Linking Customer Retention to Intelligent Technology: An Optimization Approach

Li-Shang Yang
International Graduate School of Management
University of South Australia
Adelaide, Australia
lishang.yang@msa.hinet.net

Chaochang Chiu
Department of Information Management
Yuan Ze University
Chungli, Taiwan
imchiu@saturn.yzu.edu.tw

Arthur Preston
International Graduate School of Management
University of South Australia
Adelaide, Australia
apreston@powerup.com.au

Abstract

Marketing managers in the telecommunication sectors are confronted with considerable complexity. They have to make decisions about the optimum combination of products or offerings, customer groups and the means of interacting with potential customers. Further, in saturated markets such as mobile telephony, it is increasingly important to retain customers potentially to churn. On the optimal campaign planning, this research describes how the customer survey was conducted for those potentially churning customers based on which an optimal campaign planning was followed. This research engages with the subjects of customer retention from the perspective of a major mobile operator in Taiwan. Customers’ preferences with C&C (campaign offer and communication channel) were predicted and input for further analysis for target selection optimization. These models was proved novel in an organizational prototype project suggesting that the use of the hybrid of data mining and optimization approaches can be effective for target selection.

1. Introduction

According to the ITU world telecommunication indicator database [1], the mobile telephone penetration rate in Taiwan had reached 97% at the end of 2001. This is the second highest level in the world and has resulted in a highly competitive market place for providers. It is both difficult and costly for the operators to acquire new customers when a market demand is so highly saturated. So, one of the key success factors for the operators is their ability to retain high value customers. To do this they need to understand, in depth, the behavior of customers and to use this information effectively to support successful retention campaign activities.

Artificial Intelligence (AI) technologies offer a way of achieving this knowledge. These have a number of underlying reference disciplines such as biology, neurology, psychology, statistics, and computer science, and have increasingly been applied to various types of problems. Because of advances in IT technologies, the power of AI approaches has increased dramatically.

These techniques offer the business community a broad set of tools capable of addressing problems that are much difficult, or impossible, to solve using the more traditional technologies from statistics and operations research. However, these AI technologies are not commonly employed in business generally, largely because many organizations are not adequately prepared to capitalize on a technology driven environment.

This paper focuses on customer retention from the perspective of the major mobile operator. Its purpose is to explore the applicability of AI techniques in the customer retention campaign design process. A hybrid of data mining and optimization techniques is developed for target selection in order to retain valuable customers. The paper has five following sections. The first provides a brief review of the literature on customer retention, customer equity, customer data analysis approach, target selection and budget allocation, artificial intelligent techniques and genetic algorithms. The seconds describes the background of the decision problems of the mobile operator. In the third, the methodology used to is described and this is followed by the results, a discussion including limitations of the study and suggested areas for future research.

2. Literature Review

2.1 Customer Retention

A traditional marketing approach advocates the quest for market share dominance through mass marketing techniques and focuses on new customer acquisition. This approach has guided managers for decades in planning and implementing their marketing strategy. However, attention is drawn by some researchers to the inadequacies of the traditional marketing approach, and this has led to the birth of relationship marketing (RM). RM advocates supplier-customer interaction and the maintenance of long-term relationships with a focus on customer retention. Based on their consulting experience, Reichheld and Kenny [2] claimed that a 5% increase in retention rate led to profit swings of 25% to 80%. Ahmad and Buttle suggest that practitioners need to move beyond gaining a larger market share, to satisfying market segments and developing brand preference in order to secure lasting customer patronage and profitability [3]. They believe that firms should consider integrating customer retention into...
their strategic market planning processes and set it as one of their primary goals.

2.2 Measurement of Customer Equity in Retention

Even after having accepted the importance of customer retention, managers are left with the issues of which customers to retain and what is the appropriate metric on which to base decisions. Some researchers stress the importance of customer value as a measurement of customer retention[3][4]. Others take a more long term view and have proposed the “Customer Lifetime Value” (CLV), also referred to as Customer Equity [5], as a method of measuring customer value. The basis of a CLV approach is the “excess of a customer’s revenues over time over the company costs of attracting selling, and servicing that customer” [6]. Berger and Nasr [7] present mathematical models for determination customer lifetime value and their models have been often referenced by researchers [3][8]. Although in theory CLV is a useful form of measure, in practice it is difficult to implement. The difficulty lies in the lifetime construct because lifetimes are variable. In this study, the customer value measurement was based on the literature and additionally according to our domain manager’s experience and data availability. For example, an age variable was used instead of estimating a lifetime for each customer.

