Applying Fuzzy Data Mining to Telecom Churn Management

Kuo-Hsiung Liao and Hao-En Chueh*

Department of Information Management, Yuanpei University,
No. 306, Yuanpei Street, Hsinchu 30015, Taiwan
liao@mail.ypu.edu.tw, hechueh@mail.ypu.edu.tw

Abstract. Customers tend to change telecommunications service providers in pursuit of more favorable telecommunication rates. Therefore, how to avoid customer churn is an extremely critical topic for the intensely competitive telecommunications industry. To assist telecommunications service providers in effectively reducing the rate of customer churn, this study used fuzzy data mining to determine effective marketing strategies by analyzing the responses of customers to various marketing activities. These techniques can help telecommunications service providers determine the most appropriate marketing opportunities and methods for different customer groups, to reduce effectively the rate of customer turnover.

Keywords: Telecommunications industry, Churn management, Fuzzy theory, Data mining, Marketing activity.

1 Introduction

According to statistics produced by the Institute for Information Industry of Taiwan, mobile phone users in Taiwan numbered 26 million as of the second quarter of 2009. This statistic indicates that the popularization rate of mobile phones in Taiwan has already reached 100 % and that mobile communications service is one of the most indispensable information and communication services in contemporary life.

Additionally, the portable line number service recently provided by the government permits users to change their telecommunications provider without changing their original mobile line number. This has made it easier for users to change their telecommunications service providers in consideration of their own best interests. To maintain market share and profitability, telecommunications service providers implement various policies and management mechanisms in an attempt to retain customers and avoid serious customer churn problems [1, 15, 16, 17, 20, 21].

A common churn management procedure is to analyze information about past churners, build a model for prediction of customer churn, use this prediction model to determine from current customer information those customers that are likely to churn, and then commence various marketing activities or events. However, whether

* Corresponding author.

R. Chen (Ed.): ICICIS 2011, Part I, CCIS 134, pp. 259–264, 2011.

churn management strategies have been successful must be examined according to whether or not customer churn rate has truly been reduced. Thus, to assist telecommunications service providers in successfully retaining potential churners, this study used fuzzy data mining techniques [3, 6, 10, 11, 12, 14, 24, 29] to analyze the responses of customers to various marketing activities and thus determine effective marketing strategies.

2 Literature Review

Numerous previous studies have proposed methods for assisting telecommunications service providers in solving problems of customer churn management. [2, 5, 6, 7, 8, 9, 13, 18, 19, 22, 25, 26, 27, 28]. Xia et al.[28] proposed the study of architecture using support vector machines to churn prediction model, and with a variety of data mining technology framework to compare the customer churn predictive model, including neural networks, decision-making trees, logistic regression, and Bayesian classifier, the experiments confirmed the structure using support vector machines churn out of the prediction model, the prediction accuracy than other data mining techniques to the prediction accuracy of better.

Tsai et al. [25] presented the research in the use of data mining techniques to find association rules that may affect the customers of the loss of an important factor, then these factors and the use of decision tree to construct the neural network technology a customer churn prediction model, and applied to telecommunications value-added services MOD (multimedia on demand) customer churn prediction. The study also experiments confirmed that pre-use association rules to carry out factor analysis of customer churn predictive model selection accuracy of the analysis of association rules is better than no prediction model. In addition, the study also confirmed that the experimental use of decision tree-based prediction model than the structure of neural network prediction model has better prediction results.

The main purpose of the past studies described above was to build an effective customer churn prediction model to forecast which customers are likely to churn. The provision of this information could then assist telecommunications service providers in organizing various marketing activities or events. However, from the perspective of telecommunications service providers, determining which customers are likely to churn does not guarantee that providers can successfully retain these potential churners and thus reduce their customer churn rate. Rather, effective marketing activities are essential for customer churn management.

3 Fuzzy Data Mining

Fuzzy set theory is proposed by L. A. Zadeh, professor of the University of California at Berkeley in 1965 [29]. Suppose we have a universe of discourse, X , and the elements of X is x_i , i.e. $X = \{x_i\}$. Then, some of unclearly attributes can be represented as fuzzy sets. Assume that A is a fuzzy set defined on X , then the degree of membership of an element belonging to the fuzzy set A can be expressed as a membership function μ_A, and its value is normalized to between 0 and 1. The membership function is shown as the follows [29]:

$$\mu_A : X \to [0,1],\ 0 \le \mu_A(x_i) \le 1,\ x_i \in X\ . \tag{1}$$

There are three basic operators of fuzzy sets: union, Intersection, and complement.

Two fuzzy sets A and B defined on X, then the symbol of the union operator is $\mu_{A \cup B}(x)$, and the definition of the union operator is as follows:

$$\mu_{A \cup B}(x) = max(\mu_A(x), \mu_B(x)),\ x \in X\ ; \tag{2}$$

Two fuzzy sets A and B defined on X, then the symbol of the intersection operator is $\mu_{A \cap B}(x)$, and the definition of the intersection operator is as follows:

$$\mu_{A \cap B}(x) = min(\mu_A(x), \mu_B(x)),\ x \in X\ ; \tag{3}$$

A fuzzy set A defined on X, then the symbol of the complement operator is $\mu_{\bar{A}}(x)$, and the definition of the complement operator is as follows:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x),\ x \in X\ . \tag{4}$$

In this study, the fuzzy sets can be used to assist in dealing linguistic means and avoiding the boundary shape problem.

