University of Wollongong

Research Online

Faculty of Engineering and Information Sciences - Papers: Part A

Faculty of Engineering and Information Sciences

1-1-2013

A bi-level decision model for customer churn analysis

Ya Gao
University of Technology Sydney

Guangquan Zhang
University of Technology Sydney, Guangquan.Zhang@uts.edu.au

Jie Lu
University of Technology Sydney, Jie.Lu@uts.edu.au

Jun Ma
University of Wollongong, jma@uow.edu.au

A bi-level decision model for customer churn analysis

Abstract

This paper develops a bi-level decision model and a solution approach to optimizing service features for a company to reduce its customer churn rate. First, a bi-level decision model, together with its modeling approach, are developed to describe the gaming relationship between decision makers in a company (service provider) and its customers. Then, a practical solution approach to reaching solutions for the bi-level-modeled customer churn problem is developed. Finally, experiments and case studies are conducted to illustrate the bi-level decision model and the solution approach.

Keywords

Disciplines

Engineering | Science and Technology Studies

Publication Details

A Bi-level Decision Model for Customer Churn Analysis

Ya Gao[1], Guangquan Zhang[1], Jie Lu[1] and Jun Ma[1,2]

1. Decision Systems & e-Service Intelligence (DeSI) Laboratory
Centre for Quantum Computation & Intelligent Systems
School of Software
Faculty of Engineering and Information Technology
University of Technology, Sydney,
PO Box 123, Broadway, NSW 2007, Australia
2. SMART Infrastructure Facility, University of Wollongong
Wollongong, NSW 2522, Australia
yagao7411@gmail.com,{guangquan.zhang, jie.lu}@uts.edu.au, jma@uow.edu.au

Abstract. This paper develops a bi-level decision model and a solution approach to optimizing service features for a company to reduce its customer churn rate. First, a bi-level decision model, together with its modeling approach, are developed to describe the gaming relationship between decision makers in a company (service provider) and its customers. Then a practical solution approach to reaching solutions for the bi-level modeled customer churn problem is developed. Finally, experiments and case studies are conducted to illustrate the bi-level decision model and the solution approach.

Keywords: bi-level programming, decision modeling, customer churn, optimization

1 Introduction

The business environment today is characterized by high competition and saturated markets. In this environment, companies increasingly derive revenue from the creation and enhancement of long-term relationships with their customers (Coussement, et al 2010). Reinartz and Kumar (2003) have investigated and found that it is more profitable to keep and satisfy existing customers than it is to constantly attract new customers who are characterized by a high attrition rate. Torkzadeh et al. (2006) suggested that the cost of gaining a new customer could be around twelve times more than the cost of retaining an existing one. Moreover, retained customers usually produce higher revenues and margin than new customers (Reichheld and Sasser 1990). In such circumstances, customer churn, which is defined as the propensity of customers to cease doing business with a company in a given time period, has become a significant problem and is one of the prime challenges encountered by many industries worldwide (Chandar, Laha and Krishna 2006).

For many non-traditional industries and service providers which have the characteristic of being highly competitive and easy to switch between, such as telecommunications and Internet service providers, customer churn is becoming a major problem requiring vital and serious consideration. According to Wei and Chiu's investigation (Wei and Chiu 2002), the monthly predicted churn rate is about 2.2% in wireless telecommunication industries. This means that a telecommunication company loses about 27% of its subscribers every year. On the other hand, acquiring new customers is much more expensive than retaining existing ones (Kisioglu and Topcu 2011). Retaining current subscribers is therefore an efficient, but challenging approach to enable a company to stay in the market.

A broad range of data mining and statistics techniques have been adopted to address the customer churn problem, but there are still limitations. Decision-tree-based algorithms have been extended to determine the churn ranking, but it is possible that some leaves in a decision tree have similar class probabilities and the approach is vulnerable to noise. The neural network algorithm does not explicitly express the uncovered patterns in a symbolic, easy-to-understand way. Genetic algorithms can produce accurate predictive models, but they cannot determine the likelihood associated with their predictions (Xie, et al 2009). Other methods, such as the Bayesian multi-net classifier (Luo and Mu 2004), support vector machine, sequential patterns, and survival analysis (Lariviera and van den Poel 2004), have made competitive attempts to predict customer churn, but the error rates are still significant.

The above approaches mainly focus on searching the behavior indicators of customers who are likely to churn. Solutions or campaigns will then be identified to retain those potential churners. Our research addresses the customer churn problem from an innovative perspective: we do not search the customers who will churn; we research the reasons why customers want to churn and identify factors that could reduce customer churn rates using bi-level programming techniques. The main difference between the work presented in this paper and the existing research on customer churn is that our research seeks to aid the design of service features that will meet customers' certain preferences and thus retain them, whereas other approaches use techniques such as neural networks and Bayesian networks to predict customers' churn behaviors. They are different approaches to solving customer churn problems, yet they have the same aim of reducing customer churn rates. The analysis results show that corresponding service features for a company can be identified by the method proposed in this research to minimize the customer churn rate.

Bi-level programming techniques are used for partition control over the decision variables of a decision problem between the two levels of decision entities. The decision entity at the upper level, known as 'leader', will influence or induce the behaviour of the decision entity at the lower level, called 'follower', but will not completely control the follower's actions. In addition, the follower achieves its aim of maximizing or minimizing a mathematically formatted objective function under certain constraint functions faced by followers (Amat and McCarl 1981). In such a bi-level decision situation, the decision entity at each level has individual payoff functions, but the decision of the upper level is global. Therefore, the final solution of a bi-level decision problem reflects the leader's goal and also considers the reaction of the follower (Bard 1998). The investigation of bi-level decision problems is strongly motivated by real world applications, and bi-level programming techniques have been

applied with remarkable success in different domains such as decentralized resource planning (Yu, Dang and Wang 2006), the electric power market (Hobbs, Metzler and Pang 2000), logistics (Zhang and Lu 2007), civil engineering (Amat and McCarl 1981), and road network management (Feng and Wen 2005; Gao, Zhang and Lu 2007).

