Knowledge-Based Systems 28 (2012) 97–104

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

# Predicting customer churn through interpersonal influence

Xiaohang Zhang [a,*], Ji Zhu [b], Shuhua Xu [c], Yan Wan [d]

[a] School of Economics and Management, Beijing University of Posts and Telecommunications, 1318 Main Building, 10 Xitucheng Road, Haidian District, Beijing 100876, China
[b] Department of Statistics, University of Michigan, 439 West Hall, 1085 South University Ann Arbor, MI 48109-1107, USA
[c] School of Economics and Management, Beijing University of Posts and Telecommunications, 1319 Main Building, 10 Xitucheng Road, Haidian District, Beijing 100876, China
[d] School of Economics and Management, Beijing University of Posts and Telecommunications, 506 Mingguang Building, 10 Xitucheng Road, Haidian District, Beijing 100876, China

ARTICLE INFO

*Article history:*
Received 24 July 2011
Received in revised form 5 December 2011
Accepted 5 December 2011
Available online 13 December 2011

*Keywords:*
Churn prediction
Interpersonal influence
Propagation
Telecommunications
Social network

ABSTRACT

Preventing customer churn is an important task for many enterprises and requires customer churn prediction. This paper investigates the effects of interpersonal influence on the accuracy of customer churn predictions and proposes a novel prediction model that is based on interpersonal influence and that combines the propagation process and customers' personalized characters. Our contributions include the following: (1) the effects of interpersonal influence on prediction accuracy are evaluated while including determinants that other researchers proved effective, and several models are constructed based on machine learning and statistical methods and compared, assuring the validity of the evaluation; and (2) a novel prediction model based on interpersonal influence and information propagation is proposed. The dataset used in the empirical study was obtained from a leading mobile telecommunication service provider and contains the traditional and network attributes of over one million customers. The empirical results show that traditional classification models that incorporate interpersonal influence can greatly improve prediction accuracy, and our proposed prediction model outperforms the traditional models.



## 1. Introduction

Preventing customer churn is an important task for many enterprises, especially in matured industries, including telecommunications [15,22,1] and finances [42]. Achieving it requires churn prediction, which is defined as identifying customers who tend to switch to other service providers. Findings from previous research can be categorized according to two aspects. First, churn determinants are analyzed and verified using customer behaviors in various industries. Some attributes, including customer satisfaction, switching costs, customer demographics, tendency to change behavior, and service usage, have been found to be common churn determinants [15,22,1]. Second, researchers have proposed prediction models based on machine learning methods, including the decision tree, neural network, and support vector machine [18,41,8,7], or statistical methods, including logistic regression, survival analysis, and Markov chain [21,24].

Though many practices benefit from these valuable results, one limitation exists. Many researchers have assumed implicitly or explicitly that a customer's decision to switch service providers is independent of other customers' decisions. Most prior research focuses exclusively on individual customers, without accounting for any interpersonal influence [9], as they measure each customer's perception and behavior independently. Typically, many explanatory variables are collected on each customer and used in multivariate prediction modeling. In reality, customers' behaviors not only depend on their own perceptions and subjective desires but also interplay with each other. Thus, customers' choices are interdependent [45].

This paper investigates the effects of interpersonal influence on the accuracy of predicting customer churn and proposes a novel prediction model based on interpersonal influence that combines the propagation process and customers' personalized characters. Our contributions include the following: (1) the effects of interpersonal influence on prediction accuracy are evaluated while including traditional attributes (i.e., customers' personalized characters) that other researchers proved to be effective, and several models are constructed based on machine learning and statistical methods and compared, assuring the validity of the evaluation; and (2) a novel prediction model based on interpersonal influence and information propagation is proposed. The dataset used in the empirical study was obtained from a leading mobile telecommunication service provider and contains the traditional and network attributes of over one million customers. The empirical results show that traditional prediction models incorporating interpersonal influence can greatly improve prediction accuracy, and our proposed prediction model outperforms the traditional models.

* Corresponding author. Present address: 427 West Hall, 1085 South University Ann Arbor, MI 48109-1107, USA. Tel.: +1 7347096295, mobile: +86 13910390890; fax: +86 62282039.

*E-mail addresses:* zhangxiaohang@bupt.edu.cn (X. Zhang), jizhu@umich.edu (J. Zhu), xushuhua@bupt.edu.cn (S. Xu), wanyan@bupt.edu.cn (Y. Wan).

0950-7051/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2011.12.005

This paper is organized as follows. Section 2 discusses several related works. Section 3 provides the churn determinants used in our models. In Section 4, several classification models are separately constructed based on different attributes and methods, and a propagation-based model is proposed. Section 5 discusses the experimental results and their implications. Section 6 gives our conclusions.

## 2. Literature

Customer churn prediction can be regarded as a classification problem, in which each customer is classified into one of two classes, churn or non-churn. Machine learning and statistical methods are the most widely used approaches for classification problems. Popular churn prediction methods include logistic regression [33], decision trees [18], neural networks [41], support vector machines [8], and evolutionary algorithms [3]. However, many previous studies solely focused on customers' individual attributes when using these models. Though some studies accounted for interpersonal influence [9,37], they did not combine it with customers' personalized attributes.

