

A Conceptual Model for Combining Enhanced OLAP and Data Mining Systems

Muhammad Usman and Dr. Sohail Asghar
*Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology,
Islamabad, Pakistan,
usmanspak@yahoo.com, sohail.asghar@szabist-isb.edu.pk*

Simon Fong
*Faculty of Science and Technology,
University of Macau, Macau SAR,
ccfong@umac.mo*

Abstract— Online Analytical Processing (OLAP) was widely used to visualize complex data for efficient, interactive and meaningful analysis. Its power comes in visualizing huge operational data for interactive analysis. On the other hand, data mining techniques (DM) are strong at detecting patterns and mining knowledge from historical data. OLAP and DM is believed to be able to complement each other to analyze large data sets in decision support systems. Some recent researches have shown the benefits of combining OLAP with Data Mining. In this paper, we reviewed the coupling of OLAP and data mining in the literature and identified their limitations. We proposed a conceptual model that overcomes the existing limitations, and provides a way for combining enhanced OLAP with data mining systems. Furthermore, the proposed model offers directions to improving cube construction time and visualization over the data cube.

Keywords- Data Mining, Enhanced OLAP, Performance Enhancement, Visualization Techniques

I. INTRODUCTION

OLAP technology refers to a set of data analysis techniques to view the data from different sources in different dimensions, interactively, for decision support. The growing complexity and volumes of the data to be analyzed impose new requirements on OLAP systems [1].

According to [2] and [3], OLAP systems have rapidly gained momentum in both the academic and research communities, mainly due to their allowance for quick, multi-dimensional analysis and for their visualization capabilities. According to [4] the main reasons for this rapid development are as follows;

- a) Able to obtain quick answers of complex queries on their huge operational data.
- b) Efficient and interactive analysis on organizational historical data.
- c) Can be specially designed for supporting retrieval of higher level summary information from detailed data.
- d) Offer better performance for aggregate queries.

Efficiently visualizing multi-dimensional data sets produced by scientific processes is becoming increasingly challenging and is attracting the attention of a wide community of researchers [5]. With this increase in popularity along with rapid development the user's expectations and demands from the existing OLAP systems

have also been amplified [6]. That is therefore a strong need to improve the performance and enhance the visualization of the current OLAP systems.

It has been observed that research endeavors on OLAP performance and OLAP visualization progressed separately in the past. With the coupling of these enhancement techniques, OLAP functionality can be improved [7]. However, the growing complexity and volumes of the data to be analyzed impose new challenges on OLAP systems [1]. To fulfill these increasing needs, a number of OLAP enhancement techniques have been proposed in the literature [3,4,6,8,9,10,11,12,13,14,15,16]. Although some of the previous research work was on the enhancement of OLAP systems but none was intended towards an integrated enhancement of OLAP performance and visualization. One of the main contributions in this paper is the proposed model for the integrated enhancement of OLAP. It is enabled by shortening the cube construction time and intensifying visualization of cube data with the help of cube visualization tool.

The other contribution is on combining OLAP and data mining. According to [4], data mining is comprised of analytic techniques such as *Association* [17], *Classification* [18], *Clustering* [19] and *Trend analysis* [20]. Some previous works [6, 8, 10 and 21] indicated the likelihood and interest of coupling OLAP and data mining. However, [22] introduced a new concept of integrating OLAP and data mining called Online Analytical Mining (OLAM). A literature review is presented in the following section.

On the basis of these observations and limitations, we proposed a conceptual model for the combination of enhanced OLAP and data mining systems. The proposed model, firstly, enhances the power of OLAP with respect to its performance and visualization. Secondly, it integrates the enhanced OLAP with a data mining technique of hierarchical clustering.