Bi-level decision problems have been studied for decades. A large part of the research on bi-level decision problems has centered on linear bi-level decision problems, for which nearly two dozen algorithms have been developed. The well-known ones include the Kuhn-Tucker approach (Bard 1998), the Kth-best algorithm (Zhang, Lu and Dillon 2006), the Branch-and-bound algorithm (Lu, Shi and Zhang 2006b), and heuristic based approaches (Murata and Ishibuchi 1995). For bi-level decision problems with non-linear formats for either the objective functions or constraints functions from the leader or follower, particle swarm optimization is an emerging and competitive solution strategy that has the advantages of high computation efficiency and effectiveness (Gao, Zhang and Lu 2009) . Current research into bi-level programming techniques includes nonlinear (Gao, Zhang and Lu 2009), multi-leader (Zhang, et al 2010), multi-follower (Lu, Shi and Zhang 2006a), multi-objective (Feng and Wen 2005), and fuzzy bi-level decision problems (Zhang and Lu 2007).

A frequently occurring decision situation is when a bi-level decision problem involves multiple followers. In this kind of decision situation, the leader's decision will be affected not only by the followers' individual reactions but also by their relationships. The relationships among these followers are mainly reflected by the sharing of objective functions, constraints and variables (Lu, Shi and Zhang 2006a). For the different relationships among them, we need different models to describe them and different algorithms to obtain solutions. The decision problem addressed by this research falls into the category of bi-level decision problems with multiple followers.

To investigate the hierarchical nature of customer churn problems, this research develops a bi-level decision model for customer churn problems. This decision model analyzes the gaming relationship between a service provider and its customers. The service provider plays the role of leader by initially designing service features, following which customers act as followers by reacting accordingly (churning or not). This decision model fully considers the sequential decisions and the mutual influences between a service provider and its customers. That is, both the service provider and its customers aim to optimize their own objectives, but their decisions are related to each other in a hierarchical way.

The methodologies proposed in this paper can be applied to address customer churn problems in many industries, such as telecom, internet service, and even catering. However, to format a bi-level decision model for a specific industry, the corresponding industry background and knowledge are required. This research takes the telecom industry, where customer churn is a frequent and well-recognized problem, as an application field to illustrate the proposed methodologies.

This paper is organized as follows. Following the introduction in Section 1, Section 2 introduces related definitions and theories of bi-level decision techniques. Section 3 presents a bi-level decision model together with a modeling approach to describe the decision relationships between service providers and their customers in the customer

churn analysis. To solve the problems defined by this bi-level decision model, Section 4 provides a solution approach to optimize service features to retain customers. Section 5 employs a numerical example and a case study for experimentation. Finally, conclusions and further studies are outlined in Section 6.

2. Preliminaries

2.1 Bi-level Decision Models

A bi-level decision problem can be viewed as a static version of a two-player (decision entity) game. The decision entity at the upper level is termed the 'leader', and at the lower level, the 'follower'. The control for the decision variables is partitioned amongst the decision entities who seek to optimize their individual objective function.

Bi-level programming typically models bi-level decision problems in which the objectives and constraints of both the upper and lower level decision entities (leader and follower) are expressed by linear or non-linear functions. A definition of a bi-level decision problem has the following format (Bard 1998):

For $x \in X \subseteq R^n, y \in Y \subseteq R^m,\ F: X \times Y \to R^1,\ and\ f: X \times Y \to R^1,$

$$\min_{x \in X} F(x, y)$$
$$\text{subject to } G(x, y) \leq 0$$
$$\min_{y \in Y} f(x, y) \tag{1}$$
$$\text{subject to } g(x, y) \leq 0$$

where the variables x, y are called the leader's decisions variables and follower's decision variables respectively, $F(x,\ y)$ and $f(x,\ y)$ are the leader's and follower's objective functions, $G(x,\ y)$ and $g(x,\ y)$ are the leader's and follower's constraint functions.

This model aims to find a solution to the upper level problem $\min_{x \in X} F(x, y)$ subject to its constraint $G(x, y) \leq 0$. For each value of the leader's variable x, y is the solution of the lower level decision problem $\min_{y \in Y} f(x, y)$ under its constraint $g(x, y) \leq 0$.

The above model describes a classical bi-level decision problem with one leader and one follower. Many real world bi-level decision situations may involve many followers. A bi-level decision problem with one leader and L followers is defined as:

For $\ x \in X \subseteq R^m, y_1 \in Y_1 \subseteq R^{n_1}, \dots, y_L \in Y_L \subseteq R^{n_L}, Y = Y_1 \times Y_2 \times \dots \times Y_L,\ F,\ f_i :$
$X \times Y \to R^1,\ i = 1, \dots, L, L \geq 2$
$$\min_{x \in X} F(x, y_1 \dots, y_L,)$$
$$\text{subject to } g_j(x, y_1 \dots, y_L) \leq 0, j = 1, \dots, p$$
$$\min_{y_i \in Y_i} f_i(x, y_1 \dots, y_L), i = 1, \dots, L \tag{2}$$

$$\text{subject to } q_{ik}(x, y_1 \ldots, y_L) \leq 0, k = 1, \ldots, i_k, i = 1, \ldots, L$$

In the problem defined by (2), both the leader and the multiple followers have individual controlling variables, objectives and constraints.

To model a real world bi-level decision problem, we need to identify the hierarchy decision relationship among decision makers. Although we term them 'leader' and 'follower', there is no leading or following relationship between a leader and a follower. The decision maker who decides first is called the leader while the second decision maker is called a follower. It is more like a gaming relationship. Once a leader and a follower are identified, their decision variables (the factors on which they make decisions), the objectives, and constraints are the subsequent requirements for formatting and modeling the bi-level decision problem.

To solve a modeled bi-level decision problem, popular methods are the Kuhn-Tucker approach (Bard 1998), the Kth-best algorithm (Zhang, Lu and Dillion 2006), the Branch-and-bound algorithm (Lu, Shi and Zhang 2006b), and heuristic based approaches (Murata and Ishibuchi 1995).

2.2 A Particle Swarm Optimization Algorithm for Multi-follower Bi-level Decision Problems

To solve the decision problem defined by (2), a particle swarm optimization (PSO) algorithm (Gao, et al 2009) was developed. In the PSO algorithm, we first sample the leader-controlled variables to obtain candidate solutions for a leader. We then use the PSO algorithm, together with the Stretching technology (Parsopoulos and Vrahatis 2002), which is used to prevent local solutions, to obtain the follower's response for every choice of the leader. Thus, a pool of candidate solutions for both the leader and the follower is formed. By pushing every solution pair towards the current best solutions, the solution pool is updated. Once a solution is reached for the leader, the Stretching technology (Parsopoulos and Vrahatis 2002) is used to avoid local optimization. We repeat this procedure by a pre-defined count and reach a final solution.

