♦ Prediction of Subscriber Churn Using Social Network Analysis

Chitra Phadke, Huseyin Uzunalioglu, Veena B. Mendiratta, Dan Kushnir, and Derek Doran

In today's world, mobile phone penetration has reached a saturation point. As a result, subscriber churn has become an important issue for mobile operators as subscribers switch operators for a variety of reasons. Mobile operators typically employ churn prediction algorithms based on service usage metrics, network performance indicators, and traditional demographic information. A newly emerging technique is the use of social network analysis (SNA) to identify potential churners. Intuitively, a subscriber who is churning will have an impact on the churn propensity of his social circle. Call detail records are useful to understand the social connectivity of subscribers through call graphs but do not directly provide the strength of their relationship or have enough information to determine the diffusion of churn influence. In this paper, we present a way to address these challenges by developing a new churn prediction algorithm based on a social network analysis of the call graph. We provide a formulation that quantifies the strength of social ties between users based on multiple attributes and then apply an influence diffusion model over the call graph to determine the net accumulated influence from churners. We combine this influence and other social factors with more traditional metrics and apply machine-learning methods to compute the propensity to churn for individual users. We evaluate the performance of our algorithm over a real data set and quantify the benefit of using SNA in churn prediction. © 2013 Alcatel-Lucent.

Introduction

The mobile industry has seen tremendous growth in the last decade. Today's mobile users are more informed and savvy about the different handsets available and the services being offered via mobile devices such as voice, short message service (SMS), multimedia messaging service (MMS), data and video services. Mobile penetration is reaching high saturation levels [2, 17] in both developing and developed countries. It is well accepted that it costs more to acquire new customers than to retain existing ones and therefore service providers are looking at ways to reduce churn through innovative

Bell Labs Technical Journal 17(4), 63–76 (2013) © 2013 Alcatel-Lucent. Published by Wiley Periodicals, Inc.
Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/bltj.21575

Alcatel·Lucent

retention campaigns. The key to maximizing the value for a service provider is to be able to detect potential churners before they leave the service and target them exclusively for such campaigns.

Users churn for a variety of reasons. Satisfaction with the service quality is definitely important, but not the sole reason. A survey report by the Yankee Group [10] cites competitive pricing, network service quality, discounts, and promotions as some of the primary reasons a user may decide to leave a service provider. The influence of social factors in leaving a service provider, however, has not yet received significant consideration. It is natural to believe that when a person leaves a service, he also impacts the social circle around him with his actions. For example, his friends and family, who may be with the same service provider, may now not be able to enjoy free calls to him typically offered by some type of calling plan for family and friends and therefore may be incentivized to switch providers as well. Social pressures to adopt new technology may also encourage users to move to a provider that has the fastest data access or popular handsets. A famous example of this is when the iPhone* was first introduced in the AT&T network, and an exodus of subscribers from other service providers moving to AT&T.

Classic customer churn prediction and analysis techniques focus on the quality of experience for the mobile service user. Preliminary studies have shown, however, that members in the social circle of a subscriber also influence the subscriber to churn [3]. In this paper we examine how to integrate the social relationships among subscribers to enhance churn prediction. We propose a new methodology to quantify the strength of social ties between users, and then apply these tie-strengths to a novel influence propagation model across a mobile call graph. The total influence, representing the social pressures a user faces to churn, is then incorporated into an enhanced machine learning churn prediction algorithm. We apply our approach to data provided by a large service provider and demonstrate the utility of incorporating social network analysis (SNA) features for churn prediction.

Panel 1. Abbreviations, Acronyms, and Terms

CDR—Call detail record
CRM—Customer relationship management
MMS—Multimedia messaging service
SMS—Short message service
SNA—Social network analysis

This paper is outlined as follows. We start with a discussion of the general concepts of churn and social network analysis (SNA) and explain the differences between traditional SNA work and applicability in the telecom domain. We look at some of the related work that has been done in this area. Next, we outline our three-step model for using SNA to predict churn for telecom subscribers. Lastly, we present the results of the application of our model in a large telecom providers' network.