Propagation models have also been widely used to describe the dynamic processes of viral marketing [11,36,19], trust formation [23,25], risk propagation [31], epidemic diffusion [10], and computer virus infection [34]. Popular propagation models include the SI, SIR, and SIS models [16], which are based on different diffusion mechanisms. In our proposed propagation model, we partially employ the SI model and combine its diffusion process with customers' personalized characters. In the SI model, each infected individual can pass information (or influence) to his/her neighbors in a network in which recovery is not possible, i.e., an infected individual continues transmitting information indefinitely.

To our knowledge, little research has been conducted on using propagation models to predict customer churn. Dasgupta et al. [9] proposed a spreading activation-based technique (SPA) that predicts potential churners by examining the current set of churners and their underlying social networks. Using the underlying topology of the customer contact network, the SPA initiates a diffusion process with the churners as seeds. Essentially, it models a "word-of-mouth" scenario, in which a churner influences a neighbor to churn, and the influence spreads from that neighbor to another neighbor, and so on. At the end of the diffusion process, the amount of influence received by each node is inspected. Using a threshold-based technique, a node that is currently not a churner can be declared to be a potential future churner based on the influence it accumulated. However, many propagation models assume that customer behaviors (i.e., churn or non-churn) depend solely on interpersonal influence, not accounting for customers' personalized attributes, i.e., customers are believed to be homogeneous. In reality, customers may behave differently, though they have similar neighbors or interpersonal influence. For example, if a customer has high switching costs, strong influences from other churn neighbors may not lead him or her to churn.

## 3. Determinants of customer churn

For clarity and comparison, we classify customer churn determinants into two categories: network and traditional attributes. Network attributes measure interpersonal influence and describes each customer's local topology in customer contact network and his or her relationships with their neighbors. The other attributes are included in the traditional attributes category, which has been frequently discussed in previous research.

All attributes used in our study were obtained from the telecommunication service database of a mobile telecommunication company. The database stores certain customer-related information and call detail records (CDR), which include the labels of calling and called customers and the start time and duration of the transaction. Our attributes differ from those used in previous research, in which data are based on questionnaires. However, actual customer transactions or billing data may fully represent customer's actual future decisions better than survey data [1]. Moreover, to paint a complete picture of interpersonal influence, the ideal dataset would have measurement of direct communication between customers [17], which can be extracted from the CDR.

### 3.1. Traditional attributes

To assure the validity and credibility of the evaluation of interpersonal influence on prediction accuracy, we include traditional attributes in our study. We focus on five prominent drivers in this literature: customer satisfaction, switching barrier, service usage, price sensitivity, and previous anomaly behavior. We do not discuss traditional attributes in detail, as they have been frequently discussed and used in previous research. Table 1 shows the descriptions of traditional attributes and their supporting references.

### 3.2. Network attributes

We represent interpersonal influence through network attributes, including neighbor composition, tie strength, similarity, homophily, structural cohesion, influence degree of neighbors, and order-2 neighbors (see Table 2).

#### 3.2.1. Neighbor composition

Customers who have direct links with customer $i$ are defined as customer $i$'s neighbors. Neighbor composition is important because different types of neighbors (i.e., churn neighbors or non-churn neighbors) may have different influences. Moreover, external neighbors who belong to different service providers may either explicitly influence customer churn by conveying their perceptions of satisfaction or implicitly influence customers through their choices of service provider.

#### 3.2.2. Tie strength

Tie strength measures the intensity of contact between a pair of customers. Dasgupta et al. [9] show that tie strength can improve churn prediction accuracy in the telecommunications industry. In this study, we quantify tie strength by the total number of calls made between a pair of customers over the studied period, as prior studies [35,9].

#### 3.2.3. Similarity

In this paper, we use the notation of similarity to denote overlap in a local neighborhood, as in other studies [17,35]. Similarity between customers $i$ and $j$ is defined as $C_{ij}/(k_i + k_j - C_{ij} - 2)$, where $k_i$ and $k_j$ denote the number of neighbors of customers $i$ and $j$ respectively, and $C_{ij}$ denotes the number of common neighbors of customers $i$ and $j$ [35].

#### 3.2.4. Homophily

Homophily is the tendency of customers to associate and bond with others who are similar. The presence of homophily has been discovered in a vast array of network studies [32]. In this paper, homophily refers to "churn homophily", which implies that customers with similar characters are more likely to influence each other when making churn decisions. The following measures churn homophily between customer $i$ and his or her neighbor $j$:

