Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 35 (2008) 739–754
Expert Systems with Applications
www.elsevier.com/locate/eswa

Business intelligence approach to supporting strategy-making of ISP service management

Sheng-Tun Li [a,*], Li-Yen Shue [b], Shu-Fen Lee [c]

[a] Institute of Information Management, National Cheng Kung University, 1 Ta-Hsueh Road, Tainan 701, Taiwan
[b] Department of Information Management, National Kaohsiung First University of Science and Technology, 2 Jhuoyue Road, Nanzih District, Kaohsiung City 811, Taiwan
[c] Department of Information Management, Cheng Shiu University, Chengching Road, Niaosong Township Kaohsiung County 833, Taiwan

Abstract

The recent deregulation of telecommunication industry by the Taiwanese government has brought about the acute competition for Internet Service Providers (ISP). Taiwan's ISP industry is characterized by the heavy pressure for raising revenue after hefty capital investments of last decade and the lack of knowledge to develop competitive strategies. To attract subscribers, all ISP dealers are making an all-out effort to improve their service management. This study proposes a Business Intelligence process for ISP dealers in Taiwan to assist management in developing effective service management strategies. We explore the customers' usage characteristics and preference knowledge through applying the attribute-oriented induction (AOI) method on IP traffic data of users. Using the self-organizing map (SOM) method, we are able to divide customers into clusters with different usage behavior patterns. We then apply RFM modeling to calibrate customers' value of each cluster, which will enable the management to develop direct and effective marketing strategies. For network resource management, this research mines the facility utilization over various administrative districts of the region, which could assist management in planning for effective network facilities investment. With actual data from one major ISP, we develop a BI decision support system with visual presentation, which is well received by its management staff.

Keywords: Internet service provider; Service management; Business intelligence; Data mining; Decision support systems

1. Introduction

Following the wave of global liberalization on telecommunication, the Taiwanese government deregulated the telecommunication industry in 1997. Soon after, a host of ISP (Internet Service Provider) companies entered the market to provide value-added services over the network infrastructure and compete for customers. Over the years, network bandwidth has been upgraded from narrowband to broadband, and the network population has grown substantially. It has since grown from 3% of national households at the end of 1996 to 38% at the end of 2002, and reaching 43% in June of 2006 (Institute for Information Industry, 2006). The growth shows a very steep curve in the first few years and a somewhat slower pace lately. During this period of fast growth, one can imagine the fury of ISP dealers in trying to offer various products with different fee schemes to attract subscribers. Initially, pricing strategies did work in recruiting new subscribers. However, after these years, the marketing emphasis may have shifted from product and cost orientation to that of customer needs. Users have started realizing that service quality may be more important than slight differences in fees. At the same time, ISP management has also realized the well-known fact that, the cost of developing a new customer is five to seven times the cost of retaining an existing one (Wayland & Cole, 1997). Management discovered that the need of

* Corresponding author. Tel.: +886 6 2757575x53126; fax: +886 6 2362162.
E-mail addresses: stli@mail.ncku.edu.tw (S.-T. Li), ly_shue@ccms.nkfust.edu.tw (L.-Y. Shue), doris@csu.edu.tw (S.-F. Lee).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.07.049

network stability, data security, usage convenience, and personal preference should be high on the customers' service agenda. This is also evidenced by McCue's (2006) ISP satisfaction survey, which indicated that reliability is the most important factor for 78% of business customers (McCue, 2006), and more than half of them indicated they would switch to another ISP to get improved reliability. Thus, how to develop a service management strategy, which recognizes the shift from market share to percentage of life-long customers, has become a major issue for all ISP companies. The service management for the Telecom industry, in general, consists of customer and product management and resource management (Ericsson, 2005). It aims to ensure that customers experience quality and perceive the value of services delivered, and improve operational readiness for short time-to-market of new innovative services, as well as enhances utilization of existing network facilities. In order to develop a relevant management strategy in this increasingly competitive ISP market, management must understand customers' needs and preferences and network facility utilization, before any proactive actions for customer care can be devised.

Users' needs and preference may be expressed in terms of usage patterns. However, the nature of ISP industry, where users and management may never see each other face to face, makes it impossible to develop a traditional in-depth mutual understanding. Currently, the most common means in servicing customers in ISP industry in Taiwan is through call center. Over the years, it was found that a call center is a rather passive way to service customers; it basically waits for customers to call to present problems. The most management can do with a call center is to call customers to understand the reasons for switching to another company. In addition, management has found the following problems with a call center. Firstly, staff of a different shift may not be able to respond properly to the customers who call at an earlier shift. Secondly, it is difficult to market effectively, because customers are usually not in a happy mood when they call. Thirdly, it is difficult to measure the effectiveness of each individual staff. Overall, most ISP providers in the nation today are lacking the knowledge of their customers' network usage behaviors; they are not able to raise the profile of customer loyalty. As Oracle stated, customer loyalty might be the only sustainable competitive advantage in this very challenging economical time (Oracle Corporation, 2006). Thus, at this stage, any company who knows how to deal with customers effectively will have the definite strategic edge over others. In addition to customers' usage patterns, the network resource management is also an important issue that can benefit both customers and the company. With proper management of network resources, an ISP company must plan for resource allocation according to users' needs of geographical nature, which will aid in achieving better cost effectiveness.

The objective of this research is to propose a Business Intelligence (BI) process for the ISP industry in Taiwan, which could assist management in developing effective service management strategies. The ISP industry in Taiwan is characterized by the heavy pressure for raising revenue after hefty capital investments in the last decade and the lack of knowledge to develop effective competitive strategies. The process applies data mining, visualization, and RFM customer value modeling as the underlying methodologies to identify various knowledge patterns. These patterns form the basis for discovering business intelligence, which includes the identification of VIP status, characteristics of different usage group, users' monetary contribution, and network facility utilization. A BI decision support system is developed with MVC (Model-View-Controller) architecture to facilitate the intelligence dissemination. The performance of the system is empirically and subjectively evaluated by the company staff. The remaining structure of this paper is as follows. In Section 2, we present a review of underlying methodologies that are utilized in the study. Section 3 describes various phases of the proposed BI process. In Section 4, we implement BI process with actual data from the company and describe the development of a BI decision support system. The system evaluation is described in Section 5, and Section 6 concludes this paper.

2. Review of methodologies

This part provides a brief review of the concepts of BI as well as the methodologies that are utilized in various phases of this study.

2.1. Business intelligence

The effective management and leverage of data represent both the greatest opportunity and the most difficult challenge for most enterprises. Gartner's 2006 CIO survey showed BI as their highest rating technology issue; as they focus on projects that enable users to positively affect financial and business performance (Gartner, 2007). BI is a set of concepts, methods, and processes (Maria, 2005) to improve business decisions, which use information from multiple sources and apply experience and assumptions to develop an accurate understanding of business dynamics. It integrates the analysis of data with decision support system to provide information to people throughout the organization in order to improve strategic and tactical decisions. With appropriate BI a company will be able to develop intelligent decision support systems to gain the competitive advantage of the industry (Davis, 2002). It has been applied to many areas that are related to the enterprise management process, and some of them have formed their own systems with specific characteristics. Typical application scopes include: ERP (enterprise resource planning), CRM (customer relationship management), HRM (human resource management), SCM (supply chain management) and E-business (Xie et al., 2001). Buytendijk (2001) has reported that, based on a study from 2001 to

2006, enterprises that apply BI had achieved two to three times return of investment more than those who do not (Buytendijk, 2001).

The ever-changing business environment has warranted that the BI process is a continuous and systematic process, which consists of identifying what information is needed, how it should be gathered, how it should be organized and stored, who should have access to it, and reviewing how the management has exploited the knowledge to gain significant competitive advantages. There are a number of process models (Pirttimaki & Hannula, 2003) that have been presented, and it seems that they all share some common elements; although each process may have different number of phases due to a somewhat different emphasis. We will present the BI process models that have a direct bearing on the development of our model.

