Scientia Iranica E (2012) 19 (3), 849–855

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
www.sciencedirect.com

Customer lifetime value prediction by a Markov chain based data mining model: Application to an auto repair and maintenance company in Taiwan

C.-J. Cheng [a,1], S.W. Chiu [b,2], C.-B. Cheng [c,*], J.-Y. Wu [c,3]

[a] Department of Information Engineering, Kun Shan University, 949 Dawan Rd., Yongkang Dist., Tainan City, Taiwan
[b] Department of Business Administration, Chaoyang University of Technology, 168, Jifeng E. Rd., Wufeng District, Taichung, 41349, Taiwan
[c] Department of Information Management, Tamkang University, 151 Yıngzhuan Rd., Tamsui Dist., New Taipei City, Taiwan

Received 17 March 2011; revised 5 October 2011; accepted 21 November 2011

KEYWORDS
Customer lifetime value;
Auto repair and maintenance industry;
Data mining;
Markov chain;
Decision tree;
Neural networks.

Abstract The present study attempts to establish a framework for computing customer lifetime values for a company in the auto repair and maintenance industry. The customer lifetime value defined in this study consists of the current and future values of a customer, which involve an estimation of lifetime length, future purchasing behavior and the profit associated with each behavior of the customer. The proposed framework contains three groups of techniques to obtain these estimates from historical customer transactions. The first group includes a logistic regression model and a decision tree model to estimate the churn probability of a customer and to, further, predict the lifetime length of the customer. The second group comprises a regression analysis to identify the critical variables that affect a customer's purchasing behavior, and a Markov chain to model the transition probabilities of behavior change. Finally, the third group contains two neural networks to predict the profits contributed by a customer under various purchasing behaviors. The proposed framework is demonstrated with the historical customer transactions of an auto repair and maintenance company in Taiwan.
© 2012 Sharif University of Technology. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

The automobile industry in Taiwan used to be protected under a high tariff imposed on imported cars by the government, before 2002. After joining the World Trade Organization (WTO) in year 2002, and opening the market to foreign automobiles, Taiwan's auto industry has encountered the impact of declining sales. In 2005, 445,000 vehicles produced by local makers were sold, but the number dropped to 234,000 in 2009 [1]. Together with economic decline and high fuel prices during past years, the entire auto industry of Taiwan is pessimistic about the future sales of new cars.

In Taiwan, car dealers also operate auto repair and maintenance factories (referred to as the original brand maintenance factory), providing major services to cars of their own brand. As the demand of new cars is weak, repair and maintenance services have become the major source of profit for car dealers. However, these factories still need to face competition from more than 8600 garages that provide the same services with lower prices. In general, when a car is within its warranty period, the owner will return to the dealer's factory for maintenance, but when the warranty is expired, 42% of the owners would choose different garages [2]. Such high customer defection has a great impact on the profitability of companies in this industry. On the other hand, a study by Reichheld and Sasser [3] pointed out that, in the auto service industry, reducing customer defection by 5% would boost profits by 30%. Thus, it is important for an auto repair and maintenance company to take action to retain customers.

* Corresponding author. Tel.: +886 2 26215656x3332; fax: +886 2 26209737.
E-mail addresses: cjcheng@mail.ksu.edu.tw (C.-J. Cheng), swang@cyut.edu.tw (S.W. Chiu), cbcheng@mail.tku.edu.tw (C.-B. Cheng), Jun8954145@gmail.com (J.-Y. Wu).
[1] Tel.: +886 6 2727175x65101; fax: +886 6 2050207.
[2] Tel.: +886 4 23323000x4379; fax: +886 4 23742331.
[3] Tel.: +886 2 26215656x2645; fax: +886 2 26209737.
Peer review under responsibility of Sharif University of Technology.

Production and hosting by Elsevier

1026-3098 © 2012 Sharif University of Technology. Production and hosting by Elsevier B.V. All rights reserved.
doi:10.1016/j.scient.2011.11.045

The ability to identify profitable customers and build long-term loyalty with them is a key factor in today's highly competitive business environment. To achieve this goal, companies have adopted the concept of Customer Relationship Management (CRM) as a business strategy to integrate their sales, marketing and services across multiple business units and customer contact points. Under the concept of CRM, customers are not equal and, thus, it is unreasonable for the company to provide the same incentive offers to all customers. Instead, companies can select only those customers who meet certain profitability criteria based on their individual needs or purchasing behaviors [4]. Precise evaluation of customer profitability is a crucial element for the success of CRM [5]. However, evaluation of customer profitability is not an easy task, because it involves prediction of future contributions by the customer. The major factors in consideration of this prediction include: for how many years in the future will the customer stay with the company, and how much will the customer contribute to the company each year? The time length that a customer will stay with a company is referred to as the lifetime of this customer, and the profitability attributed to a customer is called the lifetime value (LTV) of the customer.

