Towards Collaborative Multidimensional Query Recommendation with Triadic Association Rules

Sid Ali Selmane, Laboratoire ERIC, Université Lyon 2, Lyon, France
Omar Boussaid, Laboratoire ERIC, Université Lyon 2, Lyon, France
Fadila Bentayeb, Laboratoire ERIC, Université Lyon 2, Lyon, France

ABSTRACT

This paper describes a new personalization process for decisional queries through a new approach based on triadic association rules mining. This process exploits the decision query log files of end users and follows these five steps: (1) generation of a triadic context from the multidimensional query logs of OLAP query analysis server; (2) mapping the triadic context into the dyadic one; (3) computation of (conventional) dyadic association rules; (4) generation of triadic association rules through a factorization process of dyadic ones and convey a richer semantics. The aim of the personalization approach which is based on triadic rules is to recommend new decision queries to OLAP end users sharing some common properties. This paper aims at helping this class of users by recommending them personalized OLAP queries that they might use in their future OLAP sessions. To validate the approach, the authors developed a software prototype called P-TRIAR (Personalization based on TRIadic Association Rules) which extracts two types of triadic association rules from decision query log files. The first type of triadic rules will serve to the recommending queries by taking the collaborative aspect of OLAP users into account. The second type of triadic rules will enrich user queries. Preliminary experiments were conducted on both real and synthetic datasets to assess the quality of the recommendations in term of precision and recall measures, as well as the performance of their on-line computation.

Keywords: Collaborative OLAP, Decisional Query, Personalization, Three-Dimensional Data, Triadic Association Rules, Triadic Concept Analysis

INTRODUCTION

Recommender systems have become extremely common in recent years, and are applied in a variety of applications such as music, news, research articles, search queries, social tags, and products in general. One of the goals of recommender systems is to help users navigating large amounts of data. Existing recommender systems are usually categorized into content-based methods and collaborative filtering methods (Adomavicius & Tuzhilin, 2005). Applying recommendation
technology to multidimensional databases is an emerging and promising topic, especially for recommending decision queries (Chatzopoulou, Eirinaki, & Polyzotis, 2009); (Stefanidis, Drosou, & Pitoura, 2009). It is of particular relevance to OLAP domain analysis which is inherently tedious since the user has to navigate through large data cubes to find valuable information.

OLAP systems users indeed formulate multidimensional queries to meet their specific needs for decision support. OLAP tools are known to be intuitive as their end users are not necessarily computer scientists. However, the large volume of data and the complexity of analytical queries which involve a lot of aggregations make this task of analysis more difficult to users. So it seems necessary to provide them solutions best suited to their way of thinking through methods of recommendation and enrichment of their analytical tasks. These methods are known under the name of personalization. In this paper, we propose a new personalization process of analytical queries to help users in their decision making. We are particularly interested in collaborative recommendation and enrichment of users’ decision queries based on their query log files.

Personalization research works which exploit query log files use in most cases frequent itemsets (Khemiri & Bentayeb, 2013) and association rules (Veloso, de Almeida, Gonçalves, & Meira, 2008). However, the large number of obtained frequent itemsets and association rules makes the task of personalization more difficult due to their volumetrics. Contrary to these approaches, the work we propose in this paper is based on another type of rules, more compact, called triadic association rules (Biedermann, 1997). These rules convey a richer semantics than conventional ones as they include a condition, in addition to the premise and the conclusion. Our personalization process is composed of five steps:

1. Modeling users query log files of OLAP servers by a triadic context. This triadic context is composed of the users set, the queries set, and the attributes set (descriptors and measures) in the SELECT clause and a ternary relation between these three sets;
2. Mapping triadic (tridimensional) context into a dyadic (bidimensional) one. This is done by flattening the set of users over the set of attributes;
3. Computing dyadic association rules \((\text{premise} \rightarrow \text{conclusion})\);
4. Generating triadic association rules \((\text{premise} \rightarrow \text{conclusion} | \text{condition})\) through a factorization process of dyadic ones;
5. Exploiting these triadic association rules for personalization.

To validate our collaborative recommender system in OLAP environment, we develop a software prototype \(P-\text{TRIAR}\) (Personalization based on TRIadic Association Rules). \(P-\text{TRIAR}\) extracts two types of rules from users query log files. The first ones will serve in query recommendation process by taking the collaborative aspect of users into account during their analysis sessions. This recommendation process will be carried out by the user communities discovered across multiple links between them. The second type of the extracted rules aims at enriching user queries by recommending to them relevant attributes that allow creating new decision queries. Preliminary experiments were conducted to assess the quality of the recommendations in term of precision and recall, as well as the performance of their on-line computation.

RUNING EXAMPLE

In this section, we illustrate our personalization process, through a simplified example, from \(PUBS\) data warehouse (Figure 1). \(PUBS\) concerns the analysis of turnover (TO) and quantity (Qty) of books sold. These measures are observed over the following dimensions: \(Titles\) (title_id,
An Interactive Spatial Decision Support System Enabling Co-Located Collaboration using Tangible User Interfaces for the Multiple Capacitated Facility Location Problem
www.igi-global.com/article/an-interactive-spatial-decision-support-system-enabling-co-located-collaboration-using-tangible-user-interfaces-for-the-multiple-capacitated-facility-location-problem/133209?camid=4v1a

The SccoB Process: An Integration of the Exploitory and Exploratory Processes Through a Self-Sustaining Process of Knowledge Creation
www.igi-global.com/article/the-sccob-process/215351?camid=4v1a

From Human Decision Making to DMSS Architecture
www.igi-global.com/chapter/human-decision-making-dmss-architecture/8061?camid=4v1a