Churn and Social Network Analysis Over Mobile Call Graphs

Telecom providers generate a call detail record (CDR) for every call made to or from one of its subscribers' mobile phones. A CDR includes details of the call including caller and callee identifiers, call date, call time, call duration, and the location of the cell towers for call initiation and termination among other details. The CDRs are typically used by billing systems to determine the call tariff for billing purposes. However, a CDR also provides the information needed to construct a call graph of the service providers' user base. In such a mobile call graph, the nodes represent subscribers of the service provider, and a link between two nodes indicates that one or more calls were placed between the two subscribers. In our paper, the term mobile call graph or call graph implies such a graph constructed from the CDRs. We use the term user and subscriber interchangeably to denote the nodes of the mobile call graph. The social circle of a user can be viewed as those nodes that are reachable from the user, via links in the call graph.

Churn is the phenomenon where users switch from one service provider to another. Users can also churn implicitly if they stop using a service for a long period of time such that the probability that they resume using the service in the future is very low. At the point of leaving the service, the service provider may offer incentives to the user with the hope of retaining him. But often that moment is too late since the user has already made a decision to churn. A potential churner may exhibit telltale signs such as reduced calls and other service usage over time or reduction in the frequency and amount of top-ups in the case of pre-paid users. By monitoring such behaviors it may be possible to detect users who are likely to churn. If these potential churners can be proactively contacted with a targeted campaign, the service provider improves the likelihood of retaining the user.

Humans, being social creatures, are influenced by the actions of others in society. We pay close attention to the actions of our immediate social circle and seek to follow them when it suits our interests. As a result, when someone leaves a service provider, it raises the prospect that others in his social circle may also follow. Therefore, analyzing and understanding a phenomenon like churn in a social context is a key factor in preventing many users from churning together in a service provider's network.

Social network analysis (SNA) involves the study of the behavior of people in a large social group or network. Traditionally, in online social networks such as Facebook* or Twitter*, group membership defines the reason why people in that group are bound together. For example, the members of "Cyclers International" could be cycling enthusiasts, cycle-gear retailers, or sports reporters. In such networks, one has purview into the detailed nature of the communications among the different members of the group. By mining the communication patterns among members of these groups, one can quickly establish insights into who might be considered an influencer and who is a follower within the group, as well as the hierarchy of social relationships. This analysis can be done by analyzing the structure of the graph and applying the concepts of degree, closeness, and between-ness centrality [11]. Telecom networks, however, are different. Social relationships are not explicit; they have to be discovered using the information available within the call detail record (CDR).

In this paper we address the following two questions: First, given call detail records, is it possible to realize and quantify the social connection or strength of ties between subscribers? Second, once social connections are inferred, can they enhance our predictive power for churn events? In both cases, the answer is *yes*, since we show that CDRs can be used to infer subscriber link attributes that can be combined together into a novel measure of social tie-strength. We then formulate a diffusion process over the call graph, whose links (edges) are weighted according to the previously computed social strength. We show that our diffusion process yields more accurate propensities of subscribers to churn, in particular, when coupled with another algorithm that uses only individual subscriber attributes.

Related Work

Churn prediction is an important area of focus for data analytics for telecom providers. This is evident in the fact that in 2009 the ACM Conference on Knowledge Discovery and Datamining (KDD) hosted a competition on predicting mobile network churn using a large dataset posted by Orange Labs [13]. The TreeNet* prediction model from Salford Systems [16] won the Gold Prize for highest predictive accuracy. A study by Moser et al. [12] compared various machine learning and classification algorithms for churn prediction. The focus of this research was to study each subscriber in isolation to analyze their usage, billing, and service data [6]. Others [7] have used complaint data to predict churn. A study by Nanavati et al. [3] showed that as the number of neighbors who churn increases, the probability of churn for an individual also increases, thereby providing evidence that social relationships play an influential role in affecting churn in an operator network. Richter et al. [14] took a different approach to apply social factors in churn prediction by identifying social groups based on call records and then modeling churn prediction separately for each group. Studies have shown that social relationships influence user behaviors with respect to buying patterns [18]. Social analysis is also used in other areas such as predicting human mobility, in urban planning, and in predicting purchasing behavior [1, 9].