Microsoft's process (Vitt et al., 2002) is perhaps the simplest BI process ever presented. Microsoft defines a BI cycle as a progression from Analysis to Insight to Action and finally to Measurement. At first, the data target is identified and collected from as many sources as possible; then, proper analysis of data may yield information of certain characteristics. These characteristics need to be investigated to see if they could lead to some business insights, and based on the degree of significance of these insights, various suggestions may be made as to how business policies/operations should be re-formulated. The Action phase is to take these suggestions and direct them to relevant management levels/positions for implementation. After the implementation, measurements should be taken to understand if the intended effects, identified in the insights phase, have been achieved. The results of these measurements would normally provide clues of targets for further analysis and thus start a new cycle again. Thomas Group Inc., a strategic intelligence consulting company in US, developed a slightly different version of BI cycle (Thomas, 2001) with six phases: Planning and Direction, Data Collection, Information Processing and Storage, Analysis and Production, Dissemination, and Intelligence Users and Decision makers. One feature of this process is the emphasis of the "needs" driven cycle. The "needs" are the alternative expressions of business objectives. The second feature of this process is the Dissemination phase, where the emphasis is to interpret results of analysis in terms of business contexts, so that knowledge insights may be revealed. The revealed knowledge must be further translated into clear and understandable business policies/operations, and it should be distributed, through the application of IT technologies, to a wide range of management levels/positions for implementation. A similar process is proposed by Novintel Inc. (Viva Business Intelligence Inc., 1998), a Finnish international company specializing in providing BI and competitive intelligence services and products. The Novintel's process (Viva Business Intelligence Inc., 1998) consists of eight phases: Need Analysis, Observing and Monitoring, Collecting Information, Structuring and Elimination, Analysis, Communication, Storing, and Utilization and Feedback. This process stresses, in particular, the importance in Observing and Monitoring possible information sources, so that correct data sources may be identified for uncovering the true knowledge of the "needs" in the first phase. The second emphasis is the preprocessing of raw data through structuring and elimination, which is to ensure data validity and soundness, so that meaningful analysis can be carried out with statistical methods.

2.2. SOM model for data mining

The SOM neural network is one of the most popular unsupervised neural network models, which simultaneously performs a topology-preserving projection from the data space onto a regular two-dimensional grid (Kohonen, 1990). A basic SOM network composed of an input layer, an output layer, and network connection layer. The input layer contains neurons for each element in the input vector. The output layer consists of neurons that are located on a regular, usually two-dimensional grid and are fully connected with those at the input layer. The network connection layer is formed by vectors, which are composed of weights in the input and output layer. The neurons in the map are connected to adjacent ones by a neighborhood relation dictating the topological structure of the neurons. Each neuron i is represented by a reference vector$w_i = [u_{i1}, u_{i2}, \ldots, u_{in}]^{\mathrm{T}}$, where n is the number of neurons in the input layer. Given an input vector $x \in R^n$, the neurons in the map compete with each other to be the winner b or the best-matching unit (BMU), which is closest to the input vector in terms of some kind of dissimilarity measure such as Euclidean distance,

$$\|x - w_b\| = \min_i\{\|x - w_i\|\} \tag{1}$$

During training session, weights of neurons are topologically arranged in the map within a certain geometric distance and are moved toward the input x using the' self-organization' learning rule as represented in formula (2).

$$w_i(t+1) = w_i(t) + \eta h_{bi}(t)[x(t) - w_i(t)] \tag{2}$$

where $t = 0, 1, 2, 3, \ldots$ is the time lag, η is a small positive learning rate and $h_{bi}(t)$ is the neighborhood kernel around the BMU at time t (Kohonen, 1997). $h_{bi}(t)$ can be defined as formula (3)

$$h_{ci}(t) = h(\|r_c - r_i\|, t) \tag{3}$$

where $r_c, r_i \in R^2$ are the location vectors of neurons c and i, respectively, and when $\|r_c - r_i\|$ increases, $h_{ci}(t)$ decreases to zero gradually. This leads to local relaxation or smoothing effects on the weight vectors of neurons in the neighborhood of the BMU. Therefore, similar input vectors are grouped into a single neuron or neighboring ones in the map when learning is accomplished. The learning algorithm of the SOM network is stated as follows (Li & Shue, 2004).

Step 1: Initialize randomly weights of neurons, learning rate, and winner neighborhood.
Step 2: Determine the winner using formula (1).
Step 3: Update the weights of the neighborhood of the winner using formula (2).
Step 4: Decrease the neighborhood of the winner using formula (3).
Step 5: Repeat Steps 2–4 until learning is accomplished.

In the past, Sung and Sang (1998) have applied data mining, SOM, and RFM techniques to a real-world application for analyzing customer behavior. RFM data extracted from the data mart are used extensively for our analysis. Abidi and Ong (2000) and Vesanto and Alhoniemi (2000) adopted two phases for data mining clustering strategies. The first phase used SOM and the second used K-means for clustering analysis.

2.3. RFM customer value model

One important marketing concept in running today's business is the recognition of value of each individual customer. This concept is guiding today's business in developing personalized and value-added products and services. The value of a customer is defined as the profit resulting from payment of a transaction; however, we normally like to evaluate a customer's contribution over a longer period with more transactions. Customers, on the other hand, will normally attach a quality level to a company by comparing the received value with the expected one and then decide if they want to continue to transact with the company, how much, and how often. Thus, the amount and frequency of transaction may indicate the potential contribution of a customer, which is what is looked for in value management. The common model for measuring customers' value is the RFM (Recency, Frequency, Monetary) model (Stone, 1989), which is made up of three major factors: recency, frequency, and monetary. Recency stands for the time of the last transaction, frequency stands for the number of transactions, and monetary value stands for the value of transactions. In Table 1, we provide the comprehensive definition made by Hughes (1994) and summarize two major RFM implementation models proposed by Hughes (1994) and Stone (1995), respectively. For our study with ISP industry, due to the unique characteristics of transaction with network via IP traffic, we prefer Hughes, 1994 RFM implementation model.

3. A BI process for ISP service management

Currently, the service management of ISP industry in Taiwan is characterized by the development of price inducing schemes for attracting new subscribers and aggressive management of call centers to handle customer relations. Pricing schemes are made up of combination of upload and download speed, peripheral services, and payment arrangements. In earlier years, most customers were very fee-conscience and pricing schemes played a very important role in the development of service management strategy. After these years however, it seems that small differential in fees is no longer a major consideration when choosing an ISP. Similarly, call center was one of the necessary facilities in early ISP days when problems were numerous. Part of it was due to users' unfamiliarity with internet services, and they required instant assistance by service personnel. Part of it at that time was due to the unreliability of facilities, and service personnel must be available to accept complaints to reduce the impact of customer relations to a minimum. These problems, however, have been greatly reduced after these years, because users have become much more experienced and the reliability of facilities has been greatly improved after heavy capital investments over the period. It is the heavy capital investment of the past few years that has forced management to refocus its business goal away from recruiting customers to enhancing business revenue. To achieve this goal, management must be able to create new revenue from existing customers by recognizing their potential monetary contribution, and take proactive

Table 1
RFM definition & implementation

	Authors	Recency	Frequency	Monetary
Definition	Hughes (1994)	The number of days between the last purchase and the time of analysis. The smaller the number, the higher the probability of next purchase	The number of purchases during a period of time. The higher the frequency, the higher the loyalty and value of a customer	Total amount of purchase during a period of time. The higher the amount, the higher the value of a customer
Implementation	Hughes (1994)	Divide the sorted purchase dates into five equal intervals; then assign a weight 5 to the first 20%, 4 to the next 20%, and so forth	Divide the highest purchase count into five equal intervals; then assign a weight 5 to the first 20%, 4 to the next 20%, and so forth	Divide the total purchase amount into five equal intervals; then assign a weight 5 to the first 20%, 4 to the next 20%, and so forth
	Stone (1995)	For the latest purchase, assign a weight 24 if it is within 3 months, 12 if it is between 3 and 6 months, 6 if it is between 6 and 9 months, 3 if it is between 9 and 12 months, and 0 if it is longer than 12 months	Purchasing frequency × 4 points	Purchasing amount × 10% (with the highest being 9)

actions accordingly. While, most of ISP management realized that they need to provide value-added products to customers, they, in general, do not have the knowledge to develop effective plans.

Based on the present needs of ISP management in Taiwan, we referred to above processes and present our BI process in Fig. 1. This process emphasizes immensely on the processing of raw data and the modeling of valid data, because the data is IP traffic in K-bytes and is voluminous. In addition, most managing personnel in the ISP industry in Taiwan were trained in network related engineering fields. Consequently, they are generally unfamiliar with modern managing methodologies. We thus decided to present appropriate methodologies for each phase. The six phases consist of: knowledge identification, data collection, data preprocessing, modeling, analysis and evaluation, and dissemination. The Knowledge Identification phase takes the targeted business aims as inputs and proceeds to identify the types of knowledge that are needed for achieving them. For example, the former could be achieving maximum effects of a new product, and the latter could be who will be most interested in the new product. The Data Collection phase would then decide where and how the data will be collected and stored. Among the major tools for collecting IP flows, the most widely used one is MRTG (Multi Router Traffic Grapher) (Kemper, 1997), which is used in this research. Because the acquired data could be voluminous; we propose the application of Data Warehouse for storing raw as well as processed data. The data preprocessing phase carries out validity checking of raw data to ensure data integrity for further analysis. In the Modeling phase, we propose Data Mining methodology for analyzing the processed data, which includes multidimensional model, OLAP (On-Line Analytical Processing), and SOM neural network. Customer value model is also included in this phase. The Analysis phase examines the results of modeling in business context, and tries to discover relevant BI that is important for achieving the original business aims. The very last phase of the cycle is the Dissemination, which is designed to share the uncovered intelligence through appropriate channels and to imbed the intelligence into the system for actual implementation. At the same time, this phase collects feedback from users, which will be translated into business aims of the next cycle of the process.

