Tianjiao Liu
EAP508
Analytical Summary
10/05/2016

An Analytical Summary of Tak-chung Fu (2010)

After reading lots of papers about times series Tak-chung Fu found that though there are abundant researches in this field in last decade but those papers are so complicated, which may hamper the entry of interested researchers. Therefore in this article, he wanted to do a comprehensive review about the literatures of current time series data mining researches and serve as a glossary for interested researchers to have an overall picture on the current time series data mining development and identify their potential research direction. The author provided fundamental knowledge about the field in macroscopic way and also demonstrated some problems that need further study. This is a great help for those who are interested in time series but didn't have much knowledge about that. By this way the author lowered the entry line of the research field of time series data mining and pointed out some general topic in the field.

Focused Analysis of Time Series

One of the key concept in this article is time series, which can be a type of dataset or a method of analysis. The author didn't give a clearly definition perhaps because he assumed that people who read this paper in the statistics field could understand time series without any explanation. Firstly on page 164 he use "an important class of temporal data objects" to limit the data type to temporal data. And in the same page he use "a collection of observations made chronologically" to clarify how time series data are formed by single observation. And then he use "a whole instead of individual numerical field" to explain the usefulness of time series data that it can be a replacement data form for research datasets. Then the author listed some related researches such as subsequence searching and dimensionality reduction in the following paragraph to expand the discussion about time series and give a further developed knowledge about time series.

Language Pattern Analysis

Terms for time series

an important class of temporal data objects

a collection of observations made chronologically

time series data

Terms for common segmentation method

Fixed-length segmentation of a time series

Even segmentation of a time series

Common segmentation method

I think at the beginning of the section of specific topic the author will use larger terms with more information and more modifiers to give a description of the key concept to help the reader have a better understanding of the concept and notice that there may be a specific terms for the

concept and in the following part the term will be discussed. And in the discussion part the author move to use short term to make the article easy to read and not too wordy since the reader has already know what the concept about. And also the part we should focus is changing from the concept to the discussion of it. Long terms will surpass the role of the discussion. That means we should use long terms first to give information and modification of the concept and avoid using those long terms in the following part of the discussion to make the article simple and clear.

Pursuing a Line of Inquiry

The two concept I'm interested in is segmentation and visualization. The author discuss the segmentation from three different aspect. It is a preprocessing step for variety of data mining, a trend analysis techniques and also a discretization problem. From what I learn now I only use segmentation to pretreat data and make it smaller to be use by software. His discussion broadened my understanding of segmentation. And the other concept is visualization. In the article there has a brief intro of some different visualization tools such as VTT and VizTree (p.170), which I didn't know before. A further understanding of these concept can help me in future learning in my field about data processing and prediction visualization. In many practice cases about time series what we need is the prediction or the trend. Though statistical analysis give the researcher an obvious result after all the work, statisticians usually serve for the clients who know little about the professional statistical. So that we need to visualize the data and prediction in an intuitional way that even outsiders will understand easily. For a simple example, we want to predict the future sales volume for a specific drug in some area. After gathering the data of previous sales we will do analysis based on these data but the result we get maybe just a prediction dataset. Then it is the time to visualize the prediction to show the future trend of the sales to the clients and help them make a decision about the strategy of further promotion or drugs developing. For this simple situation may be some line charts comparing previous sales and the predictions will enough. But in business field there will be a space for visualization of specific problems because the problems may be various with more and more demand in data mining and big data. I think this concept should be helpful in my career.