Unlike these previous studies, our work focuses on a hybrid model where we exploit the strength of social relationships to define how churn influence is propagated in social networks and use that to predict the churn using machine-learning methods. Instead of using a single link attribute, we compute social strengths by synthesizing many link attributes at once using a novel tie-strength quantification formula. Furthermore, we employ a diffusion model for churn influence to enhance our prediction accuracy. To the best of our knowledge, such a methodology has not been previously used for churn prediction.

SNA: Enhanced Churn Prediction Methodology

The methodology behind our SNA-enhanced churn prediction model can be summarized by the following three steps:

- *Quantification of tie-strength.* This involves construction of a call graph and quantifying the strength of social ties among different users based on multiple attributes.
- *Influence propagation model.* A model for churner influence propagation in a call graph is defined and the overall influence that is accumulated at the nodes is computed.
- *Application of machine learning techniques to combine traditional and social predictors.* Service performance and usage metrics, billing, customer relationship management (CRM) data like customer support call data and demographic information are combined with socially relevant predictors and social influence and this aggregated user information is fed into a classification algorithm to predict future churners.

The details of each of these steps are described in the subsequent subsections.

Quantification of Tie Strength

Individual calling features, such as the frequency of calls between two users or total monthly call time, are not good indicators of social tie strength. For example, if a strong social relationship exists between two users, we expect that an attribute such as the number of calls made between them is high. Alone, however, one cannot conclude that a high number of calls between two users suggests the presence of a strong social connection (consider, for example, a customer calling their bank frequently for financial transactions). Therefore, in order to provide a more accurate view of the social ties between users, we consider a collection of calling attributes to compute tie strength.

It is important to note that calling features are associated with the direct interactions between users and are *not* associated with any global or structural attributes of the mobile call graph. This distinction is important because, while the value of calling attributes is affected by social ties, global and structural attributes of the call graph are not. Thus, our tie-strength algorithm and churn prediction model does not rely on structural graph properties such as node centrality, closeness, or between-ness. However, within our formulation we use the concept of degree-centrality to normalize the link attributes to reflect this property in the diffusion process described further below.

The attributes of a connection between two users in a mobile call graph are typically correlated and are measured using different scales. This makes it difficult to be able to combine them in a meaningful way. To overcome the scaling difference, we normalize each observation of attribute x_i by dividing it by $|x_i|$, where $|x_i| = \sqrt{\sum_{k=1}^{d} x_{ik}^2}$. This operation rescales each attribute to a single unit length. We associate n attributes with each edge and define a real positive function $w(x)$ that computes the social tie strength. Let $x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$ be the weighted sum of these normalized calling attributes. The values of the weights α_1 and α_2 are tuned empirically using a training data set that maximizes the lift computed for detecting churners using accumulated influence. We use the performance measure of first decile lift to evaluate our results. Lift, in general, is defined as: For any fraction $0 < T < 1$, the ratio between the number of churners among the fraction of T subscribers that are ranked highest by the proposed system, and the expected number of churners in a random sample from the general subscriber pool of equal size. The influence diffusion process will be defined in the section following. The social tie-strength $w(x)$ is defined to be a monotonically increasing function of x given by $w(x) = 1 - \exp(-x/\varepsilon^2)$.

Clearly, $w(x)$ is restricted to the interval [0,1], with the parameter ε controlling the rate of saturation. Our exponential formulation is based on the assumption that once a strong social connection is manifested there is a high probability that an "idea" (such as churning) will be transferred through the edge from one subscriber to another. The value of ε, like the weights, is learned from the training data set such that the lift is maximized.

Influence Propagation Model

The quantification of tie-strength, as outlined in the previous section, helps understand which users are closely connected to each other and therefore more likely to be influenced by each other's behaviors. We use this information to define an influence propagation model where we quantify how influence travels from a churner to his social circle and how much influence is retained by a recipient. There are many models for information flow in a network. We consider a receiver-centric model for influence propagation wherein the receiver of the influence decides how much of the influence to retain. If the receiver is a close friend of the sender, reflected in a higher tie-strength than some other neighbors, then, it is natural to assume that he will be influenced by the senders' actions more than a random friend or colleague he may have social relationships with. Thus, the total amount of retained influence in our model is relative to the social tie-strength that the receiver has with the sender, in relation to the tie-strengths with all his neighbors. We believe that this is a closer representation of real-world social dynamics in comparison to sender-centric models where the sender decides how much influence to pass on to each of his neighbors.