2.3 Approaches to Analysing Customer Data

Demographic and behavioral data are very useful in describing customers and have an important place in direct marketing decisions but they provide little insight into understanding the needs that motivate and shape the purchase process [9].

As Compared with the more ‘objective nature’ of demographic and behavioral data, needs-based data reflects more psychologically abstract dimensions of consumer segmentation. The use of buying motive and benefit data are appropriate inputs for segment-specific product decisions, promotion decisions, and target marketing decisions [10]. Segmenting markets by consumption patterns can be quite insightful for understanding the customer mix, as by classifying customers into usage categories; management can design appropriate strategies for each market segment [4]. In the current study, the needs-oriented data about consumer preferences on different retention offers and communication channels were obtained by survey. The experimental framework (Figure 2) was designed using two approaches in building C&C models, with and without segmentation. Then the prediction performance of two approaches was compared.

2.4 Target Selection and Budget Allocation

Most of the literature on target selection focuses on the issues of response model of direct mail. This study focuses on target selection of multiple channels. The process of target selection in our experiment is defined as a prior process before building the response model of direct mail. In our experiment, direct mail target customers were determined by the C&C prediction model and target selection model. However, a customer who prefers a specific channel does not necessarily mean he/she will respond to the content delivered by the channel. If a response model is expected, it can be built after finishing the target selection process. Related research on response models can be referenced in [11-13].

In general, target selection in customer retention tends to maximize customer equity under resource restriction. This resource restriction associates with the retention process lies in promotion budgets, call center labor capacity, or other medias channel capacity. Berger and Nasr present a model focusing on promotional budget allocation decisions between acquisition and retention in an organizational level. They assume that a promotional budget has been already set, and they address the optimal allocation of this preset budget [14].

In the current study the decision problems were being made where the retention promotion budget was either undetermined or determined; and the goal was to optimize the budget allocation in an individual level.

2.5 Artificial Intelligence Techniques

Increasingly, work is becoming “knowledge oriented”, with a growing number of “knowledge workers” in organizations required to work with information-gathering, summarizing and interpreting - in order to make decisions. AI can play an important role as a powerful analytical tool to this knowledge oriented working environment.

With the improvements in storage and processing capacity of computers in recent years, artificial intelligence (AI) derives its power to offer the business community a broad set of tools capable of addressing problems that are much harder to solve using traditional techniques from statistics and operations research [15].

Another terminology related with AI technique is data mining. Data mining brings together ideas and techniques from a variety of fields that are similar to AI. Statisticians, artificial intelligence researchers, database administrators, and marketing people use different words to mean the same thing and the same words to mean different things [16]. Data mining frequently uses a variety of AI and statistical techniques for solving problems, but the approach of data mining is more grounded in the methodology than just in the techniques. Data Mining, or AI, is good at solving tasks related with classification, estimation, prediction, affinity grouping, clustering and description. Typical examples of these techniques are Genetic Algorithms, Neural Networks, Decision Trees, Automatic Cluster Detection, Fuzzy Logic, Market Basket Analysis, Memory-Based Reasoning, and Rule-Based
Companies use information in their customer relationship management (CRM) strategies. Data warehouse and data mining techniques allow companies to collect, store, analyze and manipulate enormous volumes of data. This can be important for marketers trying to provide better service to retain more customers than competitors. A survey of 40 companies in the UK shows that few companies are adapted in using such approaches and that there is a need for more applied research using AI techniques in business.

2.6 Genetic Algorithms and Target Selection

Roberts and Berger [18] discuss the most frequently used models to select targets for direct mail. Among these are multiple regression analysis, multiple discriminate analysis, log-linear modeling, and chi-square regression automatic interaction detection (CHAID). More recently target selection has used a neural network model [13]. Hurley et al. [19] suggest some potential marketing research can employ genetic algorithms. Targeting and optimization of distribution channels are among them. Bhattacharyya [20] presents a genetic algorithm-based approach for obtaining models in the area of direct mail. This paper is unusual in its application of genetic algorithms to the process of target selection although doing so is still within the scope of a response model to direct mail.

Traditional use of optimization methods assumes that a 'best solution' can be generated. These methods are not well suited to less structured problems, where the desired objectives are not known in advance [21]. Many real world decisions problems are unstructured. In less structured problems, the conditions indicating the existence of problems are not defined; there is no best methodology to solve these problems, and the criteria for choosing the optimized decisions are not clear [22]. A desired characteristic of a decision support approach is to provide diverse alternatives for consideration. A Genetic Algorithm (GA) is good at generating diverse alternatives in such unstructured problems. GAs can also deal with problems that incorporate nonlinearity, discontinuity, uncertainty, complexity, and other demands which make the application of traditional search and optimization methods inappropriate [21]. In this study, target selection in customer retention is characterized by searching for alternatives of budget allocation in offer-channel-customer combinations to achieve the optimization of objectives under multiple constraints.