Below is the outline of this algorithm:

Algorithm A: Generate the response from a follower

Step 1: Input coefficients of x;

Step 2: Sample N_f candidates for followers and the corresponding velocities;

Step 3: Initiate the follower's loop counter $k_f = 0$;

Step 4: Record the best particles for the followers;

Step 5: Update the velocities and positions;

Step 6: $k_f = k_f + 1$;

Step 7: If $k_f \geq MaxK_f$, which is the pre-defined maximum number k_f, or the solution changes for several consecutive generations are small enough, then we use the stretching technology (Parsopoulos and Vrahatis 2002), which is a technology to prevent local solutions, to obtain the global solution and go to Step 8. Otherwise go to Step 5;

Step 8: Output the response from the follower.

Algorithm B: Generate the optimal strategy for a leader

Step 1: Sample N_l particles of decision variable x, and the corresponding velocities;

Step 2: Initiate the leaders' loop counter $k_l = 0$;

Step 3: For k−th particle, k =1,…, N_l, calculate the optimal responses from l – th follower by Algorithm A, $l = 1, \ldots, L$;

Step 4: Calculate the objective value for the leader of every particle in (2);

Step 5: Record the best previously visited position for each particle;

Step 6: Move current particles by the best positions;

Step 7: $k_l = k_l + 1$;

Step 8: If the sum of the differences between the current optimal responses and the previous found optimal responses is small enough or $k_l \geq MaxK_l$, which is the pre-defined maximum number of k_l, we use Stretching technology for the current leader's solutions to obtain the global solution. Otherwise, go to Step 3.

Above is the outline of the particle swarm optimization algorithm to reach a solution for a bi-level decision problem with multiple followers. This algorithm will be used in subsequent sections to solve a customer churn problem.

3   A Bi-level Decision Model for Customer Churn Analysis

This section presents a modeling approach to developing a bi-level decision model for costumer churn analysis.

Bi-level decision making is to describe decision situations where two decision parties are involved. One is called the leader, and the other is called the follower. The leader makes the decisions first. Based on the choices from the leader, the follower makes his or her decisions. Both the leader and the follower aim to maximize their own objectives under certain constraints subject to but without having to obey the other party's decision. However, since the choices from either party will influence the achievements of the other party's objectives, they both need to consider the choice from the other. In a customer churn problem, a service provider makes decision first by designing plan features. The customers then decide if they would like to remain or

leave based on their own needs. The service provider cannot force the customers to be loyal to it. The customers cannot force the service provider to change their services either. However, if the services are not competitive enough, the customer may churn. Thus the service provider will lose the market. To stay in the market, the service provider needs to consider the requirements from its customers. This customer churn problem is thus a typical bi-level decision problem. To establish the bi-level decision model for describing customer churn problems, we need to determine the objectives and constraints for the service provider (leader) and the customers (followers) by mathematic formulas. This section presents a modeling approach to establishing the objectives and constraints of this bi-level decision model for a customer churn problem.

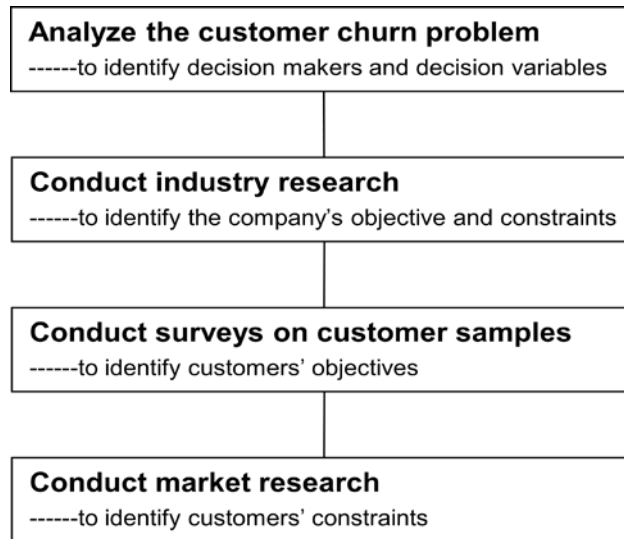Four steps are employed to conduct the modeling approach, as illustrated in Fig. 1.

Fig. 1. The steps in conducting the modeling approach.

The first step is to analyze a customer churn problem to identify the decision makers and the decision variables. Analyzing the customer churn problem from the perspective of decision making, we identify two kinds of decision makers in the problem: a service provider and the current customers. We illustrate the decision making process in Fig. 2. The service provider makes decisions on factors of plan features such as service rates, pre-paid or post-paid, while the customers decide whether they will churn or remain with their current service provider. Once the service provider issues a specific plan with a certain set of plan features, the customers then compare this plan with those of the competitors to decide whether they will switch or not. The service provider and its customers make sequential and independent decisions, but their decisions will have an influence on each other. This kind of decision problem has typical bi-level decision making features. The service provider acts as a leader, while the customers react as followers. We note that in this bi-level decision modeled customer churn problem, there is not only one follower