4. Design and implementation of the BI process

In the following sections, we will demonstrate the application of the proposed BI process to a major ISP company in Taiwan. This company was originally the nation's sole enterprise in telecommunication and was only recently privatized. Its management is struggling to compete with newly established companies who are eroding its market share and is very keen to develop service strategy through the understanding of the "needs" of users.

4.1. Knowledge identification

The management of this company is very much interested in developing a service management strategy, which can boost business revenue through providing value-added products to its customers. They believe very strongly that personalized service is the way to grow business revenue further, because it will foster long-term loyalty of customers, when will then lead to increased sales of value-added products. They further identified the knowledge they will need, which includes: network usage patterns of individual

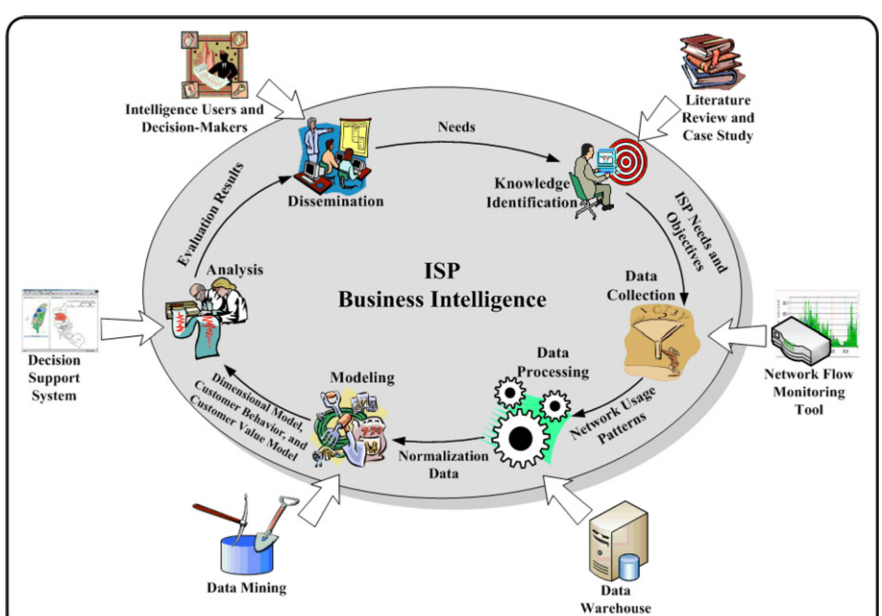


Fig. 1. The proposed ISP BI process.

customer, network usage patterns of the region, revenue contribution of customers, and network facilities utilization. The network usage patterns of individual customer should reveal network usage over 7 days of a week and 24 h of a day, along with the usage intensity. This usage pattern will allow management to develop the knowledge of VIP status of users and initiate meaningful business dialogue with individuals. The usage patterns of customers of a region should reveal the grouping of users and behaviors of each group, which can help management formulate marketing strategies by targeting selected groups. In addition, it will form the basis for understanding the potential revenue contribution of each group. Lastly, the facility utilization among geographical regions will lend management an important piece of knowledge in achieving cost effectiveness.

4.2. Data collection

ISP customers' raw data consists of socio-demographic data, records of call data, IP traffic log, logging authorization data, application system, and system record. Socio-demographic data is recorded at the time customers fill in the application form. Records of call data contain source and destination of IP address, TCP port number, URL address, etc. IP traffic log contains switch-router IP address, customer account number, and input and output traffic log per five minutes. The logging authorization data includes customer account number, log-in and log-out time, facility name of logging in, and IP address. Application system data is generated when customers make use of WWW, e-mail, FTP service, etc. Finally, the system record is generated by routers. With the cooperation of the company, we selected a region of southern Taiwan that consists of several districts for this study. The IP flows in K-bytes were collected every 5 s over 7 weeks using MRTG (Kemper, 1997). We make use of a timer program to transform network flow records into a SQL database. Table 2 shows the contents of the database, which contains fields of ADSL_phone, Log time, the average input K-bytes per five minutes (Avg_In), the average output K-bytes per five minutes (Avg_Out), the largest input K-bytes in the interval (Max_In), and the largest output K-bytes in the interval (Max_Out), respectively. The final count of data is 41.7 million.

4.3. Data preprocessing

With the big volume of raw data, we need to process them to ensure its validity for later use. Through the socio-demographic data provided by the administration, we found that there are 10.3 million valid data. These data must be normalized to avoid inconsistency during the mining process, because different user may be with a different scheme and hence different bandwidth. We apply the formula defined in Eq. (4) to transform data to achieve normalization. In the formula, we take the ratio of customer's IP flow to his/her scheme stipulated bandwidth, Customer_NetUsage/Customer_ Bandwidth, and compare it with a selected Threshold_rate, which can be set at 1%, 5%, 10%, or other rates, as shown in Table 3. The setting of the Threshold_rate depends on the conceptual purpose in the modeling phase. The technical personnel of the company indicates that threshold rate at 1% will be sufficient to indicate customer's intention to use network facilities.

$$
\begin{aligned}
&\text{IF}(\text{Customer\_NetUsage}/\text{Customer\_Bandwidth}) >= \text{Threshold\_rate}\\
&\text{THEN Threshold\_rate\_record} = 1\\
&\text{ELSE Threshold\_rate\_record} = 0
\end{aligned}
\tag{4}
$$

4.4. Modeling

With the normalized records, we construct a data warehouse with multi-dimensionality to facilitate the analysis of customers' behavior. We then applied SOM network to segment customers into different homogeneous clusters and select the one that can best exhibit customers' behavior patterns. We further modify the RFM model to evaluate

Table 2
Records of network flows in the database (in K-bytes)

ADSL_Phone	LogTime	Avg_In	Avg_Out	Max_In	Max_Out
07-xxxxx90	2003/10/1 08:35:00	184	25	1489	1489
07-xxxxx90	2003/10/1 08:30:00	215	56	1546	2275
07-xxxxx90	2003/10/1 08:25:00	612	10	2275	654
07-xxxxx90	2003/10/1 08:20:00	584	2	2179	125

Table 3
Normalized records

ADSL_Phone	···	Max_In	Max_Out	In 1%	···	In 50%	Out 1%	···	Out 50%
07-xxxxx90	···	1489	1489	0	···	0	0	···	0
07-xxxxx90	···	1546	2275	0	···	0	0	···	0
07-xxxxx90	···	2275	654	1	···	0	0	···	0
07-xxxxx90	···	2179	125	1	···	0	0	···	0

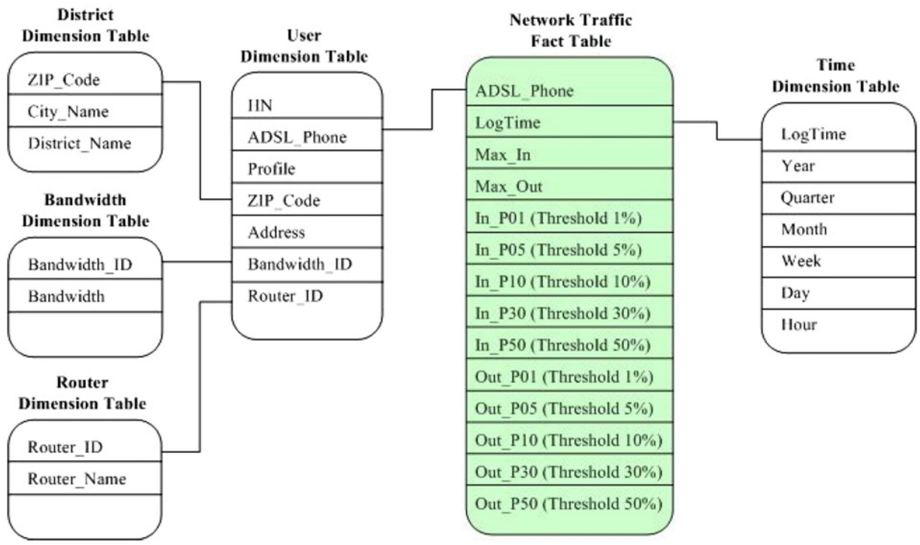


Fig. 2. The dimensional model.

potential value of customers of each cluster. The detailed modeling processes are as follows.