In the current study, the target selection is achieved by genetic algorithm that involves decision making with a budget allocation and offer–channel combinations.

3. Background of ABC Company

Six operators provide mobile service in Taiwan. Table 1 shows that the number of Taiwan mobile users grew tremendously between 1998 and 2002 - almost 500%. Over this period, average revenue per user (ARPU) has fallen by 33%, minutes of use per user (MOU) grew for the first two years, and has been steady for the last three years. Obviously, this is a demand-saturated market. As for the past several years, mobile operators have competed for market share, that is, the acquisition of new customers has been the major market strategy.

<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>June, 2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penetration Rate %</td>
<td>21.46</td>
<td>52.24</td>
<td>80.24</td>
<td>96.55</td>
<td>102.69</td>
</tr>
<tr>
<td>MOU</td>
<td>689</td>
<td>828</td>
<td>884</td>
<td>883</td>
<td>415</td>
</tr>
<tr>
<td>APRU</td>
<td>11725</td>
<td>9118</td>
<td>8492</td>
<td>7892</td>
<td>3838</td>
</tr>
</tbody>
</table>
However, the competitive environment has forced the ABC to adjust their marketing strategies and refocus their business strategies on retaining - rather than acquiring - customers.

In order to understand the defection status of their customers, the ABC company has implemented a churn prediction model based on the framework for Customer Retention process shown in Figure 1 [23]. The process shows that data in data warehouse can be used to analyze customer behavior from several perspectives, including segmentation, profitability, satisfaction, and defection (churn). A churn prediction model is a logical output as it is important to understand which customer is likely leaving, and why, so that retention initiatives can become more focused. A churn prediction model aims to aid in the development of more accurate estimate information for use in subsequent retention campaigns. The focus in this paper is on the next step process - retention action plan.

The company intends to retain customers identified as having a high likelihood of churn. To do this the following considerations have to be met:

1. The output name list of churn prediction model, normally around 5000 customers, is too large to be managed and needs to be reduced. The basis for the reduction is unclear.
2. A retention offer which meets customers’ preferences and hence achieves a higher retention needs to be formulated.
3. The desire is to announce the retention offer through the communication channel which mostly meets customers’ preferences.
4. Budgetary allocation for the retention program is complex because of different offer-channel combinations.

The company initiated a prototype project to explore solutions to the above issues.

4. Methodology

4.1 Research Experiment Framework

The research framework adopted the steps in Figure 2:

1. A survey to collect information on customers’ preferences on campaign offers and communication channels.
2. The integration of customer preference data from survey, and customer data from data warehouse, to build a customer C&C predictive model.
3. Two approaches were designed. Approach I involved building a segmentation model first and then building C&C models for each segment. The major premise behind this needs-based segmentation approach was that by segmenting customers into various user groups, management could develop appropriate strategies for each market segment [10]. Approach II involved building the C&C model without considering segmentation. This is a much quicker way to build C&C model.
4. Comparison of the predicted performance of the models, and choosing the model with best performance.
5. Application of the models to a customer population to provide a prediction of the preference C&C for each customer.
6. Using data predictions, to build a target selection model to evaluate and differentiate budget-offer-channel combinations. The model will search for a series of good (optimal) solutions under resources constraints, and optimize some presetting business objectives.
7. To take action on retention campaign programs targeting customers selected from the target selection model.
8. Evaluation of the campaign performance from data in data warehouse.
Steps 7 and 8 were not undertaken in this study.

4.2 Campaign Offer and Communication Channel (C&C) Model

4.2.1 Data Mining Methodology

Fayyad’s process [24] was followed as shown in the Figure 3.

4.2.2 Data Collection and Preprocessing

A population of 4481 (February, 2002) customers with a high churn propensity were provided by ABC company. Six hundred (600) customers were randomly surveyed by mail. Three main questions were asked:

- Which retention offer do you prefer?
- Which communication channel do you prefer to receive our notice of promotion activity?
- What are potential reasons will cause you to decide to leave your current operator?

Five retention offers were designed by the company. These offers are related with handset subsides, free traffic minutes, prepaid discounts, special gifts. Four communication channel options were presented including email, DM mailing with monthly invoice, short message, and outbound call from sales representatives. Two hundred and sixty one (261) questionnaires (43.5%) were returned within fifteen days of mailing.