$$H_{ij} = \sqrt{\sum_p w_p (a_{ip} - a_{jp})^2}, \tag{1}$$

**Table 1**
Traditional attributes and their descriptions.

| Attributes | Descriptions | References |
|---|---|---|
| *Customer satisfaction* | | |
| Prices | The price of local, long-distance and roaming calls | [1,12,15,22,21,30] |
| Brand* | Different usage and tenure levels, enjoying different servicing levels from the operator | |
| Key accounts level* | Different customer groups who own different levels of contributions and rapid growth in contributions and enjoy different service levels and degrees of concern | |
| Functions of the handset | The number of functions of the mobile device, including GPRS service, multimedia message service, Java application, and wireless application protocol | |
| Complaints | Number of times a customer files a complaint with a customer service center regarding problems with billing or service quality | |
| *Switching barrier* | | |
| Loyalty points | Amount of credits a customer earned, which are redeemable for a wide variety of goods and services, including retail gifts and coupons | [1,4,22] |
| Feedback gift | Remaining months of feedback from the service provider | |
| Discount level | Discount customers enjoyed, which come from the promotion feedback | |
| Lucky number** | Whether the telephone number contains a lucky number | |
| *Service usage* | | |
| Monthly call counts | Average monthly call counts | [22,33,21] |
| Monthly payment | Average monthly payment for mobile service | |
| Short message | Average monthly times of short message usage | |
| Diversity of service usage | The number of value-added services the customer uses | |
| Activeness of service usage | The number of days customers use the value-added service | |
| Duration of subscription | Duration of subscription with the present service provider | |
| *Price sensitivity* | | |
| Chang billing suite | Average monthly frequency of changing the billing suite | [2,40] |
| Free call duration | Ratio of free call duration to total call duration | |
| IP call duration | Ratio of IP call duration to total long call duration | |
| Frequency of paying | Average monthly frequency of paying bills | |
| *Previous anomaly behavior* | | |
| Change of fees | Change of recent monthly fees | [38,39,15,18] |
| Change of call counts | Change of recent monthly call counts | |
| Anomaly status | Number of months when customer is in anomaly status, such as temporal arrear, in previous months | |
| Calling behavior** | Whether there are calling behaviors | |

* Denotes nominal attribute.
** Denotes binary attributes.

**Table 2**
Network attributes and their descriptions.

| Attributes | Descriptions | References |
|---|---|---|
| *Neighbor composition* | | |
| Outside neighbors | Total number of neighbors in different operators' networks | [20,17,27,28] |
| Ratio of outside neighbors | Ratio of neighbors in different operator's networks | |
| Churn neighbors | Total number of churn neighbors | |
| Non-churn neighbors | Total number of non-churn neighbors | |
| Ratio of churn neighbors | Ratio of churn neighbors | |
| *Tie strength* | | |
| Tie strength to churn neighbors | Average of call counts to/from churn neighbors | [6,9,35] |
| Tie strength to non-churn neighbors | Average of call counts to/from non-churn neighbors | |
| *Homophily* | | |
| Homophily to churn neighbors | Average homophily to churn neighbors. If there are no churn neighbors, it is 0 | [29,32] |
| Homophily to non-churn neighbors | Average homophily to non-churn neighbors. If there are no non-churn neighbors, it is 0 | |
| *Similarity* | | |
| Similarity to churn neighbors | Average similarity to churn neighbors. If there are no churn neighbors, it is 0 | [17,35,37] |
| Similarity to non-churn neighbors | Average similarity to non-churn neighbors. If there are no non-churn neighbors, it is 0 | |
| *Structural cohesion* | | |
| Cohesion of churn neighbors | Structural cohesion of churn neighbors. If there are no churn neighbors, it is 0 | [26,43,37] |
| Cohesion of non-churn neighbors | Structural cohesion of non-churn neighbors. If there are no non-churn neighbors, it is 0 | |
| *Order-2 neighbors* | | |
| Order-2 churn neighbors | Number of order-2 churn neighbors | [17,9] |
| *Influence ability of neighbors* | | |
| Influence of churn neighbors | Average degrees of churn neighbors | [13] |
| Influence of non-churn neighbors | Average degrees of non-churn neighbors | |

where $a_{ip}$ denotes standardized attribute $p$ of customer $i$ and $w_p$ denotes the weight of attribute $p$, which reflects the importance of this attribute in evaluating churn homophily. Except for categorical attributes, all other continuous attributes defined in Table 1 are used to compute homophily. A variety of methods can measure attribute weight, including online or batch optimizers, class projection and mutual information. Wettschereck et al. [44] provided a good review on this subject. In this study, we employ Gini reduction to measure the weight for each attribute.

We obtain a class probability distribution $p^s_{\text{churn}}$ and $p^s_{\text{non-churn}}$ for the target attribute (churn or non-churn) in the complete dataset $S$, where $p^s_{\text{churn}} + p^s_{\text{non-churn}} = 1$. The Gini index [5] is defined as:

$$\text{Gini}(S) = 1 - \left(p^s_{\text{churn}}\right)^2 - \left(p^s_{\text{non-churn}}\right)^2. \tag{2}$$

For the continuous attribute $p$, the weight $w_p$ is defined as:

$$w_p = \text{Gini}(S) - \frac{|S_1|}{|S|}\text{Gini}(S_1) - \frac{|S_2|}{|S|}\text{Gini}(S_2), \tag{3}$$

where the dataset $S$ is separated into two parts, $S_1$ and $S_2$, by one cut point chosen from all possible values of attribute $p$ based on the criterion that $w_p$ is the largest. The weight of one attribute reflects the importance of this attribute in predicting customer churn.

#### 3.2.5. Structural cohesion

Interpersonal influence on churn is highly dominant in a social group with high structural cohesion [37]. In this study, the structural cohesion of a group is measured by the ratio between $e$ and $k(k-1)/2$, where $k$ denotes the number of members in the group and $e$ denotes the number of actual links among group members [43].

#### 3.2.6. Order-2 neighbors

In a network model, an entity is typically influenced most by those directly connected to it, but the entity is also affected, though to a lesser extent, by those farther away [17]. In this study, we account for the effects of order-2 churn neighbors. A churn customer who has direct links to at least one non-churn neighbor of customer $i$ and no direct links with customer $i$ is defined as customer $i$'s order-2 neighbor.

#### 3.2.7. Influence ability of neighbors

Except for relationships between a customer and his/her neighbors, influence effects also depend on the neighbors' own influence abilities. Goldenberg et al. [13] believe that the number of social ties represents the traits of influential people, and they use it to identify social hubs. In this study, we also employ the number of social ties to represent a customer's influence ability.

