World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:12, No:07, 2018

Distributed Perceptually Important Point Identification for Time Series Data Mining

Authors : Tak-Chung Fu, Ying-Kit Hung, Fu-Lai Chung

Abstract : In the field of time series data mining, the concept of the Perceptually Important Point (PIP) identification process is first introduced in 2001. This process originally works for financial time series pattern matching and it is then found suitable for time series dimensionality reduction and representation. Its strength is on preserving the overall shape of the time series by identifying the salient points in it. With the rise of Big Data, time series data contributes a major proportion, especially on the data which generates by sensors in the Internet of Things (IoT) environment. According to the nature of PIP identification and the successful cases, it is worth to further explore the opportunity to apply PIP in time series 'Big Data'. However, the performance of PIP identification is always considered as the limitation when dealing with 'Big' time series data. In this paper, two distributed versions of PIP identification based on the Specialized Binary (SB) Tree are proposed. The proposed approaches solve the bottleneck when running the PIP identification process in a standalone computer. Improvement in term of speed is obtained by the distributed versions.

Keywords : distributed computing, performance analysis, Perceptually Important Point identification, time series data mining
Conference Title : ICDM 2018 : International Conference on Data Mining
Conference Location : Dublin, Ireland
Conference Dates : July 23-24, 2018