Winter 12-5-2005

# Application of Business Intelligence Techniques in China Telecom

Haigang Li

# Application of Business Intelligence Techniques in China Telecom

Li Haigang
School Of Management, Shanghai Jiao Tong University, P.R. China, 200052
hglihg@yahoo.com.cn

**Abstract:** In this paper, we have developed a data warehouse and an OLAP( on-line analytical processing ) based framework that has been used for customer profiling and comparison. The system architecture for OLAP and data warehouse based calling behavior profiling and multilevel multidimensional pattern analysis is introduced. First, the method how customer profiles can be represented as data cubes is described, then, the architecture of a profiling engine is presented, finally, the process of using the engine to compute profiles and calling patterns is discussed and an application case of China telecom is studied.

## I. Introduction

In the last decade, China's telecom industry has made great progress. The reopening of market and the lifted ban of government monopoly changed its whole situation and brought it development and innovation of great significance. Nowadays its scope of business has become increasingly extensive, and more and more services are becoming available. On the other hand, the development of basic institution doesn't match the increasing demand. And the expense on 3G has been heavy burden to the telecom industry. Then they turn to some basic aspects, such as lowering the operation cost and maintaining the service of their current customers.

Though to support the commercial operations, telecom companies have spent plenty of money and energy installing automatic systems, the real problems are always more complicated than the systems. For instance, they do know who their customers are and who may be interested in their new service. But in fact, they are not able to tell which customer are those who can bring them more profit, or which kind of service is more profitable, or which province should be paid more attention.

With the deepening of informization in China, especially the popularization of the data analytical software, the data accumulation of the users has come to a certain scale. But most of the users just do some simply or partial enquiry dealing with the data. In some sense, data may be burden rather than resource to them. For the business intelligence applications can do the analysis and mining of large volumes of transaction data to support business managers in making informed decisions, the situation can be improved.

## II. Relative Concepts

Proceedings of the Fifth International Conference on Electronic Business, Hong Kong, December 5-9, 2005, pp. 101 - 104.

As a technology of transforming the information of a company to useful resource as well as knowledge, business intelligence gathers, manages and analyzes large amounts of data on a company's customers, products, services, operations, suppliers, and partners and all the transactions between them. Business intelligence applications require the analysis and mining of large volumes of transaction data to support business managers in making informed decisions. Typically, these applications involve extracting data from operational databases, transforming the data, and loading it into data warehouses. The data in the warehouse is then input to a variety of reporting, querying, on-line analytical processing (OLAP) and data mining tools, which generate patterns, rule, or models used by business managers for making decisions. These decisions are then fed back in the form of business actions such as product recommendations or fraud alerts (see Fig. 1).

The overall architecture of an OLAP system comprises a hierarchy of basically three data models: the physical model, the logical model and the presentation model. The central logical cube model defines the concept of a cube and its corresponding operations. The physical model deals with how the cubes are stored or indexed for efficient access. The presentation model is concerned with grouping several logical cubes, defined as parts of one (or more) underlying cube(s), in one presentation entity. The mapping between these levels ensures independence and this is achieved through the use of the intermediate logical model.

## III. Application of Crm Based on Olap in Telecom

### III. 1 Overview

In a telecommunication network, hundreds of millions of call detail records are generated daily. Business intelligence applications such as fraud detection and churn analysis require the collection and mining of these records on a continuous basis. An important component of many of these applications is customer profiling, which aims to extract patterns of behavior from a collection of transaction records, and the comparison of such patterns. In telecommunication applications, for instance, a customer's calling behavior is represented by the composition and periodic appearance of the callees (the persons called); time-windows (when calls are made); and duration (how long calls last). Trough e-commerce applications, a customer's shopping behavior could be represented by ads viewed, products selected, products bought, time-windows, price, etc. The techniques for customer profiling and comparison are very similar. Nowadays, Business Intelligence has attracted more and

more people's attention. In Europe and US, it even has been an essential equipment of the personal computer. And BI is believed to play a more and more important roll in our life in the future.

