Review: Effective and Robust Mining of Temporal Subspace Clusters

Layla Pournajaf
Department of Mathematics and Computer Science
Emory University
January 22, 2013

Summary  Hardy Kremer et. al. in the paper ”Effective and Robust Mining of Temporal Subspace Clusters” proposed an approach for clustering multivariate time series. Since this type of data can be categorized as high dimensional data, they exploited the existing methods such as subspace clustering in their solution. However, they pointed out two major shortcomings of using this method including dealing with misalignment of time series and noisy values. To overcome misalignment, they suggested to shift time series individually within a threshold to find the maximum compactness of time series in a cluster. To deal with noise, they check the validity of every deviated value by testing if it follows a normal distribution with a variance based on the compactness threshold. Then the authors proposed an iterative clustering algorithm that fulfills the compactness requirement and checks the validity of all values at the same time that minimizes the cost of each cluster which is the average pairwise time series shift distance in that cluster.

Positive Points  The paper highlights the shortcomings of the existing methods which gives a good understanding of the new problem. While elastic distance metrics are popular in time series, the need for other metrics such as compactness and cluster diameter defined in this paper seems necessary especially in high dimensional temporal clustering problems.

Negative Points  The paper can not be considered well-written or well-detailed. Problem definition could be more clear introducing some applica-
tions or examples. Moreover, while assumptions are stated in details, the
paper does not provide enough justifications or proofs for the claims.
The efficiency of the method is only shown by comparing it to other meth-
ods, but there is no analysis on time complexity. For example, it does not
show how the heuristic algorithm finds the best shifting values to achieve the
required compactness which is important with regard to quality and time
complexity.

Questions

How does the method guarantee the compactness maximization? Is it
efficient regarding to the all possible combinations of shiftings?

What is the complexity of the algorithm regarding the time and space?