ISSN:2231-0711

J.Sreedhar et al | IJCSET(www.ijcset.net) | March 2016 | Vol 6, Issue 3, 98-100

# Online Analytical Report Recommender System

J.Sreedhar
*Professor & Principal*
*Department Of CSE*
*ECET , Hyderabad*

M.Neelima
*SR.Software Engineer*
*UGS-Hyderabad*

D.Ramesh
*Associate Professor*
*JNTUHCEJ-Hyderabad*

**Abstract— This project is a Recommender System for OLAP Reports, KPI's,Scorecards and Business Dashboards.OLAP (Online Analytical Processing) is a powerful technology which performs multidimensional analysis of business data and lays a foundation to many kinds of business applications like Business Performance Management, Planning, Budgeting, Forecasting, Financial Reporting, Analysis, Simulation Models, Knowledge Discovery, and Data Warehouse Reporting. OLAP enables end-users to perform ad hoc analysis of data in multiple dimensions, thereby providing the insight and understanding they need for better decision making.**

**However It is very time consuming and requires lot of research and ground work to come up with the list of Business Questions, Metrics, KPI's, Scorecards,Reports and Dashboards , Report Layouts etc. Our proposed framework does all the groundwork at back end and displays the proposed list of KPI's, Reports and Dashboards to each user based on his Past History, Data Mart Accessibility, Frequently Accessed Reports and Seasonal Trends etc. It also provides Ranking to Metrics and KPI's as per Users History. Users can easily select from the proposed list and access the reports then and there.**

**Keywords— OLAP, Online Analytical Processing, Data-warehousing Reports and Dashboard, Business Analytics,KPI's, Scorecards, Key Performance Indicators, Decision Making Analytics, Recommender System for OLAP Reports**

## I. INTRODUCTION

Managers see information as a critical resource and require systems that let them exploit it for competitive advantage. One way to better use organizational information is via online analytical processing and multidimensional databases (MDDBs). OLAP and MDDBs present summarized information from company databases. They use multidimensional structures that let managers slice and dice views of company performance data and drill down into trouble spots. For over a decade, proponents have touted these tools as the ultimate executive information system, but most of the hype comes from product vendors themselves. Based on our experience with several OLAP tools, we have developed a more pragmatic approach to the design of multidimensional information systems that lets managers make the most of their companies' information assets.

Multidimensional analysis helps organizations extract maximum value from their corporate data. It transforms volumes of data into information about the business, allowing users to analyse information in a business context — comparisons of things such as product or channel performance, in light of other important factors like regions, customers, and time. With a multidimensional view, users can quickly gain insight into business performance and trends.

OLAP (Online analytical processing) is computer processing that enables a user to easily and selectively extract and view data from different points of view. For example, a user can request that data to be analysed to display a spreadsheet showing all of a company's beach ball products sold in Seattle in the month of September, compare revenue figures with those for the same products in March, and then see a comparison of other product sales in Seattle in the same time period. To facilitate this kind of analysis, OLAP data is stored in a multidimensional database Whereas a relational database can be thought of as two-dimensional, a multidimensional database considers each data attribute (such as product, geographic sales region, and time period) as a separate "dimension." OLAP software can locate the intersection of dimensions (all products sold in the Eastern region above a certain price during a certain time period) and display them. Attributes such as time periods can be broken down into sub-attributes. OLAP can be used for data mining or the discovery of previously undiscerned relationships between data items. An OLAP database does not need to be as large as a data warehouse, since not all transactional data is needed for trend analysis. Using Open Database Connectivity (ODBC), data can be imported from existing relational databases to create a multidimensional database for OLAP.

## II. APPROACH OF OLAP REPORT RECOMMENDER SYSTEM

Even though OLAP Offers easy and efficient ways to create reports for complex user queries, users who are not familiar with Data Warehouse may overlook queries that retrieve interesting data, or they may not know what parts of the Data Warehouse provide useful information. Whenever Business Reviews and Analytics are required and Business Decisions have to be made, it is a repetitive task for Business Analysts and Managers to do rigorous discussions and finalize on the Reports, Dashboards, KPI's and Scorecards to be showcased etc. This issue clearly hinders data exploration, and thus reduces the benefits of using an OLAP system. To address this important problem, we draw inspiration from the successful application of recommender systems in social networks based on web data.Recommender Systems aim to provide users with recommendations about items that people with similar

tastes and preferences have liked in the past. Collaborative Filtering is the dominant technique for Recommender Systems; it relies on the opinions expressed by other users. The premise is simple: If a user A has similar querying behaviour to user B, then they are likely interested in the same data. Hence, the reports of user B can serve as a guide for user A.