Customer data were extracted from data warehouse, Details for individuals included demographic information (age, geographic code, etc.), rate plan, transaction apply items, tenure, credit record, churn score, complaint records and usage information.

A preliminary statistical analysis was performed to enhance familiarity with the data and to detect and remedy any missing values or outliers that might distort the subsequent analysis.

4.2.3 Data Mining Approach

With the initial analysis completed, knowledge discovery algorithms were applied to the data. The aim was to statistically identify the relationship between variables and dependent variables so that significant variables could be selected, derived and transformed, and input into next modeling process. The total number of independent variables used for C&C modeling was 61 while one dependent variable (preference on retention offers or communication channels) existed.

The data was then divided into a training set and a testing set. This was done because it was important to ensure that our models do not merely memorize the patterns in the training set. The testing set confirms findings are valid and can be generalized to enable predictions to be made on new data. In the modeling process, decision tree and neural network training was applied to the prediction model.

In the modeling Approach I, the clustering algorithm - k means was used. Before the C&C modeling for each subgroup, the relationship of clusters with retention offers by independence test was examined. The purpose was to test if customer preference of retention offer was significantly different between subgroups. If so, the attributes of customers between subgroups is different and hence their preference of retention offer is different from other subgroups. If this was the case it would be appropriate to build a C&C model for each subgroup. If not, then a segmentation approach was not necessary, and Approach II was more appropriate.

The model performance of the decision tree and neural network methods was assessed and the superior approach retained. Then, the developed model was applied to the population to predict the preferred campaign offer and
4.3 The Optimization Model for Target Selection

With thousands of high churn propensity customers, the selection of retention targets is a difficult problem in practice. However, a GA can be used to optimize objective functions within constraints, and select retention targets under the following situation:

(1) When the budget is undetermined, to provide several budget allocation alternatives with combinations of offer-channel-customer.

(2) When budget is determined, provide several alternatives with combinations of offer-channel-customer.

4.3.1 The Genetic Algorithm

GAs are good for solving optimization problems. The basic goal of optimization tasks involves finding one, or a series, of very good (optimal) solutions from among a very large number of possible solutions. Optimization problems involve three components: a set of problem variables, a set of constraints, and a set of objectives. These components can be transformed and operated by a GA.

GAs solve problems by borrowing a technique from nature evolutions. GAs use Darwin’s basic principles of survival of the fittest, mutation, and cross-over to create solutions for problems. A GA consists of five main components:

The chromosomal representation, initial population, fitness evaluation, selection, cross-over and mutation.

A simple GA operation cycle is:

1. Generate initial random population
2. Evaluate fitness of current population
3. Select chromosomes based on fitness for reproduction
4. Perform cross-over and mutation give new improved population
5. Repeat 2-4 until stop criteria is met, the stop criteria may be set by number of trials, minutes of running time, a stop formula function or else.

A GA searches intelligently through the possible permutations. This is much more practical than searching through all possibilities. A GA allows a good solution to a problem quickly rather than waiting for the absolute best solution. Unlike many mathematical techniques, solution times with GA are usually highly predictable. Solution time is usually not radically affected as the problem gets larger, something which is not always so with the more traditional techniques. However this heuristic technique does not guarantee optimal solutions, users must often settle for “near optimal” solutions. [15]

4.3.2 Problem Formulation

4.3.2.1 Chromosome Representation

Each chromosome represents a possible solution for a selected group of customers. The individual genes with each chromosome are represented by a binary alphabet; 0 indicates that a particular customer is not selected, whereas a 1 would indicate that a particular customer is selected. To illustrate this, assume we have five existing customers, then the chromosome

11010

represents where customer 1, 2, 4 are selected and customer 3, 5 are unselected. In a case of one thousand customers, our chromosome will have a length of 1000 genes.

4.3.2.2 Fitness Function

Fitness evaluation involves defining an objective or fitness constraints against which each chromosome is tested for its fitness, a high fitness value would indicate a better solution than a low fitness value. In this study, the fitness function was formulated as follows:

Maximize \( CVS (customer\ value\ score) = \sum_{i,j} W_i \cdot q \_CLV + W_i \cdot q \_Tenure + W_i \cdot q \_Payment\ _Overdue + W_i \cdot q \_Age \_Now \cdot X_{ij} \cdot \ldots \cdot \beta \) subject to the following constraints:

\[ Tcost \leq MaxBudget \]
\[ \sum_{i,j} PcusNum_{ij} \leq MaxPcusNum \]
\[ 0 \leq X_{ij} \leq 1, X_{ij} \text{ is an integer} \]