4.4.1. Multi-dimensional modeling

In developing a multi-dimensional model, we adopt the method proposed by Kimball (1996) to, first, define business process and grain, and then, define the dimension table that contains time, user, location, and facts. The relational model and the dimension tables are shown in Fig. 2. Each of the tables is described in the following:

(a) Network Traffic Fact Table contains customer network usage attributes: ADSL_Phone, LogTime, Max_In/Max_Out, In_P (01/05/10/30/50), and Out_P (01/05/10/30/50). The primary key is constituent of ADSL_Phone and LogTime, and the observation value includes Max_In, Max_Out, and different threshold value.
(b) District Dimension Table contains ZIP_Code, City_Name, and District_Name; where ZIP_Code is the primary key. The hierarchical relationship among attributes is defined in order as City_Name $\supset$ District_Name.
(c) Bandwidth Dimension Table contains attributes of Bandwidth_ID and Bandwidth.
(d) Router Dimension Table contains both Router_ID and Router_Name.
(e) User Dimension Table contains HN, ADSL_Phone, Profile, ZIP_Code, Address, Bandwidth_ID, and Router_ID. The primary key is HN and the foreign keys are ZIP_Code, Bandwidth_ID, and Router_ID.
(f) Time Dimension Table has attributes consisting of LogTime, Year, Quarter, Month, Week, Day, and Hour. The primary key is LogTime. The hierarchical relationship among attributes is Year $\supset$ Quarter $\supset$ Month $\supset$ Week $\supset$ Day $\supset$ Hour.

4.4.2. Mining customer behavior

We referred to the customer behavior model built with detailed dialogue records in the telecommunication industry (Berry & Linoff, 2000; Li et al., 2006) and used network flow records as the model's input to establish customer behaviors of network usage for this study. Fig. 3 represents the usage pattern of one user. This figure is developed using the concept/hierarchy classes of AOI (attribute-oriented induction) (Han et al., 1992; Li et al., 2006), which generates the concept hierarchy of time attributes that consist of 5-min, hourly, daily, and weekly data. The value in each cell represents the count of network flow that exceeds 1% of

	Sun.	Mon.	Tues.	Wed.	Thu.	Fri.	Sat.
00-01	0	0	0	0	0	0	0
01-02	0	0	0	0	0	0	0
02-03	0	0	0	0	0	0	0
03-04	0	0	0	0	0	0	0
04-05	0	0	0	0	0	0	0
05-06	0	0	0	0	0	0	0
06-07	0	0	0	0	0	0	0
07-08	0	0	0	0	0	0	0
08-09	0	25	25	15	14	17	0
09-10	0	58	59	57	32	24	0
10-11	0	35	24	49	38	16	0
11-12	0	35	40	45	31	23	0
12-13	0	24	8	26	7	10	0
13-14	0	51	67	54	41	23	0
14-15	0	35	42	42	8	1	0
15-16	0	30	33	33	13	26	0
16-17	0	21	38	35	24	26	0
17-18	0	29	40	23	22	20	0
18-19	0	1	2	2	1	0	0
19-20	0	0	0	0	0	0	0
20-21	0	0	0	0	0	0	0
21-22	0	0	0	0	0	0	0
22-23	0	0	0	0	0	0	0
23-24	0	0	0	0	0	0	0

Fig. 3. Network usage behavior of one user.

flow threshold. The maximum value for a cell is 84 over the 7-week data collection period ((60 min/5 min) $\times$ 7 weeks = 84). This figure clearly indicates that the usage patterns of this particular customer were concentrated over the regular working days and hours.

In order to improve the readability of modeling results, we adopted pixel-oriented techniques (Oliveira & Levkowitz, 2003) to provide visualization with color and brightness. The basic idea of pixel-oriented techniques is to map each data value to a colored pixel and present the data values belonging to one variable in separate windows. Pixel-oriented visualization techniques maximize the amount of information represented at one time without any overlap. They effectively preserve the perception of small regions of interest while still maintaining the global view. The network usage patterns of Fig. 3 are transformed into color through the formula in Eq. (5), and the result is shown in Fig. 4. It is obvious that Fig. 4 can better enhance readers' perceptions and understanding than Fig. 3.

$$ColorCode = \text{“}FF\text{”} + Hex\left(255 \times \left(1 - \frac{nfValue}{Max\_nfValue}\right)\right) + Hex\left(255 \times \left(1 - \frac{nfValue}{Max\_nfValue}\right)\right) \quad (5)$$

where $nfValue$ is the value of the cell and $Max\_nfValue$ is the maximum of the cell, i.e. 84.

4.4.3. Customers segmentation

We use the two-dimensional format to exhibit users' network usage pattern, 24 h versus 7 days, and there are 168 attributes in total for each customer. In order to apply SOM to develop customer groupings, we tried a several cluster numbers to decide the most appropriate one. With the help from the company's ISP senior manager, we concluded that the two-dimensional design 3×3 is more effective in distinguishing one group from another. The clustering results of nine groups are visually presented in Table 4. The brightness in cells ranges from light to heavy to represent different degree of network usage from light to heavy. By collectively examining the brightness, time interval, and day that characterize network usage of each group, we name them as: Group 1: Midnight medium usage group (6.84%), Group 2: Weekend group (2.56%), Group 3: Overall heavy usage group (5.98%), Group 4: Midnight light usage group (12.82%), Group 5: Mid-day medium usage group (4.27%), Group 6: Office hour heavy usage group (3.42%), Group 7: Overall light usage group (47.01%), Group 8: Office hour light usage group (8.55%), and Group 9: Office hour medium usage group (6.84%).

4.4.4. Customer value modeling

Based on the customer usage groupings found in the last part, we applied RFM to evaluate customers' potential monetary contribution. The concept of purchasing behavior from the original RFM is somewhat different from that of network usage behavior, because the former represents transaction of actual purchasing and payments, and the latter may not be regarded as transaction of network service until certain traffic threshold is reached. Thus, we refer to the survey conducted by TWNIC (2004) to define this threshold. The survey indicates that, on an average, the longest continuous network usage interval in Taiwan is between 2 and 3 h. The company's experienced ISP personnel suggest that this network flow threshold be defined as 1% of stipulated bandwidth over three continuous hours. As a result, we modified the original definition of three factors R, F, and M of the RFM model in the following to suit the ISP environment; while the division of 5 intervals and the weight assignments of RFM implementation remain the same.

R: time of most recent traffic that lasts for 3 h and its network flow exceeds the threshold.

F: usage counts over 7 weeks; with each usage lasting for 3 h and its network flow exceeding the threshold.

M: cost per network traffic unit, which is calculated as network-monthly-rental/monthly-network-traffic.

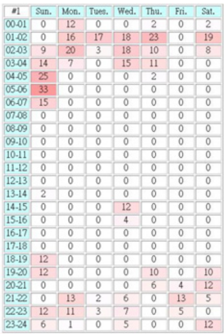
For this study, our analysts had decided that R will not be taken into consideration, because our purpose is to study customers' behaviors that should last over a longer period and has little to do with the most recent usage.

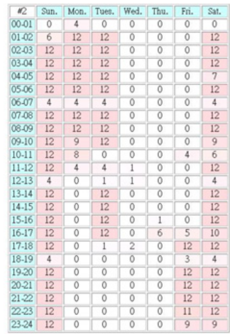
With the definition of F and M, one can see that a higher F means a higher demand on the network usage, and a higher M means a higher usage cost per traffic unit. Thus different combinations of F and M value would imply different business intelligence. For example, customers with high F and low M value tend to be frequent users with low value schemes, and those with low F and high M value are

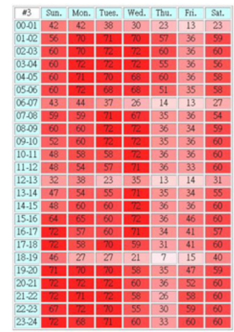
	Sun.	Mon.	Tues.	Wed.	Thu.	Fri.	Sat.
00-01	0	0	0	0	0	0	0
01-02	0	0	0	0	0	0	0
02-03	0	0	0	0	0	0	0
03-04	0	0	0	0	0	0	0
04-05	0	0	0	0	0	0	0
05-06	0	0	0	0	0	0	0
06-07	0	0	0	0	0	0	0
07-08	0	0	0	0	0	0	0
08-09	0	25	25	15	14	17	0
09-10	0	58	59	57	32	24	0
10-11	0	35	24	49	38	16	0
11-12	0	35	40	45	31	23	0
12-13	0	24	8	26	7	10	0
13-14	0	51	67	54	41	23	0
14-15	0	35	42	42	8	1	0
15-16	0	30	33	33	13	26	0
16-17	0	21	38	35	24	26	0
17-18	0	29	40	23	22	20	0
18-19	0	1	2	2	1	0	0
19-20	0	0	0	0	0	0	0
20-21	0	0	0	0	0	0	0
21-22	0	0	0	0	0	0	0
22-23	0	0	0	0	0	0	0
23-24	0	0	0	0	0	0	0

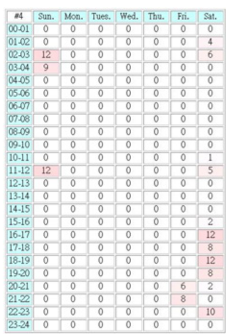
Fig. 4. Visualization of network usage behavior.