It is observed that by using the proposed model we can enhance the existing OLAP systems in terms of performance and visualization and can get a higher degree of overall improvement. To the best of the authors' knowledge, there is no such model which proposes a solution to both improve the performance and visualization capabilities for OLAP systems. Our experimental results show that the cube construction time can be reduced remarkably with the use of clustered data tables instead of relational tables. Our paper is

organized as follows: Section 2 presents the literature review. Section 3 elaborates the proposed model for combining enhanced OLAP and data mining. Finally, section 4 concludes and indicates possible future research directions.

II. LITERATURE REVIEW

The purpose of this literature review is to identify the limitations of automatic schema generation process. The related topics span across three major themes: (1) OLAP enhancement, (2) Visualization enhancement and (3) combining OLAP with data mining.

A. OLAP Enhancement

In this section, we review various significant OLAP enhancement techniques in the literature. The major motivation is performance improvement and visualization enhancement techniques. Firstly, we discuss some of the performance improvement techniques followed by the visualization enhancement techniques.

According to Goil et al. [8], parallel computing can be used as a way to enhance the performance of OLAP. It has been identified in the research work that typically OLAP operations are responsible for running complex queries on the underlying data, so these operations increase data processing time. The introduction of high performance parallel computers can lessen the computational time by distributing the tasks among various computer resources. The authors proposed an algorithm for Cube construction on parallel computers using distributed memory. The limitation of the work is the usage of a small and simple data set for the experiment. If the data set is large then a mismatch in the order of the dimensions can occur, as the multidimensional array doesn't allow the memory's use of their sort method to be efficient.

Moving on with the performance enhancement quest, Papadias et al. [9], reported on the efficient OLAP operations in Spatial Data Warehouses. Authors identified the problem that the computation of spatial data is expensive as online processing of such complex spatial data is inapplicable. They suggested a data structure, named aR-Tree, which combines a spatial index with the materialization technique. The strength of their work was the identification of the position of the objects in space in which the groupings and the hierarchies among the dimensions were unknown at design time. The work done is limited only to Spatial Data Warehouses and it also does not give any indication of OLAP storage types such as ROLAP and MOLAP for Spatial data warehouses. The dataset used by the authors was found to be insufficient for the acceptance of this technique for all spatial cases

In addition to the OLAP enhancement work, Asghar et al. [6], proposed a functionality-enhancement technique using Growing Self-organizing Neural Networks. This technique proposes the integration of data mining with OLAP by passing on mined data to the OLAP engine for a more focused analysis and hence adding intelligence to the OLAP system and thereby helping its users in sharp analysis of data. The major limitation is the manual schema generation of hierarchical clusters for the OLAP engine to perform

operations. Furthermore, the data set has to be in a special format for the generation of hierarchical clusters using Growing Self-Organizing Map (GSOM) method. The enhancement architecture provided by the authors does not satisfactorily answer for complex data sets. No indication was given as to how the hierarchical clusters generated from the dataset will be visualized.

A few months later, a new contribution in the performance improvement by associating OLAP and data mining was proposed. Messaoud et al. [10] suggested an enhanced OLAP operator based on the Agglomerative Hierarchical Clustering (AHC). The main problem identified is OLAP limitation of aggregation and summarization of complex objects like text, images, sounds and videos. The operator called Operator for Aggregation by Clustering (OpAC) is able to provide significant aggregates of facts referred to complex objects. The strength of the work is the combination of OLAP and data mining techniques as data mining can discover knowledge from both simple and complex data. The major limitation of the technique is that for n individuals to classify, the AHC generates n hierarchical partitions but it does not give guidance as to which best partition is to be chosen. Data miner has to decide the number of clusters that corresponds to the context and to the goal of his/her analysis.

B. Visualization Enhancement

So far we discussed the performance improvement techniques, now we move our review onto some major OLAP visualization enhancement work. For the purpose of visualization enhancement, Maniatis et al. [4], recommended the Advanced Visualization of OLAP. They suggested the Cube Presentation Model (CPM), which can be naturally mapped with an advanced visualization technique from the Human-Computer interaction area, called Table Lens. They formalized the simultaneous presence of more than one query which was done in two layers, Presentation Layer and Logical Layer. The limitation of the work is that the visualization technique used for the screens is restricted only to one aspect of visualization enhancement at a time.

