View Materialization in a Data Cube: Optimization Models and Heuristics

Vikas Agrawal, Fayetteville State University, USA
P. S. Sundararaghavan, The University of Toledo, USA
Mesbah U. Ahmed, The University of Toledo, USA
Udayan Nandkeolyar, The University of Toledo, USA

ABSTRACT

Data warehouse has become an integral part in developing a DSS in any organization. One of the key architectural issues concerning the efficient design of a data warehouse is to determine the “right” number of views to be materialized in order to minimize the query response time experienced by the decision makers in the organization. We consider a bottleneck objective in designing such a materialization scheme which has the effect of guaranteeing a certain level of performance. We examine linear integer programming formulations, and develop heuristics and report on the performance of these heuristics. We also evaluate heuristics reported in the literature for the view materialization problem with a simpler objective.

Keywords: data cube; data warehouse; view materialization

INTRODUCTION

In today’s fast-paced, ever-changing and wants-driven economy, information is seen as a key business resource to gain competitive advantage (Haag, Cummings & McCubbrey, 2005). Effective use of this information requires good decision support systems. Most decision support systems require reliable and elaborate data backbone which needs to be converted into useful information. With the widespread availability and ever-decreasing cost of computers, telecommunications technologies, and Internet access, most businesses have collected a wealth of data. However, that is only the first and easy step. Many firms are becoming data rich but remain information and knowledge...
Accessing data from a data cube, if not materialized, can be a time consuming and resource intensive process. A data cube consists of many views with existing interrelated dependencies among themselves (such view is also known as a cuboid or a query). If such a view is stored, it is denoted as a materialized view. The problem of quick and easy access to the data cube may be alleviated by an efficient selection of a set of views to be materialized. Since not all views in a data cube may be materialized due to constraints imposed on the system, selecting the right set of views to materialize is an integral part of the design of data cube and its associated views. An efficient design will dramatically reduce the execution time of decision support queries and hence prove pivotal in delivering competitive advantage.

Many researchers have studied the problem of selecting the “right” set of views to be materialized in a data cube in order to minimize decision support query response time. The problem is generally described as the materialized view selection (MVS) problem, which has the objective of minimizing the access time subject to constraints on either the number of views that may be materialized or the storage space that may be used for materialization of views (Gupta & Mumick, 2005; Harinarayan, Rajaraman & Ullman, 1996, 1999). In this article we have worked on several variants of the MVS problems and have solved these optimally as well as using heuristics. Our specific contributions may be summarized as follows:

- We have presented a linear integer programming formulation for two versions of the MVS problem with a bottleneck objective, which minimizes the maximum weighted access time experienced by any class of users.
- We have developed heuristics for the above two problems and reported on their performance.
- We have also presented linear integer programming formulations for the general MVS problem reported in Harinarayan et
Related Content

The Quality of Data Representations Developed by NonExpert Designers: An Experimental Study
[www.igi-global.com/article/quality-data-representations-developed-nonexpert/51125?camid=4v1a](www.igi-global.com/article/quality-data-representations-developed-nonexpert/51125?camid=4v1a)

An MDA Approach and QVT Transformations for the Integrated Development of Goal-Oriented Data Warehouses and Data Marts
Jesús Pardillo, Jose-Norberto Mazón and Juan Trujillo (2011). *Journal of Database Management* (pp. 43-68).
[www.igi-global.com/article/mda-approach-qvt-transformations-integrated/49723?camid=4v1a](www.igi-global.com/article/mda-approach-qvt-transformations-integrated/49723?camid=4v1a)
Interactive Query Expansion with Automatically Generated Category-Specific Thesauri
[www.igi-global.com/chapter/interactive-query-expansion-automatically-generated/30274?camid=4v1a](www.igi-global.com/chapter/interactive-query-expansion-automatically-generated/30274?camid=4v1a)

Information Technology in the Practice of Law Enforcement
Susan Rebstock Williams and Cheryl Aasheim (2006). *Cases on Database Technologies and Applications* (pp. 81-103).
[www.igi-global.com/chapter/information-technology-practice-law-enforcement/6206?camid=4v1a](www.igi-global.com/chapter/information-technology-practice-law-enforcement/6206?camid=4v1a)