## 4. Models

In this study, we employ two types of models. The first are classification models, which are built based on machine learning or statistical methods. We use the classification models to examine whether interpersonal influence (i.e., network attributes) can improve prediction accuracy. The second is our proposed propagation model that combines a propagation process and customers' personalized characters.

### 4.1. Classification model

To evaluate the effects of interpersonal influence on prediction accuracy, we construct three types of classification models. The first (CM1) focuses solely on the traditional attributes shown in Table 1, the second (CM2) focuses solely on the network attributes shown in Table 2, and the third (CM3) focuses on the combination of traditional and network attributes. For each model, three popular methods are adopted, including logistic regression (LR), decision tree (DT) and neural network (NN) methods. The tool used to train the models is SAS Enterprise Miner. For comparison, each method for all models uses the same configuration. The logistic regression configuration is a logit link function. The Gini reduction is the splitting criteria for a decision tree. A multilayer perceptron neural network structure is chosen with one hidden layer of five neurons.

### 4.2. Propagation model

In this study, we propose a propagation model that combines a propagation process and customers' personalized characters. Fig. 1 describes the prediction algorithm.

This algorithm must specify two rules: the IR and SR. Both rules can have various forms. The IR can be a pre-specified probability threshold above which customers enter $S'$, or it can be a top-$n$ rule ($n$ must be specified), i.e., customers whose churn probabilities are the top-$n$ largest enter $S'$. The SR can be a fixed number of iterations or satisfied when the number of iterations chosen to maximize prediction accuracy is reached. We provide a general architecture in which rules can be specified on the basis of solving problems. Our proposed propagation model accounts for both interpersonal influence (network attributes) and customers' personalized attributes (traditional attributes), as the CM3 models used in the algorithm are based on these two types of attributes. Moreover, this model takes advantage of the propagation method and is a dynamic prediction process because the algorithm is iterative and changes the churner set and further interpersonal influence in each loop.

## 5. Experiments

### 5.1. Dataset description

In this study, the experimental data were obtained from December 2008 to June 2009 from a leading mobile service provider and comprise more than one million customers. The

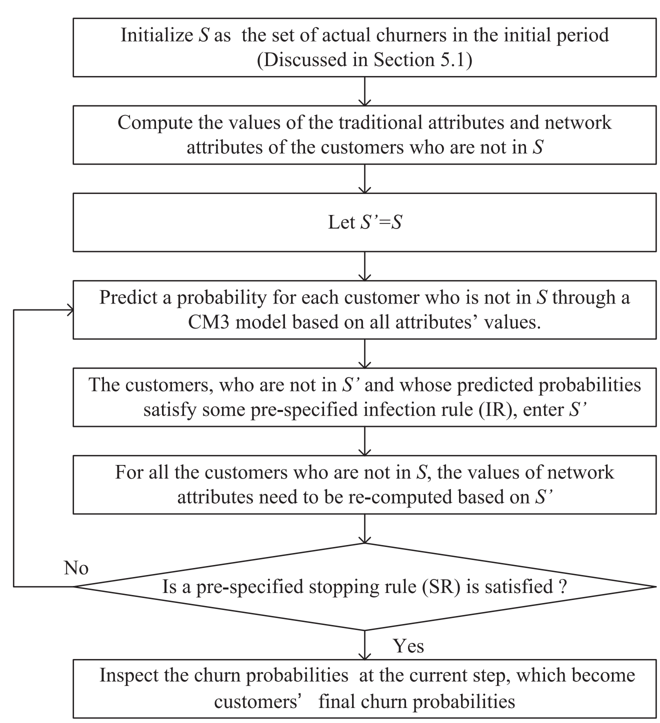


**Fig. 1.** The prediction algorithm for the propagation model.

customers who had a normal status and still used services in March 2009 were defined as the "basis customers" for prediction. Some basis customers churned from April 2009 to June 2009, during which churn status was identified if the customers never used services in the period and their contacts were terminated by the mobile provider. In the dataset, approximately 4% of the basis customers churned during the period. Data from December 2008 to March 2009 were used to compute the values of traditional and network attributes. The actual churn customers in the period are defined as an initial churner set to serve as the basis for computing the network attributes' values. The basis customer dataset was used to evaluate the effects of network attributes and test the prediction accuracy of our proposed propagation model. Moreover, to evaluate the effects of network attributes, the basis customer dataset was randomly divided into two parts, following a 60/40 proportion, to train and test the classification models separately.

### 5.2. Measures of prediction accuracy

In this study, we used the same measures to quantify prediction accuracy, as in Dasgupta et al. [9]. These measures are popular in evaluating prediction models and are often used in customer churn prediction (e.g., [3]. The first measure is the lift curve, which plots the fraction of all churners having a churn probability above the threshold against the fraction of all subscribers having a churn probability above the threshold. The lift curve indicates the fraction of churn customers that can be caught (retained) if a certain fraction of all subscribers are contacted. As an operator's customer services center has a fixed number of personnel to contact some fraction of all subscribers, the lift curve can be useful in determining how to allocate limited resources. The second measure is the AUC, which is equivalent to the area under the ROC curve. This measure is a probability that a randomly chosen churner ranks higher than a randomly chosen non-churner; an AUC = 1.0 indicates that the classes are perfectly separated, and an AUC = 0.5 indicates that the list is randomly shuffled. The third measure is the "hit rate", which is the number of correct "churn" predictions as a percentage of the total number of nodes labeled "churn". A low hit rate implies a large number of "false positives".

