Chapter 72
High Order Time Series Forecasting Using Fuzzy Discretization

Mahua Bose
University of Kalyani, India

Kalyani Mali
University of Kalyani, India

ABSTRACT

In recent years, various methods for forecasting fuzzy time series have been presented in different areas, such as stock price, enrollments, weather, production etc. It is observed that in most of the cases, static length of intervals/equal length of interval has been used. Length of the interval has significant role on forecasting accuracy. The objective of this present study is to incorporate the idea of fuzzy discretization into interval creation and examine the effect of positional information of elements within a group or interval to the forecast. This idea outperforms the existing high order forecast methods using fixed interval. Experiments are carried on three datasets including Lahi production data, enrollment data and rainfall data which deal with a lot of uncertainty.

1. INTRODUCTION

Forecasting is one of the challenging tasks in the area of Time Series Analysis. Timely and accurate forecasting has a significant role in decision-making in the area of economy, ecology, climatology, employment sector, agriculture, tourism, transport sector, etc. Advance knowledge about future event not only supports policy planning but also helps in disaster management to reduce the loss of life and property.

Decisions that involve factor of uncertainty of the future, fuzzy time series models have been found one of the most effective methods of forecasting. In last twenty years, some methods have been developed on fuzzy time series to make prediction in many research areas, such as enrollments, weather, stock price, agricultural production, terrorism, tourism.

DOI: 10.4018/978-1-5225-5643-5.ch072

The influence of interval partitioning on forecast results has received a considerable amount of attention in the studies of Fuzzy time series. Discretization is a process of grouping the values of the attributes in intervals in such a way that the discernibility is preserved. In the crisp discretization (Lenarcik and Piasta, 1992; Nguyen and Skowron, 1995; Nguyen and Nguyen, 1998) intervals are formed by the adjacent values of the cuts, within the dynamic ranges of the corresponding attributes. Fuzzy discretization technique proposed by Roy and Pal (2003), the positional information of the samples within an interval is taken into consideration. Following the idea of fuzzy partitioning (Roy and Pal, 2003), interval is defined in this study.

The main motivation behind this proposed work is to investigate the effect of positional information of each element within a static interval on forecast accuracy. In order to prove the superiority of the proposed scheme, "w-step fuzzy predictor" algorithm (Singh, 2009) is applied with the modified intervals defined in this study. In addition to that, intervals created by automatic clustering technique (Chen and Tanuwijaya, 2011) are also used directly to the "w-step fuzzy predictor" algorithm. Performance of the "w-step fuzzy predictor" algorithm, for different orders, is evaluated in terms of execution time requirement and forecast accuracy in both of the above cases and compared with the previous work where the author (Singh, 2009) used fixed length partitions. It is observed that suggested technique yield better results than other two models.

This paper is organized as follows: Section 2 summaries previous works in this direction. In Section 3, the basic concepts of fuzzy time series and the fuzzy forecasting techniques with fuzzy time series is discussed. Section 4 explains automatic clustering technique. Section 5 presents the methodology along with the basic concept of fuzzy discretization. Section 6 shows comparison of the experimental results. Conclusion is presented in Section 7.

2. PREVIOUS WORKS

Fuzzy set theory introduced by Lotfi A. Zadeh (1965, 1975), is based on the notion of partial containment. Fuzzy characteristic function is related to vagueness. In fuzzy logic degree of membership of a variable has a truth value that ranges between 0 and 1. The advantage of fuzzy forecasting is that it can handle problems using numerical as well as linguistic information such as low, moderate, high, very high, etc.

Among the earlier important research works on prediction were generated by Sugeno and Tanaka (1991) Wang and Mendel (1992). Song and Chissom (1993a, 1993b, 1994) developed the time variant and time invariant models for fuzzy time series forecasting. Further, many researchers, Chen (1996), Sullivan and Woodall (1994), Kim and Lee (1999), Chen and Hwang (2000), Huarng (2001), and Tsai and Wu (2001), worked on the development of various models of fuzzy time series forecasting and its implementations.

Problem of student enrollment using high order fuzzy time series models were addressed by Hwang, Chen, and Lee (1998), Tsai and Wu (1999), Chen (2002). Song (2003) introduced an average autocorrelation function as a measure of the dependency between fuzzy data. Forecast results were improved by Lee and Chou (2004) using redefined universe of discourse and partition of intervals and Li and Chen (2004) using recursive partitioning of universe of discourse level by level. Chen and Hsu (2004) proposed a method to forecast enrollments by arranging intervals of discourse according to the frequency of their occurrences.