Since the similarity of customer behavior can be represented from different angles, they compare calling patterns derived from customer calling behavior profiles, rather than comparing profiles directly. For example, some calling patterns might be similar in terms of the volume (total number or duration) of calls to a set of callees, others might be similar in terms of the time windows when these calls were made. Their objective, therefore, is to enable the comparison of calling patterns along multiple dimensions and at multiple levels of the dimension hierarchies. This type of multidimensional, multi-level pattern extraction and comparison is facilitated through the use of OLAP. Like many existing efforts, they take advantage of OLAP technology for analyzing data maintained in data-warehouses. In particular, they perform large-scale data mining on an OLAP based computation platform.

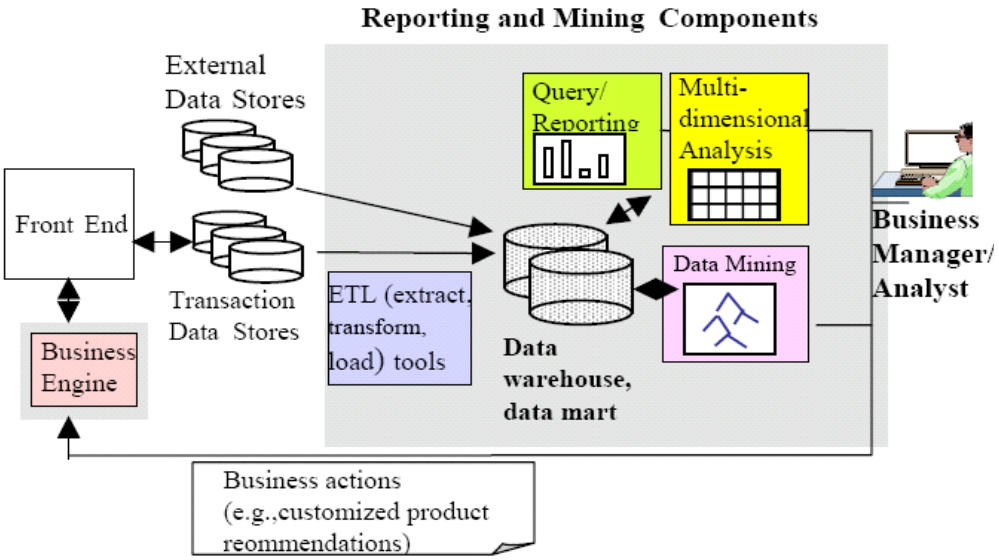


Fig. 1: A typical architecture for business intelligence

### III. 2 The Construction of the MASA System

Application of information technology today in the telecom industry can be concluded like this: first, the prevalent instrument and software used nowadays cannot meet with the increasingly highly required service quality, and then, 0.2 billion telecom customers in China promise a profitable market, third, along with China's entering in the WTO, the telecom industry in China will confront strong competitors from all over the world. A chance but a challenge for state-owned telecom companies arises. In general, the telecom industry will burgeon in the near future and this could bring both opportunities and challenges.

ABC Mobile Communications (AMCC) has fully adopted the advanced business intelligence solutions provided by Informix, which aid it constructed the mobile communication market and consumer conduct analysis system (MASA). The development and application of MASA finds its origin from the urgent needs required by company's operation, which contains data mining (DM) analysis on historic data, kinds of report forms showing the status quo, dependable information when making a decision, and the support to operation system and net maintenance, etc. According to the demand of the enterprise, the data warehouse in AMCC is the data-analyzing warehouse based on calling records. It contains list of VIP, assaying customers' credits (namely the Fraud Inspection), the calling stream, analyzing efficiency of the selling channels, assessing the effect of market competition and measures used in operation, etc.

In the process of operating the system, we construct the data warehouse that is independent from the payment recording system. The memory cells of data warehouse use the dynamic server options for advanced decision support system, the extraction and load of data adopt the loading system that is especially devised for applications of the data warehouse, data analytical and information visiting part use OLAP which is based on web techniques as well as multiple stratum client/server structure, meanwhile it can be shown directly through the combination with map information system. The system structure can be seen like fig. 2. First, AMCC does market segmentation by synthesizing and analyzing the different business behavior of consumers, and make relative product and consumption section to conduct cross-sale. Second, it Segments consumer and adjusts consumer structure and make marketing strategy by synthesizing and analyzing the accumulative total consume and payment behavior of consumers. Third, it can know information of sales channel and control the sales channel. Finally, the company uses the system to adjust the net

according to net operation analysis result. Fig. 3 shows the business development analysis result of AMCC.

