THE CHURN MANAGEMENT FOR TELECOM MARKET USING THE KNOWLEDGE DISCOVERY IN DATABASE

Tae Hyup Roh, [1] InGoo Han[1] Won Ho Jang[2]

Department of Management Engineering, Korea Advanced Institute of Science and Technology[1]
130-012, Seoul, South Korea
Department of Information System, Korea Telecom, 158-050, Seoul, South Korea[2]

ABSTRACT

As the market competition becomes keen, and value of each customer is various, constructing a churn management is coming to the front for winning over the customer, developing service and products for customer satisfaction and segmenting customer for one-to-one target marketing. We apply KDD methodologies to the telecom market, one of the fiercest competition markets, and suggest prototype and logic for churn management model. This churn management model includes churn probability using NN, decision tree, logistic regression tools, and segment customer according to customer segmentation matrix. In addition to this, marketing strategy is also presented with reference to real telecom market data.

KEYWORDS

Churn Management; KDD; CRM; Customer Segmentation

1. Introduction

The strategic importance of managing customer relationships both drives and is driven by technology. (Puckey, 1999) In particular, this applies to data and the increasingly sophisticated and useful ways in which data is used to model relationships and to drive contact strategies. The rapid emerging field of knowledge discovery in database (KDD) known as data mining is driven by a mix of daunting practical needs and strong research interest. Current computing and strong technology is rapidly outstripping society's ability to make meaningful use of the torrent of available data in marketing field. (Fayyad, 1996)

At first, fundamental objective of churn management as a part of customer relationship management is, to extract the features of churn. According to the features, the next step is to construct a churn management model and to identify most profitable customers and assign value to the target customers. The other is to provide insight into marketing efforts aimed at promoting the retention or acquisition of a customer. According to Anderson consulting, present customer churn rate reaches an average 25 % of wireless operator's customer base in Europe. The cost to bring in each new customer is approximately US $480, which means providers must retain their customers longer in order to recoup the initial acquisition costs.

There are three types of churn (when customer switch vendors or cancel service altogether): unavoidable churn, involuntary churn and voluntary churn (Modisette, 1999). Unavoidable churn occurs when a customer dies or moves out the provider's operating area. Involuntary churn occurs when a subscriber fails to pay for service and as a result the provider terminates service. Termination of service due to theft, fraudulent service acquisition or fraudulent usage is also classified as involuntary churn. The primary focus is on voluntary churn – service termination on the part of the customer when leaving one operator for another. This three classification also can be divided into four categories defined by two variables according to Schmitt (1999). : Financial/non-financial churn: Financial churn is defined as bad-debt subscribers who leave, while non-financial churn refers to paying customers who leave. Voluntary/involuntary churn: This variable refers to whether customers have left by their own choice.

Involuntary, Financial churn	Voluntary, Financial churn
Bad-debt customers who were most likely remove from the subscriber list for lack of payment.	Bad-debt subscriber who decided to defect, but who might rehabilitates through carefully planned retention strategy.
Involuntary, Non-financial churn	Voluntary, Non-financial churn
Paying customers, who died, moved or were otherwise permanently removed from the subscriber list.	Paying customer who defected because they were unhappy with the service, dissatisfied with the price, got a better offer, or left for any number of other reasons.

Table 1 Four Basic Types of Churn

Poor performance and limited coverage, price of service, competitor's promotion of enhanced service and satisfaction with customer service, these kinds of reasons should be taken into deep consideration, because these reasons attribute to company's negligence and could be avoided. Churn management's fundamental objectives are first to extract the features of churn. According to the features, the next step is to construct a churn management model and to identify most profitable customers and assign value to the target customers. The other is to provide insight into marketing efforts aimed at promoting the retention or acquisition of a customer.

2. Real Application

2.1 Needs for Churn Management

In the Korean long distance call market, the third operator-started service at December 1999, which means long distance call market encounters complete competition period. Although this market is decreasing due to wireless communication market's growth, three operators' competition will be accelerated. In spite of second operator's change, Korean long distance market has ma intained a monopoly system, whereas second operator's market share is under about 10 %. However, new participant's cost reduction promotion strategy will make this market fiercer and as time goes on, the number of churners is expected to become higher. Not only the purpose of customer retention, but also customer segmentation for target marketing is essential, which can be used for new product development and promotion. Through the churn management, company achieves to prevent customer churn, to assure stable income and to reduce bad credit. Above all, scientific and analytical customer-base marketing is possible.