Followed by the above advanced visualization technique, Sifer [11] introduced a visual interface technique for exploring OLAP data with coordinated-dimension hierarchies. They identified the problem that a number of text based and visual interfaces for querying multi-dimensional data exist but many of these interfaces are not applicable to OLAP. The reason is the lack of support for the use of dimensional hierarchies for selection and aggregation. The main contribution is the progressive view coordination interface which provides better support for query refinement than existing interfaces, by helping users decide the next query step with intermediate result overviews, and also helping the users to change a previous selection decision with retained selection context views. The strength of their work is the implementation of SGViewer visualization tool for their technique using a web log dataset. The limitation is that a very simple and small (only 3 dimensions) web log dataset was used for the illustration and, to use this technique the data has to be converted into Structured Graph Format

Extensible Mark-up Language (XML) document. Therefore, it is a specific data format conversion limitation.

Voss et al. [12], adopted another way and suggested an extension in an advanced tool (CommonGIS) for the highly interactive visual exploration of spatial data. The identified problem was that the CommonGIS tool lacks the connection to OLAP warehouse to be a complete Business Intelligence (BI) application. They explored how to connect the tool to OLAP warehouses as another source of multi-dimensional data and designed architecture for the extension of CommonGIS. The strength of their work is that they allowed support of MDX (Multi-Dimensional Expression) for the manipulation of cube data using the application builder of the tool. The limitation of their work is the manual script writing process to convert client data into a compatible format so that the data is incrementally transmitted from the client source to the OLAP data cube.

Scotch et al. [13] proposed a tool for Spatial OLAP, called SOVAT (Spatial OLAP Visualization and Analysis Tool). The authors identified the problem that for community health research a comprehensive and thorough analysis is needed for effective public health evaluation. The major contribution of their work is the development of tool for community health assessments. The strength of the tool is that it combines all necessary features of OLAP and GIS (Geospatial Information Systems) into one system to support comprehensive community health decisions. The limitation of this system is that it allows the visualization of spatial data and only permits the analysis on numeric data; however while performing analysis on spatial databases the researchers require tool support for not only number data but also other complex data types such as images, videos and maps. The community health researchers are more interested in the patterns or areas of an image for evaluation purposes, rather than exploring numerical analysis of the data using OLAP.

Cuzzocrea et al. [3] introduced a hierarchy-driven compression technique for the advanced visualization of multi-dimensional cubes. This technique relied on the facility of generating a “semantic-aware” compressed representation of two-dimensional OLAP views. The strength of this particular technique is that it is most suitable for the handheld-devices where compression in the data cubes is required. The limitation is that it only supported two-dimensional OLAP views but, in real scenarios, there are multiple dimensions and facts in the dataset. The technique is more suitable for MOLAP scenarios and not very effective in the case of ROLAP, where multiple attributes exist in the relational dataset and the hierarchies are more complex within each dimension.

In 2005, [14] presented a new visual interactive exploration technique for OLAP. This work is similar to our work here in terms of OLAP user facilitation. This technique and our proposed conceptual model, allows users who have novice knowledge of OLAP technology to explore and analyze OLAP data cubes without generating sophisticated queries. The following year, [15] proposed a framework for querying complex multidimensional data with a major effort

applied to transforming irregular hierarchies to make them navigable in a uniform manner.

Lastly, in 2008, [16] introduced a comprehensive visual exploration framework which implements OLAP operations in form of powerful data navigation. Users can explore data using a variety of interactive visualization techniques under the proposed framework. Research communities have often neglected the data-visualization problem in the past [3]. But we note that a considerable amount of work is recently being performed by various research communities on the visualization of OLAP data which indeed can play a very vital role in effective analysis of data. Our work in this paper is similar to the above mentioned contributions as it also improves the performance of OLAP in terms of cube construction time.