The use of customer lifetime values in marketing, for segmenting customers or formulating strategies, has been found in literature. Hwang et al. [6] and Kim et al. [7] proposed a customer lifetime value computation model considering the past profit contribution, potential benefit, and defection probability of a customer. They also covered a framework for analyzing customer value, and segmenting customers based on their values, and applied their approach to formulating marketing strategies for a wireless communication service company. Shih and Sohn [8] used three clustering methods (k-means, self-organizing map, and fuzzy k-means) to segment stock trading customers to determine differentiated commission rates based on customer potential values. Shih and Liu [9] proposed using a collaborative filtering technique via customer lifetime value and customer demand to develop a product recommendation system. Chan et al. [10] also used a customer lifetime value to predict the company long-term return on investment.

In particular, Chan [11] also adopted the concept of the customer lifetime value in segmenting customers of an automobile retailer. He formulated a genetic algorithm that combined customer segmentation with campaign decisions where the fitness function of the algorithm was defined as a customer's lifetime value. Computation of the lifetime value was based on the potential profit that is obtained from the customer, by adopting a certain campaign strategy in which the adoption of a strategy was considered as a probability distribution. However, how such a probability distribution could be obtained was not discussed, neither did the estimation of profit generated from a campaign strategy. The missing link between the customer lifetime value computation and the historical customer transaction data in Chan's approach reduces the practical usefulness of the approach. Thus, the present study attempts to develop a clear framework of customer lifetime value computation for the auto repair and maintenance industry, which can be used as a road map for the industry to compute customer lifetime values.

The length of the lifetime of a customer depends on his loyalty to the company. This study predicts the lifetime of a customer by a loyalty measure, which is estimated based on customer demographic data and historical transaction records. Data mining techniques are employed to obtain the loyalty measure of a customer. To estimate the contribution of a customer during his lifetime with the company, the purchase frequency and profit generated from each visit are further predicted. A Markov chain is used to model the possible purchasing frequency of a customer in this study, while neural network approaches are used to estimate the profit generated from customer purchases. Our approach is formulated as a step-by-step framework that synthesizes various data mining techniques, including regression analysis, decision tree and neural networks, to provide a roadmap for a company to compute the lifetime values of its customers. The proposed customer lifetime value computation framework is applied to an auto repair and maintenance company in Taiwan.

2. LTV prediction model

The basic concept of customer lifetime value computation is based on the Net Present Value (NPV) received from a customer over his lifetime of transactions with the company (e.g., [12–14]). The common formulation of the NPV-based model is as follows:

$$LTV = \sum_{t=0}^{n} \pi(t) \cdot \frac{1}{(1+r)^t}, \tag{1}$$

where $\pi(t)$ is the profit contributed by a customer at time t, r is the interest rate, and n is the total period of projected life of a customer staying with the company.

Hwang et al. [6] expanded the NPV-based model to comprise, not only the projected value of a customer, but also his past profit contribution (called current value). The present study adopts the same concept, and computes the lifetime value of the kth customer by the following equation:

$$LTV_k = CV_k + FV_k, \tag{2}$$

where CV_k and FV_k denote the current and future values of the kth customer, respectively. Computation of the current value is simple. It is the present value of past profits contributed by a customer. Let the number 0 denote the current time point, and −1, −2, etc. denote one period, two periods, etc. before the current time. Considering that the bank interest rate is usually changed every year, the present study suggests computing the current value of the kth customer as:

$$CV_k = \sum_{t=-m}^{-1} \pi_k(t) \prod_{d=t}^{-1} (1 + r(d)), \tag{3}$$

where $\pi_k(t)$ is the profit contributed by customer k at period t, $r(d)$ is the bank interest rate at period d, and m is the number of periods before the current time point with which we compute a customer's current value.

The prediction of future value is more difficult and complicated. The computation of the future value of a customer involves two dimensions, namely:

1. The profit that the customer will contribute in a certain period in the future.
2. The length of time that the customer will stay with the company.

Traditionally, the future profit contributed by a customer is a projection of the customer's past contribution. However, the purchasing behavior of a customer may change in the future and, thus, make such a projection inaccurate, as it is purely based on the customer's historical behavior. Considering the uncertainty of customers' future behavior, Colombo and Jiang [15] developed a stochastic RFM model to rank customers in terms of their expected contribution. Hwang et al. [6] predicted a customer's potential value by estimating the probabilities of cross selling and the profit generated from it. The cross selling probability was computed by decision tree models or neural networks, while the generated profit was a deterministic value. However, the assumption of potential cross selling in the future, by Hwang et al. [6], is not suited to the auto repair and maintenance industry, as the services provided by this industry are less diversified. Thus, this study seeks other customer behavior from the transaction database to find the variables that are more suitable for describing the customer's purchasing behavior.