(customer), but many. Thus, unlike a classical bi-level decision problem defined by (1), where only one leader and one follower are involved, the customer churn decision problem falls into the category of one leader, multi-follower bi-level decision problems as defined by (2).
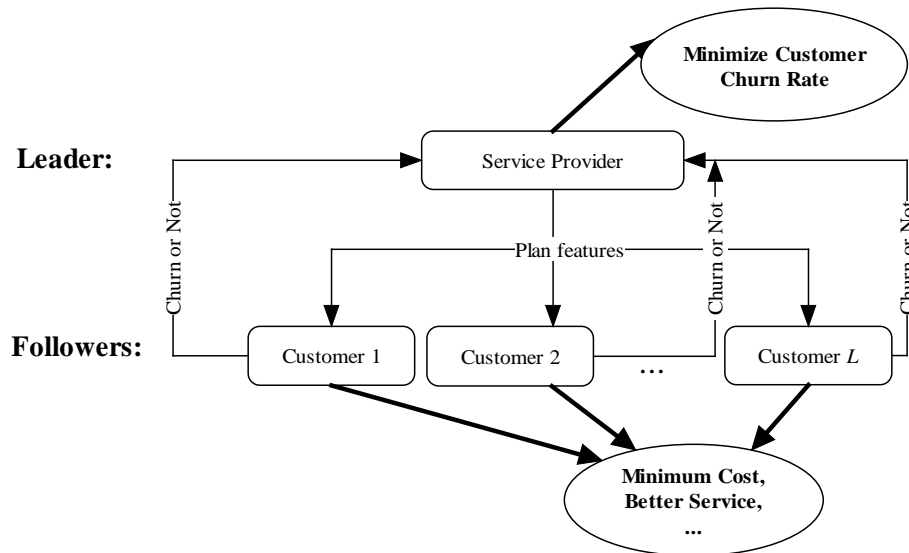

Fig. 2. The outline of the bi-level decision process for a customer churn problem.

The second step is to conduct industry research to identify the objective and constraints faced by the leader. In the bi-level decision modeled customer churn decision problem, the objective of the leader (service provider) is to make the whole churn rate as low as possible, while its decisions are within certain constraints such as budgets, profit expectation, and technical limitations. The objectives of the followers (customers) vary: either they seek minimal cost, clear signals, better services, or a combination of any of these. We conducted industry research in this step to identify the objectives and constraints.

The third step is to obtain the objectives of the followers. When evaluating followers' objectives, we normally evaluate customer satisfaction based on different service features. Sometimes, this kind of information is not available in our database, thus we need to conduct surveys to discover the mathematical formula of customer satisfaction defined by service features. For example, a question in the survey might be: which pricing unit do you prefer? The candidate's answer can be: 10 seconds, 30 seconds, or 60 seconds. If 7 of 10 surveyed customers choose 30 seconds, the plan feature of 30 seconds has the satisfaction degree of 0.7. In a customer churn problem, a customer, as a decision maker, decides whether or not she or he will churn for certain reasons. They may seek lower fees, better service, or have a combination of reasons. To discover the objectives for a certain group of customers, we can use the existing data or conduct surveys on samples of customers. In our surveys, we asked the respondents to give rankings, such as very satisfactory, satisfactory, or not

satisfactory, to the services they currently receive or to select the plan features they would prefer. Based on the survey data, we use the regression method to obtain the mathematic definition of the relationship between customer satisfaction ranking and service features.

The fourth step is to conduct research on the current market to understand the current market situation faced by customers. The reason a customer churns is direct; that is, there are other service providers who can help them better achieve their objectives. These providers may have lower fees, a stronger network, or better service. The constraints faced by an existing customer are other available service providers and their detailed service features. We need to conduct market research to obtain this information.

From the above four steps on the modeling approach, we develop a bi-level decision model for customer churn problems as follows:

For $(x_1, \dots, x_m) \in X, y_1, \dots, y_L \in Y = \{0,1\}$, F, $f_i : X \times Y \to R^1, L \geq 2$

$$\min_{x \in X} F(x_1, \dots, x_m, y_1 \dots, y_L,) = y_1 + y_2 + \dots + y_L + \eta(y_1 \dots, y_L) \qquad (3.1)$$

$$\text{subject to } g_j(x_1, \dots, x_m) \leq 0, j = 1, \dots, p \qquad (3.2)$$

$$\max_{y_i \in Y} f_i(x_1, \dots, x_m, y_i) = \begin{cases} f_i, (y_i = 0) & f_i \leq Q \\ Q, (y_i = 1) & f_i > Q \end{cases}, i = 1, \dots, L \qquad (3.3)$$

$$\text{subject to } q_{ik}(x_1, \dots, x_m, y_i), i = 1, \dots, L \qquad (3.4)$$

Explanation:

Variables:

$x_1, \dots, x_m$: The specific plan features, such as x_1 dollars a month for one particular plan.

y_i: The decision on churning or not for the ith customer. If $y_i = 1$, the ith customer decides to churn. If $y_i = 0$, the ith customer decides to remain with the current service provider.

Constants:

L: The number of customers within a certain customer group. Since different customer groups have particular patterns and attributes, we study these groups individually. Here L is not the number of all customers, but the number of customers within a certain group.

Q: The current biggest value of customer satisfaction function f_i, which can be achieved by other available service providers in the current market. This value is obtained through market research in the fourth step in the modeling approach.

Functions:

(3.1) is the objective of the service provider (the leader). It means the service provider aims to minimize the customer churn rate as a whole. $\eta(y_1 \dots, y_L)$ represents

the swarm effects on churn, which are mutual influence effects with certain customer groups. In most situations, $\eta(y_1 \ldots, y_L)$ is 0, meaning that customers will only base their decision to churn or not on the services provided.

(3.2) describes the specific constraints faced by the service provider, such as space constraints, stack ability constraints, loading and unloading rules, warehouse efficiency, load stability, and other constraint limitations.

(3.3) describes the objective for the ith customer. It could be to minimize mobile fees, to obtain better services, or to fulfill other customer satisfaction functions. The customer satisfaction functions can be discovered by conducting customer surveys, as described in the third step of the modeling approach.

(3.4) are the constraints encountered by the i th customer. Specifically, the constraints faced by the ith customer when they aim to achieve their objective is the current market situation in which other service providers can help to achieve the customer's objective to a certain degree.

In most circumstances, it is not direct to retrieve a mathematical formulation of f_i in (3.3) from the customer surveys, which are described in the third step in the modeling approach. Unlike a classical bi-level decision problem, where both the objectives and constraints are defined by clear mathematical formulas, this customer churn problem is a formless bi-level decision problem. To solve this problem, we use the regression method to simulate the relationship between plan features and customer satisfaction. Existing bi-level programming techniques are then applied to solve the customer churn problem.

In the above bi-level decision modeled customer churn problem, the service features from other service providers often vary from the choices offered by one service provider. It is a kind of multi-leader multi-follower bi-level decision problem. When helping one particular service provider to reduce its customer churn rate, we usually take the choices from other service providers as a constant input to the constraints; thus, the problem becomes a one leader multi-follower bi-level decision problem.