2.2 Research Method and Model Construction

2.2.1 Data Collection

The research data for nationwide churn modeling is collected from one telecom company's data warehouse. Targeting (maintaining subscriber: 0, churner: 1) is to forecast the churn probability. Churn means that one exiting long distance call subscriber moves out to different long distance competitor. Data's raw variables and derivative variables consist of 8 demographic variables, 11 additional service variables, and 14 bill-related variables of 6 months. To extract raw data set, random sampling method is used, and all churners' time base is grounded on the request of bill on July 1999, and key is based on line. Raw data set is composed of subscribers' table (7,001 lines) and churners' table (50,534 lines),

2.2.2 Model Construction

Raw data set is composed of 6 groups, and about 90% of subscribers consist of two groups, general subscriber and corporate one. As those two group's derivative variables and usage patterns are different, churn model is built as two model, churn-person model, and churn-corporate model. Table 2 shows composition of raw data set's subscriber group. After pair matching processing, experiment data set is composed of 2 subsets: training set and validation set.

	Frequency	Percent
General	30,878	61.1
Gov.Office	230	.5
Foreigner	15	.0
Small company	5,020	9.9
Corporation	13,797	27.3
Group	149	.3
Others	434	.9
Total	50,523	100.0%

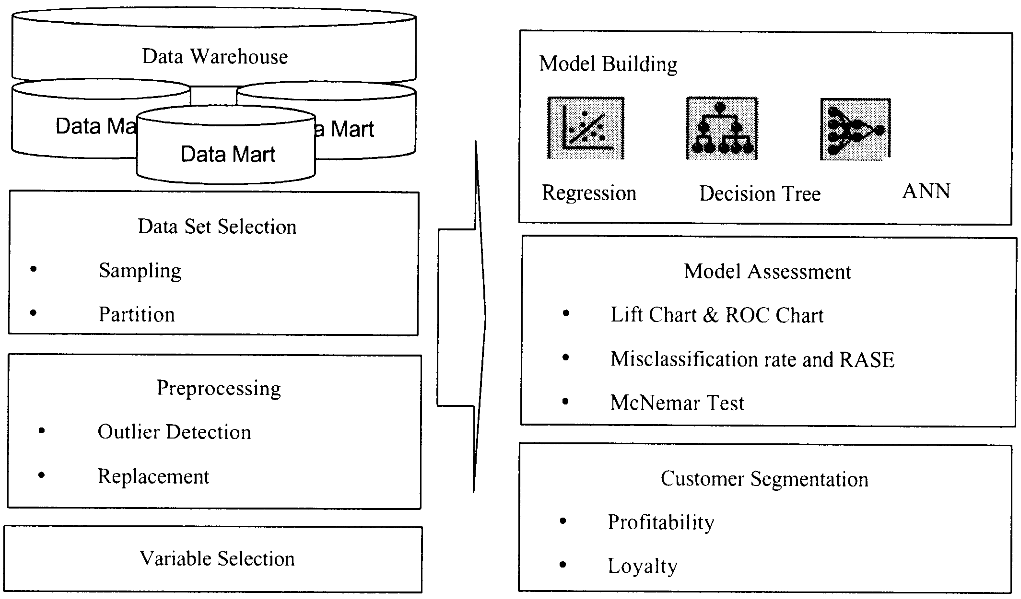
Table 2 Composition of Raw Data Set

	Churn-Person Model	Churn-Corporate Model
Training Set	1,064	384
Validation Set	633	265
Total	1,697	649

Table 3 Composition of Experiment Data Set

Training set is used for learning the feature's pattern and building a proper model. Validation set, known as holdout set, is for assessing the fitness of constructed model, and this set does not use in the training. Table 3 describes the composition of two-churn models' experiment data sets. Training data set gets through the data preprocessing which is two steps: Outlier detection and replacement. Outlier is a extreme value of each feature, and it makes a model skew or deviate from real data. This model considers the upper and lower bound of five-sigma as outlier, and if a feature is continuous variable, this outlier is replaced by mean value of feature, and if it is a discrete feature, by a mode value of feature. For the efficiency and comprehensiveness of model, input variable is selected in advance instead of using all variables. As variable selection method, R^2 method is used after independent t - test. Churn model is considered as a binomial question, so regression model (logit model), decision tree (C4.5, entropy reduction method, Quinlan 1993), and ANN are used for modeling method. Each model is assessed by 4 different methods: Lift chart, ROC chart, Misclassification rate and RASE (Root Average Sum of Error) and McNemar test. According to the result of assessment, one model is selected and used for customer segmentation matrix that has two axes, profitability and loyalty. Churn probability is tansformed to loyalty index by the form of (1 - churn probability). Figure 1 shows the flow of churn model construction.