Data mining is defined as use of automated or semi-automated method from a large number of data collections to extract the potential, unknown, meaningful and useful information or patterns [3, 6, 10, 11, 12, 14, 23, 24]. The main technologies of data mining include classification, prediction (trend) analysis, cluster analysis, association rules analysis, sequential patterns analysis and so on.

This study uses ID3 decision tree algorithm [23] for the last customer for the call center do in order to retain customers reflect the results of various marketing activities to identify successful and effective customer retention strategy to reduce customer churn. ID3 is a widely used algorithm of classification task. Classification is the process of mining a classifier from a set of pre-defined training data that can describe and distinguish data classes or concepts, such that the found classifier can assign a class or concept to a new un-defined data. In general, classification (mining a classifier) involves three major tasks: data representation, which represents data in machine-readable structures, classifier construction, which constructs a classifier from a set of training data, and classifier evaluation, which evaluates classifier accuracy with a set of testing data and in terms of various evaluation functions. Classification has been popularly applied on insurance risk analysis, credit approval, medical diagnosis, etc. Under the previous literature can also be found using the decision tree algorithm to solve the problem of churn management has a good effect [25].

4 Experiment and Results

The experimental dataset used in this study came from the randomly sampled customer retention activities and the responses of customers of a telecommunications company in Taiwan whose contracts were due to expire between June and July 2008.

From the customers whose contracts were due to expire in June and in July 2008 respectively, 400 customers were randomly selected from each of the following groups: customers with monthly bills of NT$ 0 ~ NT$300, customers with monthly bills of NT$301 ~ NT$800, and customers with monthly bills of NT$801 ~ NT$1000. Each group of 400 customers in the different bill amount ranges were then divided further into two subgroups of 200 customers each. Customer retention marketing programs were implemented by sending direct mail (DM) and through telemarketing. During this retention marketing process, customers could choose the marketing programs that they wanted. The finally marketing results of the activities recorded in Table 1. To category various customer groups to effective marketing, in this study, churn rate will be converted into a fuzzy set [29] called effective marketing, $\mu_{EM}(x)$,

$$\mu_{EM}(x) = \begin{cases} 1.0 & x < 10 \\ 0.8 & 10 \le x < 30 \\ 0.6 & 30 \le x < 50 \\ 0.4 & 50 \le x < 70 \\ 0.2 & 70 \le x < 90 \\ 0 & 90 \le x \end{cases}, \tag{5}$$

where x is the churn rate.

The established marketing model [23] is shown as in Figure 1. Through the marketing model that, for the customers whose monthly bills from NT$801 to NT$1000, regardless of their contractual maturity date in June or July, when the telephone marketing is used, effective marketing is the extent of up to 0.8; but If mailing DM is used, then the degree of effective marketing is 0.2 only. For the customers whose

Table 1. The result of customer marketing activities

Customer group	Contract expires	Bill payment	Marketing method	Churn rate	Effective marketing
1	June	NT$0 ~ NT$300	Telecom marketing	83%	0.2
2	June	NT$0 ~ NT$300	Sending DM	75%	0.2
3	June	NT$301 ~ NT$800	Telecom marketing	52%	0.4
4	June	NT$301 ~ NT$800	Sending DM	96%	0
5	June	NT$801 ~ NT$1000	Telecom marketing	13%	0.8
6	June	NT$801 ~ NT$1000	Sending DM	87%	0.2
7	July	NT$0 ~ NT$300	Telecom marketing	82%	0.2
8	July	NT$0 ~ NT$300	Sending DM	78%	0.2
9	July	NT$301 ~ NT$800	Telecom marketing	52%	0.4
10	July	NT$301 ~ NT$800	Sending DM	96%	0
11	July	NT$801 ~ NT$1000	Telecom marketing	26%	0.8
12	July	NT$801 ~ NT$1000	Sending DM	89%	0.2

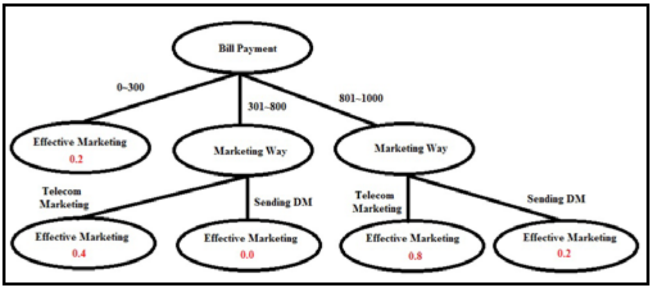

Fig. 1. Telecom marketing model

monthly billing amount from NT$ 0 to NT$300, you must use the same telephone marketing, the degree of effective marketing is 0.4; if mail DM is used, then it is completely not effective marketing. In addition, for the customers whose monthly bills from NT$ 0 to NT$301, no matter what kind of marketing, the degree of effective marketing is very low, only 0.2.