### 5.3. Results of the classification models

Fig. 2 shows the prediction results of the three classification models. For all methods (DT, LR and NN), the combination models (CM3) outperform both network attributes-based models (CM2) and traditional attributes-based models (CM1). The CM3 model especially outperforms other models in the LR method. The hit rate increases by approximately 4%, and lift increases by more than 10% in the 10th percentile. The results illustrate that combining traditional and network attributes can improve prediction accuracy. Moreover, some differences exist among the results of the three methods. For the NN method, the CM2 prediction accuracy is close to the CM3 accuracy, both of which are better than CM1's. Conversely, for the DT method, the CM1 prediction accuracy is close to the CM3 accuracy, both of which are better than CM2's. For the LR method, however, the CM1 and CM2 results are close, and both of them are worse than CM3's. Thus, the CM1 and CM2 models have different strengths, and neither should be ignored. From these results, we can conclude that network attributes can improve prediction accuracy. If we compare the results of the three methods, we find that NN outperforms LR and LR outperforms DT for all three models.

Fig. 2 shows that network and traditional attribute-based models have their own advantages. To further explore the complementarity of traditional and network attributes, we analyzed the overlap of churn customers predicted by CM1 and CM2 based on the NN method. Table 3 shows the prediction results in the 10th percentile.

The first row of Table 3 is the composition (churn and non-churn) of customers only predicted using the traditional attributes-based model (CM1 only), the second row is the composition of customers only predicted using the network attributes-based model (CM2 only), and the last row is the composition of overlapping customers predicted using both models (CM1 and CM2). We can see that the prediction accuracy of the overlapping part is much higher than with either separate model. Additionally, the overlap of the churn customers predicted by both models is a small proportion, occupying only 28.14% of the test dataset. This result indicates that the characters of some churn customers can only be captured by traditional attributes, so it is difficult to predict these customers using network attributes; similarly, other churn customers can only be characterized with network attributes. This finding further demonstrates that these two attribute types should be combined to predict churn, which can improve prediction accuracy.

To overcome the shortcomings of CM1 and CM2 or to combine their advantages, we also constructed two ensemble models. The first model is the "join_mean" model, in which predictions are based on the mean of the results obtained with the CM1 and CM2 models, i.e., one customer's churn probability equals the average of the churn probabilities predicted by these two models separately. The second is the "join_max" model, in which churn probability equals the maximum of the results obtained with the two models. Table 4 shows the results in the 10th percentile. Both the join_mean and join_max models outperform the CM1 and CM2 models. However, they are still relatively poor compared with the CM3 model. The join_mean model results are close to the CM3 model's. The two ensemble models again show the importance of network attributes in improving prediction accuracy.

### 5.4. Results of the propagation model

Our proposed propagation model contains an iteration process, in which some parameters and rules must be specified. First, in our experiments, all methods (i.e., DT, LR and NN) are adopted separately by the CM3 model in each loop to predict customers' churn probabilities. Here, we call each loop one round of propagation. Second, the infection rule (IR) is a top-$n$ rule, in which $n$ refers to the number of infected customers, i.e., $n$ customers with the highest churn probabilities change to 'churn' status and enter $S'$. In the experiments, $n$ is specified to 1000, 5000 and 10,000, respectively. Third, the stopping rule (SR) is satisfied after reaching a specified number of propagations. To examine the effects of the number of propagations on prediction accuracy, we set the number of propagations to be large enough. After each round of propagations, we compute the AUC values based on the current predicted churn probabilities, with different methods and $n$ values. Strictly speaking, the results in the first round of propagations should be same as the results obtained in Section 5.3. However, there is little difference between the results, as the results in Section 5.3 come from the prediction based on the test dataset (Section 5.1), and the propagation model results are based on the prediction from the entire dataset. Because the propagation model accounts for whole relationships among customers, we cannot solely focus on the test dataset.

Fig. 3 shows the AUC values with different methods and the $n$ values. For the DT method, as the number of propagations increases, the AUC value first increases and reaches its highest value at a certain point, and then it begins to decline. The difference between the highest AUC value and the value of the first round of propagations is approximately 0.035 when $n$ is 10,000. This result illustrates that the propagation process can improve prediction