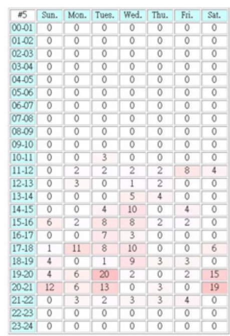
Table 4
Clustering outcome by SOM

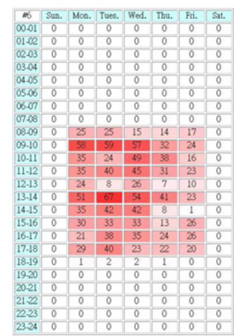

1. Midnight medium usage Group (6.84%)

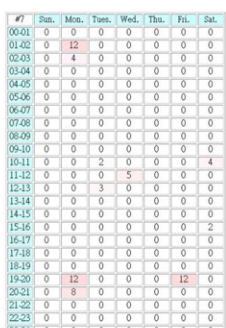

2. Weekend Group (2.56%)

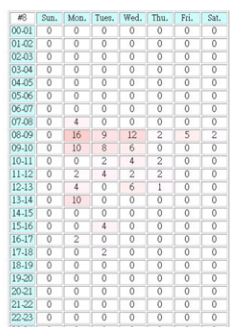

3. Overall Heavy Usage Group (5.98%)

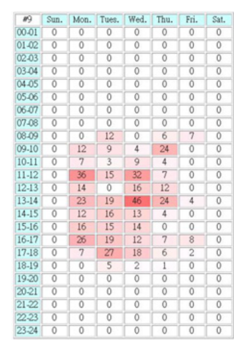

4. Midnight light usage Group (12.82%)


5. Mid-day medium usage Group (4.27%)


6. Office hour heavy usage Group (3.42%)


7. Overall light usage Group (47.01%)


8. Office hour light usage Group (8.55%)


9. Office hour medium usage Group (6.84%)

infrequent users with high value schemes. The worst case is customers with low F and low M; it means they subscribed to low value schemes and also rarely use network facilities.

4.4.5. Facility utilization modeling

In developing the facility utilization model, we first examine locations of districts and routers to define network flow and utility rate for comparison. However, different router or district may serve different number of customers; hence we develop formulas (6) and (7) as the basis for proper comparison. Formula (6) measures average network flow per unit time of a router or a district, while (7) measures average usage counts per user that exceed the threshold.

Table 5
Utilization of routers

Router	Average flow (K-bytes)	Average usage counts
Router 1	78,560	673
Router 2	23,112	1016
Router 3	339	12
Router 4	1988	63
Router 5	6220	170
Router 6	380	87
Router 7	20,362	446
Router 8	594	87
Router 9	162,402	983

Table 6
Network utilization of Kaoshiung City & County

County/City	Villages and towns/city area	Average flow (K-bytes)	Average usage counts
Kaohsiung City	–	183,111	738
	Sanmin district	13,370	966
	Siaogang district	7205	192
	Zuoying district	97,106	11,322
	Cianjin district	23,432	1350
	Cianjhen district	12,234	171
	Lingya district	4398	346
	Sinsing district	11,462	221
	Nanzih district	9909	771
	Gushan district	3990	211
Kaohsiung County	–	110,848	587
	Dashe township	199	82
	Daliao township	6483	281
	Dashu township	4818	1801
	Renwu township	0	0
	Gangshan township	25,376	387
	Linyuan township	177	137
	Meinong township	125	0
	Jiading township	4059	350
	Niaosong township	150	123
	Hunei township	4568	650
	Lujhu township	5313	164
	Cishan township	1802	69
	Fongshan city	54,244	1300
	Ciaotou township	1558	478
	Yanchao township	1970	157

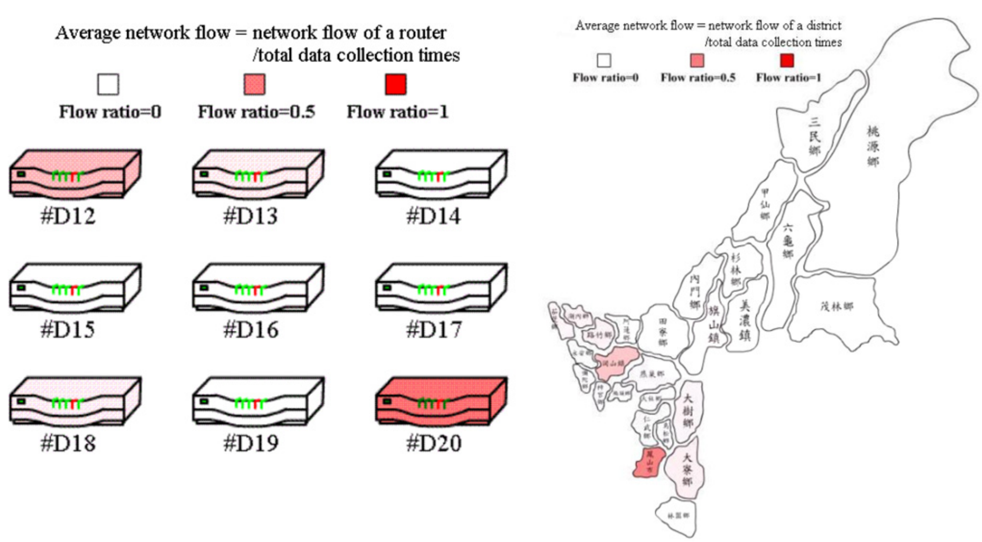
$$\text{Average network flow per time uint} = \frac{\text{network flow of a router(district)}}{\text{data collection duration}} \quad (6)$$

$$\text{Average usage counts per customer} = \frac{\text{total usage counts of a router(district)}}{\text{customers served}} \quad (7)$$

These two measurements allow us to present the utilization of nine routers of the chosen area in Table 5. These nine routers serve both Kaoshiung City, consisting of nine districts, and Kaoshiung County, consisting of fifteen districts, as shown in Table 6.

4.5. Analysis and evaluation

With the results of the modeling phase known, this phase proceeds to analyze the imbedded business knowledge, which could facilitate the development of relevant service strategies. One can examine the patterns and brightness of Fig. 4 to determine the VIP status of a customer. The cross examination of usage behaviors of groups in Table 4 could lead to the discovery of where the issues of personalized services lie and how to approach these issues. For example, the Overall Heavy Usage group is a group of great immediate value to management. This group provides management a much focused target with heavy usage for nearly all time, and management could develop and market high value-added products that fit their needs. Similarly, for the three office hour groups with light, medium, and heavy usage, they present management very focused business targets with different applications needs. On the other hand, the Overall Light Usage group happens to be the largest grouping with 47.01% of users, most of them are probably individual users, and it would certainly presents the challenge to management to conduct further analysis, so that some of them may be converted to more regular users. For resource utilization, Fig. 5 displays the utilizations of various routers and districts in colors, which allows one to quickly realize that the router #20 has the highest utilization, #12 comes in second, and is followed by #13 & #18 and the rest. In terms of districts, Daliao Township has the highest network usage in Kaohsiung county and is followed by Fongshan city, Ciaotou, Gangshan, Hunei, and Dashu townships. These high usage districts present management targets for understanding



Flow distribution of routers

Flow distribution of each administrative division

Fig. 5. Visualization of flow distribution.

their social-demographics factors that underlie the needs of network services, and project the future needs of capital investments.

The results of application of the modified RFM are giving in Table 7. This table presents the cross analysis between customers' network usage patterns and customers' potential contribution. The five intervals of F (frequency) and M (monetary) are given in the second column; with F5 being the frequency percentage of weight 5 (first 20%), M1 being the monetary percentage of weight 1 (last 20%)..., etc. The R factor was not shown in the table, because ISP specialists considered it irrelevant in this particular case. The third column of the table presents suggested marketing strategies for each group by marketing personnel, who has examined both F and M values. For each of the nine groups, their advices of strategy consist of "Active marketing", "Customer care", or "Other". "Active marketing" means the development of value-added products and aggressive promotions, "Customer care" means providing appropriate schemes to match customers' needs, and "Other" literally indicates to management that relevant strategies remain to be discovered. The percentage attaching to each strategy represents the emphasis for that particular group. Some examples of important findings are presented here. The third group, the overall heavy usage group, is characterized by 100% F5 and 100% M1, and it signifies that customers in this group are heavy users; however their monetary contributions are relatively low. This is a big surprise for the management. Hence, 100% "Active marketing" is suggested to market to them higher value services. On the contrary, customers in the overall light usage group, Group 7, is characterized by values of

