Applied Soft Computing 47 (2016) 644–652

Contents lists available at ScienceDirect

Applied Soft Computing

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

Bayesian analysis of time series using granular computing approach

Olgierd Hryniewicz, Katarzyna Kaczmarek*

Systems Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

ARTICLE INFO

Article history:
Received 8 January 2014
Received in revised form
19 November 2014
Accepted 22 November 2014
Available online 29 November 2014

Keywords:
Time series forecasting
Granular computing
Soft computing
Data mining
Bayesian methods
Linguistic summaries

ABSTRACT

The soft computing methods, especially data mining, usually enable to describe large datasets in a human-consistent way with the use of some generic and conceptually meaningful information entities like information granules. However, such information granules may be applied not only for the descriptive purposes, but also for prediction. We review the main developments and challenges of the application of the soft computing methods in the time series analysis and forecasting, and we provide a conceptual framework for the Bayesian time series forecasting using the granular computing approach. Within the proposed approach, the information granules are successfully incorporated into the Bayesian posterior simulation process. The approach is evaluated with a set of experiments on the artificial and benchmark real-life time series datasets.

1. Introduction

The Bayesian methods for the time series analysis are proven successful in many practical applications, e.g., [1–3]. Nonetheless, the Bayesian time series analysis and probability theory are not meant to process directly the information described in natural language, and according to Zadeh [4], people granulate information and operate mainly on words and propositions easy to express in natural language. Hopefully, the soft computing methods, especially data mining, enable to retrieve and process the conceptually meaningful human-consistent information from large datasets. Such information entities have given rise to the general framework of Granular Computing [5,6].

Although, time series data mining with the use of the soft computing tools has gained a lot of attention in the literature, the interdisciplinary combination of soft computing and Bayesian approach does not seem to be extensively investigated. As presented in Table 1, according to the keyword and abstract search, there are over 2.7k articles on '*Bayesian' AND 'time series*', and over 2.3k results about '*granular computing*'. However, only 1 result combing these keywords. Furthermore, the Scopus search engine returns only 6 positions for the keywords *'soft computing' AND 'time series' AND 'Bayesian'*.

The objective of this paper is to review the related work on the application of the soft computing methods for the time series analysis, and to propose the conceptual *Bayesian Granular Computing (B-GC) framework for Time Series Forecasting*. The goal is to incorporate the human-consistent soft computing methods, especially data mining and classification techniques, in the construction of the prior model probability distributions.

The proposed approach assumes employing techniques from the following research fields: the fuzzy sets theory, the knowledge discovery from sequential data, especially the time series abstraction and mining for linguistic summaries, Support Vector Machines for the classification problem and the Markov Chain Monte Carlo methods for the posterior simulation and the Bayesian inference. The approach is in line with the Generalized Theory of Uncertainty [7].

We provide an application example as a demonstration of concept and the numerical results of forecasting real-life data. It is observed that the granular computing approach may successfully support the probabilistic time series analysis and forecasting. In the opinion of the authors of this paper, the intelligent combination of the granular computing methods and the Bayesian modeling is a promising direction for the future research on the time series analysis and forecasting.

The structure of this paper is as follows. Next section summarizes state-of-the-art of the soft computing techniques supporting the time series analysis and forecasting. Section 3 recalls basic definitions related to the time series analysis and granular computing. Section 4 presents the proposed *Bayesian Granular Computing*

* Corresponding author. Tel.: +48 603692082.
E-mail address: K.Kaczmarek@ibspan.waw.pl (K. Kaczmarek).

http://dx.doi.org/10.1016/j.asoc.2014.11.024
1568-4946/

Table 1
Results from document search according to keywords, title and abstract in the Scopus database of research (http://www.scopus.com).

Search field by article title, keywords, abstract	Results
'granular computing' AND 'Bayesian" AND 'time series'	1
'soft computing' AND 'Bayesian" AND 'time series'	6
'granular computing' AND 'time series'	40
'soft computing' AND 'time series'	239
'Bayesian' AND 'time series'	2704
'granular computing'	2303
'soft computing'	9633
'time series'	127,933

(B-GC) method for Time Series Forecasting. The description of the experiments is gathered in Section 5. This paper concludes with discussion in Section 6.

2. State-of-the-art

Within this review, the focus lays on the approaches processing information that can be easily interpreted in natural language. First, we review the main developments and objectives of the time series data mining. Secondly, methods for the human-consistent time series forecasting are reviewed. Finally, we shortly discuss the Bayesian approach to the time series analysis.