Furthermore, we are using cube visualization tool, which dramatically enhances the visual interactive exploration of cube data with the help of charts, graphs and data grids at the front end. In the next section, we present some noteworthy work done on the integration of OLAP enhancement techniques.

C. Combining OLAP and data mining

Several previous studies [6, 8, 10 and 21] emphasized the likelihood and interest of coupling OLAP and Data Mining. Efficient and effective data mining integrates the concept of [22], which is the integration of OLAP and OLAM systems. OLAM is a promising direction for mining knowledge from neural networking and agglomerative hierarchical clustering. The work in this paper is a bit similar in terms of OLAP performance improvement but we have only focused the performance improvement in terms of the cube construction time. Similarly, Hua [23] proposed and developed an interesting association rule mining approach called Online Analytical Mining of association rules. It integrated OLAP technology with association rule mining methods and leads to flexible multidimensional and multi-level association rule mining. Similarly, Joseph et al. [24] presented a methodology that derives the association rule of web pages tick sequences according to the support level and confidence level of user requirements. This methodology was responsible for identifying a set of frequently accessed web pages on a website by a user. The result is the list of potential customers for a certain product or service on a target web page. Dzeroski et al. [25] combined OLAP and Data Mining in a different way to discover patterns in a database of patients. Two data mining techniques, clustering and decision tree induction were used. Clustering was used to group patients according to the overall presence/absence of deletions at the tested markers. Decision trees and OLAP were used to inspect the resulting clustering and to look for correlations between deletion patterns, populations and the clinical picture of infertility. Dehne et al. [26] studied the applicability of coarse grained parallel computing model (CGM) to OLAP for data mining. Authors presented a general framework for the CGM which allows for the efficient parallelization of the existing data cube construction algorithm for OLAP. Experimental data showed that this approach yield optimal speed up even when run on a simple

processor cluster via a standard switch. The study shows that OLAP and data mining, if combined together, can produce greater benefits in a number of diverse research areas. Our proposed work is similar in terms of combining OLAP and data mining techniques but our main focus is on a particular data mining technique known as Hierarchical Clustering. Furthermore, we are using data mining as a pre-processing step to get better understanding of data before passing it to the automatic schema builder and which then generates schema for OLAP engine. In the next section, the related data mining technique which is hierarchal clustering and its usage with OLAP is the centre of attention.

D. Comparative analysis

In the previous sections the individual papers were discussed. Now, this section presents a comparison of the proposed model with the major themes reviewed. Firstly, a substantial amount of work has been done on the enhancement of OLAP in terms of its performance and visualization. According to [6, 8 and 9], performance of OLAP systems can be improved with the help of different techniques like parallel computing, indexing and materialization.

In the area of OLAP's visualization, research communities has also paid attention and reported the work on visualization enhancement in [3, 4, 11, 12, 14, 15 and 16].

Visualization enhancement techniques like Cube Presentation Model (CPM), visual interfacing, interactive exploration, hierarchy driven compression and complex multidimensional querying were introduced. In this paper we have used cube visualization API to enhance the visualization capabilities of the data present in the cube.

Various researchers have contributed in the combination of OLAP and data mining in the past. Significant works in literature in this area include [22, 23, 24, 25 and 26]. These authors used different data mining techniques in their areas of interest and produced better results by combining OLAP with data mining. The two major techniques of data mining used in this regard were Association Rule Mining and Hierarchical Clustering. In this paper, we have used the hierarchical clustering technique for the integrated enhancement of OLAP's performance and visualization. Hierarchical clustering has been used by [4, 27, 28, 29 and 30] in conjunction with OLAP systems for a number of reasons. Some authors used this technique for query optimization, physical data organization, fast data access and reduced storage cost. The work in this paper is different from these as we are combining the hierarchically clustered data with enhanced OLAP system.