The details of our proposed influence-propagation model can be described using the following principles:

1. A churner initiates the spread of influence to his social circle immediately following his decision to churn.
2. As soon as a user receives any amount of influence, he replicates that influence and will pass the same amount of influence to all of his neighbors.
3. A user who receives influence will decide what proportion of influence to retain. The influence retained is proportional to the tie-strength of the link on which the influence is received to the sum of the tie-strengths of all his incoming links.
4. A user will not be influenced about the same churn event from the same user twice.
5. A user, however, can be influenced by many churn events originating at different nodes.
6. Influence will propagate to a maximum of three hops, beyond which the influence to churn is not meaningful.
7. The total churn influence accumulated at a node is the sum of influences received at that node, across all churn events.

These principals are encoded in the following mathematical definition. Consider a call graph G that is constructed from the CDRs.

Let n_1, n_2, n_3, n_4 be the nodes of the graph.

Consider a particular node n_i. Let t_{ij} be the tie-strength between nodes i and j. Since we consider an undirected symmetric graph, tie strength is undirectional, i.e., $t_{ij} = t_{ji}$.

Let Ni be the set of all node neighbors of node i. Then $T_i = \sum_{j \in Ni} t_{ij}$ is the sum of the tie-strengths of all links incident on node i.

Then I_{ij}, the influence received by node i from node j is proportional to the tie-strength of the link between node i and node j to the sum of the tie-strengths of all links incident on node j given by $I_{ij} = \frac{t_{ij}}{T_i} I_j$.

The total influence received by node i, I_i is the sum of the influences received from all its neighbors and given by $I_i = \sum_{j \in Ni} I_{ij}$.

We demonstrate our model with a simplified call graph as shown in **Figure 1**. Here node 1 is a churner node and the tie strengths of the various edges are marked. In this illustrative example the churner has an initial influence I. For simplicity, we also limit the maximum propagation of influence to two hops.

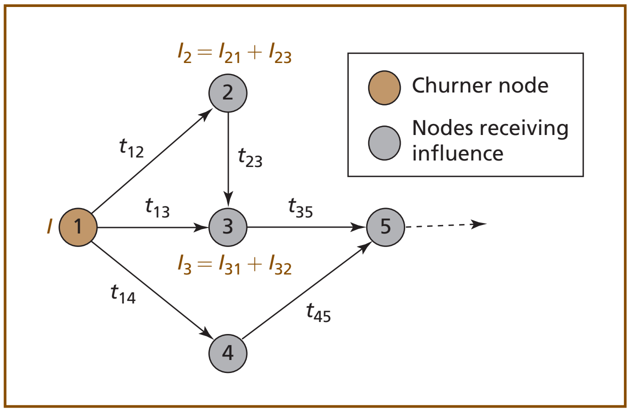


Figure 1.
Influence diffusion principles illustrated.

When influence propagation begins, the churner node 1 will send all of its influence to all its neighbor nodes 2, 3, and 4. Each of the receiving nodes will retain only a proportion of the influence received based on the relative tie-strength with the churner node to all of its incident edges. The proportion of influence that node 3 receives from node 1 is $\frac{t_{13}}{t_{13} + t_{23} + t_{35}}$. Thus node 3 receives influence $I_{31} = I * \frac{t_{13}}{t_{13} + t_{23} + t_{35}}$.

Nodes 2 and 4 will similarly receive a proportion of the influence passed by churner node 1. After nodes 2, 3, and 4 decide on the total influence they retain, they will then propagate this influence to their social ties. During this second iteration, node 3 will again receive influence about the churn of node 1 indirectly thorough node 2 via I_{32}. The total influence at node 3 is then $I_3 = I_{31} + I_{32}$. Note that node 3 will receive the influence of churn of node 1 in two ways—directly through node 1 and indirectly via node 2. However, the amount of influence it receives from both of these nodes is different because of the relative tie-strengths on each of the links. The receiving neighbors will again retain a portion of the influence and then pass it on. The propagation of influence continues until it reaches the maximum pre-configured number of hops. Similarly, node 5 will be influenced by a churn event at node 1 via nodes 3 and 4.

