Chapter 13
Real-Time BI and Situational Analysis

Maik Thiele
Technische Universität Dresden, Germany

Wolfgang Lehner
Technische Universität Dresden, Germany

ABSTRACT

In the past, data-warehouse systems served as information providers for key management members and knowledge workers; today, they are the central platform for the enterprise-wide integrated information provision. Aside from strategic analyses on historical data, this encompasses, above all, the submission of real-time data to operational processes. The clear separation of the operational and the analytical world, as it has been promoted until now, will thus become obsolete in the future. This development (i.e., the merging of the operational and the analytical world in context of Web technologies) is topic of this chapter. For this purpose, we take a closer look at current business intelligence (BI) and data warehouse trends from the perspective of both the applications and the database systems. We discuss scenarios which show the emerging trend to real-time business intelligence and present techniques that address the problems upcoming when building a real-time DWH. Additionally, we extend the concept of real-time business intelligence by the aspect of situational data analyses.

INTRODUCTION

Business intelligence and data warehousing have enjoyed immense popularity and success over the last years and now play a key role in strategic corporate decision-making. Through the evolution of web technologies, the ubiquitous presence of data and the increasing demand for current data in the recent years the requirements posed to business intelligence systems have changed: 1) On the back-end and system level this evolution raised the need for more up-to-date, so-called real-time, analyses. The real-time aspect in the context of business intelligence describes a new
processing model where every change is automatically captured and pushed into the underlying data warehouse (DWH). Thus, the data in a real-time DWH is subject to continuous changes, denoted as a trickle-feed of updates. 2) On the end-user and application level the shift from static to dynamic web applications through modern web technologies leads to a raising demand for interactive and customizable reports as well as ad-hoc integrations of arbitrary data sources. Especially the latter aspect is not compatible with today’s business intelligence and data warehouse architectures.

Some applications scenarios should strengthen the two challenges sketched above: One application scenario from the business sector is the avoidance of so-called out-of-stock items; these are situations in which goods are requested but not available for purchase any longer. In order to minimize the number of occurrences of such situations, we require historical sales transaction data as input for prediction models on the one hand but also information on the current inventory and transaction data for consistency checks and refinements of the prediction data on the other hand. Additionally, an analyst could build a situational application to see how weather is affecting sales at specific retail stores and the occurrences of out-of-stock situations. In order to do so the analyst has to integrate public information from weather Web sites together with the sales information.

In e-commerce scenarios the current behavior of customers must be analyzed with the help of click-streams and basket analyses and then compared to similar behavioral patterns of past customers. From such comparison, we can derive so-called cross-selling and/or up-selling opportunities (i.e., the chance to recommend complementary or better products and thereby increase the company’s revenue).

Another real-time scenario focuses on supporting customer calls in call centers: Employees of a call center must always have access to current customer data in order to ensure the smooth and consistent communication with clients. Such information must be gathered via all communication channels (web, e-mail, phone etc.) of a company and then made available in real-time.

Airport operators, whose business model is based on the optimal aviation management and ground handling, are highly dependent on operation information, too. A plane that is taking off or touching down can be assigned up to 1,000 attributes, including, for example, the airport weather, the plane’s load, or the passengers’ nationalities. These data are compared to past data, statistics, and predictions in order to control the mentioned processes. One example is the scheduling of ground personnel for visa processing in dependence on the passengers’ nationalities. In order to compare the performance of the own aviation management with other airports, business analysts may build a situational application which integrate third-party flight delay statistics with their own information. Furthermore, they may rely on external web services to improve and visualize their results.

These example scenarios show that the currently strict separation of operational and analytical systems on a system level, as it has been propagated until now, will thus become obsolete in the near future. As the central information provision platform within enterprises, data warehouses will thus simultaneously have to provide up-to-date information and qualitative, verified data that has been refined over the course of many processing steps. Data-warehouse systems are thus faced with a heterogeneous group of users who have different requirements in terms of freshness, query latency and the consistency and stability of data. Looking at our application scenarios, we see that a data warehouse does not only have to provide real-time data to detect out-of-stock items, but it also needs to provide data to classic applications such as controlling or marketing.

On an application level there is evident lack of support for ad-hoc integration in today’s business intelligence and data warehouse architectures. The reason for this lies in the large amount of