TABLE I
SUMMARY OF LITERATURE REVIEW

Authors	Work title	Area of Work				
		OLAP performance	OLAP Visualization	Integrated OLAP	Enhancement OLAP and Data Mining (Integrated)	Hierarchical Clustering
Göil et al.	High performance OLAP and Data Mining on paralel computers	✓	✗	✗	✓	✗
Papadias et al.	Efficient OLAP Operations in Spatia data warehouse	✓	✗	✗	✗	✗
Asghar et al.(2004)	Enhancing OLAP functionality using self-organizing neural networks	✓	✗	✓	✓	✓
Messaoud et al.	A new OLAP aggregation based on AIIC technique	✓	✗	✗	✗	✓
Maniatis et al.	Advanced Visualization for OLAP	✗	✓	✗	✗	✗
Vess et al.	Interactive visual exploration of multidimensional data	✗	✓	✗	✗	✗
Teepichetvanich et al.	Interactive visualization for OLAP	✗	✓	✗	✗	✗
Mansmann et al.	Extending visual OLAP for handling irregular dimensional hierarchies	✗	✓	✗	✗	✗
Markl et al.	Improving OLAP performance by multi-dimensional hierarchical clustering	✓	✗	✗	✗	✓
Karaynnidis et al.	CUBE File: A File Structure for Hierarchically Clustered OLAP Cubes	✗	✗	✗	✓	✓
Teodoratos et al.	Heuristic optimization of OLAP queries in multidimensionally hierarchically clustered databases	✓	✗	✗	✓	✓
Tsois et al.	Cost-Based Optimization of Aggregation Star Queries on hierarchically Clustered Data Warehouses	✓	✗	✗	✓	✓
PROPOSED WORK	A conceptual model for combining enhanced OLAP and data mining systems	✓	✓	✓	✓	✓

III. PROPOSED CONCEPTUAL MODEL

It is evident from the literature review that none of the proposed work in the past had targeted towards the integration of enhanced OLAP with data mining. The major limitation of those highlighted in the literature reviews is the deficiency of enhancement on the performance and visualization aspects of OLAP systems. During the last few years, a lot of work has been done on the enhancement of OLAP particularly on its performance and visualization. Various architectures [6], [31], [32] were proposed for the enhancement of OLAP systems. At this point the question arise that *why there is a need for conceptual model for combining enhanced OLAP and data mining?* The answer is that, though a number of enhancement architectures were proposed in the past, none of the work was intended towards focusing on the integrated enhancement of OLAP performance and visualization and there exist a strong need for it [7]. OLAP systems are gaining popularity rapidly, user's expectations and demands from the modern OLAP systems have also increased [6]. To fulfill these growing

demands of OLAP and data mining users, a model is required which can easily be deployed as a complete system and also support the integrated enhancement. Figure 1 depicts the proposed conceptual model, which not only integrate enhanced OLAP with data mining but also enforce integrated enhancement of OLAP's performance and visualization. We elaborate on the distinct components of the proposed conceptual model for the enhancement of OLAP systems. The goal of this architecture is to integrate OLAP with Data Mining and to provide integrated performance and visualization enhancement. To achieve this objective, we have used separate servers; one is for the database and the other one for the OLAP data. This model indicates two approaches to pass data to the OLAP engine. The first approach is the conventional or non clustered one where a data set is loaded in the database server through Extract, Transform and Load (ETL) process and stored in the form of a relational database. From the relational database a star schema is designed using the standard SQL queries and data is loaded into the star schema.

