



RESEARCH NOTE

Data Warehouse Interoperability for the Extended Enterprise

Aristides Triantafillakis, National & Kapodistrian University of Athens, Greece
Panagiotis Kanellis, National & Kapodistrian University of Athens, Greece
Drakoulis Martakos, National & Kapodistrian University of Athens, Greece

ABSTRACT

The purpose of this paper is to raise awareness and identify a number of challenges regarding the issue of data warehouse interoperation in Web-based collaborative environments. Adopting system and information quality as success variables, we argue that existing works fell short of addressing complex issues that relate to their refreshment and extent far beyond view maintenance solutions within single warehouses. Considering a solution that approaches warehouse refreshment as a business process in a federation of data warehouses, we define a special kind of materialized view that emanates from such an environment, provide a roadmap for implementing the appropriate warehousing architecture, and give some preliminary empirical results.

Keywords: data warehouse refreshment; electronic commerce; Federation of Data Warehouses; view maintenance

INTRODUCTION

A base requirement for the success of a data warehouse is the ability to provide decision makers with both accurate and timely consolidated information (information quality) as well as fast query response times (system quality) (Chen, Soliman, Mao, & Frolick, 2000). For this purpose, a common method that is used in practice for providing higher information and system quality is the concept of stored result sets, that is, materialized views

(MVs), where a query is more quickly answered against the MV than querying directly the base data sources. However, MVs can improve query performance only if we can manage to update them consistently (Do, Drew, Jin, Jumani, & Rossum, 1998).

Collaborative electronic commerce (Ce-commerce) augments this challenge because the data sources are not only internal, as they largely were a mere few years ago. For example, the emergence of business communities in the form of busi-

ness-to-business (B2B) exchanges means that the boundaries of organizations are more fluid than they used to be (Yang & Papazoglou, 2000). 'Extended enterprises' have to integrate far more data into a single repository originating outside the organization. A recent paper (Hammer, 2001) underlines this trend by felicitously pointing out that "streamlining cross-company process is the next great frontier for reducing costs, enhancing quality, and speeding operations" (p. 83).

The implications for data warehouse information and system quality is that we should start considering the data warehouse refreshment (DWR) as a business process that must provide explicit support for cross-enterprise collaboration. Therefore, DWR should not only be limited to MV maintenance in the case of a single warehouse (as we are accustomed to), but also support the refreshment in a federation of data warehouses. This, in turn, translates to providing a new set of algorithms and techniques to materialize views from source data that may reside in MVs of remote data warehouses, and to incrementally maintain these views. Besides, separate DWR processes in separate data warehouses augments this challenge because there might exist different maintenance policies on the MVs of interest in each DWR process. To the best of our knowledge, previous work on view maintenance has mainly considered the case of (Select-Project-Join) SPJ views in a single warehouse, while not providing insights for a data warehousing architecture in such an environment. In short, the problem addressed in this paper can be stated as follows: *"how to maintain MVs in environments where a data warehouse utilizes data from other data warehouses."*

Adopting system and information quality as success variables (DeLone &

McLean, 1992), we argue that DWR is more complex than the MV maintenance problem, and we introduce a special kind of MV that emanates from such an environment. In the next section we demonstrate the data warehouse refreshment process in this environment. We then provide the reader with a roadmap for implementing a warehousing architecture of this kind, and present some empirical results while our conclusions close the paper.

DATA WAREHOUSE REFRESHMENT PROCESS

In the case of a single data warehouse, DWR is a complex process and may be composed of asynchronous and parallel tasks. The main concept is to capture the differential changes in the data sources and to provide an efficient process for the update propagation and customization steps of the summarized data to the data marts (Bouzeghoub, Fabret, & Matulovic-Broque, 1999).

However, an integrated value system, such as a federation of DWs, necessitates a solution where the enterprise system of each participating business in the value chain should be able to communicate with other such systems in order to accomplish a task. In particular, a form of B2B exchange that has gained a lot of attention due to its potential benefits is continuous replacement planning (CRP), where a manufacturer supplies the distributors before requested. In this way multiple participating enterprises within a shared market segment can collaboratively plan and manage the flow of goods, services, and information along the value chain in a way that increases customer-perceived value and optimizes at the same time internal efficiencies (Yang & Papazoglou, 2000).

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/data-warehousing-interopability-extended-enterprise/3317

Related Content

Cooperative IDS for Detecting Collaborative Attacks in RPL-AODV Protocol in Internet of Everything

Erukala Suresh Babu, Bhukya Padma, Soumya Ranjan Nayak, Nazeeruddin Mohammad and Uttam Ghosh (2023). *Journal of Database Management* (pp. 1-33). www.irma-international.org/article/cooperative-ids-for-detecting-collaborative-attacks-in-rpl-aodv-protocol-in-internet-of-everything/324099

Common Sense Reasoning in Automated Database Design: An Empirical Test

Veda C. Storey, Robert C. Goldstein and Jason Ding (2002). *Journal of Database Management* (pp. 3-14). www.irma-international.org/article/common-sense-reasoning-automated-database/3272

Modeling Design Patterns for Semi-Automatic Reuse in System Design

Galia Shlezinger, Iris Reinhartz-Berger and Dov Dori (2012). *Cross-Disciplinary Models and Applications of Database Management: Advancing Approaches* (pp. 29-56). www.irma-international.org/chapter/modeling-design-patterns-semi-automatic/63661

Graph Applications in Chemoinformatics and Structural Bioinformatics

Eleanor Joyce Gardiner (2012). *Graph Data Management: Techniques and Applications* (pp. 386-420). www.irma-international.org/chapter/graph-applications-chemoinformatics-structural-bioinformatics/58620

BACIIS: Biological and Chemical Information Integration System

Zina Ben Miled, Nianhua Li and Omran Bukhres (2005). *Journal of Database Management* (pp. 72-85). www.irma-international.org/article/baciis-biological-chemical-information-integration/3338