The model defined in (3) is a general model which can be applied to many industries. However, the decision variables, objective functions, and constraints in (3) need to be specified when addressing the customer churn problem for a particular field. In the experiments in Section 5, we apply this model to telecom industries for customer churn problems.

4   A Solution Approach to Optimizing Business Decisions on Customer Retention

In this section, we develop an approach to solving customer churn problems modeled by (3).

Addressing the customer churn problems defined by (3), the following practical steps are adopted.

Step 1: To build a multi-dimensional database that is tailored for customer churn analysis from a data warehouse of a company for which we conduct the customer churn analysis.

Step 2: To pre-process data (such as removing noise, filter outliers, replacing missing values, and transforming variables) and to select the most relevant and important variables that are related to the target of churning.

Step 3: To segment customers by their consumption behaviors.

Step 4: To conduct customer surveys and to establish the mathematical formulas of the customer objectives defined by (3.3) by regression method from the surveyed data.

Step 5: To solve the customer churn problem (3) by the bi-level particle swarm optimization algorithm introduce in Section 2.2.

There are two aims in Step 1. One is to gain an insight to the available information from a multi-dimensional perspective to explore the original data; the other is to speed up data extraction and facilitate data inspection. For a multi-dimensional database that is tailored for customer churn analysis, related dimensions include service package, time, customer location, and customer demographics, with the detailed transaction information as the central fact table. In this step, we need to be aware that the churn information of a customer is not always directly known to us. We need to add the information concerning whether to churn or not in the multi-dimensional database. We can use the Bayesian network (Kisioglu and Topcu 2011) or the predictive method (Xie, et al 2009) to decide whether a customer has churned or will possibly churn.

In Step 2, we first clean the data to remove irrelevant information, replace missing values, and add necessary information. Often, we need to transform input variables to structure information, since the available data defined in the current database may not have exactly the data structures we need. For example, we have the options of a contract period of '12 months', '24 months', or 'null'. In the database, this information is a combination of these options. We need to transform this information as categorical variables with each value representing a combination of variables. From our practical experience, if a variable displays a highly skewed distribution, an appropriate transformation, such as standardizing or taking the log, needs to be adopted to achieve normal distribution for a better fitted model. We then analyze input variables to remove two kinds of variables. The first are the variables that are not closely related to customer churn. The second are the redundant variables that are highly correlated with those that have been identified as important variables to determine customer churning. These are all very important for establishing a correct churn model to better reflect customer churn information in the following steps.

Step 3 requires customer segmentation. Since different customer groups have different behavior profiles, we need to address each of them using certain strategies. There is no single service plan that can retain all our customers. A set of optimized features for one service plan will target one particular customer segmentation. Our task is to optimize plans for each customer group. The outline for the optimized plans for different customer groups is shown in Fig. 3. The classification method is recommended if we can clearly identify each segmentation and the classification rules. Otherwise, the clustering method is recommended for this step.
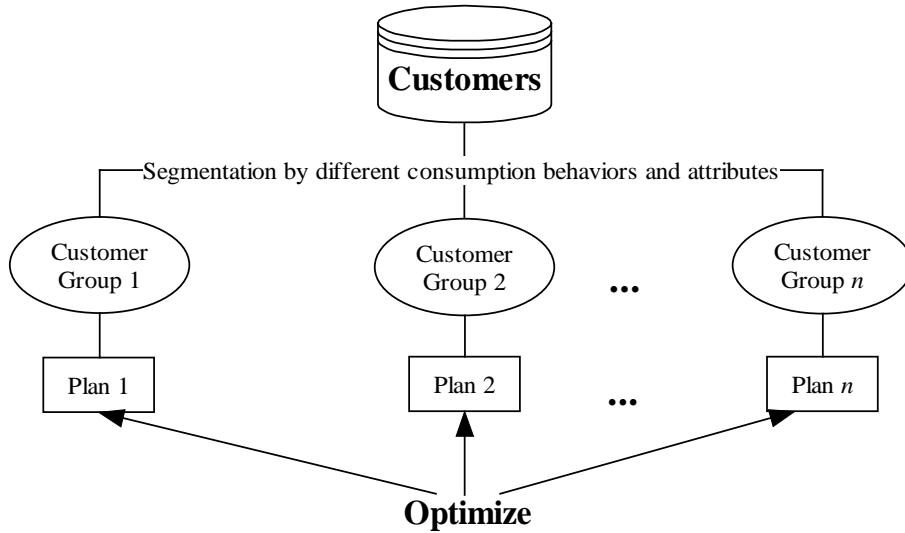
Fig. 3. The outline of customer segmentation.

To develop a mathematical definition of the customer satisfaction function in Step 4, we need first to conduct surveys to collect data. The regression method is then used to establish the mathematical formula of the relationship between customer satisfaction degrees and service features.

In Step 5, we solve the bi-level decision modeled customer churn problem. Since the formats for the objectives and constraints defined in (3) can be linear or non-linear, we adopt the particle swarm optimization algorithm in Section 2.2, which addresses bi-level decision problems with both linear and non-linear objectives and constraints to solve the churn problem.

The final solutions from this approach are sets of detailed service plan features aimed at different customer segments. Under these plan features, the overall customer churn rate will most likely reach the minimum.

5   Experiments

This section illustrates the bi-level decision model and the solution approach developed in this study by a numerical example. A case study is then conducted to illustrate the application.