In this study, we consider the probability of customer behavior change at a certain period being affected by the customer's status in the preceding period. The Markov chain is a suitable tool for modeling such conditional status change with its state transition probability model. A similar concept has been presented by Chan et al. [10] where they modeled the purchasing behavior switching probabilities of a customer by the Markov chain model. Three possible states of a customer were considered in their model, namely, potential customer, first-time customer and active customer. However, the assumption of these three states somewhat simplifies customer purchasing behavior. For example, for two active customers, there may still be different purchasing frequencies that yield different profits. The present study also uses a Markov chain to model the possible purchasing behavior changes of a customer. However, in our approach, important attributes in representing customer states are extracted from historical transaction records, rather than assumed a priori. Furthermore we estimate the profit contributed by a customer under all possible behaviors by neural network approaches to compute the expected future value of a customer.

Customer churn is a critical factor in LTV computation, because it affects the length of the service period and, hence, future profit generation. The churn rate measures the number of customers who stop using or purchasing products/services from a company. Hwang et al. [6] used the churn rate to estimate the length of periods that a customer stayed with a company, and employed data mining techniques to compute the churn probability of a customer. A similar technique is used in this study to estimate the service periods of a customer.

Consequently, the computation of the future value of the kth customer is defined as follows:

$$FV_k = \sum_{t=1}^{n_k} I_k P^t \boldsymbol{\pi}_k(t), \tag{4}$$

where:

n_k: the expected length of service periods of the kth customer, i.e. lifetime length;

I_k: the initial state vector of the kth customer, and the elements in the vector represent whether a state occurs or not:

$$I_k = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_L \end{bmatrix}^T,$$

in which $s_j, j = 1, \ldots, L$, and $s_j \in \{0, 1\}$ and $\sum_{j=1}^{L} s_j = 1$, meaning that only one state can occur at a time; $s_j = 1$ means that state j occurs, and 0 otherwise.

P: the state transition probability matrix, and:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1L} \\ p_{21} & p_{22} & \cdots & p_{2L} \\ \vdots & & \ddots & \vdots \\ p_{L1} & p_{L2} & \cdots & p_{LL} \end{bmatrix},$$

in which p_{ij} denotes the probability of switching from state i to state j. P^t is the resulting transition matrix after t periods.

$\boldsymbol{\pi}_k(t)$: a vector of the profits generated by the kth customer under all possible states at time t, and:

$$\boldsymbol{\pi}_k(t) = \begin{bmatrix} \pi_{k,1}(t) \\ \pi_{k,2}(t) \\ \vdots \\ \pi_{k,L}(t) \end{bmatrix},$$

in which $\pi_{k,j}(t)$ denotes the profit contributed by the kth customer under state j (i.e. the jth purchasing behavior) at time t. $\boldsymbol{\pi}_k(t)$ is already converted in its present value in Eq. (4).

The estimates of the expected length of service periods, n_k, the profits generated by the customer under various purchasing behaviors, $\boldsymbol{\pi}_k(t)$, and the state transition probability matrix, P, are obtained through data mining techniques and Markov chain analysis based on the customer database as depicted in Figure 1. The procedures are grouped in three functional blocks, namely, lifetime prediction, profit prediction and behavior prediction. All three functional blocks use customer demographic data and historical transaction data to construct their models. The purpose of the lifetime prediction block is to estimate the expected length of service periods. The procedure includes, first, using classification tools, logistic regression or a decision tree to model the probability that a customer will leave, and then use this probability as a loyalty measure to compute the expected length of service periods, n_k. In the behavior prediction block, a regression analysis is first conducted to find the variables that are significant in estimating a customer's contribution to the company. These variables are used to represent the state of a customer as well in the Markov chain model, and the transition probabilities of customer states are then computed from the model. Finally, in the profit prediction block, the Backpropagation Neural Network (BPN) and the Radial Basis Function Network (RBFN) are used to model the profits generated by a customer under various states (i.e. purchasing behaviors). The detailed procedures of obtaining these parameters are discussed in the next section.

3. Model parameters estimation

The approaches to determine the length of service periods, n_k, the state transition probability matrix, P, and the profit vector, $\boldsymbol{\pi}_k(t)$, are discussed in the following three subsections, respectively.

3.1. Estimation of service length

The process to determine the length of service periods, or equivalently, the expected lifetime of customer k, n_k, is obtained by computing the expected length of service periods from a churn model suggested by Hwang et al. [6]. The churn model assumes customer churn is a geometric distribution, where the expected length of time before the customer leaves can be computed with a given churn probability.