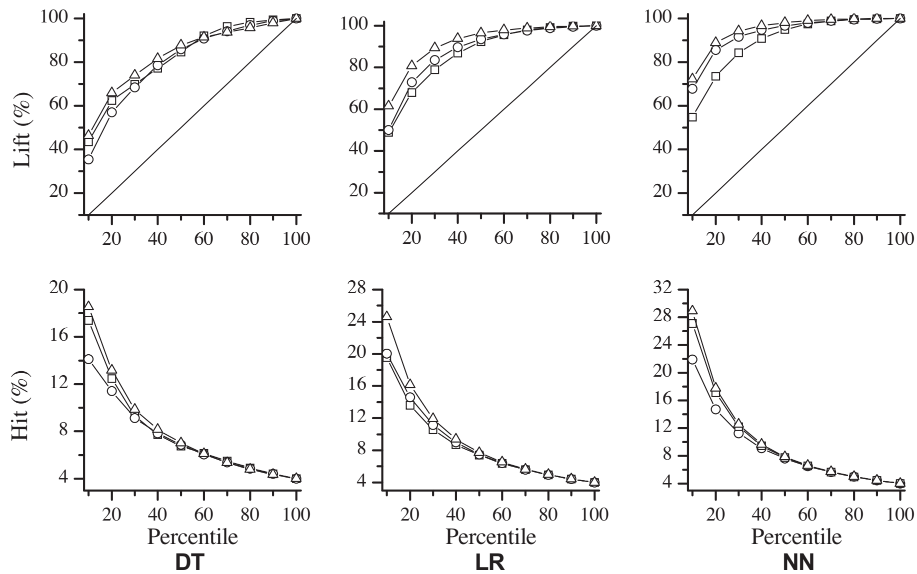


**Fig. 2.** Hit and lift curves for the three classification models: CM1 (□), CM2 (○) and CM3 (△).

**Table 3**
The overlap of the customers predicted with CM1 and CM2 in the 10th percentile.

| Model | Number of churners | Number of non-churners | Hit rate (%) |
|---|---|---|---|
| CM1 only | 4358 | 28,911 | 13.10 |
| CM2 only | 6767 | 26,502 | 20.34 |
| CM1 and CM2 | 5776 | 7249 | 44.34 |

**Table 4**
Prediction results of five model types in the 10th percentile using the NN method.

| Model | Number of churners | Number of non-churners | Hit rate (%) | Lift (%) |
|---|---|---|---|---|
| CM1 | 10,134 | 36,160 | 21.89 | 54.73 |
| CM2 | 12,543 | 33,751 | 27.09 | 67.74 |
| CM3 | 13,371 | 32,923 | 28.88 | 72.21 |
| Join_mean | 13,324 | 32,970 | 28.78 | 71.95 |
| Join_max | 12,749 | 33,545 | 27.54 | 68.85 |

accuracy. Moreover, we can adopt the number of propagations, which optimizes prediction accuracy, as the stopping rule when we predict future customer churn. The shapes of the AUC curve at different values of $n$ are also similar, which is consistent for all methods. Analogously, the AUC curves in the NN method also first increase, then decrease. The difference between the highest and first points is smaller in the NN method than in the DT. The reason for this result may be that the AUC value in the NN method at the first round of propagations reached a high level of 0.888, which results in a lack of improvement space. Additionally, the LG method results are similar to the NN and DT.

From the results of the experiments, we can conclude that: (1) the propagation model can produce better prediction accuracy than the classification models described in Section 4.1, (2) $n$ values would not significantly impact the results, and (3) the number of propagations maximizing prediction accuracy can be chosen as the stopping rule.

We also compared our proposed propagation model to the SPA model [9], whose basic prediction principle is described in Section 2. Fig. 4 shows the AUC values of the SPA model for different spreading factor values. The highest AUC value (0.6304) of the SPA model is lower than that in our model, possibly because the SPA model focuses solely on network topology and does not incorporate customers' personalized characters (i.e., traditional attributes). The experimental results in Section 5.3 show that some churn customers can only be captured by traditional attributes.

### 5.5. Discussion

In many earlier studies, customer churn prediction depends on a customer's individual characters, and others do not influence his/her churn behavior. Though customers are often assigned roles, including gender and occupation, to illustrate their behaviors in social structures, they are still independent and simply attached to their roles. Granovetter [14] gave a brilliant exposition of this concept. However, customers can be more accurately characterized from the social network perspective, in which their behavior is constrained and influenced to some extent. This study shows that incorporating network attributes into a prediction model can improve prediction accuracy (in logistic regression model, the hit rate improves from 19.57% to 24.58%, and the lift rate improves from 48.93% to 61.46% when targeting 10% of customers). In particular, some customers are more likely to be influenced by others and can only be predicted by network attributes. Of course, some churn customers can only be captured by traditional attributes. It is thus necessary to examine customer behavior from an interpersonal influence perspective and combine traditional and network attributes to predict churn.

While incorporating network attributes into classification models can improve prediction accuracy, traditional classification models are static and do not consider time-varying characteristics. Our proposed propagation model addresses this drawback. On the one hand, the proposed model is a dynamic process with multiple propagation rounds. In each round, some non-churn customers change to churn status, and they bring new influences to their neighbors in the next round. On the other hand, the propagation model combines interpersonal influence and customers' personalized characters when it decides which customers change to churn. This facet differs from some traditional diffusion models, which are often used on the condition that all nodes in the network are homogeneous (e.g., [9,36]. The propagation model provides a new perspective and tool to understand and predict customers' behaviors.