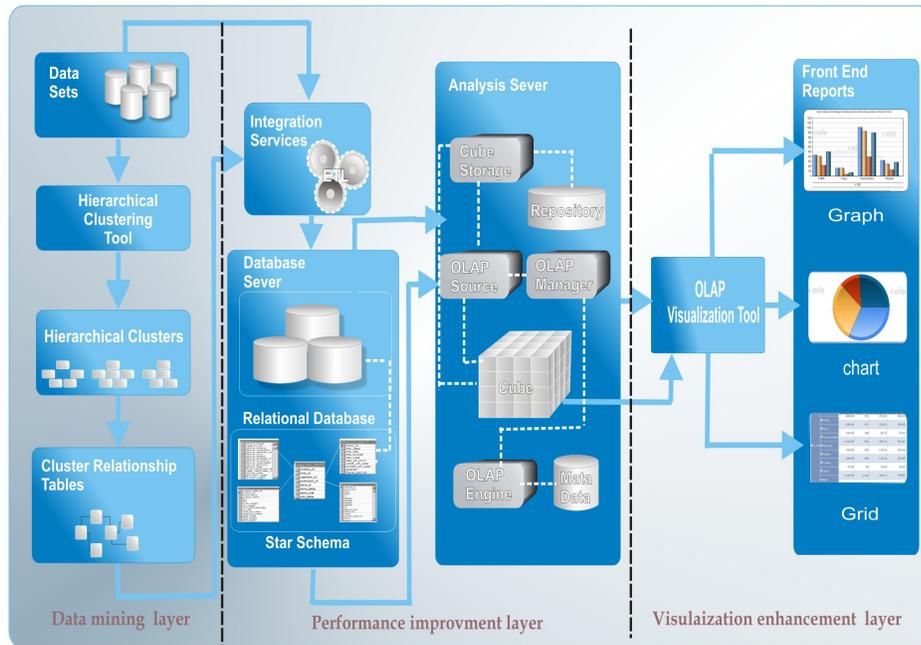


Fig 1. Proposed conceptual model for combining enhanced OLAP and data mining systems

The OLAP server takes this *star* schema as a source to construct OLAP cubes. It also provides storage and management mechanism for the cube data. At the front-end a visualization tool captures the cubes generated by the OLAP server and displays data in the form of charts, reports and tables. The model we have devised to enhance OLAP is based on hierarchical clusters generated by hierarchical clustering tool. The clusters generated by the tool are then transformed into relational tables which are fed into the database server through the integration services. Star schema is designed from the database and the source. A Front-end visualization tool is connected with the OLAP server generates cubes using star schema as a data OLAP cubes,

displaying the cube data in various views such as charts, graphs, reports and tables at the user's end. Users can perform the basic OLAP operations using the front-end tool; hence, the user can perform interactive analysis using the *drill down*, *roll up* and other standard OLAP operations.

The proposed conceptual model achieves its objective as it integrates enhanced OLAP with Data mining and uses clustered data for performance improvement in terms of cube processing time. The Visualization tool at the front end supports an interactive visual exploration of OLAP data using *drill-down*, *roll up* charts and tables which enhances the visualization capabilities of both traditional and clustered OLAP data.

IV. CONCLUSION AND FUTURE WORK

We studied the work related to enhancement of OLAP performance, visualization and the integration with data mining. Literature review reveals a fact that the functionality and visualization of OLAP systems can be extended using different techniques. But none of the previous work combined enhanced OLAP and data mining system together. Review shows that there is no single model available which deals with the enhancement of OLAP functionality and Visualization in parallel. We have proposed a conceptual model for this purpose. With this integration, we devised a conceptual model for the integrated enhancement, hence making the existing OLAP and data mining systems more useful. Hierarchical clusters have been used as a way to enhance the targeted analysis of OLAP. We have discussed the proposed model along with its components. It is concluded that functionally enhanced OLAP should be combined with data mining systems in order to achieve better performance and improved visualization from the current OLAP and data mining systems. Future work is mainly on dynamic generation of relational tables from the clustered data. In addition to this we are also working on other Data Mining techniques that can be integrated with OLAP systems to enhance its analysis capabilities.

