Abstract

Time series representation is one of key issues in time series data mining. Time series is simply a sequence of number collected at regular interval over a period of time and obtained from scientific and financial applications. The nature of time series data shows characteristics like large data size, high dimensional and necessity to update continuously. With the help of suitable choice of representation it will address high dimensionality issues and improve the efficiency of time series data mining. Symbolic Piecewise Trend Approximation is proposed to improve efficiency of time series data mining in high dimensional large databases. SPTA represents time series in trends form and obtained its values. Sign of value indicate changing direction and magnitude indicates degree of local trend. Depending on the trend of time series, it is segmented into samples of different size which are approximated by the ratio between first and last points within the segment. Each segment then represented by alphabet. The time series is thus represented as sequence of alphabets thus reducing its dimension. Validate SPTA with naïve based classification method.
References

- Lei Sun, Yujiu Yang, Wenhuang Liu, "Trended DTW Based On Piecewise Linear Approximation for Time Series Mining". in 11th IEEE International Conference on Data Mining Workshops, 2011.
- Chung, F. L., Fu, T. C., Luk, R., Ng, V., "Flexible time series pattern matching based on perceptually important points". in International Joint Conference on Artificial Intelligence Workshop on Learning from Temporal and Spatial Data, pp. 1–7, 2001.

Index Terms

Computer Science

Information Sciences

Keywords
Trend based Approach for Time Series Representation

Data mining  Time series representation  Time series  piecewise trends
Approximation