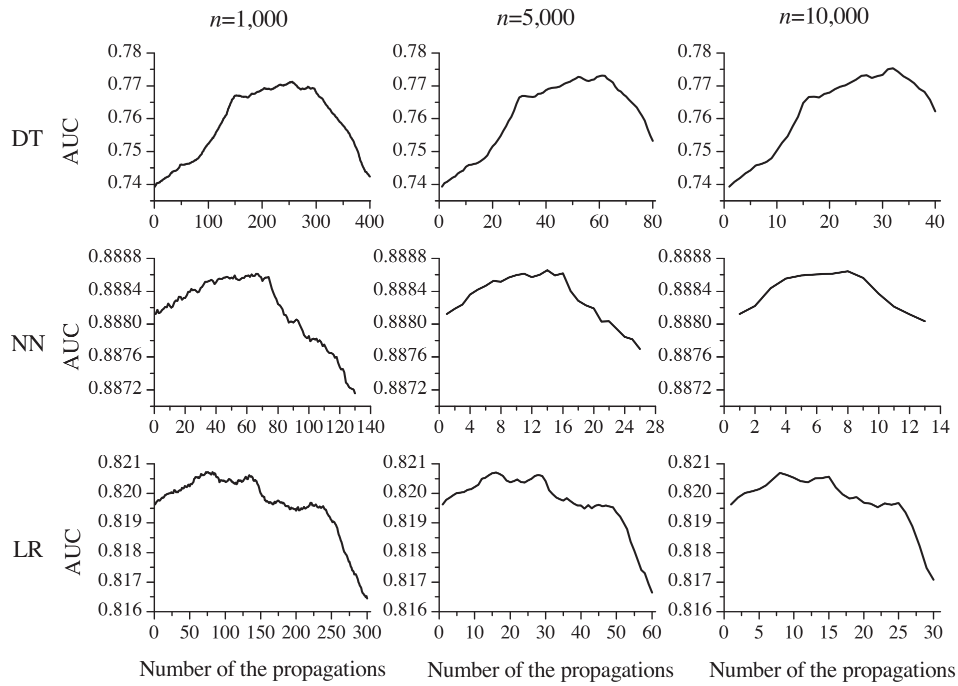


**Fig. 3.** AUC values of the propagation model for different methods and $n$ values.

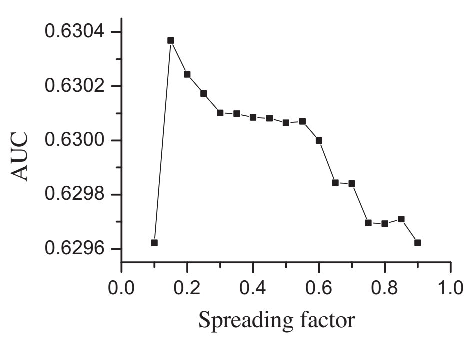


**Fig. 4.** AUC values of the SPA model for different spreading factor values.

Our proposed propagation model can also predict other customer behaviors, including service adoption, in which the implementation process is the same. The model can also be applied to other industries. The big challenge in its implementation is collecting direct communication data between customers. Fortunately, many Internet and financial companies own customer network data. We believe that the propagation model has a wide application scope.

## 6. Conclusion

In this paper, we investigate the effects of interpersonal influence on the accuracy of predicting customer churn. Interpersonal influence is represented by network attributes that refer to interactions among customers and the topologies of their social networks. We compared the prediction results of traditional attributes-based models, network attributes-based models and combined attributes models and found that incorporating network attributes into predicting models can greatly improve prediction accuracy. In particular, churn behavior for some customers can only be distinguished using network attributes. Network attributes can thus be useful complements to traditional attributes. Moreover, we proposed a novel prediction model based on the propagation process that accounts for interpersonal influence and customers' personalized characters. The empirical results show that the proposed propagation model outperforms traditional classification models.

## Acknowledgments

We thank the editor and two referees for many helpful comments. This work was partially supported by the Project 70901009 of National Science Foundation for Distinguished Young Scholars, and the Youth Research and Innovation Program 2009RC1027 in Beijing University of Posts and Telecommunications.

## References

[1] J.-H. Ahn, S.-P. Han, Y.-S. Lee, Customer churn analysis: churn determinants and mediation effects of partial defection in the Korean Mobile Telecommunications Service Industry, Telecommun. Policy 30 (2006) 552–568.
[2] K.L. Ailawadi, B. Harlam, An empirical analysis of the determinants of retail margins: the role of store-brand share, J. Marketing 68 (2004) 147–166.
[3] W.H. Au, K.C.C. Chan, X. Yao, A novel evolutionary data mining algorithm with applications to churn prediction, IEEE Trans. Evolut. Comput. 7 (2003) 532–545.
[4] R. Bolton, P. Kannan, M. Bramlett, Implications of loyalty program membership and service experiences for customer retention and value, J. Acad. Mark. Sci. 28 (2000) 95–108.
[5] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and Regression Trees, Wadsworth, Belmont, 1984.
[6] J.J. Brown, P.H. Reingen, Social ties and word-of-mouth referral behavior, J. Consum. Res. 14 (1987) 350–362.
[7] B.-H. Chu, M.-S. Tsai, C.-S. Ho, Toward a hybrid data mining model for customer retention, Knowl. Based Syst. 20 (2007) 703–718.
[8] K. Coussement, D. Van den Poel, Churn prediction in subscription services: an application of support vector machines while comparing two parameter-selection techniques, Expert Syst. Appl. 34 (2008) 313–327.
[9] K. Dasgupta, R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjea, A.A. Nanavati, A. Joshi, Social ties and their relevance to churn in mobile telecom networks, in: Proceedings of 11th International Conference on Extending Database Technology: Advances in Database Technology, ACM, New York, 2008, pp. 668–677.
[10] Z. Dezso, A.L. Barabasi, Halting viruses in scale-free networks, Phys. Rev. E 65 (2002), doi:10.1103/PhysRevE.65.055103.
[11] P. Domingos, Mining social networks for viral marketing, IEEE Intell. Syst. 20 (2005) 80–82.