Table 7
Customer value and marketing strategy for each group

Group	Customer value		Marketing strategy suggestions
1. Midnight group with medium usage (6.84%)	F5: 12.5%	M5: 0%	Active marketing: 75%
	F4: 87.5%	M4: 0%	Customer care: 25%
	F3: 0%	M3: 25%	
	F2: 0%	M2: 25%	
	F1: 0%	M1: 50%	
2. Weekend group (2.56%)	F5: 100%	M5: 0%	Active marketing: 100%
	F4: 0%	M4: 0%	Customer care: 0%
	F3: 0%	M3: 0%	
	F2: 0%	M2: 0%	
	F1: 0%	M1: 100%	
3. Overall heavy usage group (5.98%)	F5: 100%	M5: 0%	Active marketing: 100%
	F4: 0%	M4: 0%	Customer care: 0%
	F3: 0%	M3: 0%	
	F2: 0%	M2: 0%	
	F1: 0%	M1: 100%	
4. Midnight group with light usage (12.82%)	F5: 0%	M5: 0%	Active marketing: 20%
	F4: 33.3%	M4: 20%	Customer care: 20%
	F3: 33.3%	M3: 26.7%	Others: 60%
	F2: 6.7%	M2: 40%	
	F1: 26.7%	M1: 13.3%	
5. Mid-day group with medium usage (4.27%)	F5: 0%	M5: 0%	Active marketing: 20%
	F4: 60%	M4: 0%	Customer care: 40%
	F3: 20%	M3: 60%	Others: 40%
	F2: 0%	M2: 40%	
	F1: 20%	M1: 0%	
6. Office hour group with heavy usage (3.42%)	F5: 100%	M5: 0%	Active marketing: 100%
	F4: 0%	M4: 0%	Customer care: 0%
	F3: 0%	M3: 0%	
	F2: 0%	M2: 25%	
	F1: 0%	M1: 75%	
7. Overall light usage group (47.01%)	F5: 0%	M5: 40%	Active marketing: 0%
	F4: 0%	M4: 30.9%	Customer care: 10.1%
	F3: 25.5%	M3: 18.1%	
	F2: 38.2%	M2: 5.5%	Others: 89.9%
	F1: 36.3%	M1: 5.5%	
8. Office hour group with light usage (8.55%)	F5: 0%	M5: 0%	Active marketing: 30%
	F4: 60%	M4: 30%	Customer care: 50%
	F3: 30%	M3: 30%	
	F2: 10%	M2: 30%	Others: 20%
	F1: 0%	M1: 10%	
9. Office hour group with medium usage (6.84%)	F5: 80%	M5: 0%	Active marketing: 90%
	F4: 20%	M4: 0%	
	F3: 0%	M3: 10%	Customer care: 10%
	F2: 0%	M2: 60%	
	F1: 0%	M1: 30%	

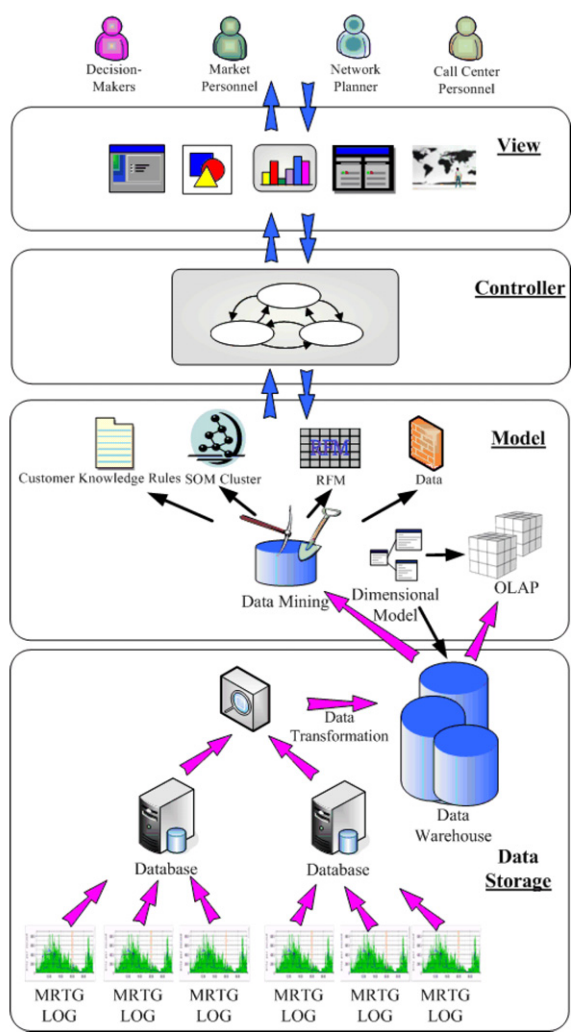


Fig. 6. System framework of the proposed BI-DSS.

less frequent categories F1, F2, F3, and high M categories F5 and F4, which indicates that, although, they are good monetary contributors, they are nonetheless very infrequent users. For customers of this group, ‘‘active marketing’’ is not an issue, instead, a small portion of them requires ‘‘customer care’’, and the remaining 90% is in ‘‘Other’’ category, which indicates that further investigation is needed to find new strategy that can convert them to frequent users. Another example that is different from the two is the midnight medium usage group, Group 1, which is characterized by high F categories 12.5% F5, 87.5% F4 and low M categories 50% M1, 25% M2 and 25% M3, thus, the suggestions are 75% active marketing and 25% taking care of customers’ needs.

4.6. Dissemination

In this phase, after taking into consideration the present decision structure of the company, we feel very strongly that the development of a BI decision support system can better achieve the purpose of sharing and imbedding intelligence discovered in the previous phases; because of the fact that most managerial personnel are unfamiliar with BI methodologies. One key issue of successfully designing such a system is the seamless integration of system components with the information and intelligence. For this particular case, the framework of our decision support system for the company is shown in Fig. 6. It is based on software architecture MVC (Model-View-Controller), which separates data model, user interface, and control logic into three distinct components (Krasner & Pope, 1988; Lalonde & Pugh, 1990; Singh et al., 2002). In the system, the Model part is business logic layer that consists of various models for extracting business intelligence of different nature. The View part is the presentation layer and also serves as the user interface of the system, and the ‘‘Controller’’ is the control layer that serves to communicate with View and Model and present the results in the form required by the ‘‘View’’.

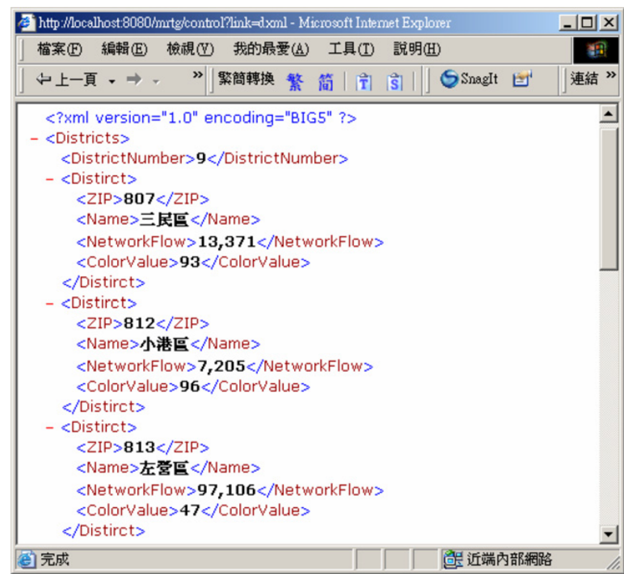


Fig. 8. The XML document of district flow.

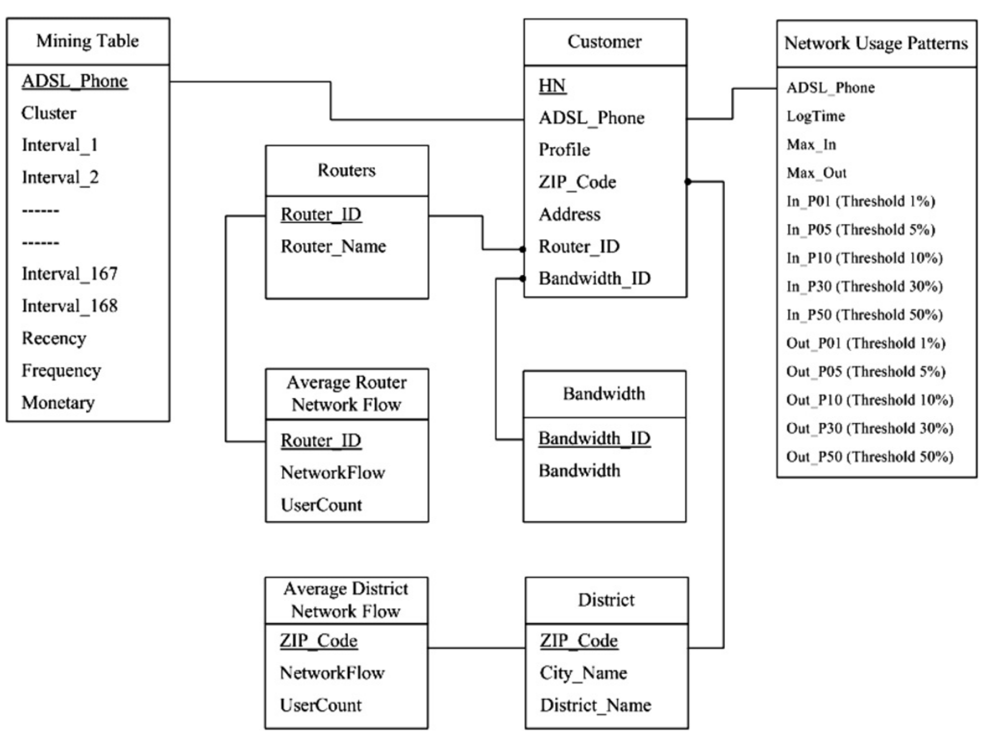


Fig. 7. The E-R model of BI-DSS.