Fig. 3 the business development analysis

Fig.4 analysis of customer cluster

Fig.4 shows the analysis result of customer cluster.

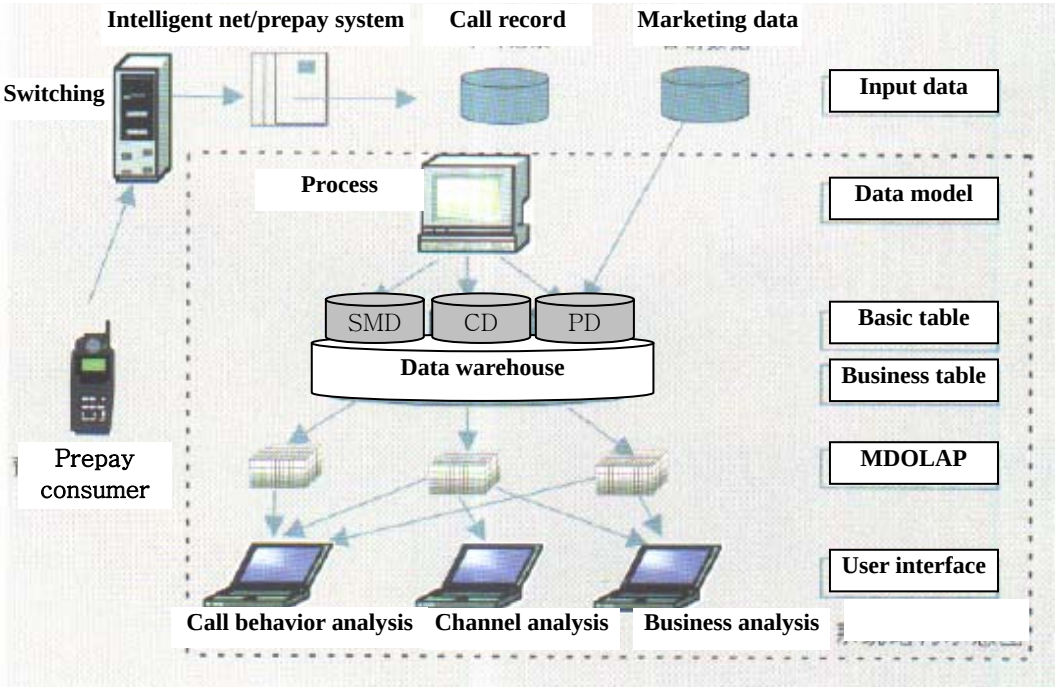


Fig. 2 MASA structure of BI system for prepay
SMD: small message data, CD: calling data, PD: payment data

### III. 3 Profiling Engine

To create and update customer behavior profiles, hundreds of millions of call records must be processed every day. The profiling engine is built on top of an Oracle-8 based data-warehouse and Oracle Express, an OLAP server. Customer behavior profiles and other reference data are persistently stored in the warehouse and handled in the OLAP multidimensional database (MDB) as data cubes. The profiling engine architecture and flow of data is shown in Figure 6. From the calling profile cubes, individual customer based, multilevel and multidimensional calling patterns are derived. The similarity of calling patterns belonging to different customers, or belonging to the same customer but for different profiling periods, can then be computed. The profiling engine can be used by a fraud detection application to generate alarms when suspicious events are detected (e.g., a call exceeds some threshold, an abnormal calling pattern occurs; or a pattern similar to a known fraudulent one occurs). An investigator can then examine the case database to determine whether a fraud indeed occurred.

