Chapter 7
Dynamic View Management System for Query Prediction to View Materialization

Negin Daneshpour
Amirkabir University of Technology, Iran

Ahmad Abdollahzadeh Barfourosh
Amirkabir University of Technology, Iran

ABSTRACT

On-Line Analytical Processing (OLAP) systems based on data warehouses are the main systems for managerial decision making and must have a quick response time. Several algorithms have been presented to select the proper set of data and elicit suitable structured environments to handle the queries submitted to OLAP systems, which are called view selection algorithms to materialize. As users’ requirements may change during run time, materialization must be viewed dynamically. In this work, the authors propose and operate a dynamic view management system to select and materialize views with new and improved architecture, which predicts incoming queries through association rule mining and three probabilistic reasoning approaches: Conditional probability, Bayes’ rule, and Naïve Bayes’ rule. The proposed system is compared with DynaMat system and Hybrid system through two standard measures. Experimental results show that the proposed dynamic view selection system improves these measures. This system outperforms DynaMat and Hybrid for each type of query and each sequence of incoming queries.

INTRODUCTION

OLAP is defined as online analytical processing system to answer the multidimensional queries (Dehne et al., 2008; Lawrence & Rau-Chaplin, 2008; Ravat et al., 2008). Multidimensional queries are complex and operate on huge amount of data, furthermore; these queries are used to managerial decisions in decision support systems (DSS) and data mining.

Multidimensional structures are used to decrease query response time. Multidimensional
structures, data cube, are the structures of the Data warehouses to represent data sources.

To achieve analytical process of queries, data cubes store data in different summarization degree related to the aggregation function type. When we have multidimensional data, we can construct a lattice of cuboids which contains data in different level of summarization. The cuboid which stores data in the minimum level of summarization is called “base cuboid” and another cuboid which stores data in the maximum level of summarization is called “apex cuboid”.

Data cubes are pre-computed and stored in data warehouses in the form of materialized views to improve query response time. Data cube computation is time and money consuming and various researches have been done to improve query response time based on parallel processing, index selection and view selection (Agrawal, Chaudhuri, & Narasayya, 2000; Agrawal, Chaudhuri, Kollar, Marathe, Narasayya, & Syamala, 2004; Agrawal, Narasayya, & Yang, 2004; Asgharzadeh Talebi et al., 2008; Chaudhuri, 1997; Le et al., 2007; Taniar et al., 2008; Taniar & Wenny Rahayu, 2002a, 2002b, 2002c, 2002d, 2004).

We focus on view selection techniques which are the main issue to construct data warehouses (Ahmed et al., 2007; Aouiche et al., 2006; Aouiche & Darmont, 2009; Choi et al., 2003; Gong & Zhao, 2008; Gupta, 1997; Gupta & Mumick, 2005; Harinarayan et al., 1996; Hung et al., 2007; Kalnis et al., 2002; Kotidis & Roussopoulos, 1999, 2001; Lawrence & Rau-Chaplin, 2008; Mahboudi et al., 2006; Nadeau & Teorey, 2002; Phan & Li, 2008; Ramachandran et al., 2005; Shah et al., 2006; Shukla et al., 1998; Valluri et al., 2002; Xu et al., 2007; Zhang et al., 2003).

Other important issues in data warehousing are: multidimensional design methodologies, partitioning methods, refreshment mechanisms, building XML data warehouses, and warehousing XML documents (Bellantreche et al., 2009; Chen et al., 2010; Maurer et al., 2009; Romero & Abello, 2009; Rusu et al., 2005, 2006, 2009).

Three choices for view materialization are reported (Han & Kamber, 2006):

1. **Base Cuboid Materialization:** In this choice, only base cuboid, which can be used to answer all multidimensional aggregated queries, is pre-calculated and materialized.

2. **Full Materialization:** In this choice, all of the cuboids, which are answers to all multidimensional aggregated queries, are pre-calculated and materialized.

3. **Partial Materialization:** In this choice, the proper subset of the whole set of possible cuboids which is the answers to some multidimensional aggregated queries, is selected, pre-calculated and materialized.

The first choice requires minimum amount of memory to materialize a view which is sufficient to calculate the answers of all multidimensional aggregated queries; but it could be slow and leads to expensive process of computing multidimensional aggregates on the fly.

The second choice has quick response time; but it requires huge amounts of memory space in order to store all of the pre-computed cuboids and it will be time consuming processes to maintain all of them.

The third choice, which could be more efficient, presents an interesting trade-off between storage space and response time.

There are several works that present techniques for partial materialization which we call view selection techniques (Ahmed et al., 2007; Aouiche et al., 2006; Aouiche & Darmont, 2009; Asgharzadeh Talebi et al., 2008; Choi et al., 2003; El-Helw et al., 2009; Gong & Zhao, 2008; Gou et al., 2006; Gupta, 1997; Gupta & Mumick, 2005; Hanusse, 2009; Harinarayan et al., 1996; Hung et al., 2007; Kalnis et al., 2002; Kotidis & Roussopoulos, 1999, 2001; Lawrence & Rau-Chaplin, 2008; Lin & Kuo, 2004; Liu et al., 2008; Mahboudi et al., 2006; Nadeau & Teorey, 2002; Phan & Li, 2008; Ramachandran et al., 2005; Sapia, 1999;
Related Content

**Big Data and National Cyber Security Intelligence**

**A Survey on Database Performance in Virtualized Cloud Environments**
[www.igi-global.com/article/survey-database-performance-virtualized-cloud/67571?camid=4v1a](www.igi-global.com/article/survey-database-performance-virtualized-cloud/67571?camid=4v1a)

**Vertical Fragmentation in Databases Using Data-Mining Technique**
[www.igi-global.com/article/vertical-fragmentation-databases-using-data/1812?camid=4v1a](www.igi-global.com/article/vertical-fragmentation-databases-using-data/1812?camid=4v1a)

**An Encryption Methodology for Enabling the Use of Data Warehouses on the Cloud**
[www.igi-global.com/article/an-encryption-methodology-for-enabling-the-use-of-data-warehouses-on-the-cloud/215005?camid=4v1a](www.igi-global.com/article/an-encryption-methodology-for-enabling-the-use-of-data-warehouses-on-the-cloud/215005?camid=4v1a)