Abstract: Adaptive Trace of Multi-Dimensional Clusters by Predicting Data Streams

Nam Hun Park[1] and Kil Hong Joo[2*]

[1] *Dept. of Computer Science, Anyang University, 102 Samsungli, Buleunmyun, Ganghwagun*
Incheon, Korea, 417-833
nmhnpark@anyang.ac.kr

[2] *Dept. of Computer Education, Gyeongin National University of Education, San 6-8*
Seoksudong Manangu Anyangsi, Gyeonggi, Korea, 430-040
khjoo@ginue.ac.kr

Abstract

Clustering multi-dimensional data streams has always been a serious challenge in the field of data mining. In a real-life application, data mining methods for data streams have to effectively extract the on-going change of a data stream with respect to all the subsets of the dimensions of the data stream. In this paper, a subspace clustering method with predicting supports of clusters is proposed. The current statistics of data elements in the partitioned range are maintained. For the range of elements with high density, the range is partitioned to detect the detailed boundary of clusters. To identifying the recent change of a data stream quickly, the support of elements is carefully monitored and predicted to determine partitioned ranges to become clusters. By predicting the change of supports, the on-going change of a data stream can be reflected in real-time. The proposed method is comparatively analyzed by a series of experiments to identify its various characteristics.

Acknowledgement