Figure 1 Flow of Churn Model



2.3 Experimental Results

2.3.1 Churn – Person Model

For this experiment, ANN model is slightly more dominant than two other models. Its misclassification rate is 12.42% in the validation data set, which is better than regression (15.48%) and decision tree (15.01%) and this result is also applied to RASE.

Tool	Root ASE	Valid: Root ASE	Misclassification Rate	Valid: Misclassification Rate
Neural Network	0.2614	0.3018	0.0961	0.1242
Regression	0.2676	0.3237	0.1001	0.1548
Decision Tree	0.2838	0.3533	0.1055	0.1501

	NN & REG	REG & TREE	NN & TREE
N	633	633	633
Chi-Square		.040	1.163
Asymp. Sig.		.842	.281
Exact Sig. (2-tailed)	.049		

Table 4 Model Assessment and McNemar Test of Churn - Person Model

Statistically, as seen in Table 4 , it is hard to say that ANN always dominates the decision tree model. One can be certain only that ANN might dominate regression model according to the McNemar test that is used for nonparametric tests. However, ANN model is supported by lift chart and ROC chart. Lift chart shows the actual correct rate to each probability density percent that is sorted by the most probable churn rate. The lift chart (Figure 2) shows an approximately 95% hit ratio in the first 30% of subscribers, lined up by churn probability.

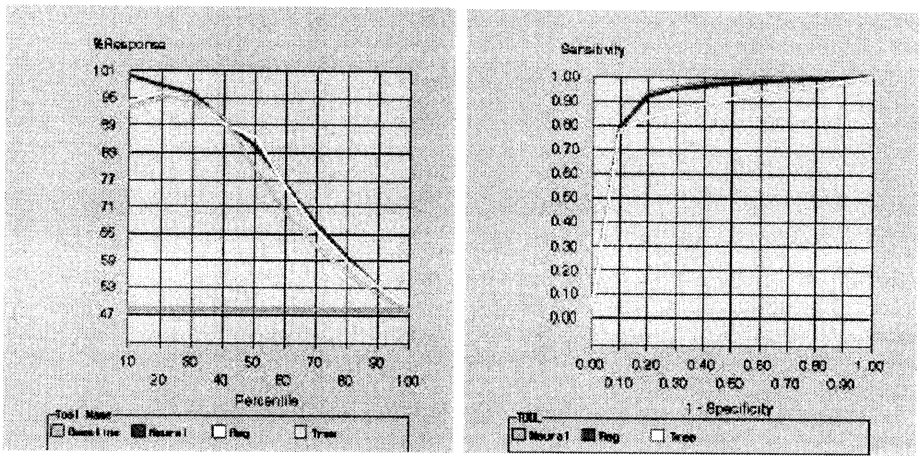


Figure 2 Lift chart and ROC Chart: Churn – Person Model

To the basis of Figure 2's ROC chart, ANN shows better forecasting power (sensitivity) from 0.1 of (1- specificity). Sensitivity means that an actual churner is correctly classified as a churner and specificity means an actual non-churner is classified as a non-churner. ANN model uses MLP(multi-layer perceptron), and churn-person model is best fitted 8 input variables and 4 nodes in the hidden layer. Continuous variables are obtained through standard deviation standardization. In the hidden layer, logistic function as activation function, Linear-general function as combination function and bias are used. 8 variables (Residential sub-area, Subscribing duration, Average wireless call charge per year, Maximum wireless call charge per year, Age, Average local call charge per month, Hometown, Residential Area) are selected as input variables.. District (residential sub-area) is the most critical variable to classify the churn probability. As nominal variable such as district code, and hometown code has various values but small size of sample, it is grouped like sub-group form "group variable". DISTRICT is grouped into 5 sub-group, STATE, 6 sub-groups, and HOME, 5 sub-groups. Subscribing duration, average and maximum wireless call charge per year, age and also can be used for explanation.

2.3.2 Churn – Corporate Model

ANN model and regression model show similar forecasting power in the Churn - Corporate model, 12.12% and 12.45% by each. That is supported by the McNemar test (Table 6), which shows the dominance of ANN and regression model rather than decision tree, this reason might attribute to small size of experiment data set. Although ANN model and regression model shows similar results, ANN model is slightly better than regression model according to RASE, lift chart and ROC chart results. In general, its hit ratio is similar to Churn - Person model and all three training data set is well trained. Also used experiment data set is half of the data set.

