=1}^{k} \partial x_{l-(k-i-1)}}{k} \tag{1}$$

where k is set to 3.
The number of clusters (linguistic labels, information granules) considered in the clustering procedure done for these one-dimensional data are set independently and are equal to r and p for the space of amplitude and change of amplitude, respectively. The clusters come with well-formed semantics and as such could be conveniently labeled as positive *high* amplitude, negative *low* amplitude, positive *small* changes of amplitude, etc. Note that the linguistic labels in the space of amplitude $\boldsymbol{X}$ and the space of changes of amplitude ∂X are formed independently from each other.

(ii) Formation of linguistic terms in the two-dimensional space of data $\boldsymbol{X} \times \partial\boldsymbol{X}$. Here assuming a certain number of information granules set to c, the clustering is carried out for the two-dimensional data vectors $[\boldsymbol{x}, \partial x]$

The choice of one of the approaches has an impact on further processing. In the first approach, as the linguistic labels are independent, a Cartesian product of the data $\boldsymbol{X} \times \partial\boldsymbol{X}$ in which a composite description is sought involves rp pairs of linguistic descriptors, say *positive low* amplitude and negative *small* change of amplitude. In other words, the vocabulary of linguistic terms that is used in further interpretation or processing of time series involves altogether $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_r$ and $\partial\boldsymbol{A}_1, \partial\boldsymbol{A}_2, \ldots, \partial\boldsymbol{A}_p$

linguistic terms (fuzzy sets). Recall that A_i and ∂A_j are the linguistic terms, say fuzzy sets (with their membership functions) generated as a result of running the method of fuzzy clustering.

$$\boldsymbol{A}_i(\boldsymbol{x}) = \frac{1}{\sum_{j=1}^{c} \left(\frac{\|\boldsymbol{x}-\boldsymbol{v}_i\|}{\|\boldsymbol{x}-\boldsymbol{v}_j\|} \right)^{2/(m-1)}} \tag{2}$$