5.1 A Numerical Example

Suppose we are solving the costumer churn problem for telecommunications company A. Company A makes decisions on x_1 and x_2, where x_1 is the prepaid monthly fee for a pre-paid mobile calling plan, and x_2 is the service ranking, defined by the quality of the provided service which is quantized as a real number. x_1 and x_2 are the main determining factors which play major roles in determining the degrees of customer satisfaction. Through a survey, we obtain the data on degrees of customer satisfaction and the determining factors of x_1 and x_2. Using the multi-linear regression method, we obtain the mathematical function to define the relationship between x_1, x_2 and customer satisfaction degree function f as: $f(x_1, x_2) = -3x_1 + 4x_2$. Through market investigation, we know that in the current market situation, the best available value of f that can be achieved from other existing telecommunications service providers is 45. The model is detailed as follows:

For $(x_1, x_2) \in X, y_1, \dots, y_L \in Y = \{0,1\}$, F, $f_i : X \times Y \to R^1$,

$$\min_{x \in X} F(x_1, x_2, y_1 \dots, y_L,) = y_1 + y_2 + \cdots + y_L \qquad (4.1)$$
$$\text{subject to } x_1 \geq 23 \qquad (4.2)$$
$$30 \geq x_2 \geq 20 \qquad (4.3)$$
$$\max_{y_i \in Y} f_i(x_1, x_2, y_i) = \begin{cases} -3x_1 + 4x_2, (y = 0) & -3x_1 + 4x_2 \leq 45 \\ 45, (y = 1) & -3x_1 + 4x_2 > 45 \end{cases} \qquad (4.4)$$

$$i = 1, \dots, L$$
$$\text{subject to } y_i \in \{0,1\} \qquad (4.5)$$

We apply the bi-level particle swarm optimization algorithm introduced in Section 2.2 on this problem formatted by (4), and the solution of $(x_1 = 25, x_2 = 30)$ has been reached for Company A. In an ideal and experimental environment, this solution makes the objective value of Company A reach the minimum of 0. However, in a real world application, customer churn rates are influenced by many other factors beyond the service provider's control. No matter how perfect a service plan is, there are always customers who leave. Therefore, the most reasonable recommendation we can give to Company A is that, if Company A sets the prepaid monthly fee as 25 and gives the service a ranking of 30, its customer churn rate for the prepaid customer group will decrease.

We take this example to illustrate the bi-level decision model on customer churn analysis. In a real world mobile plan designing process, more complicated modeling and calculations are needed. In the next section, we will use a case study to illustrate the application.

5.2 A Case Study

In this section, we use the solution approach developed in this study to optimize mobile plans to retain customers for company B. B is a large telecommunication company in Australia, and the data used to build the bi-level decision model for its customer churn analysis is taken from its customer management database, which

includes both the churn information and calling history. For the purpose of company data protection, we have modified the original data and display only parts of it.

Step 1: Based on Company B's data warehouse, we build a multi-dimensional database which is tailored for customer churn analysis. The structure of this multi-dimensional database is illustrated in Fig. 4:
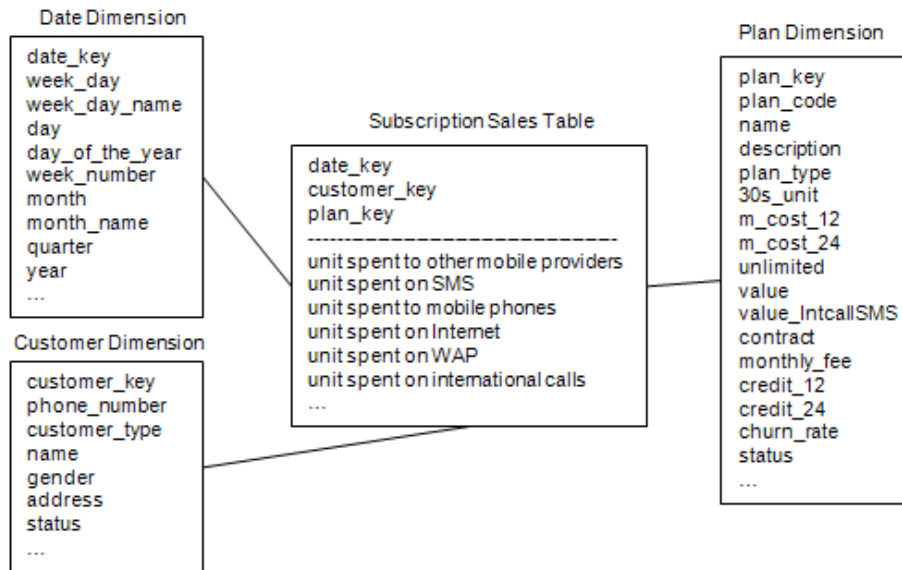

Fig. 4. The structure of the multi-dimensional database.

In Fig. 4, the "Subscription Sales Table" is a fact table on customer subscriptions, while the tables of "Date Dimension", "Customer Dimension", and "Plan Dimension" are the main dimensions for the fact table. The explanations of some items listed in the tables from Fig.4 will be explained when they are used in the following steps.

Step 2: Through inspection on the data that needs to be populated in the multi-dimensional database developed in Step 1, the following major data cleansing and data transformation is conducted:

- The value of 'NULL's in relevant fields are replaced by '0' or certain values calculated based on business rules in Company B.
- From the available data, we note that a customer consumes only a few items, whilst other items have consumption values of null. The data has a typical distribution of large positive skewness. To achieve normal distributions, log transformations are conducted on the data.
- We structure and transform the information on the options of a contract period of '12 months', '24 months', or 'null' as 0 (replacing 'no contract' in the database) / 1 (12M/24M) / 2 (12M/24M/M2M) / 3 (12M) to conform the data source in the original database.

Through variable correlation analysis by the forward stepwise regression method in SAS 9, some variables such as "measure_unit", "plan_name", which are not closely related to "churn" are removed.

Step 3: We use the mobile calling behavior of customers as profile indicators to do the segmentation. Parts of the data are listed in Table 1:

Table 1. Summary of the calling features for customer segmentation.

Unit spent on other mobile providers	Unit spent on SMS	Unit spent on mobile phones	Unit spent on Internet	Unit spent on WAP	Unit spent on international calls
6600	263	25680	0	1	0
5640	4	20880	21035	0	0
15960	45	13320	8893	0	0
1680	1	4620	0	0	0
26070	111	28530	93521	0	0
12000	186	12270	828	746	0
18420	65	53580	123430	13893	5640
2580	0	7350	0	4	60

Since there is no assumption imposed for the customers, we use the clustering method to segment customers into groups. The following variables play important roles in separating the groups:

- consumption on internet visits
- calling domestic numbers from service providers other than Company B
- consumption on voice mail

Two groups of customers, Group 1 and Group 2, have been identified. Group 2 has a larger spend on the above 3 items compared to Group 1. Most members in Group1 and Group 2 have different plans and thus have different plan features. Focusing on each group, we use the bi-level decision model developed in this study to provide recommendations on corresponding plan features to achieve customer retention.