In web collaborative filtering systems, a user-item matrix approach is used to generate recommendations. More specifically, each user is represented as an item vector, where the values of the vector elements correspond to the user's preferences for each item (such as movie ratings, purchased products, read articles, etc.) The similarities between users in this representation can be easily computed using vector similarity metrics. Given the most similar users and their preferences, the collaborative filtering system can subsequently predict, what items will interest each user, and generate item recommendations

This approach have to be modified as per the functionality of OLAP Reporting. In OLAP Reporting, the "items" of interest are data records, and the users access these items indirectly by creating reports which contains attributes pulled-in from Cubes and the data is accessed from Data-warehouse. Thus, even though the users' behaviour is identified by the set of Reports they create, their interest lies on the data they retrieve. This complicates the evaluation of similarity among users based on their Reports alone, since it is no longer obvious whether they are interested in the same "items". This raises an important issue that needs some consideration. The similarity between users can be expressed as the similarity between the fragments of their queries or, alternatively, the data that they retrieve. This is not as straightforward, since a query fragment or a tuple might have different levels of importance in different user sessions. Thus, we must be able to create implicit user profiles that model those levels of importance, in order to effectively compare the users. Finally, contrary to the user-based collaborative filtering approach, the recommendation to the users have to be in the form of Reports, since those actually describe what the retrieved data represent. Thus, we need to "close the loop" by first decomposing the Reports to SQL Queries and then into lower-level components in order to compute similarities and make predictions, and then re-construct them back to SQL queries and then back to Reports in order to recommend them.. All these issues make the problem of interactive database exploration very different from its web counterpart.

In this paper, we present our work in the development of a Report recommender system for a Data warehouse. We first discuss an abstract framework that conceptualizes the problem and defines specific components that must be instantiated in order to develop a solution. Based on this framework, we develop a solution that transfers the paradigm of collaborative filtering in the context of relational queries (OLAP Reports). The recommended solution can be implemented using existing technology, and is thus attractive for a real-world deployment.

## III. How Does ReportRecommender System Work?

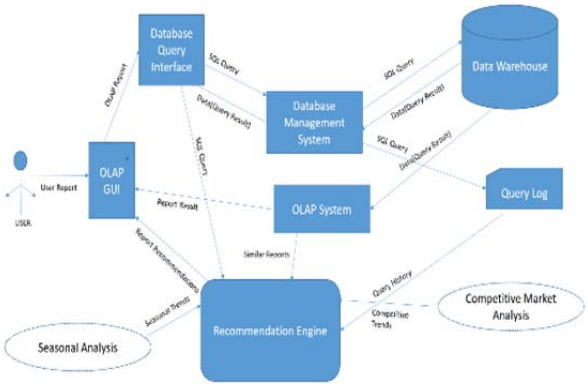


Fig: 1 Process Flow Diagram

High Level Process flow is described in the Figure 1

- User(Report Analyst/Managers/End Users) access the OLAP GUI(Graphical user Interface) and generates an OLAP Report
- This OLAP Report automatically has its SQL Query generated at the backend. This SQL Query is processed through a Database Management System and required results are fetched from a Data Warehouse.
- The OLAP System receives the resultant set of data and sends back the result to OLAP GUI.
- Recommendation Engine takes inputs from OLAP System and Query Log File and comes up with Report Recommendations. It also considers inputs from the data retrieved from external sources (Competitive Market Analysis and Seasonal Analysis). It implements various algorithms suggested in section 1 and comes up with the set of recommended Reports. These recommendations are sent to OLAP GUI for display.

## IV. Techniques Used In Olap Report Recommender Systems

In this section, we discuss significant work in the areas of recommender systems based on user ratings and other techniques. With Significant development in the area of Web Mining Engineering Domain, a lot of sophisticated techniques like Probabilistic Latent Semantic Analysis,Association Rule Mining, Robust Collaborative Filtering,Keyword Matching, KNearest Neighbor algorithms, K-Means clustering and matrix factorization are recently utilized to address web usage mining by the researchers. Our Current Recommender system commonly uses content based filtering and collaborative filtering system algorithms. Content based filtering system generally generates recommendation based on the pre-defined or pre constructed user profiles by comparing the similarity of data accessed by those Users. Collaborative filtering system makes recommendations by using the rating of current user for items or products or web pages via referring other user's preference that is closely similar to the current user. As of today, CF system has been widely adopted in web recommendation applications. But here in OLAP Reports

we did not have an option of Rating a Report. Based on these techniques we need to come up with an Unified algorithm which automatically derives Rating to a Report based on Frequency of Access, Relevance in Dashboard, Number of key Metrics and KPI's it holds etc. The Key factor to consider here is Data Security. Only the authorized Reports should be visible for each User. This Unified algorithm should effectively support User Authentication and Data Confidentiality. We can also implement the user-based collaborative filtering techniques proposed in the Web context. In this work, the architecture of a Collaborative Query Management System targeted at large-scale, shared-data environments is outlined. As part of this architecture, the data mining techniques such as clustering or association rulescan be applied to the query logs in order to provide the users with query suggestions. By Implementing these techniques and algorithms an OLAP Report Recommender System can be developed.