The influence propagation phenomenon will be repeated for every churn event that occurred in the call-graph. At the end of the diffusion processes, each of the nodes in the call graph will have a net amount of influence gathered due to all the churn events. This net influence to the node is the influence metric that will be used as one of the predictors for churn.

Classification for Churn Prediction

The final step to predict churn using SNA is to incorporate the influence events above into a traditional machine learning algorithm to predict customer churn. However, influence is only one of the social metrics used as a predictor. We also consider other social metrics such as number of neighbors who are churners, total calls to churners, and the distance (in hops) to the nearest churner. All of these metrics are computed from the call graph constructed using the CDRs. Churn prediction is a classification problem with a binary target variable representing churn and no churn events. In our approach, we focus on predicting the probability to churn rather than just labeling the subscriber as a potential churner or non-churner, so that subscribers can be ordered from high to low churn likelihood. With such ordering, a retention campaign can target a limited number of subscribers with the highest likelihood to churn.

A number of classification algorithms such as logistic regression, decision trees, and random forests [4, 8] can be used for binary classification problems. Such supervised learning algorithms start with a training data set, where the value of the target variable is known for the individual cases. The algorithms use the training data to learn the relationship between the predictor variables and the target variable to create a prediction model. Then, this learning model is used to score new samples where the target variable is not known. Our predictor variables include traditional subscriber-level metrics such as service usage, billing, CRM data such as calls to customer support, outcome of complaints, demographic data, and a number of SNA-based metrics such as social influence and the number of hops from a churner.

The key steps for the successful application of a particular classifier are pre-processing, feature selection, and the tuning of the model parameters.

Important pre-processing steps include imputing the missing values, removing correlated variables, and removing variables that do not provide any useful insight towards predicting the target variable. Model parameters are specific to the particular classifier being used whether that is a logistic regression or an ensemble of decision trees.

Ensemble decision tree classification has been shown to predict churn with high accuracy [8]. This classification approach adapts well to the churn prediction problem because decision trees are particularly suited to handle mixed variable types, missing values, and the large numbers of samples in the churn prediction training dataset. The ensemble decision tree algorithm adds decision trees to the prediction model in an iterative fashion so that mislabeled classifications from previous iterations are corrected. For our implementation, we use an ensemble decision tree method referred to as stochastic gradient boosting [5]. Stochastic gradient boosting iterates by building regression trees that are able to fit to the prediction error of the existing ensemble at each step. Further details of stochastic gradient boosting can be found in [5].

Experimental Results

To test our SNA model for churn prediction, we applied our approach to a dataset provided by a large telecom service provider. The data consists of subscriber contract and churn information for over half a million subscribers, and the CDR data includes all calls made over a period of two months. The CDR and usage information also contained information on whether the callee was with the same service provider (referred to as in-network) or with another service provider (referred as off-network).

Most service providers have incentives for calling people within their network such as free calls or reduced rates to incentivize users' friends to join their network. The ratio of a user's calls within the network and outside the network has a high social connotation. If a user is making the majority of his calls outside the network, it is very likely that he will consider switching to a different provider in the future, to save costs. Thus, in addition to the usage predictors such as mean number and minutes of calls, and range of calls, we also created additional variables reflecting the ratio of usage and bills that were to off-network users versus in-network users. Additional socially relevant predictors such as hops to the nearest churner, number of directly connected churners, and number of calls to churners were also computed. The data set was limited in that it did not contain any location or demographic information or any information about customer service complaints and dropped calls. It was also limited to voice calls only so usage data about additional services such as SMS was not available to be used as a predictor in our model.

An undirected call graph using two months of data was created and, using our proposed algorithm, the social tie-strength was computed for all of the links. The rationale behind using an undirected graph is that irrespective of the direction in which the call was initiated, once a communication link is established, information will flow in both directions. Calling attributes were normalized so as not to penalize users who may have churned during that period. To compute the social tie-strength we used three attributes viz., 1) the number of calls placed between two users, 2) the total duration of calls between two users, and 3) the neighborhood overlap of the two connected users. The neighborhood overlap measures the proportion of neighbors that two users have in common with each other. Typically, the higher the number of common friends, the greater the probability that two users are socially well connected. **Figure 2** demonstrates the behavior of tie-strength as a function of the attributes for $\varepsilon = 1$. We used equal unit weights in our formulation for tie-strength. The variability that is observed for each attribute is a result of the other two attributes that are used to determine the tie-strength. Specifically, the attributes "call time" and "number of calls" are highly correlated and their similar behavior is dominant in determining the tie-strength. The "neighborhood overlap" is another important factor but it is not correlated with the other two attributes. Therefore the tie-strength exhibits more variability as a function of the "neighborhood overlap."