### III. 4 Calling Pattern Cubes

A calling pattern cube is associated with a single customer for representing the individual calling behavior of that customer. Multiple calling pattern cubes may be generated to represent a customer's calling behavior from different aspects. They may be based on volumes or probability distributions; and they may be materialized (defined as variables) or not (defined as formulas). In our design, probability-based calling pattern cubes are derived from volume-based ones.

A volume based calling pattern summarizes a customer's calling behavior by counting the number of calls of different duration in different time-bins. Represented as cubes, they are commonly dimensioned by time, duration and dow (day of week), and in addition, for those related to outgoing calls, dimensioned by callee, and for those related to incoming calls, dimensioned by caller. Their cell values (measures) represent the number of calls. For example, a calling pattern might express that there were 300 short calls, 100 medium calls, and 50 long calls in the mornings of the profiling period, from one specific phone number to another. A probability distribution based calling pattern represents a customer's calling behavior with probability distributions. For example, a calling pattern from one phone number to another might express that 70% of the calls in the morning were short, 20% were medium, 10% was long. Cubes representing different probability measures can be derived from a profile cube or a volume-based pattern cube. For efficiency as well as consistency, it is only necessary to store profile cubes persistently in the data warehouse. Calling patterns, either based on volume or probability, can be derived on the fly (at analysis time) using the OLAP engine for computation. This shows the simplicity, and yet the power, of OLAP for customer profiling.

## IV. Conclusion

Telecommunication prepay satisfies the benefits of operation company and user. For operation company, prepay can avoids much debt from consumer, as for consumer, prepay is the best choice for those low consumption and low credit. In our approach, the OLAP engine actually serves as a scalable computation engine. From a performance point of view, it supports indexed caching, reduces database access dramatically and extends main memory based reasoning. From a functionality point of view, it allows us to deliver powerful solutions for profiling, pattern generation, analysis and comparison, in a simple and flexible way. We have developed the multi-dimension and multilevel cube similarity comparison formalism for comparing customers calling patterns in terms of cube manipulation.

## References

[1]. Nikos Karayannidis, Aris Tsois, Panos Vassiliadis, and Timos Sellis, "Design of the ERATOSTHENES OLAP Server".
[2]. Qiming Chen, Umeshwar Dayal, and Meichun Hsu, "OLAP-Based Data Mining for Business Intelligence Applications in Telecommunications and E-commerce".
[3]. TONG Xing, CHEN Shan-shan, XI Jian-chun, "Research and Application of Business Intelligence in the Telecom Enterprise".
[4]. http://www.chinabi.net.

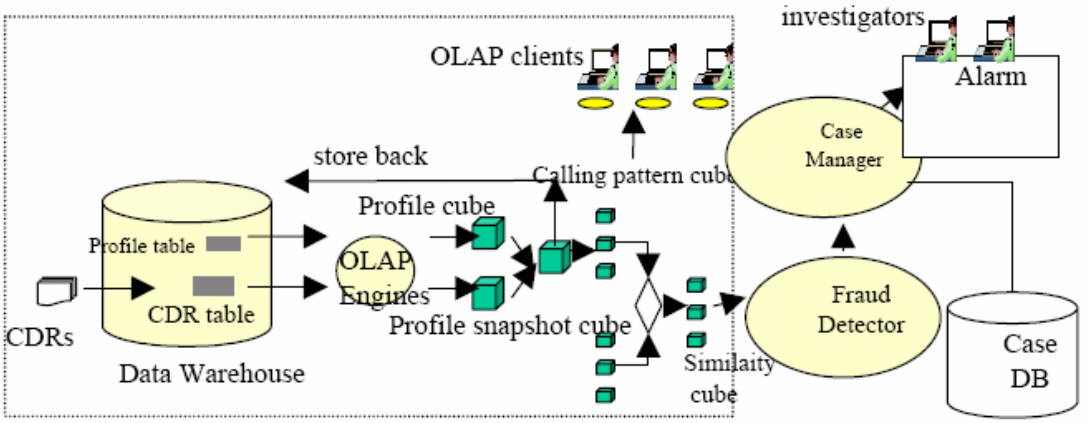


Fig. 6 Data-warehouse/OLAP based profiling engineering for fraud detection