Data warehouse solutions for customer relationship management

N. Radenkovic̀
University in Priština, Kosovska Mitrovica, Serbia
e-mail: natasa_radenkovic@yahoo.com

In today's highly competitive business environment, CRM (customer relationship management) systems, which provide the framework for analyzing customer profitability and improving marketing effectiveness, have become an indispensable component in enterprise information systems. Typically, CRM activities include data analysis, campaign design, response analysis of customer data. To effectively support such activities, a data warehouse (which is a repository that integrates information from multiple operational data sources) must be developed to act as the backbone of CRM systems. A data warehouse is a core part that determines the performance of CRM systems and quality of CRM services.

Keywords: customer relationship management (CRM), data warehouse (DWH), operational data storage (ODS).

Introduction

Customer relationship management (CRM) as an advanced marketing method which combines customer contentment management, marketing strategies and IT technologies, has become an irreplaceable component of information systems found in numerous enterprises. Effective support to these processes requires a data warehouse (repository which integrates data from multiple operational resources) that serves as a "backbone" of the entire CRM system [1]. Data warehouse is the core of a CRM system which determines the features of the entire system and the quality of CRM services. A data warehouse for CRM must be customer-oriented with a unique consideration of data such as customers' demographic data, purchasing habits, return information on purchased goods etc. In this paper, we will explain some of the issues that arise during the implementation of a customer-oriented data warehouse and suggest an architecture which resolves some of the said issues.

1. Conceptual model

New business solutions have largely contributed to the evolution of data warehousing. When it comes to customer relationship management, we must emphasize that the information is fundamental. An organization cannot be successful in customer relationship management unless it possesses high-quality, timely and accurate information. If we want to possess such information we will need a customer-oriented data warehouse. Hence, Todman's conceptual

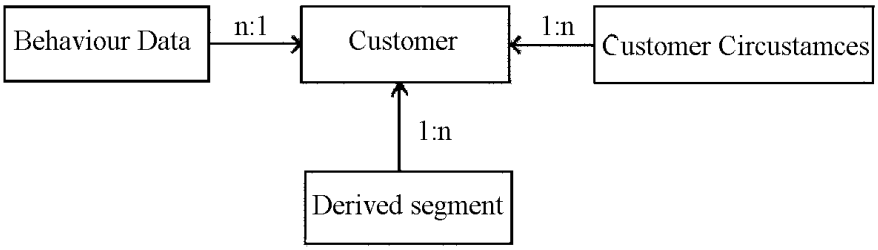

Fig. 1. General conceptual model by Todman

model [2] was utilized in this paper, as it proposes data warehousing comprised of the following information.

1. *Customer behavior data* are historical records of transactions or customer contact details. They represent a chronological outline of customers' behavior in relation to the enterprise.

2. *Changing Customer Circumstances Data* usually are attributes in the customer data, e. g.: address, marital status, age, telephone or bank account number. These can be classified into a category of slowly-altering data [3]. Alterations of these data are not as frequent as alterations in behavioral data but they have considerable impact on the semantics of queries. For example, changing of the customer's address may entail changing of customer's sales region, but the past sales volumes of this customer should not be transferred to the new region. If this becomes the case, it would result in the distorted sales totals per region.

3. *Derived customer segment data* are data calculated from attributes of the above data categories, which serve to segment the customer base, e. g., the lifetime value, the propensity to churn, the possibility for cancellation. If an enterprise intends to track the progress of its customers during their lifecycle, it also has to track these derived data over time. These external data can be integrated in the customer database and utilized for planning and envisaging.

This leads to a conceptual customer model as depicted in Fig. 1 [2], which has divided a single object Customer into four sub-entity types. The original object Customer is in an 1:n relationship with the three new data categories. This approach takes into consideration the above mentioned requirements related to customers analyses over periods of time. Therefore, besides all mentioned dimensions, a time dimension must also be included. Consequently, the queries on this data structure become more complex due to the fact that a historical presentation of dimensional data is necessary.

In addition to issues the time dimension puts on the customer dimension, there are further issues related to customer data which can grow into a serious challenge to CRM and data warehousing:

— customer data in different operational data sources cause data inconsistency;

— exponential growth of customer data reduces the efficiency of data warehousing.

2. Implications of multiple customer contact environments

When a customer-oriented model is being implemented, enterprises have a goal to align their activities with parameters derived from customer data collection. These data are collected

by sales over multiple sales channels (e. g. direct and indirect sales) and by different types of contacts between customers and enterprises (e. g. fax, email, postal service etc.). Thus, the data are spread into different databases and applications within the enterprises, e. g. in marketing department database, sales department, accounting or ERP system (Enterprise Resource Planning). Consequently, in most cases, this partial information is not integrated and therefore inconsistent. If these data are not carefully processed before loading into the warehouse, the quality of the data will be questionable.

Due to inconsistency, the enterprise loses its "unique overview" of the customer i. e. unique and integrated set of customer/client data. For example, different departments within an enterprise are often unaware that they have been providing service to the same customer or doubled customer data contain contradictory attribute values. These occurrences are common and mostly represent the reasons why a customer had been inadequately served. The increased expenses (e. g. due to multiple contacts with a customer) or loss of sales (due to inadequate service, customer turns to another salesperson) then become apparent.