Step 4: We design customer surveys based on the data analysis in the multi-dimensional database. Table 2 lists parts of the corresponding plan features, together with the churn rate, which are used in the survey. In Table 2, plans are not identified since plan features may be common among some plans which are close to one another but have slight differences. For example, a series of plans with common features and aimed at the same users may have been put into the market at different times.

Table 2. Summary of the plan features and churn rates.

Plan feature titles	Plan features										
30s_unit	1	1	1	1	1	1	1	0	0	0	0
m_cost_12	240	360	600	840	1200	1188	1548	588	708	948	588
m_cost_24	480	720	1200	1680	2400	2376	3096	1176	1416	1896	1176
unlimited	0	0	0	0	0	1	1	0	0	0	0
value	20	30	50	70	100	99	129	500	700	900	500

value_IntcallSMS	0	0	0	0	0	100	180	0	0	0	0
contract	1	1	1	1	1	2	2	2	2	2	3
monthly_fee	20	30	50	70	100	99	129	49	59	79	49
credit_12	64	84	148	184	228	0	0	0	0	0	0
credit_24	120	160	290	415	515	0	0	0	0	0	0
Churn_rate	4%	3%	3%	3%	2%	4%	7%	3%	2%	3%	8%

The explanations of the column names in Table 2 are listed below:

30s_unit: if the pricing unit is 30 seconds for a plan, the value for this variable is 1; otherwise, if the pricing unit is 60 seconds, the value is 0.

m_cost_12: the minimum cost per 12 months for the plan.

m_cost_24: the minimum cost per 24 months for the plan.

unlimited: if customer has unlimited usage for the plan, the value is 1. Otherwise, the value is 0.

value: the value that a customer can spend under this plan.

value_IntcallSMS: the value that a customer can spend on international calling and SMS by this plan.

contract: the option of the contract. If the plan has no contract, the value is 0; if the plan has the contract period options between 12 months and 24 months, the value is 1; if the plan has the contract period options among 12 months, 24 months and pay the fee month to month, the value is 3; if the contract period is 12 months, the value is 3.

monthly_fee: cost per month for this plan.

credit_12: the money a customer can spend on buying a mobile phone over 12 months.

credit_24: the money a customer can spend on buying a mobile phone over 24 months.

churn_rate: The average monthly customer churn rate in history for plans having this feature.

Based on these plan features, two surveys were conducted in Group 1 and Group 2 respectively, and the corresponding objective functions for the two customer groups identified. The following bi-level decision models defined by (5) and (6) describe the bi-level modeled telecommunications customer churn problems for Group 1 and Group 2 respectively.

The bi-level decision model for Group 1 is as follows:

For $(x_1, \ldots x_n) \in X, y_1, \ldots, y_L \in Y = \{0,1\}$, $F, f_i : X \times Y \to R^1$,

$$\min_{x \in X} F(x_1, \ldots x_n, y_1 \ldots, y_L,) = y_1 + y_2 + \cdots + y_L \tag{5.1}$$

$$\text{subject to } x_1 \geq 23 \tag{5.2}$$

$$30 \geq x_2 \geq 20 \tag{5.3}$$

$$600 \geq x_3 \geq 0 \tag{5.4}$$

$$\max_{y_i \in Y} f_i(x_1, x_2, x_3, y_i)$$

$$= \begin{cases} 1.2x_1 + 0.3x_2 + 0.06x_3, (y_i = 0) & 1.2x_1 + 0.3x_2 + 0.06x_3 \geq 160 \\ 160, (y_i = 1) & 1.2x_1 + 0.3x_2 + 0.06x_3 < 160 \end{cases} \tag{5.5}$$

$$i = 1, \ldots, L$$

subject to $y_i \in \{0,1\}$ (5.6)

The bi-level decision model for Group 2 is as follows:

For $(x_1, \ldots x_n) \in X, y_1, \ldots, y_L \in Y = \{0,1\}$, F, $f_i : X \times Y \to R^1$,

$$\min_{x \in X} F(x_1, \ldots x_n, y_1 \ldots, y_L,) = y_1 + y_2 + \cdots + y_L \tag{6.1}$$

subject to $x_1 \geq 26$ (6.2)

$\qquad x_2 \in \{0,1,2,3\}$ (6.3)

$\max_{y_i \in Y} f_i(x_1, x_2, y_i)$

$$= \begin{cases} 0.025x_1 + 0.3x_{21} + 0.32x_{22}(y_i = 0) & 0.025x_1 + 0.3x_{21} + 0.32x_{22} \geq 12 \\ 12, (y_i = 1) & 0.025x_1 + 0.3x_{21} + 0.32x_{22} < 12 \end{cases} \tag{6.4}$$

$i = 1, \ldots, L$

subject to $y_i \in \{0,1\}$ (6.5)

where x_{21} and x_{22} are dummy variables of x_2, representing the values when $x_2 = 1$, and $x_2 = 2$, respectively. $x_1, \ldots x_n$ in (5) and (6) are the principal components with the original variables defined in Table 2.

We need to point out that we deliberately modified the coefficients of the bi-level decisions models (5) and (6) to avoid publishing confidential information about the telecommunication company in which we conducted the case study.

Step 5: We apply the bi-level particle swarm optimization algorithm in Section 2.2 on the above models, and find the following:

- For Group 1, the small spending users, the refund to spend on buying a new mobile phone has a positive influence on churn rate. This means that more refunding is likely to cause customers to leave us. This is understandable, because a frugal user on a low budget is unlikely to be interested in buying an expensive, new mobile phone and is therefore likely to be frustrated by being given a refund on items they are not interested in.
- For Group 2, the large spending users, there are two factors which are likely to cause customer churn: contract status and the minimum cost for a duration of 24 months. A contract of either 12 months or 24 months is likely to retain large spending customers, and the churn rate has a slight positive tendency with the minimum cost imposed on the following 24 months.

For the reason given in the numerical example, we provide the following practical suggestions to Company B from the above findings instead of giving specific numbers to the decision variables and objectives:

- To reduce the churn rate in small-spending customers, marketing strategies should be focused on existing services rather than providing funding to customers to spend on new and expensive mobile phones.
- To reduce the churn rate in large-spending customers, strategies should include offering a mid-term contract of 12 months and carefully designing

the requirements on the minimum cost imposed on the following 24 months.

