Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 285 (2016) 52–72

FUZZY sets and systems

www.elsevier.com/locate/fss

Linguistic characterization of time series ☆

Vilém Novák

University of Ostrava, Institute for Research and Applications of Fuzzy Modeling, NSC IT4Innovations, 30. dubna 22, 701 03 Ostrava 1, Czech Republic

Received 15 September 2014; received in revised form 13 June 2015; accepted 17 July 2015

Available online 11 August 2015

Abstract

The goal of this paper is to provide an overview of applications of special soft computing theories — the fuzzy transform and fuzzy natural logic — to analysis, forecasting and mining information from time series. The focus is especially placed on the ability of methods of fuzzy natural logic to provide information in sentences of natural language. Our approach is based on the decomposition of time series into three components: the trend-cycle, seasonal component and noise. The trend-cycle is extracted using the F-transform, and its course is characterized by automatically generated linguistic description. The latter is then used to forecast the trend-cycle. The trend-cycle can be, furthermore, decomposed into trend (general tendency) and cycle. The former is computed again using the F-transform. Moreover, the F^1-transform makes it possible to estimate the direction of the trend, which can then be characterized by expressions of natural language (stagnating, slightly increasing, sharply decreasing, etc.). Finally, we focus on selected problems of mining information from time series. First, we suggest an algorithm for finding intervals of monotonous behavior and then show how the theory of intermediate quantifiers (a constituent of fuzzy natural logic) and generalized Aristotle's syllogisms can be applied to automatic summarization of information on time series.

Keywords: Time series decomposition; F-transform; Fuzzy natural logic; Mining information from time series; Forecasting; Intermediate quantifiers

1. Introduction

The goal of this paper is twofold: first, to put together results scattered over several papers ([34,35,37] and elsewhere) on applications of *fuzzy transform* (*F-transform*) and *fuzzy natural logic* (FNL) for analysis, forecasting and linguistic characterization of time series. Second, it brings new results in the area of mining information from time series with the aim of providing the obtained information in sentences of natural language.

Our approach is based on the assumption that the time series can be decomposed into three components: the trend-cycle, seasonal component and noise. A detailed analysis, moreover, reveals that the trend-cycle can be fur-

☆ The paper has been supported by the projects KONTAKT II, LH 12229 "Research and development of methods and means of intelligent analysis of time series for the strategic planing problems" and partially also by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

E-mail address: Vilem.Novak@osu.cz.

http://dx.doi.org/10.1016/j.fss.2015.07.017

0165-0114/

ther decomposed into trend and cycle. The F-transform is applied to extraction of the trend-cycle and can be also applied to extraction of the trend itself. It has, however, more potential. Namely, it is possible to estimate the direction of the trend in an arbitrary time interval (i.e., whether it is increasing, decreasing or stagnating) using the F^1-transform, which provides estimation of an average value of a first derivative of a given function over a specified domain (see [17]). This can be very useful because the direction of the trend of the time series may not be clear even when viewing the graph. It should be emphasized that the methods described in this paper are theoretically well justified (see [33,38,39]).

The fuzzy natural logic is applied in the following tasks:

(a) Generation of linguistic description of the trend-cycle and computation of its forecast. It was demonstrated in [35] that the obtained precision of our forecasting technique is fully comparable with precision of the top professional systems such as ForecastPro®. In addition, our system also provides linguistic comments explaining how forecast of the trend-cycle was obtained. For example, when examining the forecast of a time series summarizing the number of accidental deaths, we are able to generate conditional sentences such as the following:

If the average number of accidental deaths in the recent 2 years is more or less low and the average decrease in the number of deaths in the same period is extremely big, the average increase in deaths is rather medium.

(b) Generation of linguistic characterization of the direction of the trend (obtained using the F^1-transform). Typical comments are the following:

In the months 1–5, the trend is slightly decreasing.

(c) Segmentation of time series into subintervals with monotonous behavior. A typical result is as follows:

In months 1–3, the trend is more or less decreasing, followed by 5 months of stagnation. In the final months 9–12, the trend is sharply increasing.

(d) Summarization of information on time series using formal theory of intermediate quantifiers (linguistic expressions such as *many, most, almost all, few*). We can summarize information either about one time series or about a set of them (typically several tens or hundreds).

In most *periods is* slightly decreasing *trend followed by* clearly increasing *one.*

Note that this is one of the typical tasks when mining information from time series (cf. [8]). We also propose to apply the theory of generalized Aristotle's syllogisms, which makes it possible to deduce new information based on information that has already been found before.

The paper is structured as follows. In the next section, we provide an overview of basic concepts of the F-transform and fuzzy natural logic that will be used later. Section 3 contains revamped presentation of our methods for forecasting and the trend evaluation of time series introduced earlier. Section 4 contains new results and is devoted to selected problems of mining information from time series. First, we suggest algorithm for finding intervals of monotonous behavior and in the second part, we show how the theory of intermediate quantifiers (constituent of fuzzy natural logic) and generalized Aristotle's syllogisms can be applied to automatic summarization of information on time series. Section 5 is the conclusion, in which we outline directions for further research.

2. Preliminaries: soft-computing techniques for time-series processing

In this section, we will briefly review two main soft-computing techniques that have been successfully applied to time series analysis and forecasting: the F-transform and special methods of fuzzy natural logic.