In the framework, the data storage is for the MRTG Log data, which were collected and loaded into MS-SQL 2000 database. We then transform, filter, and normalize these original data; before they become valid for building a data warehouse via MS-SQL Data Transformation Service (DTS). Fig. 7 shows the E-R Model, its entity tables are identified with users, network flow, bandwidth, router, average router network flow, districts, average districts network flow, and mining table. These tables are related to each other with relevant attributes as shown.

The Model part consists of various models mentioned above for extracting business intelligence of different nature. The Controller part is conducted by Java Servlet, whose major functions include user authorization, initiating mining operation, and organizing results for presentation in View. For example, Fig. 8 displays the districts network flow XML format with Flash MX 2004 for visualizing results in the View.

In the View module, the user interface is implemented with JSP that cooperates with JavaScript; its symbol set is encapsulated in Flash objects. One example is given in Fig. 9, which presents visualization of router and districts network flow distribution.

Fig. 10 presents the visualization results of potential VIP customers' characteristic and discriminative rules (Li et al., 2006), which could greatly help decision makers analyze customer VIP status by city, district, or HN.

The customer service management function allows users to analyze customer network behavior based on time interval or cluster grouping. The interface is via time interval, cluster list, customer contact list, or user profile. Fig. 11 presents the network behavior for user ADSL_Phone = 07-xxxx651 in

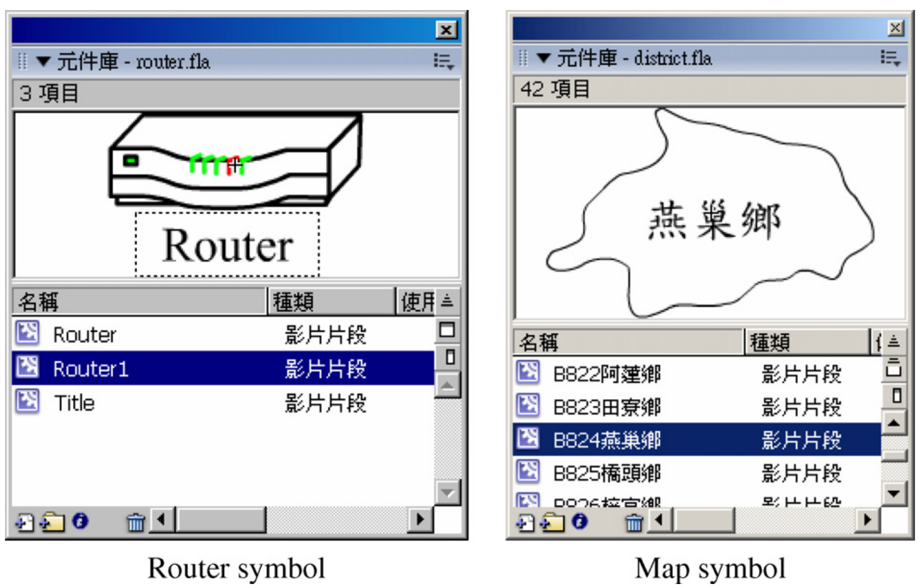


Fig. 9. Symbol sets of BI DSS.

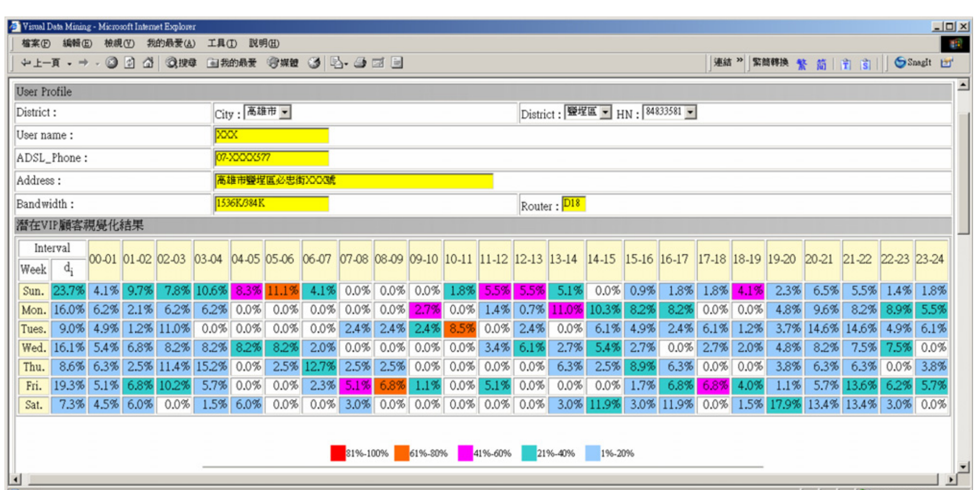


Fig. 10. The visualization interface of potential VIP customers' rules.

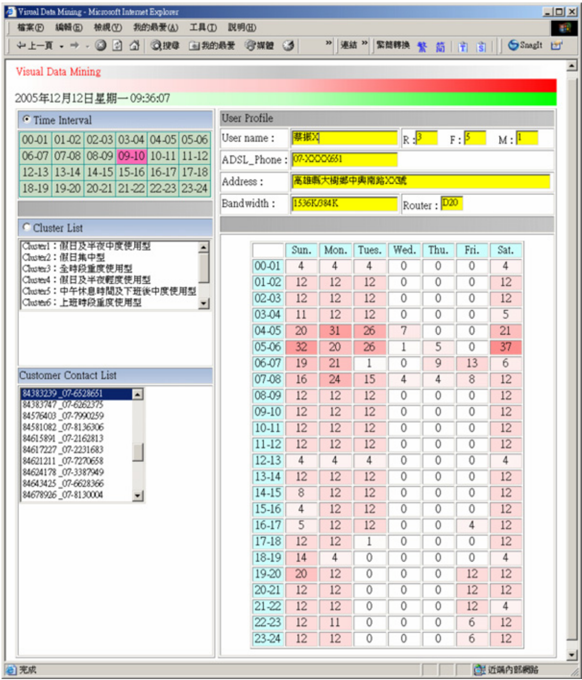


Fig. 11. Customer service management interface by time interval.

the interval 9–10 a.m. Fig. 12 presents the network behavior for user ADSL_Phone = 07-xxxx699 who is of the office hour group with heavy usage.

5. System evaluation

This paper draws from the literature (Papamichail & French, 2005; Zeleznikow & Nolan, 2001) to conduct the system evaluation of this BI DSS, and evaluation methods include both empirical, and subjective aspects. Empirical aspect aims to measure the performance of the system and its users, e.g. whether the decision makers make significantly better of faster decisions to improve decision quality. The Subjective aspect aims to measure the utility of the system, e.g. whether the system addresses an important problem, or whether the system meets the needs of its users and how well its interface is designed. We employ both the survey research and depth interview method for conducting the system evaluation. A two-part questionnaire was designed for subjective and empirical tests. Some important concepts in designing the questionnaire are giving below (Bailey & Pearson, 1983):

(a) Perceived utility: the user's judgment about the relevant balance between the cost and the considered usefulness of the DSS.
(b) Relevance: the degree of congruence between what the user wants or requires and what the DSS provides.
(c) Understanding of the system: the degree of comprehension that a user has about the system or services that are provided.
(d) Completeness: the comprehensiveness of the output information content.
(e) Format of output: the layout design and display of the output contents.
(f) Ease of use: the amount of effort required by the user to take advantage of the tools provided by the system.
(g) Performance: the ability of a DSS to help a DM accomplish a task more effectively.
(h) Usefulness: the extent to which an application contributes to the enhancement of the user's performance.

We used a five point Likert-type scale from 1 (Very disagree) to 5 (Very agree), with 3 being the midpoint (indifferent), for users to respond to the questionnaire. There were 11 statements in the subjective aspect and three statements in the empirical aspect. Twenty subjects took part in the evaluation, they consist of staff from decision makers, marketing, network planners, and call center of the company under study. Table 8 presents the results for the subjective aspect, and Table 9 for the empirical aspect.

