Multidimensional Data Mining

Ronnie Alves and Orlando Belo
{ronnie,obelo}@di.uminho.pt

Abstract. In many applications, data contains structured information that is multidimensional and multilevel in nature, such as e-commerce, telecommunications, retail, stocks, scientific-data, etc. Since last decade, we have been facing several research efforts on Data Warehousing and OLAP to get better view of multidimensional data, allowing a data-driven search for interesting patterns at any level of data abstraction. This strategy of searching patterns on multidimensional databases is also called OLAPing (or data cubing). The basis of multidimensional data analysis relies on effective and efficient computation of aggregating functions. Multidimensional data analysis on higher dimensional data is almost unfeasible; given the limitation of the most known cubing algorithms. So, how to enhance this data-driven search with discovery-driven features smoothing the curse-of-dimensionality problem? Besides, data as well as patterns evolve over time-to-time. Thus, how to highlight those significant changes? Multidimensional data mining (MDM) take its place helping to handle those previous issues. In summary, MDM attempts to combine ideas of cubing and mining techniques to get better mechanisms for multidimensional data analysis.

In this work we investigate query processing and mining techniques for mining multidimensional and multilevel patterns. The central statement of this research is: *"Multidimensional data mining, by aggregation-based mining, is to some extent as possible"*.

Three major issues are addressed to support this central statement:

a) It is possible to extend warehouse mining [a1, a2, a3, a4, a5]
b) It is possible to improve data cubing [b1,b2,b3]
c) It is possible to evaluate significant changes [c1,c2,c3,c4]

References [https://alfa.di.uminho.pt/~ronnie**]**

a1. Cube-based Mining Methods on Data Warehouses [ICEIS'04]
a2. General Framework for Warehouse Mining of Clickstreams [DM'04 Wessex]
a3. Web Crawler Detection through Cube-based Mining [DG'04] *JISBD'04*
a4. Mining Large Transactional Tables [EPIA'05] *LNAI*
a5. Mining Inter-transactional Patterns [JISBD'05]
b1. Cubing Maximal-Correlated Cuboids Cells [DAWAK'06] *LNCS*
b2. Incremental Cube Mining [ICEIS'07] *on review*
b3. Mining top-k Gradients in Large Databases *ongoing work*
c1. Anomaly Detection based on Signatures [ICDM'06 Leipzig] *LNAI*
c2. Framework for Mining Anomalous Patterns [DMBA'06] *ACM SIGKDD*
c3. Agregation- and Graph-based Anomaly Detecion [GFKL'07] *LNAI*
c4. Change Analysis by Dynamic Clustering Analysis [MLDM'07] *on review*