Where

- \( q, CLV \) = Customer lifetime quantile value of \( X_{ij} \)
- \( q, Tenure \) = Tenure quantile value of \( X_{ij} \)
- \( q, Payment\ _Overdue \) = The quantile value of the average invoice payment overdue days last six months of \( X_{ij} \)
- \( q, Age \_Now \) = The age quantile value of \( X_{ij} \)
- \( W_1, W_2, W_3 \) = The weight value assign to each variable
- \( X_{ij} \) = The customer \( j \) in the area where \( i \), it is the selection index of \( X_{ij} \) to retention program, if \( X_{ij} \) is selected as target then \( X_{ij} \) is set to 1, Otherwise, \( X_{ij} \) = 0

\[ Tcost = \text{The total cost to execute retention program} \]
\[ f(y) = \text{The cost function of retention offers} \]
\[ f(z) = \text{The cost function of communication channels} \]

MaxBudget = Maximum budget allocated to retention program

\( CasNum_i = \text{Number of customers were selected in the area } i \)

\( MinCasNum_i = \text{Minimum customers need to be selected as targets in area } i \)

\( PcusNum_i = \text{Number of customers targeted through personal contact channel in area } i \)

\( MaxPcusNum_i = \text{Maximum number of customers can be targeted through personal contact channel by call center} \)

In summary, the objective of the target selection was to optimize Equation 1 the customer value score (CVS), subject to budget and resource constraints.

Equation 2 is the customer lifetime value (CLV) based...
on Berger and Nasr’s definition applied in this research:

\[ CLV = C \sum_{r=0}^{\infty} \left( \frac{r^i}{(1+d)^r} \right) - M \sum_{r=1}^{\infty} \left( \frac{r^{i-1}}{(1+d)^r} \right) \]  \hspace{1cm} (2)

where

\[ C = \text{Yearly gross contribut on} \]
\[ M = \text{Promotion cost per customer per year} \]
\[ n = \text{The length in years of period} \]
\[ r = \text{The yearly retention rate} \]
\[ d = \text{The yearly discount rate} \]

Because the difficult ies in defining \( n \) for each customer, for simplifying purpose, it was assumed that \( n = 0 \), thus

\[ CLV = \sum_{j=1}^{n} (C_j - M_j) \]

and if \( M_j \) is equal in each of \( X_j \), therefore,

\[ CLV = \sum_{j=1}^{n} C_j \]

\( C_j = \text{The yearly invoice contribution of } X_j \)

Consequently a customer was judged only on current value with the yearly invoice contribution; however, the \( \text{Age} \_\text{Now} \) variable was set to reflect a customer’s future value. \( W_i = 1 \), \( W_j = 0.4 \), \( W_k = -0.2 \), \( W_l = -0.1 \) were assigned according to the company domain experts’ experience. Negative values were assigning to \( W_i \) and \( W_j \) to represent the negative impact to the customer value score.

4.3.2.3 Simple illustration

To illustrate the application of GAs to a target selection problem with a simple example, a population of five chromosomes is first randomly generated,

\[ C_1 \] 10100
\[ C_2 \] 10001
\[ C_3 \] 10100
\[ C_4 \] 10001
\[ C_5 \] 10100

If the customer value score (CVS) calculated is,

Customer 1 963
Customer 2 953
Customer 3 370

The fitness values (value of objective function) of chromosomes are:

\[ C_1 \] 10000 fitness = 1916
\[ C_2 \] 10001 fitness = 1568
\[ C_3 \] 10100 fitness = 10100
\[ C_4 \] 10001 fitness = 1333
\[ C_5 \] 10101 fitness = 1323

Consequently, the best chromosome is \( C_1 \) and \( C_2 \), are the fittest, so they would be more likely to be selected for cross-over and mutation. If a one-point cross-over was used at a randomly chosen position, say position 3, the following offspring would be produced:

\[ C_1 \] 11000 \( \Rightarrow \) \[ C_6 \] 11100

Their respective fitness values calculated are \( C_6 = 2891 \) and \( C_1 = 963 \). Therefore, a new and better solution represented by chromosome \( C_1 \) has been found.

The advantages in using a GA approach are:

1. An optimal or near-optimal solution is established in a predictable time.
2. The good alternatives can be accessed easily.
3. It is easy to apply, all that is needed is the ability to describe a good solution and provide a fitness function that can rate a given chromosome. This means that you can use a GA to solve problems that you don’t even know how to solve!