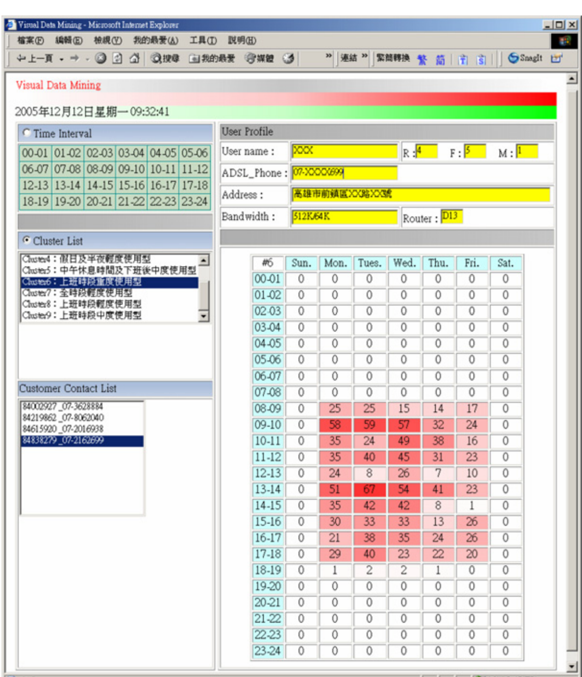


Fig. 12. Customer service management interface by cluster list.

Table 8
Evaluation of the subjective aspect

Criteria	Questions	Mean	Standard deviation
Perceived utility	9, 11	4.27	0.26
Relevance	1, 6, 8	3.82	0.41
Understanding	2	3.91	0.30
Completeness	3	3.73	0.47
Format of output	4	4.18	0.60
Ease of use	5	3.91	0.30
Performance	7	4.09	0.30
Usefulness	10	3.73	0.79

Table 9
Evaluation of the empirical aspect

Questions	Mean	Standard deviation
1. The BI DSS facilitates decision making, e.g. network development, business marketing, and so on	3.82	0.60
2. The visualization of the BI DSS expedites decision making time	4.36	0.51
3. The Web-based model of the BI DSS achieves system integration and information share effectively	4.36	0.51

The mean scores of both Tables 8 and 9 are greater than 3.0, on a scale from 1 to 5, this shows that the BI DSS has met the evaluation criteria of both subjective and empirical aspect. The mean scores of Table 8 range between 3.73 and 4.27; with the "complement" and "usefulness" being the criteria with the lowest score. A discussion with responding subjects reveals the concern that BI DSS may not provide sufficient information for users. The 'perceived utility', on the other hand, was rated the highest mean score (4.27), and this does indicate that the most subjects are positive towards the system and consider the system to be useful in developing service strategies.

6. Conclusions

This research examines the issue of developing an effective service management strategy that Taiwan's ISP management is currently facing and proposes a BI process that could assist management in discovering in-depth knowledge of customer usage behaviors and network facility utilization. With IP traffics of each individual user as data, this process emphasizes significantly on data preprocessing and subsequent modeling. We present appropriate technologies along with each phase of the process to aid management in implementation, which includes a decision support system to assist management who are otherwise not capable of integrating various methodologies together. With the cooperation of a major ISP company in Taiwan, we demonstrated the detailed implementation of this process.

The focus of this process is to discover insight knowledge of users' network usage behaviors, the characteristics of their monetary contribution, and facility utilizations; so that services that are proactive and personalized may be developed. We applied data warehouse to facilitate the retrieval of data with different dimensionality, data mining to discover network behaviors knowledge, and RFM to characterize users' monetary contribution. The mining results of nine customer clusters for the region reveal usage patterns that were unknown to the management before; hence this research does present a brand new concept to the company management in providing personalized services. These findings identify degree of usage, time of usage, and day of usage of each group and are of important information to both service department and sales department. These departments need to re-align resources of both work force and working hours. The analysis of RFM model on usage patterns provides further significant knowledge to management, which could lead them to formulate proper marketing strategies. For the network resource management, we apply data mining to find out network flow distribution among districts of the region, and map that district flow to the routers that service the region. Thus, any future investment on router and related resources will be more effectual and cost effective. For this particular case, we developed a decision support system that integrates all methodologies of the BI process together; so that management could perform analysis when they need to and on subjects they are interesting in. The result of system evaluation has indicated very positive responses from the management staff.

References

Abidi, S. S., & Ong, J. (2000). A data mining strategy for inductive data clustering: asynergy between self-organising neural networks and K-means clustering techniques. *IEEE Transactions on Neural Networks, 2*, 568–573.

Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science, 29*, 530–545.

Berry, J. A. M., & Linoff, G. (2000). *Mastering data mining: The Art and science of customer relationship management*. John Wiley & Sons.

Buytendijk, F., (2001). The BI Framework Provides Return of Integration. Gartner Group, Research Note SPA-13-8990, 11 Jully.

Davis, M. 2002. *Using business intelligence for competitive advantage*. CRMToday.

Ericsson. (2005). Enhancing telecom management. http://www.ericsson.com/products/white_papers_pdf/telecom_management.pdf, July.

Gartner. (2007). Business intelligence & information management. http://www.gartner.com/ap/bi, Sydney, Australia.

Han, J. et al. (1992). Knowledge discovery in databases: An attribute-oriented approach. In *Proceedings of international conference on very large data bases* (pp. 547–559).

Hughes, A. M. (1994). *Strategic database marketing*. Chicago, IL: Probus Publishing Company.

Institute for Information Industry. (2006). http://www.find.org.tw/0105/howmany/howmany_disp.asp?id=159.

Kemper, B. (1997). MRTG-2.8.12 logfile-format.

Kimball, R. (1996). *The data warehouse toolkit: Practical techniques for building dimensional data warehouses*. John Wiley & Sons.

Kohonen, T. (1997). *Self-organizing maps*. Germany: Springer.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE, 78*(9), 1464–1480.

Krasner, G. E., & Pope, S. T. (1988). A cookbook for using the model-view-controller user interface paradigm in Small-Talk-80. *Journal of Object-Oriented Programming, 1*(3), 26–49.

Lalonde, W., & Pugh, J. (1990). *Inside smalltalk* (Vol. II). Englewood Cliffs, NJ: Prentice Hall.

Li, S. T. et al. (2006). Enabling customer relationship management in ISP services through mining usage patterns. *Expert Systems with Applications, 30*, 621–632.

Li, S. T., & Shue, L. Y. (2004). Data mining to aid policy making in air pollution management. *Expert Systems with Applications, 27*, 331–340.

McCue, A. (2006). TalkTalk is UK's least-liked broadband provider. Broadband & ISPs.

Maria, F. (2005). Improving the utilization of external strategic information. Tampere University of Technology, Master of Science Thesis.

Oliveira, M. C. F., & Levkowitz, H. (2003). From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics, 9*(3), 378–394.

Oracle Corporation. (2006). Customer Loyalty: The New Competitive Advantage. VoIP Research Library.

Papamichail, K. N., & French, S. (2005). Design and evaluation of an intelligent decision support system for nuclear emergencies. *Decision Support Systems, 41*, 84–111.

Pirttimaki, V., & Hannula, M. (2003). Process models of business intelligence. *Frontiers of E-Business Research*, 250–260.

Singh, I. et al. (2002). *Designing enterprise applications with the J2EE platform* (2nd ed.). Addison-Wesley.

Stone, B. (1989). *Successful direct marketing methods* (4th ed.). NTC Business Books.

Stone, B. (1995). *Successful direct marketing methods*. Lincolnwood, IL: Lincoln.

Sung, H. H., & Sang, C. P. (1998). Application of data mining tools to hotel data mart on the Intranet for database marketing. *Expert System with Applications, 15*(1), 1–15.

Thomas Jr., J. H. (2001). Business intelligence-why? eAI Journal, 47–49, July.

TWNIC (2004). The survey plan of bandwidth network usage in Taiwan. http://www.twnic.net.tw/download/200307/200307index.html.

Vesanto, J., & Alhoniemi, E. (2000). Clustering for the self-organization map. *IEEE Transactions on Neural Networks, 11*(3), 568–600.

Vitt, E., et al. (2002). *Business intelligence-making better decision faster*. Redmond, Microsoft Press (p. 202).

Viva Business Intelligence Inc. (1998). Developing a business intelligence process: Viva Business Intelligence-cycle approach. Helsinki: Pro-How Paper, 2(28).

Wayland, Robert, E., Paul, M. Cole (1997). *Customer connections: New strategies for growth*. Harvard Business School Press.

Xie, W., et al. (2001). Business intelligence based group decision support system, info-tech and info-net. In *Proceedings. ICII 2001-Beijing, international conference on, 5* (pp. 295–300).

Zeleznikow, J., & Nolan, J. R. (2001). Using soft computing to build real world intelligent decision support systems in uncertain domains. *Decision Support Systems, 31*, 263–285.