Since the project to reduce customer churn with Company B has not been finalized, the results of applying the bi-level decision-making methodologies proposed in this research are yet to be seen.

The specific bi-level decision model built in this section is based on the data from Company B. Though the modeling methodology can be used for any companies on their customer churn problems, this particular model can only be applied to Company B. Since different companies have different plans and customers, their data are bound to be different. The difference between this bi-level decision model and the bi-level decision models built for other companies is that they have different coefficients. However, the methodologies presented in this paper, including both the modeling approach to developing a bi-level decision model for a customer churn problem, and the approach to solving customer churn problems described by the bi-level decision model, are applicable to other companies.

6 Conclusions and Further Study

This paper establishes a bi-level decision model and modeling approach for customer churn analysis. A practical solution approach is then developed to solve the customer churn problems defined by this bi-level decision model. Experimental results show that the bi-level decision model and the solution approach provide reasonable and effective solutions for designing service plan features for the purpose of reducing customer churn rate.

Our future research will focus on the following:

1) Besides customer churn problems, many real world decision problems, such as supply chain management, electricity market price bidding, and transportation running optimization, have bi-level decision features. However, to model and solve them by appropriate bi-level programming techniques is still a challenge, because many exceptions exist for real world decision problems. Our future research will be channeled into adapting, or extending, the existing bi-level programming techniques to real world bi-level decision problems. Currently, we are collecting data on electricity markets to build the bi-level decision models.

2) Existing bi-level programming techniques have almost exclusively focused on hierarchical decision problems whose objective and constraint functions can be generalized as specific mathematical forms. However, most real world bi-level decision problems only have information that is stored in large databases, from which it is almost impossible to generate their mathematical definitions. This situation has frequently appeared in real world problems, such as the formless objectives for customers mentioned in this study. Unfortunately, modeling and solving a formless bi-level decision problem has not received much attention, in the research literature to date. To deal with bi-level decision problems which cannot be modeled by mathematical forms will be a future focus.

7 Acknowledgment

References

Amat J., McCarl B. 1981. A representation and economic interpretation of a two-level programming problem. Journal of the Operational Research Society. 32, 783-792

Bard J. 1998. Practical bilevel optimization: Algorithms and applications, Amsterdam: Kluwer Academic

Chandar M., Laha A., Krishna P. 2006. Modeling churn behavior of bank customers using predictive data mining techniques. In National Conference on Soft Computing Techniques for Engineering Applications (SCT-2006), March, 24–26

Coussement K., Dries F. B, Van den Poel D. 2010. Improved marketing decision making in a customer churn prediction context using generalized additive models, Expert Systems with Applications. 37, 2132–2143

Feng C., Wen C. 2005. Bi-level and multi-objective model to control traffic flow into the disaster area post earthquake. Journal of the Eastern Asia Society for Transportation Studies. 6, 4253-4268

Gao Y., Zhang G., Lu J., Gao S. 2007. A bilevel model for railway train set organizing optimization. 2007 International Conference on Intelligent Systems and Knowledge Engineering. 777-782

Gao Y., Zhang G., Lu J., Wee H. 2009. Particle swarm optimization for bi-level pricing problems in supply chains. Journal of Global Optimization. 51(2), 245-254

Hobbs B. F., Metzler B., Pang J. S. 2000. Strategic gaming analysis for electric power systems: An MPEC approach. IEEE Transactions on Power Systems. 15, 637-645

Kisioglu P., Topcu Y. I. 2011. Applying Bayesian Belief Network approach to customer churn analysis: A case study on the telecom industry of Turkey. Expert Systems with Applications. 38 7151–7157

Lariviere B., Van den Poel D. 2004. Investigating the role of product features in preventing customer churn by using survival analysis and choice modeling: The case of financial services. Expert Systems with Applications. 27, 277–285

Lu J., Shi C., Zhang G. 2006a. On bilevel multi-follower decision making: General framework and solutions. Information Science. 176, 1607-1627

Lu J., Shi C., Zhang G. 2006b. An extended branch and bound algorithm for bilevel multi-follower decision making in a referential-uncooperative situation. International Journal of Information Technology and Decision Making. 6, 371-388

Luo N., Mu Z. 2004. Bayesian network classifier and its application in CRM. Computer Application. 24(3), 79–81

Murata T., Ishibuchi H. 1995. MOGA: Multi-objective genetic algorithms, Proceedings of 1995 IEEE International Conference on Evolutionary Computation, Perth, WA, Australia. 289–294

Parsopoulos K., Vrahatis M. 2002. Recent approaches to global optimization problems through particle swarm optimization. Natural Computing. 1, 235-306

Reichheld F. F., Sasser W. E. Jr. 1990. Zero defections: Quality comes to service. Harvard Business Review. 68(5), 105-111

Reinartz W. J., Kumar V. 2003. The impact of customer relationship characteristics on profitable lifetime duration. Journal of Marketing. 67(1) , 77-99

Torkzadeh G., Chang J. C. J., Hansen G. W. 2006. Identifying issues in customer relationship management at Merck-Medco. Decision Support Systems. 42(2), 1116 - 1130

Wei C., Chiu I. 2002. Turning telecommunications call details to churn prediction: A data mining approach. Expert Systems with Applications. 23,103–112

Xie Y. Y., Li X., Ngai E. W. T, Ying W. Y., (2009. Customer churn prediction using improved balanced random forests, Expert Systems with Applications. 36, 5445–5449

Yu H., Dang C., Wang S. 2006. Game theoretical analysis of buy-it-now price auctions. International Journal of Information Technology and Decision Making. 5, 557-581

Zhang G., Lu J. 2007. Model and approach of fuzzy bilevel decision making for logistics planning problem. Journal of Enterprise Information Management. 20, 178-197

Zhang G., Lu J., Dillon T. 2006. Kth-best algorithm for fuzzy bilevel programming. Proceedings of International Conference on Intelligent Systems and Knowledge Engineering, Shanghai.

Zhang G., Zhang G, Gao Y., Lu J. 2010. Competitive strategic bidding optimization in electricity markets using bi-level programming and swarm technique. IEEE Transactions on Industrial Electronics., 58(6), 2138-2146