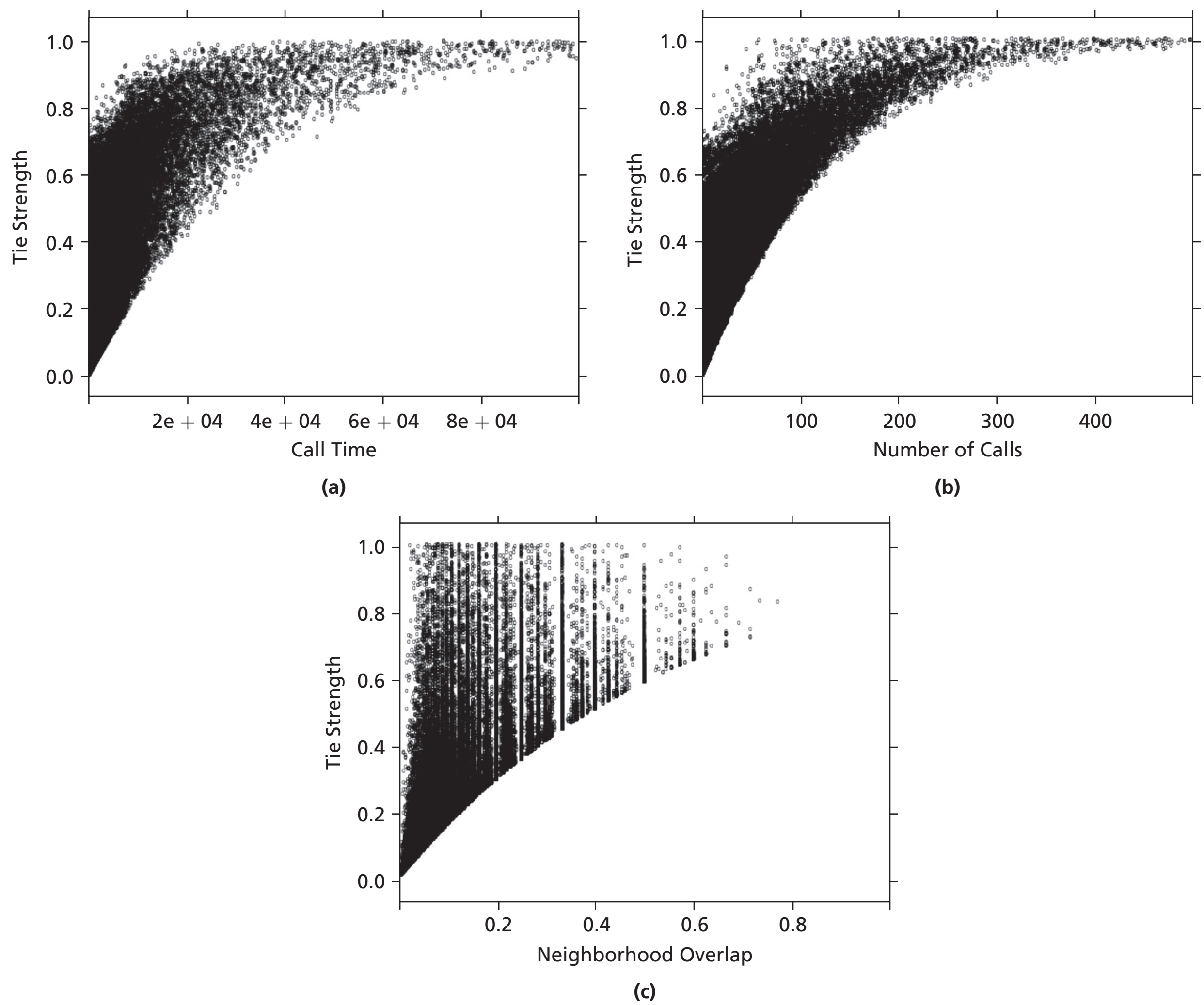
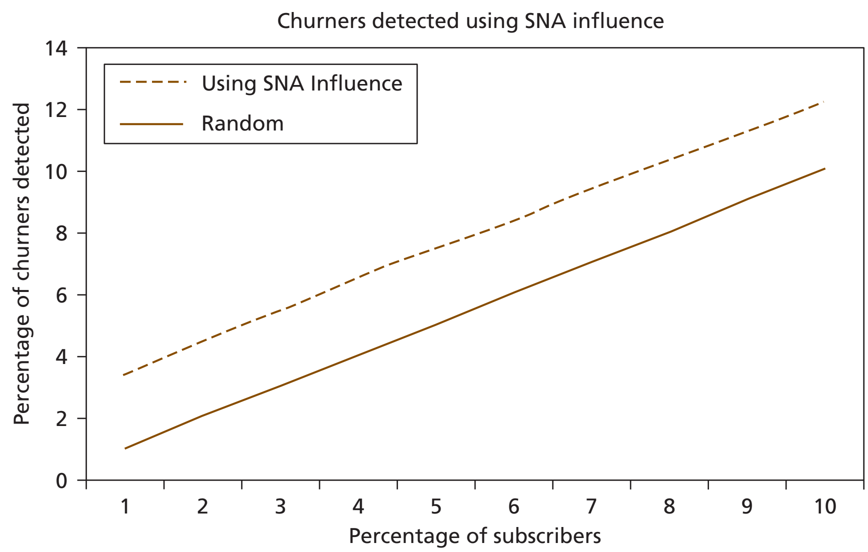