5. Experiment and Results

5.1 The Survey Result

5.1.1 Preference of retention offer

Table 2 shows the preference distribution of the retention offer and Table 3 the \( x^2 \) values for independence tests for age, education and revenue factors. In Table 3, all the \( x^2 \) values are smaller than \( x^2_{0.05} \) and \( x^2_{0.01} \). This means that there is insufficient evidence to say that the preference retention offers differ on the basis of education, age, or revenue. However, from Table 2, it is obvious that the ‘free minutes offer without prepaid’ is the most favored retention offer overall.

<table>
<thead>
<tr>
<th>Retention offer type</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handset subsidies</td>
<td>20</td>
</tr>
<tr>
<td>Free minutes offer with prepaid</td>
<td>17</td>
</tr>
<tr>
<td>Free minutes offer without prepaid</td>
<td>56</td>
</tr>
<tr>
<td>Free mobile internet minutes offer</td>
<td>2</td>
</tr>
<tr>
<td>Special designed gift</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 3 The independence test**

<table>
<thead>
<tr>
<th>Factors</th>
<th>( x^2 ) value</th>
<th>( x^2_{0.05} )</th>
<th>( x^2_{0.01} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>15.7</td>
<td>31.4</td>
<td>37.6</td>
</tr>
<tr>
<td>Education</td>
<td>25.0</td>
<td>26.3</td>
<td>31.9</td>
</tr>
<tr>
<td>Revenue</td>
<td>4.9</td>
<td>26.3</td>
<td>31.9</td>
</tr>
</tbody>
</table>

5.1.2 Preference of Communication channel

Table 4 shows the preference distribution of communication channels and Table 5 the \( x^2 \) values for the independence test with age, education and revenue. The preference of communication channel significantly depends on age at both \( x^2_{0.05} \) and \( x^2_{0.01} \), and on education at \( x^2_{0.05} \). Younger people (under 20) preferred only ‘short message’ and ‘email’, while for older people (above 50), these two forms were least preferred. The preference for personal phone contact increased with age. Also, people with higher education showed a greater preference for
‘short message’ and ‘email’ than people with lower education levels. From Table 4, ‘DM with invoice’ is the channel most often preferred, followed by ‘short message’.

Table 4 The distribution of communication channels preference

<table>
<thead>
<tr>
<th>Type of contact channels</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email</td>
<td>8</td>
</tr>
<tr>
<td>DM with monthly invoice</td>
<td>45</td>
</tr>
<tr>
<td>Short message</td>
<td>33</td>
</tr>
<tr>
<td>Personal telephone contact</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 5 The independence test

<table>
<thead>
<tr>
<th>Factors</th>
<th>$x^2$ value</th>
<th>$x^2_{0.05}$</th>
<th>$x^2_{0.01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>38.2</td>
<td>24.9</td>
<td>30.6</td>
</tr>
<tr>
<td>Education</td>
<td>24.7</td>
<td>21.0</td>
<td>26.2</td>
</tr>
<tr>
<td>Revenue</td>
<td>10.5</td>
<td>21.0</td>
<td>26.2</td>
</tr>
</tbody>
</table>

5.1.3 Potential Reasons to Churn

Communication quality and price were the top two issues that impact customer's potential churn behavior (Table 6).

Table 6 Distribution of potential reasons to churn

<table>
<thead>
<tr>
<th>Type of reasons</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfied mobile communication quality</td>
<td>32</td>
</tr>
<tr>
<td>Cheaper tariffs with friends and relatives within same operator’s network</td>
<td>15</td>
</tr>
<tr>
<td>Unsatisfied service quality</td>
<td>12</td>
</tr>
<tr>
<td>Better handset discounts from other operators</td>
<td>10</td>
</tr>
<tr>
<td>Price is higher than other operators</td>
<td>31</td>
</tr>
</tbody>
</table>

5.2 C&C Preference Model

In Approach I, before building the C&C modeling for each segment, the relationship of clusters with retention offers by independence test was examined, that is:

$H_0$: Cluster is independent from retention offer preference

$H_1$: Cluster is dependent with retention offer preference

However, the $x^2$ value of 10.61504 was smaller than the critical value of $x^2_{0.05} = 21.06$, and so the $H_0$ was accepted. Hence, it was decided not to build C&C model for each subgroup. This means that Approach II is applied to building the C&C model in this study.

By integrating the preference data from the survey with customer data in the data warehouse, and discarding some records with important missing data, the C&C model was built. Lift [25] was used to measure the prediction performance of the model. Lift values in Table 7 show that the prediction rate was improved using the models.