## V. CONCLUSION AND FUTURE WORK

In this paper we proposed an OLAP Report Recommender System which helps OLAP Users to select the required report from the list of proposed reports quickly thus saving time and effort overheads. This System provides recommendations specific to each User based on his Role, Past History, Data Marts accessible to him etc. Appending these Recommender Systems to existing OLAP Reporting Tools greatly enhances their capabilities and eases lot of Overheads making the tools more Efficient. The Future work would be to enhance and implement these Algorithms to also identify thedata relevant to Future Predictions, Competitorsand Market Trends etc. and provide recommendations based on that, thus creating a Robust Report Recommender System.

## REFERENCES


[1] Deshpande, M., Karypis, G.: Item-based top-n recommendation algorithms. ACM Trans. Inf. Syst. 22(1), 143–177 (2004)

[2] Greco, G., Greco, S., Zumpano, E.: Collaborative filtering supporting web site navigation. AI Commun. 17(3), 155–166 (2004)

[3] Khoussainova, N., Balazinska, M., Gatterbauer, W., Kwon, Y., Suciu, D.: A case for a collaborative query management system. In: CIDR 2009: Proceedings of the 4th biennal Conference on Innovative Data Systems (2009)

[4] Singh, V., Gray, J., Thakar, A., Szalay, A.S., Raddick, J., Boroski, B., Lebedeva, S., Yanny, B.: Skyserver traffic report - the first five years. Microsoft Research, Technical Report MSR TR-2006-190 (2006)

[5] Koutrika, G., Ioannidis, Y.: Personalized queries under a generalized preference model. In: ICDE 2005: Proceedings of the 21st International Conference on Data Engineering, pp. 841–852 (2005)

[6] Adomavicius, G., Kwon, Y.: New recommendation techniques for multicriteria rating systems. IEEE Intelligent Systems 22(3), 48–55 (2007)

[7] Dresner H J, Hostmann B, Tiedrich A H and Buytendijk F: 'Magic Quadrants for Business Intelligence, 1H04', Research Report, Gartner (April 2004).

[8] Nauck D, Spott M and Azvine B: 'SPIDA - a novel data analysis tool', BT Technol J, 21, No 4, pp 104-112 (October 2003).

[9] Nauck D, Klawonn F and Kruse R: 'Foundations of Neuro-Fuzzy Systems', Wiley, Chichester (1997).

[10] H. A. Mahmoud, H. J. Moon, Y. Chi, H. Hacigumus, D. Agrawal, and A. El Abbadi, "Cloudoptimizer: multitenancy for i/o-bound olap workloads," in EDBT, 2013, pp. 77-88.

[11] J. Dean and S. Ghemawat, "Mapreduce: simplified data processing on large clusters," Commun. ACM, vol. 51, no. 1, pp. 107-113, 2008.

[12] H. You , "A Knowledge Management Approach for Real-time Business Intelligence" , in 2nd International Workshop on Intelligent Systems and Applications (ISA) , pp.1 -4

[13] W. Zhao, W. Dai and K. Yang , "The relationship of business intelligence and knowledge management" , Information Management andEngineering (ICIME), 2010 The 2nd IEEE International Conference on , pp.26 -29

[14] D. Zeng, H. Chen, R. Lusch, and S.-H. Li, "Social media analytics and intelligence," Intell. Syst. IEEE, Vol. 25, no. 6, pp. 13-16, 2010.

[15] . P. Melville, V. Sindhwani, and R. Lawrence, "Social media analytics: Channeling the power of the blogosphere for marketing insight," Proc WIN, 2009.


## ABOUT THE AUTHORS


J. Sreedhar obtained his PhD in Computer Science & Engineering from JNTUK-Kakinada,India. He has received M.Tech.(Computer Science and Technology) from Andhra University. He has been published and presented good number of Research and Technical papers in International Journals , International Conferences and National Conferences. Presently Working as Professor and Principal in ECET, Hyderabad.. His main research interests are NLP,IRS, Data Mining. and Compiler Design,FLAT,Computational Linguistics.

M. Neelima is a IT Professional in UGS, Hyderabad, India. She has received M.Tech.(Computer Science and Engineering) from JNTUH Hyderabad. Presently Working as an Sr.Software Engineer in UGS, Hyderabad. Her main research interests are NLP, IRS , Data Mining.

D.Ramesh obtained his PhD in Computer Science & Engineering from JNTUK-Kakinada,India. He has received M.Tech.(Computer Science and Engineering) from JNTUH-Hyderabad,India. He has been published and presented good number of Research and Technical papers in International Journals , International Conferences and National Conferences. Presently Working as Associate Professor in JNTUHCEJ, Hyderabad.. His main research interests are Image Processing,IRS, Data Mining.