To estimate the churn probability, two different approaches are used, namely, logistic regression, and a decision tree. The approach that has the better prediction accuracy is selected to estimate the churn probability. The decision tree model for our churn prediction problem is built using the C5.0 decision tree software of Quinlan [16]. The input variables to the logistic regression model or the decision tree are determined, based on the attributes of the data tables in the customer demographic and customer transaction databases. The attributes of the data tables in these two databases are presented in Table 1.

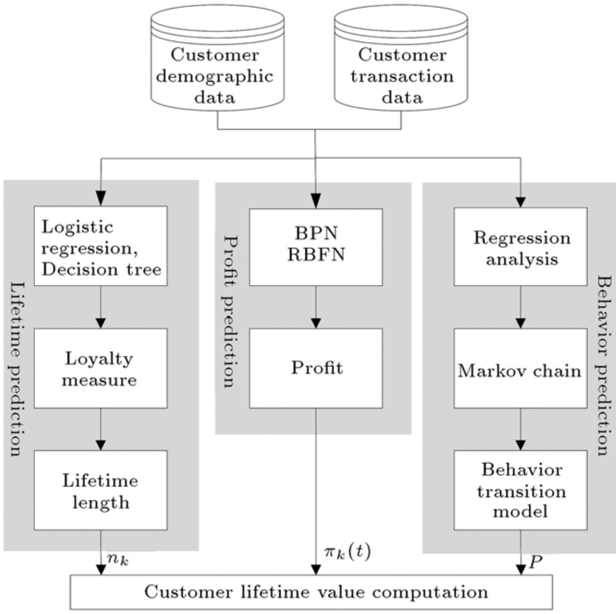

Figure 1: Customer lifetime value prediction framework.

Table 1: Original data tables in databases.

Panel A: (Customer database) Customer data table	Panel B: (Transaction database) Purchase record table	Panel C: (Transaction database) Repair and maintenance record table
Customer code	Customer code	Customer code
Customer name	Car model	Customer name
Gender	Price	License tag number
Contact address	Year of production	Car model
Permanent address		Year of production
Zip code		Engine displacement
Home phone		Date of visit
Company phone		Parts
VAT number		Labor hours
		Amount

Most auto repair and maintenance shops in Taiwan do not demand customers to provide detailed demographic data, and the data maintained in the transaction database is generally incomplete. For example, in the customer data table, excepting name, gender and contact phone of the customer, the rest of the data is generally missing; moreover, the age, occupation and income of the customer are absent. In the repair and maintenance record table, parts used for repairing and labor hours consumed are not input into the table though there are part of the attributes; in face, only the total monetary amount of the repair is recorded. Such incomplete data constrains the performance of data mining tools.

Due to the limited information provided by these data tables, the variables we can extract from the databases are gender of customer, engine displacement of the car (which is the volume swept by an internal combustion engine and which is used to indicate the size of the car), average expense per visit, visiting frequency (average number of visit per year), number of cars bought at the company, and age of car. By a pretest, we found that the gender of a customer and engine displacement are not significant in determining customer loyalty. Thus, the input variables of our logistic/decision tree model are average expense per visit, visiting frequency, number of cars bought at the company, and age of car. These variables, together with the status (churn or not-churn) of the customer, form the attributes of an example in the training data set for the logistic regression or decision tree models.

The criteria to distinguish churn/not-churn customers usually depend on the characteristics of the industry under study. For example, Xia and Jin [17] defined a churn customer in the cellular phone service industry as a customer who no longer uses any service provided by the carrier; and Tsai and Chen [18] considered that churn customers in a multimedia-on-demand industry are those who terminate the service or are withdrawn by the company due to non-payment of service.

For the case of the auto repair and maintenance industry investigated by this study, churn customers are defined as customers who have not used any service of the company in the past three years, as suggested by experts in this industry. On the other hand, a not-churn customer is one who has used the company's service every year over the past three years. The training examples compiled from the database are used to construct the churn/not-churn customer classification model by logistic regression or decision tree, and the churn probability of a customer is estimated from the output of the model by feeding this customer's attribute values into the model.

The expected lifetime of a customer is then computed, based on the estimated churn probability. Let p_k^c denote the probability that the kth customer will leave the company, i.e. a churn probability or churn rate. Furthermore, let y be the length of time before a customer churn, then, y follows a geometric distribution [6]. That is:

$$\text{Prob}\{y = n\} = p_k^c(1 - p_k^c)^{n-1}. \tag{5}$$

Thus, the expected service time length is computed by the following equation:

$$n_i = E[y] = 1/p_k^c. \tag{6}$$

3.2. Estimation of state transition probability

The state of a customer refers to the behavior of a customer at a certain time, which determines the customer's contribution to the company, and it is represented by a set of variables related to the customer. To identify the variables that are critical in representing the status of a customer in calculating his future contribution, a regression analysis based on customer transaction data is conducted. Those variables that are significant in determining the contributions by customers are used to describe a state. Various states are then defined in terms of different levels of these variables.

