



Fuzzy Clustering Time Series based on structural components

Ledys Llasmin Salazar Gomez*
Universidad de Valparaíso, Valparaíso, Chile - ledys.salazar@uv.cl

Rodrigo Salas
Universidad de Valparaíso, Valparaíso, Chile - rodrigo.salas@uv.cl

In the process of clustering time series main idea is to generate clusters with similar characteristics, so that greater accuracy in clustering is obtained. The present study describes the process of fuzzy clustering time series based on structural components. Techniques are implemented such as the Discrete Wavelet Transform (DWT) and Fuzzy c-means algorithm. The DWT allows mapping of the series in time domain and frequency and Fuzzy c-means algorithm allows the fuzzification of the series and obtaining clusters. The proposed algorithms are implemented on both synthetic data and real data from the actual data is considered pollution records of MP 10 for different times of day for 30 days. The clustering process, particularly the Fuzzy c-means, provides the ability to group time series according to the membership values thrown by the algorithm, where the maximum value determines cluster membership. Additionally, the process generates clusters characterized by a set of series belonging to the same stochastic process. Implementation, simulation and validation was performed on synthetic and real data, there was tested the hypothesis, finding that improves clustering algorithm on time series which have been transformed in the DWT in string comparison have not been transformed. Additionally, cluster validation indices realize the amount of time series must contain each cluster, so that the group is more accurate.

Keywords: Time series; Wavelets; Fuzzy c-means; cluster.