Tool	Root ASE	Valid: Root ASE	Misclassification Rate	Valid: Misclassification Rate
Neural Network	0.1672	0.3273	0.0296	0.1212
Regression	0.2029	0.3592	0.0458	0.1245
Decision Tree	0.2384	0.3869	0.0701	0.1698

	REG & NN	NN & TREE	REG & TREE
N	265	265	265
Chi-Square			4.033
Asymp. Sig.			.045
Exact Sig. (2-tailed)	.344	.002	

Table 6 Model Assessment and McNemra Test of Churn - Corporate Model

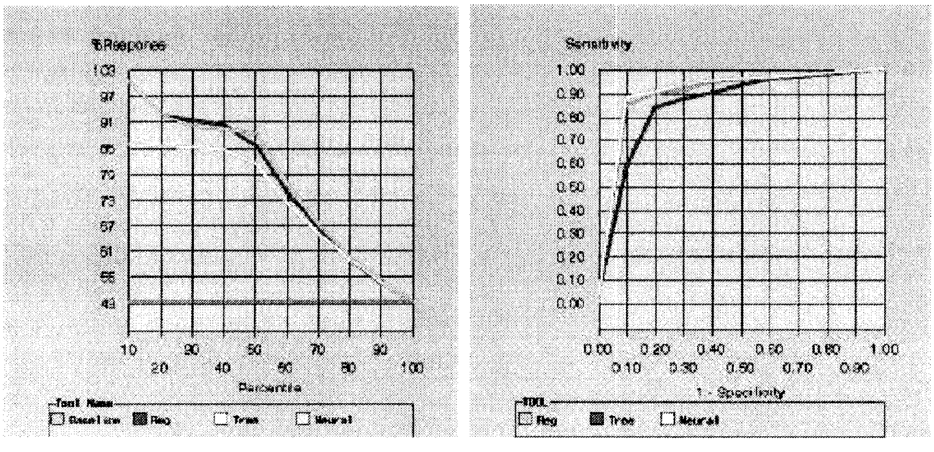


Figure 3 Lift chart and ROC Chart: Churn - Corporate Model

Lift chart is slightly down compare to Churn - Person model in spite of similar forecasting power. In the interval 20~40%, regression model is better than ANN model, but overly ANN model overwhelms slightly. 6 Variables (Residential sub-area, Average wireless call charge per year, terminating call forwarding, Maximum wireless call charge per year, Subscribing duration, residential Area) are selected as input variable using the stepwise variable selection method. Similar to Churn - Corporate model, residential sub-area is critical variable which R^2 is 0.692, but the other variables' effect is smaller than Churn - Corporate model. ANN Model is composed of 1 input layer (6 variables), 1 hidden layer (5 nodes) and 1 output layer, Terminating call forwarding is newly entered into the input model and HOME, subscribing duration, maximum wireless call charge per year variables are deleted.

2.4 Customer Segmenting and Marketing Strategy

As each customer has different value and different risk, companies should know the distribution of their customers. Customer scoring is one method to present customer's position. Customer score is the forecasting index that based on customer's historical purchase pattern (Birkhead,1999). This index uses two factors: One is customer's profitability, which is defined as customer's average telephone usage charge. The other is loyalty, which can be transformed as a form of (1 minus churn probability). Customer score gives company high quality information and it is used for customer segmentation criterion. Figure 4 shows the long distance call subscribers' position, which is divided into 9 different groups by dotted lines. Person model's profitability range is from 0 won to 150, 000 won, meanwhile corporate model is from 0 won to 300,000 won as double as person model, but distributions of customer is similar. With the effect of data-driven modeling, customer position is concentrated on two extreme sides (loyalty 0 and 1).