For our case study, the regression model is constructed as follows:

$$E = a + b \cdot g + c \cdot f + d \cdot \delta + e \cdot y, \tag{7}$$

where the dependent variable is the average expense by the customer, a is the constant, g is the gender of the customer, f is the customer's visiting frequency, δ is the engine displacement of the customer's car, y is the age of the car, and b, c, d, and e are coefficients. The regression analysis results (Table 2) indicate that significant variables include visiting frequency, engine displacement, and age of car in which engine displacement is a fixed value and the age of the car will regularly increase. Thus, visiting frequency is the ideal and only variable to define the state of a customer. We also found that among all customers using the company's service, it is rare that a customer visits the company more than four times a year as indicated by statistics in Figure 2. Therefore, the states of a customer are defined as when visiting the company; zero, one, two, three, four and/or above, times in a year, and are denoted as states 0, 1, 2, 3, and 4, respectively.

The state transition probabilities used in the Markov chain are conditional probabilities, given the previous state. The transition probability matrix is obtained by estimating the probability of a state under the condition of its previous state. The transition probability from state i to state j is obtained by computing the probability of moving to state j by a customer, given he was in state i in the preceding year. Assume that the number of visits by a customer in a year is a Poisson distribution. For a customer who was in state i in the preceding year, the probability that he will visit the company x times in the coming year is:

$$f(x) = \frac{e^{-\lambda_i}\lambda_i^x}{x!}, \quad i = 0, \ldots, 4, \tag{8}$$

where λ_i is the mean of the distribution, given that the number of visits in the preceding year is i.

The value of λ_i is estimated with data from the customer transaction database. By finding the average number of visits per year by customers, subject to their preceding state, we are able to estimate the value of λ_i. The method is to pick any two consecutive year records of a customer from the transaction database, and to count the numbers of visits in those two years. All customers that visited the company i times in the first year are grouped together and serve as a sample to estimate λ_i. Let $q_{v,i,t+1}$ denote the number of customers who visited the company i times in year t, and v times in year $t+1$. The estimates of λ_i can be obtained by the following equation:

$$\hat{\lambda}_i = \frac{\sum_t \sum_{v=0}^{V} q_{v,i,t+1} \cdot v}{\sum_t \sum_{v=0}^{V} q_{v,i,t+1}}, \tag{9}$$

where V is the maximum number of visits in a year among all customers, and the range of t depends on how many years of records are kept in the transaction database. The estimation result is shown in Table 3 where state 1 is defined as zero visits (i.e. $i = 0$), and state 2 is referred to as one visit (i.e. $i = 1$), etc. The result in Table 3 also supports the argument that a customer who purchased frequently from a company in the past will come back more frequently in the future too.

By employing Eqs. (8) and (9), the transition probability matrix for our case study is obtained as follows:

$$P = \begin{bmatrix} 0.144 & 0.279 & 0.270 & 0.175 & 0.132 \\ 0.220 & 0.333 & 0.252 & 0.127 & 0.068 \\ 0.148 & 0.283 & 0.270 & 0.172 & 0.127 \\ 0.087 & 0.213 & 0.260 & 0.211 & 0.229 \\ 0.032 & 0.110 & 0.189 & 0.217 & 0.452 \end{bmatrix}.$$

The element (0.144) in the first row and column in the above P matrix is the probability of the transition from state 0 to state 0, which is obtained by calculating $f(0)$ in Eq. (8) with $\lambda_0 = 1.938696$ from Table 3, while the element (0.279) in the first row and second column, denoting the probability of moving from state 0 to state 1, is obtained by calculating $f(1)$ with λ_0, and is the probability that the customer will visit the company one time in the coming year, given that he did not visit the company in the previous year. The remaining elements are obtained in the same manner.

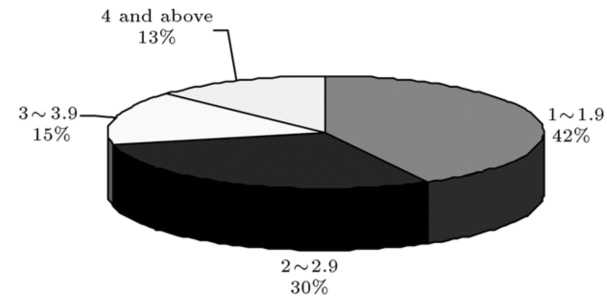

Figure 2: Distribution of the average number of visits per year by customer.

Table 2: Regression analysis of customer contribution.