3. Implications of exponential growth of customer data

If we want to assess behavior of customers, we must track their behavior from the moment they started purchasing or using our enterprise's services, until their retirement. In that case, the enterprise must collect and store data related to all transactions a customer had made in the highest level of details, despite the fact that the quantity of data for analyses will grow exponentially. The amount of data has negative influence on the performance and the response time of the data warehouse. Depending on the type of analyses that needs to be conducted in the warehouse, these characteristics (performance and response time) may be more or less important:

1) overall CRM analyses, e. g. for customer segmentation, implies access to a large quantity of data but the results are not necessarily immediately required;

2) applications of operational CRM may require access to certain values in derived segments of object Customer only but the results have to be produced in near real time because they are required by centers for direct relations, such as in e. g. call centers or web portals.

As seen above, the data amount and the response time are the two conflicting criteria in data warehousing. Since the existence of both criteria is justified, they have to be covered in an adequate concept of data warehousing.

4. Customer-oriented data warehouse architecture

Depending on which CRM operations have to be conducted in a data warehouse, the architecture of a warehouse should fulfill some of the basic requirements, specified in the following table.

Proposed data architecture must be capable of fulfilling all these opposing, even contradictory requirements of data warehousing, considering its utilization and it present an extension of [4]. ODS (Operational Data Storage) — extracts data from the transaction systems via the data interface several times per day. In this first-level ETL step (Extract, Transform, Load) — data are barely copied and slightly altered so that the ODS performance is not decreased. Important element is the influence of time on the data because it is vital for creating of chronological data reporting in a later stage. In a customer-oriented

	Analytical CRM	Operational CRM
Usage of customer relevant data	Extensive analyses, expectation of complex queries	Standard reporting queries
Requirements	Detailed data	Response to a query must be received in near real time, pre-aggregated data

model, this is crucial. Operational data storage contains a short history of low-level data. Depending on computer's capabilities, the ODS may contain information for a period of one month, e. g. for the current month.

In the second-level ETL step, data is periodically transferred from the ODS (Operational Data Storage) into DWH (Data Warehouse) or data warehouse retrieves the data from the OSD. In this short step, filtering, harmonization, aggregation and enrichment of data take place, and comprise a major part of data warehouse resources.

Only data warehouse contains a chronological presentation of detailed information on customers' behavior but also the calculated values i. e. derived information related to customers. Todman's general conceptual model served as a basis for customer-oriented architecture of an enterprise. DWH provides updated high-level detail data and derived data required by analytical applications such as analytical CRM.

Important characteristic of this architecture is a possibility to return summarized data from DWH into the operational base of ODS. Thus, a large number of standard queries, which could block the resources for more complex queries, can be transferred from DWH into ODS.

In an idealized architecture OSD could return updated data back into operational system by maintaining consistency between the operational systems but that is another topic for discussion (Fig. 2).

Another important characteristic of this architecture is that the OSD can provide both updated data and data derived by applications of operational CRM. This approach improves DWH performance because it can be freed from a large quantity of standard queries that cannot be conducted on a standard operational warehouse. In this architecture, the applications of operational CRM are intentionally not counted to operational systems because

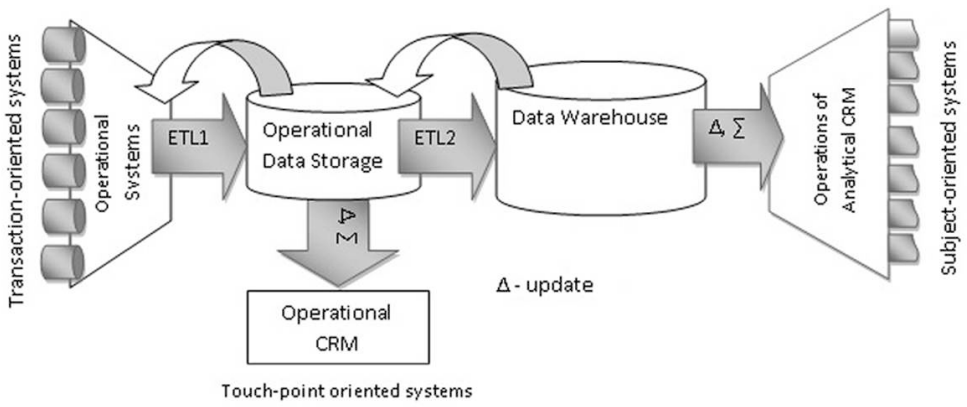

Fig. 2. Architecture of the Data Warehouse for CRM

they usually represent data interfaces of standard databases designed with possibility to be adjusted by existing data infrastructures. Therefore, in such architecture, we do not consider them as transaction-centric systems regardless of their transaction-oriented nature.

Conclusion

This paper explained the influence of implementing a customer/client relationship management system on data warehousing and presented an expanded data warehouse architecture designed for the needs of that system. Also, the goal of this paper was to emphasize it was not sufficient to change the organization of an enterprise or purchase new software when a system for customer management was implemented and when enterprise decides to focus on the customer. When a customer becomes the center of enterprise's interest, as stated above, there are major implications on the data used by management in decision making process. These changes have to be incorporated in a customer-oriented data warehouse. The purpose of architecture presented in this paper is to provide consistency of the data related to customer and to improve performance of execution of queries in data warehousing.

References

[1] McKnigh W. The CRM-Ready Data Warehouse. `http://www.dmreview.com`

[2] Todman C. Designing a Data Warehouse — Supporting Customer Relationship Management. Prentice Hall PTR, 2000.

[3] Kimball R., Ross M. The Data Warehouse Toolkit. Wiley Computer Publ., 2002.

[4] Inmon W.H. Building the Data Warehouse. John Whiley & Sons, 1996.

Received for publication 14 April 2010