[12] B. Galitsky, J. Rosa, B. Kovalerchuk, Discovering common outcomes of agents' communicative actions in various domains, Knowl. Based Syst. 24 (2011) 210–229.
[13] J. Goldenberg, S. Han, D.R. Lehmann, J.W. Hong, The role of hubs in the adoption process, J. Marketing 73 (2009) 1–13.
[14] M. Granovetter, Economic action and social structure: the problem of embeddeness, Am. J. Sociol. 91 (1985) 481–510.
[15] A. Gustafsson, M.D. Johnson, I. Roos, The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention, J. Marketing 69 (2005) 210–218.
[16] H.W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (2000) 599–653.
[17] S. Hill, F. Provost, C. Volinsky, Network-based marketing: identifying likely adopters via consumer networks, Stat. Sci. 21 (2006) 256–276.
[18] S.-Y. Hung, D.C. Yen, H.-Y. Wang, Applying data mining to telecom churn management, Expert Syst. Appl. 31 (2006) 515–524.
[19] C. Kaiser, S. Schlick, F. Bodendorf, Warning system for online market research – identifying critical situations in online opinion formation, Knowl. Based Syst. 24 (2011) 824–836.
[20] M. Kilduff, W. Tsai, Social Networks and Organizations, SAGE, London, 2003.
[21] H.-S. Kim, C.-H. Yoon, Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market, Telecommun. Policy 28 (2004) 751–765.
[22] M.-K. Kim, M.-C. Park, D.-H. Jeong, The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services, Telecommun. Policy 28 (2004) 145–159.
[23] Y.A. Kim, H.S. Song, Strategies for predicting local trust based on trust propagation in social networks, Knowl. Based Syst. 24 (2011) 1360–1371.
[24] Y.A. Kim, H.S. Song, S.H. Kim, Strategies for preventing defection based on the mean time to defection and their implementations on a self-organizing map, Expert Syst. 22 (2005) 265–278.
[25] Z. Kiyana, A. Aghaie, A syntactical approach for interpersonal trust prediction in social web applications: combining contextual and structural data, Knowl. Based Syst. 26 (2012) 93–102.
[26] D. Krackhardt, The ties that torture: Simmelian tie analysis in organizations, Res. Sociol. Organ. 16 (1999) 183–210.
[27] P.L. Krapivsky, S. Redner, Dynamics of majority rule in two-state interacting spin systems, Phys. Rev. Lett. 90 (2003), doi:10.1103/PhysRevLett.90.238701.
[28] R. Lambiotte, M. Ausloos, J.A. Holyst, Majority model on a network with communities, Phys. Rev. E 75 (2007), doi:10.1103/PhysRevE.75.030101.
[29] P.F. Lazarsfeld, R.K. Merton, Friendship as process: a substantive and methodological analysis, in: M. Berger, T.-D. Abel, C.H. Page (Eds.), Freedom and Control in Modern Society, D. Van Nostrand Company, Inc, New York, 1954, pp. 18–66.
[30] J. Lee, J. Lee, L. Feick, The impact of switching costs on the customer satisfaction-loyalty link: mobile phone service in France, J. Serv. Market. 15 (2001) 35–48.
[31] X. Liu, J. Zhang, W. Cai, Z. Tong, Information diffusion-based spatio-temporal risk analysis of grassland fire disaster in northern China, Knowl. Based Syst. 23 (2010) 53–60.
[32] M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: homophily in social networks, Annu. Rev. Sociol. 27 (2001) 415–444.
[33] M.C. Mozer, R. Wolniewicz, D.B. Grimes, E. Johnson, H. Kaushansky, Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry, IEEE Trans. Neural Networ. 11 (2000) 690–696.
[34] M.E.J. Newman, S. Forrest, J. Balthrop, Email networks and the spread of computer viruses, Phys. Rev. E 66 (2002), doi:10.1103/PhysRevE.66.035101.
[35] J.-P. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, M.A.d. de Menezes, K. Kaski, A.-L. Barabasi, J. Kertesz, Analysis of a large-scale weighted network of one-to-one human communication, New J. Phys. 9 (2007) 179–206.
[36] M. Richardson, P. Domingos, Mining knowledge-sharing sites for viral marketing, in: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, 2002, pp. 61–70.
[37] Y. Richter, E. Yom-Tov, N. Slonim, Predicting customer churn in mobile networks through analysis of social groups, in: Proceedings of SIAM International Conference on Data Mining, SIAM, Philadelphia, 2010, pp. 732–741.
[38] I. Roos, Switching processes in customer relationships, J. Serv. Res. 2 (1999) 68–85.
[39] I. Roos, Methods of investigating critical incidents: a comparative review, J. Serv. Res. 4 (2002) 193–204.
[40] V.P. Singh, K.T. Hansen, Market entry and consumer behavior: an investigation of a Wal-Mart supercenter, Market. Sci. 25 (2006) 457–476.
[41] G. Song, D. Yang, L. Wu, T. Wang, S. Tang, ;;A mixed process neural network and its application to churn prediction in mobile communications, in: Proceedings of IEEE International Conference on Data Mining, IEEE Computer Society, Washington DC, 2006, pp. 798–802.
[42] D. Van den Poel, B. Lariviere, Customer attrition analysis for financial services using proportional hazard models, Eur. J. Oper. Res. 157 (2004) 196–217.
[43] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[44] D. Wettschereck, D.W. Aha, T. Mohri, A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms, Artif. Intell. Rev. 11 (1997) 273–314.
[45] S. Yang, G.M. Allenby, Modeling interdependent consumer preferences, J. Marketing Res. 40 (2003) 282–294.