The next step was the application of these models to the remainder of the population, and the prediction of preference retention offers and communication channels to each customer.

After discarding records with missing data, the number of records available were 2247. This ‘predict list’ was used for target selection.

Table 7 Lift of C&C model

<table>
<thead>
<tr>
<th></th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
<th>(D) = (A)/ (C)</th>
<th>Lift (E) = (D)/(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention Offer: Handset Subsidies</td>
<td>24</td>
<td>0.18</td>
<td>11</td>
<td>0.46</td>
<td>2.56</td>
</tr>
<tr>
<td>free minutes offer with prepaid</td>
<td>24</td>
<td>0.18</td>
<td>15</td>
<td>0.63</td>
<td>3.49</td>
</tr>
<tr>
<td>free minutes offer without prepaid</td>
<td>78</td>
<td>0.58</td>
<td>73</td>
<td>0.94</td>
<td>1.61</td>
</tr>
<tr>
<td>Free mobile internet minutes offer</td>
<td>1</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Special designed gift</td>
<td>7</td>
<td>0.05</td>
<td>2</td>
<td>0.29</td>
<td>5.47</td>
</tr>
<tr>
<td>Total</td>
<td>134</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Communication Channel:

| Email                        | 11  | 0.09| 6   | 0.55 | 6.39 |
| DM with monthly invoice      | 53  | 0.41| 38  | 0.72 | 1.75 |
| Short Message                | 41  | 0.32| 38  | 0.93 | 2.92 |
| Personal telephone contact   | 24  | 0.18| 12  | 0.5  | 2.69 |
| Total                       | 129 | 1.00|     |      |      |

**1.(A) : Survey Result
2.(B) : Occurrence Probability
3.(C) : Model Predict Result
4.(D) : Model Hit Rate
5.(E) : Lift

5.3 Target Selection Optimization Model

The GA approach was applied to the optimization problems when:

- When the amount of budget is undetermined
- When the amount of budget is determined

Table 8 presents the solutions obtained for an undetermined budget situation. Four alternatives were provided with budgets from 200000 to 500000. Mutation rate was set to 0.15 and the cross over rate to 0.5. The objective function aimed to maximize the customer value score (CVS). The stop criterion was set at the objective value of not improving 0.1% in the last 3000 trials.

The ability of GA to find good solutions within a limited time is shown in Table 8. Resource constraint criteria were set for budget limit, personal phone contact capacity, DM with invoice capacity and minimum geographic customers needed to select. All the optimal solutions according to different budgets are shown in the Table 8. Alternative III generates the highest value per dollar. Managers can modify constraints to meet their specific business needs, and make their optimal decision based on the information suggested.
This is likely to only be a ‘good’ solution and not necessarily ‘optimal’. The mutation rate and cross over rate could have been adjusted to speed up the searching for optimal solution.

### Table 8 Alternatives when budget amount is undetermined**

<table>
<thead>
<tr>
<th>Budget alternatives</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVS</td>
<td>137249</td>
<td>222259</td>
<td>390056</td>
<td>453986</td>
</tr>
<tr>
<td>Retention Cost</td>
<td>199421</td>
<td>298674</td>
<td>396426</td>
<td>497524</td>
</tr>
<tr>
<td>Value per cost dollar</td>
<td>0.68824</td>
<td>0.74145</td>
<td>0.98393</td>
<td>0.91249</td>
</tr>
<tr>
<td>Time to generate this CVS</td>
<td>461</td>
<td>239</td>
<td>435</td>
<td>918</td>
</tr>
<tr>
<td>Numbers of customer selected</td>
<td>402</td>
<td>634</td>
<td>1041</td>
<td>1169</td>
</tr>
<tr>
<td>A=32</td>
<td>A=44</td>
<td>A=39</td>
<td>A=64</td>
<td></td>
</tr>
<tr>
<td>B=118</td>
<td>B=165</td>
<td>B=315</td>
<td>B=351</td>
<td></td>
</tr>
<tr>
<td>C=350</td>
<td>D=0</td>
<td>D=0</td>
<td>D=0</td>
<td></td>
</tr>
<tr>
<td>E=2</td>
<td>E=4</td>
<td>E=6</td>
<td>E=15</td>
<td></td>
</tr>
<tr>
<td>Communication mix</td>
<td>A=19</td>
<td>A=24</td>
<td>A=35</td>
<td></td>
</tr>
<tr>
<td>Channel mix</td>
<td>B=161</td>
<td>B=248</td>
<td>B=426</td>
<td></td>
</tr>
<tr>
<td>Geographical mix</td>
<td>C=43</td>
<td>C=74</td>
<td>C=100</td>
<td></td>
</tr>
<tr>
<td>T=106</td>
<td>D=179</td>
<td>D=288</td>
<td>D=480</td>
<td></td>
</tr>
<tr>
<td>N=41</td>
<td>E=54</td>
<td>E=98</td>
<td>E=114</td>
<td></td>
</tr>
</tbody>
</table>