Variable	Estimate	Standard error	t-value	p-value
Constant (a)	2117.935	650.118	3.258	0.001
Gender (g)	167.467	397.538	0.421	0.674
Frequency (f)	252.191	51.201	4.925	0.000
Engine displacement (δ)	0.775	0.197	3.933	0.000
Age of car (y)	384.502	37.069	10.373	0.000

Table 3: Estimates of λ_i.

i	0	1	2	3	4
λ_i	1.938696	1.513721	1.909567	2.437809	3.445738

3.3. Prediction of customer contribution

To calculate the contribution of a customer at a certain time in the future, we need to model the relation between the customer contribution and its influential factors. The regression analysis result in Table 2 has confirmed that the visiting frequency, engine displacement and age of car are significant in estimating a customer's contribution. However, the R^2 of the resulting regression model is only 0.052, which implies that a linear regression model is inadequate to represent the true relation between influential factors and customer contribution. Nevertheless, the information regarding significant variables provided by the regression analysis is useful in estimating customer contribution. Such significant variables are ideal candidates of inputs to nonparametric regression tools, such as neural networks. As an alternative, neural networks are more suitable for modeling customer contribution in this study, owing to their nonlinear modeling capability. Two neural networks (a backpropagation neural network and a radial basis function network) are used to model customer contribution with visiting frequency, engine displacement and age of car as input variables. In the application, the network that provides better prediction accuracy is selected to predict a customer's future contribution.

At the stage of predicting the contribution by a customer, the customer's status, i.e. the values of the three input variables, are fed into the trained neural network to obtain an estimate of his contribution from the network output. For example, to compute the possible contribution of customer k at time t, we need to input the engine displacement of his car, the age of the car will be at time t, and by assuming the frequency of visits as 1, 2, 3, and 4, one at a time, we can obtain vector $\boldsymbol{\pi}_k(t)$, with its four elements representing customer contributions under different visiting frequencies.

4. Empirical study

The case study company was founded in 1990 in Taiwan, currently with 450 employees. The company is an agency of Nissan Motor, Taiwan, and its business includes the sales and marketing of Nissan vehicles, auto part sales, car repairs and maintenance. This study attempts to assist the company in developing its marketing strategies by, first, identifying the lifetime values of individual customers. The approaches discussed in the previous sections are used to achieve this goal. The data used for analysis contain 3239 customers and 20,643 transaction records.

Following the procedures depicted in Figure 1, we start with estimating a customer's churn probability by both logic regression and the decision tree model in which the C5.0 decision tree tool [16] was used. The variables used include the following attributes of a customer: average expense per visit, average number of visits per year, number of cars bought at the dealer, and age of car. The distribution of the age of the car in the sample is shown in Table 4, and the average expense per visit by a customer is shown in Table 5.

The sample data are divided into a training data set and a test data set with a ratio of 80:20. The performance of the decision tree is slightly better than the logistic regression model, with a prediction accuracy of 78.66% for the training data set, and 77.06% for the test data set, while logistic regression yields 76.35% accuracy on the test data set. After the classification model is established, given the conditions of a customer, the model will output the churn probability of this customer.

Both the BPN and RBFN are used to predict the contribution of a customer. The sample data used to train the neural networks are compiled in a tuple; frequency of visits, engine displacement, age of car, and average expense, where the first three are input variables to the networks and the last is the output. The sample data consist of 2109 records from the transaction database in which 80% of the records are used as training samples and the remaining 20% are used as testing samples. The resulting prediction accuracy of the two neural networks is about the same; 94% in the training sample and 92% in the testing sample. BPN is arbitrarily chosen for this customer contribution prediction.

The following example illustrates the customer lifetime value prediction procedures proposed in this study. By using the trained decision tree model, the churn probability of a certain customer is computed to be 0.245758832 and, thus, the expected service length (i.e. lifetime) of this customer is estimated to be $1/0.245758832 \approx 4$ years according to Eq. (6). As shown in Eq. (2), the customer's lifetime value comprises two parts, his current and future values. To compute the current value, the contributions of this customer over past years are retrieved and the bank interest rates of the past years are also found as presented in Table 6. By Eq. (3) we can discover that the current value of this customer equals 47956.08.

By using the trained backpropagation neural network, we can predict future contributions by the customer under the five possible states in the next four years. The vector of the customer's contribution in the next t year, $t = 1, \ldots, 4$, is denoted by $\pi(t)$, and is obtained as:

$$\pi(1) = \begin{bmatrix} 0 \\ 3613 \\ 7226 \\ 10840 \\ 14453 \end{bmatrix} \qquad \pi(2) = \begin{bmatrix} 0 \\ 3547 \\ 7093 \\ 10640 \\ 14186 \end{bmatrix}$$

Table 4: Distribution of the age of car.