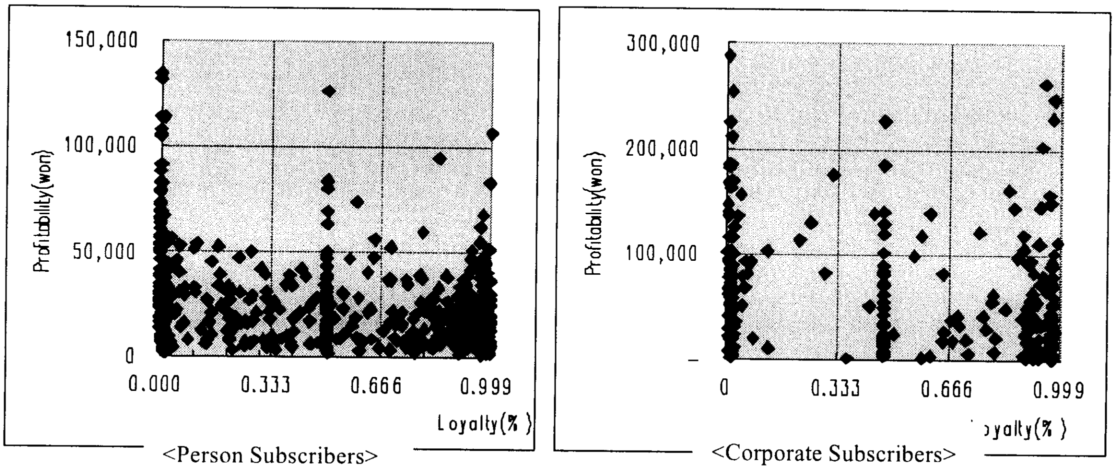


Figure 4 Distributions of Subscribers

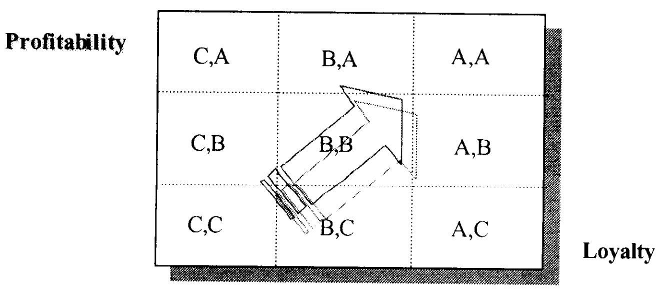


Figure 5 Customer Segmentation Matrix

Based on present scoring, marketing strategy can be summarized as two alternatives except angel marketing. One is short-term strategy and the other is long-term strategy, Companies always should concern about actively angel marketing i.e. (A,A) zone which is most profitable customer zone and also have maintained good relationship. For the short term outcome, the first target zone may be (C,A) and (C,B) zone. These groups are profitable but their loyalty is weak. If they were attracted by competitors' promotion campaign or price reduction strategy, they will leave easily. Therefore before leaving, company should reveal deep concern for these customers and suggest a competitive price plan. Secondly, targeting zone may be (A,C) and (A,B) group whose loyalty is good but profitability is relatively less. These two groups' relative inferiority might com from the price structure, As these groups have good loyalty, if their preferences are carefully examined and suggest tailored service, then their usage will go up.

3. Conclusion

As ultimate purpose is to keep good relationship with profitable customers or make customers more profitable with good loyalty, long-term strategy should include all possible customers. This strategy should be based on customer satisfaction and even more on customer movement. New products and service development are essential and company image refinement is also useful. This recognition might come out from the mutual confidence that can be accumulated by a continued good relationship.

REFERENCES

1. Birkhead, B. G. (1999). "Behavioral Segmentation system," white paper *http://www.crm-forum.com.*
2. Fayyad, U. M. (1996). "Data mining and knowledge discovery: making sense out of data," *IEEE Expert, Oct.* 20-25.
3. Korea Information Society Development Institution. (1999). "Monthly Report: Trend of Korean Information and Communication Industry," Oct. http://www.kisdi.re.kr/
4. Modisette, L. (1999). "Milking wireless churn for profit," *Telecommunications, Dedham; Feb.; Vol. 33, Iss. 2*; International edition; pg. 73, 2 pgs Meredith Lockard, "Test your retention IQ; Target Marketing," Philadelphia; Mar 1998; Vol. 21, Iss. 3; pg. 32, 2 pgs.
5. Puckey, D. (1999). "Modeling customer relationship," white paper *http://www.crm-forum.com.*
6. Quinlan, J.R. (1979). "Discovering rules by induction from large collections of examples," *Expert sytems in the micro electronic age.* Edinburgh University Press.
7. Quinlan, J.R. (1993). "C4.5 Programs for machine learning," San Mateo, Morgan Kaufmann.
8. SAS Korea Inc. (1998). "Analysis of defected customers using SAS/Enterprise Miner," *SIGDM '98*, 228-259.
9. Schmitt, J. (1999). "Churn: Can carriers cope?" *Telecommunication, Feb.* 32-33.