**1. We set constraints: maximum personal phone contact capacity ≤ 35, DM with invoice capacity ≤ 500, minimum targets per area ≥ 20.
2. Initial population is generated by random; population size is set to 200.
3. The objective function is to maximize CVS; an initial random CVS is 100633.
4. We set mutation rate = 0.15, and cross over rate = 0.5.
5. The stop criterion is the objective value without improving 0.1% for the last 3000 trials then stop.
6. The measure unit of time to generate this CVS is second.
7. The measure unit of budget and cost is Taiwan dollar (NT).
8. Retention offer mix: A is handset subsidize, B is Free minutes offer with prepaid, C is Free minutes offer without prepaid, D is Free mobile internet minutes offer, E is special designed gift.
9. Communication Channel mix: A is personal telephone contact, B is DM with invoice, C is email, D is short message.
10. Geographic mix: T is central area, S is northern area, C is central area, N is mountain area, A is easten area.

Table 9 presents solutions when fixed budget of $400,000 Taiwan dollars. A better value per dollar (1.004485) was generated in alternative III when the call center capacity constraint was released to 50. The stop criterion for alternatives II and III was not at least a 0.1% improvement over the last 3000 trials. However, a stop criterion was not set for alternative I, and it was stopped manually after a relatively good (optimal) value was generated. A comparison of time taken to reach the optimal CVS values between I and II, a better objective value was achieved as more time were taken, 702 seconds to 1037 seconds. However, more CPU resources are needed to generate the better solution. All the alternatives solutions presented in this table all started from an initial value of 100633 and achieve to the optimal solutions within reasonable time.

### Table 9 The alternatives with the determined budget**

<table>
<thead>
<tr>
<th>Budget</th>
<th>CVS</th>
<th>Retention cost</th>
<th>Value per cost dollar</th>
<th>Generations to get this value</th>
<th>Time to get this CVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>392542</td>
<td>390056</td>
<td>399109</td>
<td>85.16</td>
<td>1309</td>
</tr>
<tr>
<td>II</td>
<td>399922</td>
<td>396426</td>
<td>397327</td>
<td>46.575</td>
<td>1073</td>
</tr>
<tr>
<td>III</td>
<td>0.981546</td>
<td>0.983931</td>
<td>1.004485</td>
<td>634</td>
<td>1089</td>
</tr>
</tbody>
</table>

**1. Communication Channel mix: A is personal telephone contact, B is DM with invoice, C is email, D is short message.
2. Capacity constraint of personal telephone contact for alternative I and II is ≤ 35, for alternative III is ≤ 50.
3. We set mutation rate = 0.15, and cross over rate = 0.5.
4. An initial random CVS is 100633.

### 6. Conclusion and Suggestions for Future Research

One of the major implications of target selection in customer retention is its ability to provide marketing decisions information to evaluate different alternatives' impacts on customer equity. In this paper, a methodology for the optimization of target selection based on customer preferences, using a genetic algorithm and where the objective was to maximize customer equity is described. The study demonstrates the potential of AI technologies for the customer retention process in the mobile industry. This study used a customer needs-based approach to predict the C&C preference of customers, and a different way of applying genetic algorithm to target selection and budget allocation.

While most traditional techniques provide a single best solution, multiple solutions with high performance levels can be advantageous, especially if coupled with a faster solution search time. A genetic algorithm only use an encoding of the solution and its associated fitness value and do not utilize auxiliary information, such as derivatives or assumptions about continuity in the solution process, they can be applied efficiently to a wide range of marketing optimization problems with only minor changes to parts of the algorithm.

One limitation of the current work is the insufficient...
data to segment customers due to small scale of sampling in survey and missing data problems in the data warehouse. These problems should be avoided in a formal project implementation. Another important limitation of the work is that it did not empirically investigate the validity of models in real world, taking action on the results of models by carrying out an experiment of treatment group and control group is planned in the future work.

Further empirical investigation is needed to evaluate the effectiveness of the proposed GA method as compared with the traditional methods. Application to retention programs in other industry is needed to demonstrate the broader applicability of this methodology.

Acknowledgement

The authors would like to thank the Managing Director of the Call Centre of the ABC Company for access to data used in the study.

References