Age of car	Number of customers	Percentage (%)	Cumulative percentage (%)
≤ 1	441	13.62	13.62
(1, 2]	623	19.23	32.85
(2, 3]	555	17.13	49.98
(3, 4]	475	14.67	64.65
(4, 5]	359	11.08	75.73
(5, 6]	301	9.29	85.02
(6, 7]	215	6.64	91.66
(7, 8]	138	4.26	95.92
(8, 9]	86	2.66	98.58
≥ 10	46	1.42	100.00
Total	3239	100.00	

Table 5: Distribution of average expense per visit.

Average expense/visit (NT$)	Number of customers	Percentage (%)	Cumulative percentage (%)
0–2000	224	6.92	6.92
2001–4000	721	22.26	29.18
4001–6000	1008	31.12	60.30
6001–8000	665	20.53	80.83
8001–10000	282	8.71	89.54
10001–12000	151	4.66	94.20
12001–14000	75	2.32	96.52
14001–16000	36	1.11	97.63
16001–18000	26	0.80	98.43
18001–20000	13	0.40	98.83
20001 and above	38	1.17	100.00
Total	3239	100.00	

$$\pi(3) = \begin{bmatrix} 0 \\ 3682 \\ 7364 \\ 11047 \\ 14729 \end{bmatrix} \qquad \pi(4) = \begin{bmatrix} 0 \\ 3605 \\ 7210 \\ 10814 \\ 14419 \end{bmatrix}.$$

This customer visited the company one time in the preceding year, thus, his initial state vector is $I = [0\ 1\ 0\ 0\ 0]$ by definition. Let the expected future value of the customer in the next t year be denoted by $FV(t)$, $t = 1, \ldots, 4$. By employing the transition probability matrix, P, obtained in Section 3, and the customer's contributions in the next four years obtained above, we can compute the customer's future value in each year based on Eq. (4):

$$\begin{aligned} &FV(1) = I \cdot P \cdot \pi(1) = 5378.783, \\ &FV(2) = I \cdot P^2 \cdot \pi(2) = 6605.063, \\ &FV(3) = I \cdot P^3 \cdot \pi(3) = 7292.254, \quad \text{and} \\ &FV(4) = I \cdot P^4 \cdot \pi(4) = 7298.482, \end{aligned}$$

in which $FV(1)$ is the estimated future value that the customer will contribute in the first year of the future and $FV(2)$, $FV(3)$ and $FV(4)$ are the estimated future values of the second, third and fourth years of the future, respectively. As a result, the total future value of the customer is $FV = FV(1) + FV(2) + FV(3) + FV(4) = 26574.582$. Finally, the lifetime value of the customer is $LTV = CV + FV = 74530.662$.

5. Concluding remarks

This study has presented a framework for an auto repair and maintenance company to compute the lifetime values of

Table 6: Profits contributed by the customer and bank interest rate of previous years.

Year	−5	−4	−3	−2	−1
Profit	7 929	24 552	1 662	0	10 707
Interest rate (%)	5	2.5	1.875	1.4	1.525

its customers. The customer lifetime value defined in this study consists of the current value and the future value of a customer, which involves estimates of lifetime length, future purchasing behavior, and the profit associating with each behavior, of a customer. The proposed framework contains three groups of techniques to obtain these estimates from historical customer transactions. The proposed approach was demonstrated using the customer data of an auto repair and maintenance company in Taiwan.

In our approach, customer behavior is modeled as a Markov chain, which estimates the probability that the customer transits from one state (i.e. a certain purchasing behavior) to another. By analysis of the data of the customer transaction database, the state considered in the Markov chain model in this study involves only one variable: the number of visits by the customer in a year. The resulting single variable state in our Markov chain model is likely due to the incompleteness of the transaction database of the case study company. In general, it often requires multiple variables to well represent a customer's purchasing behavior (state). When the proposed approach is applied to other companies with more complete customer and transaction databases, multiple variables are likely to involve in representing a customer state. Multiple variables would result in the difficulty of enumerating a great deal of possible states, and make the Markov chain model unmanageable. To resolve this problem, in future study, we will be investigating simulation techniques, such as a system dynamic or Petri net, to model a customer's future purchasing behavior.

Acknowledgments

The authors are grateful to the four anonymous reviewers for their suggestions in revising this paper.