2.1. Time series data mining

The Data Mining and Knowledge Discovery adapt the soft computing techniques to provide the human-consistent description of large datasets, that can be easily understood by users. Following Zadeh [8], the Soft Computing is based on the Computational Intelligence, and together with the Hard Computing (that is based on the Artificial Intelligence) form the Machine Intelligence. The field of Computational Intelligence was formally initiated in 1994 during the IEEE World Congress on Computational Intelligence in Orlando. Computational Intelligence (CI) is defined as *a methodology involving computing that exhibits an ability to learn and/or deal with new situations such that the system is perceived to possess one or more attributes of reason, such as generalization, discovery, association, and abstraction* cf. [9].

Data mining usually leads to some generic and conceptually meaningful information entities like **information granules**. The theory of information granulation [5] is inspired by the ways in which people granulate information and reason about it. The general framework of Granular Computing [6,10–12] studies the information granulation and the information granules. It endows the information processing with a facet of human-centricity.

As defined by Zadeh [5], the information granule is a clump of points (objects) grouped together by indistinguishability, similarity, proximity or functionality. The information granules are formally described with the use of fuzzy sets, intervals, rough sets, shadowed sets, probabilistic sets, etc. Pedrycz and contributors [13,14] study the optimal granular representation of time series. In [15], the authors review the granular computing developments in general. In [16], the fast interval predictors for large-scale, nonlinear time series with noisy data using fuzzy granular support vector machines are presented.

The knowledge discovery process from sequential data may be divided into segmentation, clustering, classification of identified meaningful intervals or patterns, detecting anomalies, frequent patterns and discovery of association rules. For the review of the time series data mining methods, see e.g., Fu [17]. The comparative analysis of the time series representations and the similarity measures is provided in [18]. Batyrshin et al. [19] introduce an interesting perception-based approach to time series data mining.

Linguistic summaries in the sense of Yager [20] are an example of information granules. Linguistic summaries describe general facts about evolution of time series with (quasi) natural language e.g., *Among all increasing segments, majority are long*, and are intuitive and easily interpretable for people. Kacprzyk et al. [21–23] employ the classic calculus of linguistically quantified propositions for the development of the linguistic summaries.

Another example of the information granules are **linguistic descriptions** based on the Computational Theory of Perceptions [4] and the Systemic Functional Linguistics by Halliday. The linguistic descriptions [24–28] are characterized by more complex semantic and lexico-grammar structures than linguistic summaries, e.g., *Before the knee lesion, the gait quality is high because the gait symmetry is medium and the gait homogeneity is high* cf. [26]. In [29], the authors further develop the concept of the Granular Linguistic Model of a Phenomenon and provide basic architecture of a computational system to generate the linguistic descriptions.

Apart from the linguistic summaries and descriptions, another example of imprecise information entities mined from the time series datasets are **temporal patterns**. Höppner et al. [30] propose the approach for the identification of frequent patterns based on the Allen's temporal logic [31]. In [32], the authors focus on enhancing patterns with a context information and operating on block constraints instead of Allen's relations e.g., *If A happens before B and in the meantime we do not observe C, then we have a failure of class X*. Schockaert et al. [34] fuzzyfy Allen's temporal interval relations and propose a complete framework to represent, calculate and reason about the temporal relationships for fuzzy intervals.

Temporal patterns are used to build the **association rules**, that may be exemplified by *If exchange rate for currency X decreases slowly, the number of international travels will increase rapidly*. In [33], Agrawal et al. propose an interesting algorithm for discovery of rules in sets of items. Methods for the frequent patterns and association rules recognition have been successfully applied to describe huge datasets in different contexts, e.g., [35–39].

The primer goal of the data mining in most of the cited papers is the interpretation of huge datasets by generation of meaningful information granules. The retrieval of such information helps to increase the understanding of the dataset and may reduce the storage space. In the opinion of the authors of this paper, the generic methodology on how to benefit in statistical forecasting from the results of the time series data mining shall be further investigated in the literature.

2.2. Time series forecasting

Traditional autoregressive and moving average (ARMA) processes [40] are the most popular simple probabilistic models for forecasting, and though, very successful in applications. In the recent application [41], D'urso et al. apply the autoregressive estimates of models for the time series classification and clustering. One of the attractive features of the Box–Jenkins approach to forecasting is that the class of its processes is rich, and it is usually possible to find a process or a combination of processes, which provide an adequate description to the dataset. Nonetheless, the identification of the mathematical model and the specification of its required parameters may be time consuming and difficult to interpret for practitioners (experts of the field) involved in the forecasting process.

To alleviate this problem and increase the overall understanding of the probabilistic time series analysis, there have been proposed some **hybrid systems** supporting the model selection like e.g., [42,43]. In [44], the autoregressive and movinag average models are combined with the fuzzy rule based approach using Mamdani inference mixed with some techniques of counting in fuzzy sets. Chen et al. [45] propose an approach for the automatic generation