REFERENCES

- [1]. S. Chaudhuri and U. Dayal, An overview of data warehousing and OLAP technology, *ACM SIGMOD Record*, Vol. 26 (1997), pp.65-74.
- [2]. A. Cuzzocrea, D. Sacca and P. Serafino, A hierarchy driven compression technique for advanced OLAP visualization of multidimensional data cubes, in *Proc. of 8th Int'l Conf. on Data Warehousing and Knowledge Discovery (DaWak)*, (Springer Verlag 2006), pp.106-119.
- [3]. S. Mansmann and M. Scholl, Exploring OLAP aggregates with hierarchical visualization techniques, in *Proc. of ACM Symposium on Applied Computing* (2007), pp.1067-1073.
- [4]. Fayyad U M, Piatetsky-Shapino G., Smyth P. and Uthurusany R., "From datamining to knowledge discovery: An overview" in *Proc. of Advances in data mining and knowledge discovery*, MIT Press, pp.1-34.
- [5]. S. Goil and A. Choudhary, "High performance OLAP and data mining on parallel computers," *Data Mining and Knowledge Discovery*, vol. 1, no. 4, pp.391-417, Dec. 1997.
- [6]. S. Asghar, D. Alahakoon and A. Hsu, "Enhancing OLAP functionality using selforganizing neural networks," *Neural, Parallel and Scientific Computations*, vol. 12, no. 1, pp.1-20, March 2004.
- [7]. R. B. Messaoud, O. Boussaid and S. Rabaseda, "A new OLAP aggregation based on the AHC technique," in *Proc. of the 7th ACM Int'l Workshop on Data Warehousing and OLAP (DOLAP)*, ACM New York, 2004, pp.65-72.
- [8]. S. Goil and A. Choudhary, "High performance OLAP and data mining on parallel computers," *Data Mining and Knowledge Discovery*, vol. 1, no. 4, pp.391-417, Dec. 1997.
- [9]. D. Papadias, P. Kalnis, J. Zhang and Y. Tao, "Efficient OLAP operations in spatial data warehouse," in *Proc. of the 7th Int'l Symposium on Advances in Spatial and Temporal Databases (SSTD)*, Springer-Verlag, 2001, pp.443-459
- [10]. R. B. Messaoud, O. Boussaid and S. Rabaseda, "A new OLAP aggregation based on the AHC technique," in *Proc. of the 7th ACM Int'l Workshop on Data Warehousing and OLAP (DOLAP)*, ACM New York, 2004, pp.65-72.
- [11]. M. Sifer, "A visual interface technique for exploring OLAP data with coordinated dimension hierarchies," in *Proc. of the 12th ACM Int'l Conf. on Information and Knowledge Management (CIKM)*, ACM New York, 2003, pp.532-535.
- [12]. A. Voss, V. Hernandez, H. Voss and S. Scheider, "Interactive visual exploration of multidimensional data: requirements for common GIS with OLAP," in *Proc. of the 15th Int'l Workshop on Database and Expert Systems Applications (DEXA)*, 2004, pp.883-887.
- [13]. M. Scotch and B. Paramanto, "SOVAT: Spatial OLAP visualization and analysis tool," in *Proc. of the 38th Annual Hawaii Int'l Conf. on Systems Sciences (HICSS)*, 2005, p.165.
- [14]. K. Techapichetvanich and A. Datta, "Interactive visualization for OLAP," in *Int'l Conf. on Computational Science and its Applications (ICCSA)*, 2005, pp.206-214.
- [15]. S. Mansmann and M.Scholl, "Extending visual OLAP for handling irregular dimensional hierarchies," in *Proc. of 8th Int'l Conf. on Data Warehousing and Knowledge Discovery*, 2006, pp.95-105.
- [16]. S. Mansmann and M. Scholl, "Visual OLAP: A new paradigm for exploring multidimensional aggregates," in *Proc. of IADIS Int'l Conf. on Computer Graphics and Visualization*, 2008, pp.59-66.