References

[1] TTVMA (Taiwan Transportation Vehicle Manufacturers Association) Industry status report, http://www.ttvma.org.tw/cht/industrial-survey.php#3 (accessed 2010).
[2] Lee, W.-J. "The industry study and business strategy analysis of automobile service in Taiwan", Master Thesis, School of Management, National Central University, Taiwan 2007 (in Chinese).
[3] Reichheld, F.F. and Sasser Jr., W.E. "Zero defections: quality comes to services", *Harvard Business Review*, pp. 105–111 (Sept.–Oct.1990).
[4] Dyche, J. and Tech, J., *The CRM Handbook: A Business Guide to Customer Relationship Management*, Addison-Wesley, Reading, MA (2001).
[5] Lee, J.H. and Park, S.C. "Intelligent profitable customers segmentation system based on business intelligence tools", *Expert Systems with Applications*, 29, pp. 145–152 (2005).
[6] Hwang, H., Jung, T. and Suh, E. "An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry", *Expert Systems with Applications*, 26, pp. 181–188 (2004).
[7] Kim, S.-Y., Jung, T.-S., Suh, E.-H. and Hwang, H.-S. "Customer segmentation and strategy development based on customer life time value: a case study", *Expert Systems with Applications*, 31, pp. 101–107 (2006).
[8] Shin, H.W. and Sohn, S.Y. "Segmentation of stock trading customers according to potential value", *Expert Systems with Applications*, 27, pp. 27–33 (2004).
[9] Shih, Y.-Y. and Liu, D.-R. "Product recommendation approaches: collaborative filtering via customer lifetime value and customer demands", *Expert Systems with Applications*, 35, pp. 350–360 (2008).
[10] Chan, S.L., Ip, W.H. and Cho, V. "A model for predicting customer value from perspectives of product attractiveness and marketing strategy", *Expert Systems with Applications*, 37, pp. 1207–1215 (2010).
[11] Chan, C.C.H. "Intelligent value-based customer segmentation method for campaign management: a case study of automobile retailer", *Expert Systems with Applications*, 34, pp. 2754–2762 (2008).
[12] Berger, P.D. and Nasr, N.I. "Customer lifetime value: marketing models and applications", *Journal of Interactive Marketing*, 12, pp. 17–30 (1998).
[13] Bayon, T., Gutsche, J. and Bauer, H. "Customer equity marketing: touching the intangible", *European Management Journal*, 20, pp. 213–222 (2002).
[14] Gupta, S. and Lehmann, D.R. "Customers as assets", *Journal of Interactive Marketing*, 17, pp. 9–24 (2003).
[15] Colombo, R. and Jiang, W. "A stochastic RFM model", *Journal of Interactive Marketing*, 13, pp. 2–12 (1999).
[16] Quinlan, J.R. C5.0 online tutorial, http://www.rulequest.com (accessed 2010).
[17] Xia, G.-E. and Jin, W.-D. "Model of customer churn prediction on support vector machine", *Systems Engineering – Theory & Practice*, 28, pp. 71–77 (2008).
[18] Tsai, C.-F. and Chen, M.-Y. "Variable selection by association rules for customer churn prediction of multimedia on demand", *Expert Systems with Applications*, 37, pp. 2006–2015 (2010).

Chao-Jung Cheng received his B.A. and M.S. Degrees in Mechanical Engineering from Sun Yat-Sen University, Taiwan, in 1991 and 1993, respectively, and a Ph.D. Degree in from Department of Engineering Science at Cheng-Kung University, Taiwan, in 2005. He has been Assistant Professor in the Department of Information Engineering of Kun-Shan University, Taiwan, since 2005. His research interests include: neural networks, data mining, nonlinear dynamics, automatic control, chaos synchronization and image processing. He has (co)authored numerous papers in these areas in international journals.

Singa Wang Chiu received her M.S. Degree in Computer & Information Science from New Jersey Institute of Technology, USA, in 1988, and a Ph.D. Degree in Industrial Engineering and Management from the Chung-Yuan Christian University, Taiwan, in 2004. Professor Chiu's research interests include: operations research, production and quality management, and manufacturing systems. Her papers have been published in numerous international journals including European Journal of Operational Research, Japan Journal of Industrial and Applied Mathematics, Computers & Operations Research, Int'l Journal of Machine Tools & Manufacture, Mathematical and Computer Modeling, Applied Stochastic Models in Business and Industry, Mathematical & Computer Modeling-Dynamical Systems, Communications in Numerical Methods in Engineering, International J. of Advanced Manufacturing Technology, Journal of Manufacturing Science and Engineering, Journal of Scientific & Industrial Research, etc.

Chi-Bin Cheng is Professor in the Department of Information Management at Tamkang University, Taiwan. He received his B.S. Degree from Chung Yuan Christian University, Taiwan, his M.S. Degree in engineering from Lamar University, Texas, and a Ph.D. Degree from Kansas State University, USA. His main research interests include: soft computing, supply chain management, auctions, and negotiation agents. Professor Cheng's work has appeared in numerous international journals including: Decision Support Systems, European Journal of Operational Research, Fuzzy Sets and Systems, Computers & Industrial Engineering, Journal of The OR Society, Computers and Mathematics with Applications, International Journal of Revenue Management, and Electronic Commerce Research and Applications, among others.

Jiun-Yi Wu earned his B.S. Degree in Information Management from Shih Hsin University, Taiwan, in 2008, and his M.S. Degree from Tamkang University, Taiwan, in 2010. His research interests include: data mining, soft computing, and customer relationship management.