

A RECOMMENDER SYSTEM USING DATA STORED IN HETEROGENEOUS DATABASE

Nilesh Subhash Mali¹ and Sachin Bojewar²

Abstract: Nowadays, most large-scale commercial as well as social websites uses recommender systems. The recommender system is important and become a necessity in various applications like online shop. A recommender system uses input from various heterogeneous databases to provide aggregate recommendations for the users. These heterogeneous databases are integrated without losing the consistency of the data by maintaining the business standards; such a challenging task is to be achieved with the development of ETL (Extract, Transform and Load) model. The ETL tool whose is based on ETL process can merge the data and product purchase history of different users stored in various heterogeneous databases. So it helps to improve the decision making process of users by providing recommendations about items or products which a user may wish to utilize in future. Thus, it is help to improve and achieve the business intelligence for any system.

Keywords: Data warehouses, ETL Process, ETL tools, Business Intelligence, heterogeneous databases

I. INTRODUCTION

Data Warehouse is a collection of large store of data accumulated from a wide range of heterogeneous or heterogeneous sources and used for the improvement of business growth. Heterogeneous database system is basically integration of heterogeneous, disparate database management systems [6]. Thus for extraction, transformation of data from such a heterogeneous database system and load it into some common structure ETL (Extract, Transform and Load) tool is used.

ETL is the process of extracting data from various sources of heterogeneous database, transforming that data into desired form by applying various rules for cleaning it and finally loading or storing it into the target database. ETL process is optimized using development of ETL tool [5], as they provide connectors to common data sources and easily as well as consistently transformations of data across various data sources [3]. It includes various operations likes filtering, reformatting, sorting, joining, merging, aggregation, etc. The ETL tools also used for transformation scheduling, version control, and unified management of metadata. This transformed data is used by recommender systems.

Recommender systems are software tools that are used to provide suggestions to a collection of users for items or products that might interest them. They help to find the relevant information about items or products quickly and also increase usability with a properly designed human-computer interaction. They represent a powerful method for enabling users

¹ Department of Computer Engineering, University of Mumbai, ARMIET, Shahapur, Thane, India

² Department of Information Technology, VIT, Mumbai, India

to sort and filter through massive amounts of data involve prediction of user responses or interests about items or products. So many commercial like suggestions for books on Amazon, or movies on Netflix as well as social websites uses recommender system to increase user interaction, especially on the web.

A hybrid recommender system is generally based on content of items called as content based recommender system [2] and collaboration approach [7]. Some of the main challenges faced by the web recommender systems are like cold-start, sparsity, etc. [1], where cold start problem can be solved by combining association rules and clustering technique [4].

II. METHODOLOGY

The ETL (Extract, Transform and Load) tool is developed which uses the data stored in heterogeneous databases and convert it into one target format known as data warehouse. The working of ETL tool is totally depends upon ETL process and provides consolidated output to the recommender system. A recommender system uses integrated data from collection of heterogeneous databases to give associated result to the user. For better association result FP (Frequent Pattern) growth algorithm is used.

III. EXPERIMENTAL RESULT

ETL tool extract data from heterogeneous databases like SQL server, oracle, and access. After extraction, it transposes the data in common form i.e. in SQL server. Finally, it loads the result into data warehouse which is in the form of SQL server.