- [17]. J. Hipp, U. Guentzer and G. Nakhaeizadeh, "Algorithms for association ruling mining – a general survey and comparison," *ACM SIGKDD Explorations Newsletter*, vol. 2, no. 1, pp. 58 – 64, June 2000.
- [18]. R. C. Holte, "Very simple classification rules perform well on most commonly used datasets," *Machine Learning*, vol. 11, no. 1, pp.63-90, April 1993.
- [19]. A. K. Jain, M. N. Murty and P. J. Flynn, "Data clustering: A review," *ACM Computing surveys (CSUR)*, vol. 31, no. 3, pp.264-323, Sep. 1999.
- [20]. M. J. Shaw, C. Subramaniam, G. W. Tan and M. E. Welge, "Knowledge management and data mining for marketing," *Decision Support Systems*, vol. 31, no. 1, pp.127-137, May 2001.
- [21]. J. Han, "Towards online analytical mining in large databases," *ACM SIGMOD Record*, vol. 27, no. 1, pp.97-107, March 1998.
- [22]. J. Han, S. H. S. Chee and J. Y. Chiang, "Issues for on-line analytical mining of data warehouses," in *Proc. of the SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery (DMKD)*, Seattle, 1998, pp.2:1-2:5.
- [23]. H. Zhu, "Online analytical mining of association rules," Master Thesis, Simon Fraser University, 1998, pp.1-117.
- [24]. J. Fong, H. K. Wong and A. Fong, "Online analytical mining Web-pages tick sequences," *J. of Data Warehousing*, vol. 5, no. 4, pp.59-67, 2000.
- [25]. S. Dzeroski, D. Hristovski and B. Peterlin, "Using data mining and OLAP to discover patterns in a database of patients with Y chromosome deletions," in *Proc. AMIA Symp.*, 2000, pp.215-219.
- [26]. F. Dehne, T. Eavis and A. Rau-Chaplin, "Coarse grained parallel on-line analytical processin (OLAP) for data mining, in *Proc. of the Int'l Conf. on Computational Science (ICCS)*, 2001, pp.589-598.
- [27]. V. Markl and R. Bayer, "Processing relational OLAP queries with UB-trees and multidimensional hierarchical clustering," in *Proc. of the Int'l. Workshop on Design and Management of Data Warehouses (DMDW)*, 2000, pp.1-10.
- [28]. N. Karayannidis, T. Sellis and Y. Kouvaras, "CUBE file: A file structure for hierarchically clustered OLAP cubes," *Advances in Database Technology, LNCS*, Springer Verlag, Berlin-Heidelberg, pp.621-638, 2004.
- [29]. D. Theodoratos and A. Tsois, "Heuristic optimization of OLAP queries in multidimensionally hierarchically clustered databases," in *Proc. of the 4th ACM Int'l Workshop on Data Warehousing and OLAP (DOLAP)*, ACM New York, 2001, pp.48-55.
- [30]. K. Hann, C. Sapia and M. Balaschka, "Automatically generating OLAP schemata from conceptual graphical models," in *Proc. of the 3rd ACM Int'l. Workshop on Data Warehousing and OLAP (DOLAP)*, ACM New York, 2000, pp.9-16.
- [31]. S. Asghar, D. Alahakoon and D. Taniar, "Enhancing OLAP Functionality using Association Rules," in *Proc. of 2nd Int'l Conf. on Computational Intelligence, Robotics and Autonomous Systems (CIRAS)*, 2003, pp.1-6.
- [32]. A. Rauber and P. Tomsich, "An architecture for modular OLAP Systems: Supporting distributed and parallel query processing using cooperating CORBA objects," in *10th Int'l Workshop on Database and Expert Systems Applications (DEXA)*, 1999, pp.45.