5 Conclusions

Most of previous researches, the emphasis being given to construct a customer churn predictive model to identify in advance the list of possible loss of customers. However, the ability to identify the possible loss of potential customers does not mean that you can retain the possible loss of those customers live, in order to reduce customer churn rate, must present an effective marketing strategy. To this end, this study uses fuzzy data mining techniques to analyze the past records of results of various marketing activities to establish a marketing model. In this study, the proposed marketing model can provide companies on determining the best marketing strategies for different customer groups.

References

1. Ahn, J.H., Han, S.P., Lee, Y.S.: Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. Telecommunications Policy 30(10-11), 552–568 (2006)
2. Au, W., Chen, K.C.C., Yao, X.: A novel evolutionary data mining algorithm with applications to churn prediction. IEEE Transactions on Evolutionary Computation 7(6), 532–545 (2003)
3. Berry, M.J.A., Linoff, G.: Data mining techniques: For marketing, sales, and customer support. John Wiley & Sons, Chichester (2004)
4. Berson, A., Simith, S., Thearling, K.: Building data mining applications for CRM. McGraw-Hill, New York (2000)
5. Chu, B.H., Tsai, M.S., Ho, C.S.: Toward a hybrid data mining model for customer retention. Knowledge-Based Systems 20(8), 703–718 (2007)
6. Chueh, H.-E., Lin, N.P., Jan, N.-Y.: Mining Target-oriented Fuzzy Correlation Rules in Telecom database. In: The 2009 International Conference on Advances in Social Networks Analysis and Mining, pp. 165–199 (2009)

7. Chueh, H.-E., Lin, S.-C., Jan, N.-Y.: Mining the Telecom Marketing Information to Optimizing the Customer Retention Strategies. Accepted: ICIC Express Letters, Part B: Applications
8. Coussement, K., Van den Poel, D.: Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert Systems with Applications 34(1), 313–327 (2008)
9. Coussement, K., Van den Poel, D.: Integrating the voice of customers through call center emails into a decision support system for churn prediction. Information & Management 45(3), 164–174 (2008)
10. Dunham, M.H.: Data mining. Introductory and Advanced Topics. Pearson Education, London (2003)
11. Fayyad, U., Uthurusamy, R.: Data mining and knowledge discovery in databases. Communications of the ACM 39, 24–27 (1996)
12. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco (2001)
13. Hung, S.Y., Yen, D.C., Wang, H.Y.: Applying data mining to telecom churn management. Expert Systems with Applications 31, 515–524 (2006)
14. Kantardzic, M.: Data mining-Concepts, models, methods, and algorithms. John Wiley & Sons, Chichester (2003)
15. Keaveney, S.M.: Customer switching behavior in service industries: An exploratory study. Journal of Marketing 59, 71–82 (1995)
16. Kim, H.S., Yoon, C.H.: Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. Telecommunications Policy 28(9), 751–765 (2004)
17. Kim, M.K., Park, M.C., Jeong, D.H.: The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services. Telecommunications Policy 28, 145–159 (2004)
18. Lejeune, M.: Measuring the impact of data mining on churn management. Internet Research: Electronic Network Applications and Policy 11(5), 375–387 (2001)
19. Luo, B., Shao, P., Liu, D.: Evaluation of three discrete methods on customer churn model based on neural network and decision tree in PHSS. In: The First International Symposium on Data, Privacy, and E-Commerce, pp. 95–97 (2007)
20. Mattersion, R.: Telecom churn management. APDG Publishing, Fuquay-Varina (2001)
21. Mozer, M.C., Wolniewicz, R., Grimes, D.B., Johnson, E., Kaushanky, H.: Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. IEEE Transactions on Neural Networks 11(3), 690–696 (2000)
22. Ngai, E.W.T., Xiu, L., Chau, D.C.K.: Application of data mining techniques in customer relationship management: A literature review and classification. Expert Systems with Applications 36(2), 2592–2602 (2009)
23. Quinlan, J.R.: Induction of Decision Trees. Machine Learning 1(1), 81–106 (1986)
24. Tan, P.N., Steinbach, M., Kumar, V.: Introduction to data mining. Pearson Addison Wesley, London (2006)
25. Tsai, C.F., Chen, M.Y.: Variable selection by association rules for customer churn prediction of multimedia on demand. Expert Systems with Applications 37, 2006–2015 (2010)
26. Tsai, C.F., Lu, Y.H.: Customer churn prediction by hybrid neural networks. Expert Systems with Applications 36(10), 12547–12553 (2009)
27. Wei, C.P., Chiu, I.T.: Turning telecommunications call details to churn prediction: a data mining approach. Expert Systems with Applications 23(2), 103–112 (2002)
28. Xia, G.E., Jin, W.D.: Model of Customer Churn Prediction on Support Vector Machine. Systems Engineering-Theory & Practice 28(1), 71–77 (2008)
29. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)