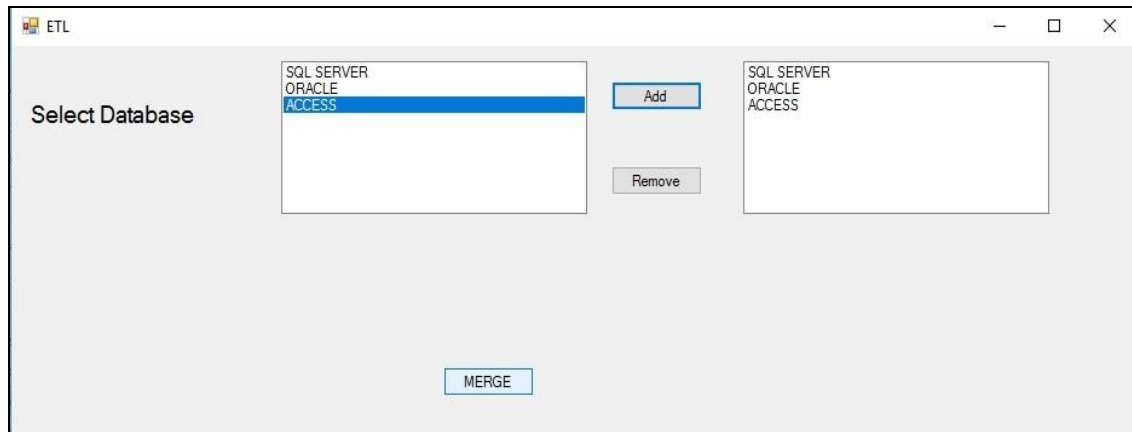
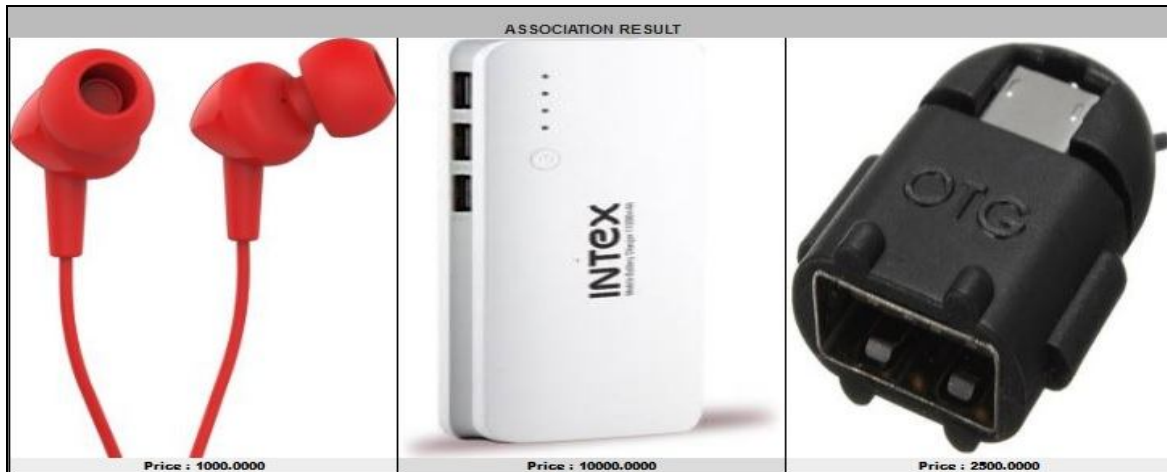


Fig. 1. ETL tool result

A recommender system uses ETL tool result to give better recommendations. Search is divided into three sections as first, it gives product suggestions based on search condition. Second, it gives the association result from data warehouse based on ETL tool for better result. Finally, it gives history of previous search products.



a)



b)



c)

Fig. 2. a) Searching result of recommender system for mobile b) Association Result c) Previous search history

IV. CONCLUSION

The experiments which we had carried out from our system with any browser can be used to accept input code. Our system takes input from integration of various heterogeneous databases using newly developed ETL (Extract, Transform and Load) tool. The recommender system is based on integrated data of purchase history of all users. It helps to improve the accuracy of system by combining the hidden relations between users and items.

V. FUTURE SCOPE

In past few years, recommendations are from few homogenous databases. In this paper, we focus on recommendations from heterogeneous databases. In future recommendations from distributed heterogeneous databases for better hybrid recommender system will be considered.

REFERENCES

- [1] Dr. Sarika Jain, Anjali Grover, Praveen Singh Thakur and Sourabh Kumar Choudhary published "Trends, Problems And Solutions of Recommender System" in IEEE International Conference on Computing, Communication and Automation (2015), pages 955-958.
- [2] Gerald Ninaus, Florian Reinfrank, Martin Stettinger and Alexander Felfernig published "Content-Based Recommendation Techniques for Requirements Engineering" in IEEE International Conference, in Karlskrona, Sweden (2014), pages 27-34.
- [3] Nithin Vijayendra, and Meiliu Lu published "A Web-based ETL Tool for Data Integration Process" in IEEE International Conference, in Sopot, Poland (2013), pages 434-438.
- [4] Hridya Sobhanam and A. K. Mariappan published "Addressing cold start problem in recommender systems using association rules and clustering technique" in IEEE 2013 International Conference on Computer Communication and Informatics (ICCCI -2013), Jan. 04 – 06, 2013, Coimbatore, INDIA.
- [5] Fang Ying-lan and Han Bing published "Design and Implementation of ETL Management Tool" in IEEE 2009 Second International Symposium on Knowledge Acquisition and Modelling, pages 446-449.

- [6] Vinit Nimkar, Kazi Fakir Mohammed and Sushopti Gawde published “Integration of Heterogeneous Distributed Database Based on E-Commerce System” in International Journal of Innovative Research in Computer Science & Technology (IJRCST), Volume-3, Issue-1, January – 2015.
- [7] Jaimeel Shah and Lokesh Sahu published “A Survey of Various Hybrid based Recommendation Method” in International Journal of Advanced Research in Computer Science and Software Engineering of Volume 4, Issue 11, November 2014, pages 367-371.