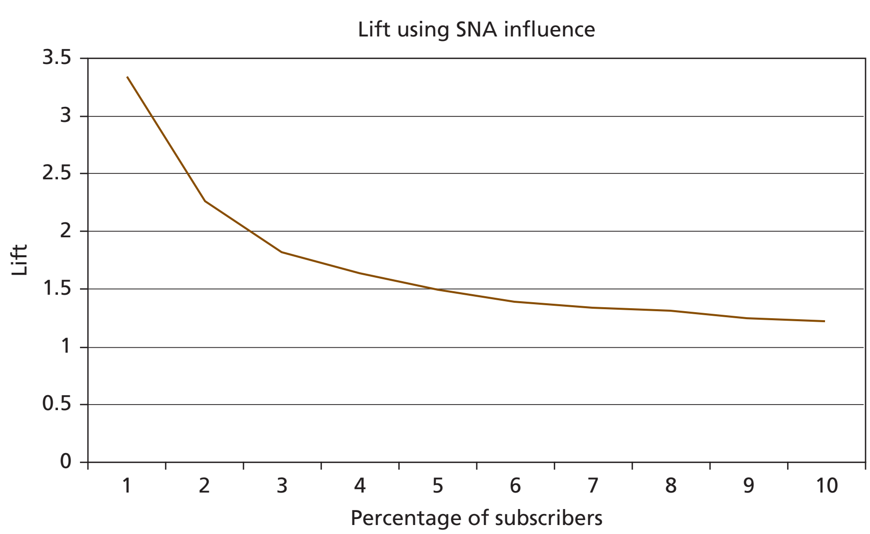
Figure 2.
Relationship between social tie strength and selected attributes.

Churners from the months in which CDRs were collected were marked on the call graph. These churners set influence propagation processes into motion with the same amount of initial influence. We limited the influence propagation to a maximum of three hops. At the end of the propagation process, each node accumulated a certain net churn influence. The users were then rank-ordered by accumulated influence and compared to determine how many future churners could be correctly identified based on the ranking of accumulated influence alone. **Figure 3** shows a comparison of this method versus one in which churners were randomly selected. Note that the X-axis shows the percentage of the total population and the Y-axis shows the percentage of churners that were correctly identified. Thus if you selected one percent of the population, in a random selection, you would detect only one percent of the churners. However, using the accumulated influence measure, one can detect 3.3 percent of the churners. This gives the top percentile lift of $3.3/1 = 3.3$. It is evident that the influence propagation process based



SNA—Social network analysis

Figure 3.
Comparison of churners using influence alone: random versus SNA shows that SNA-influence is better able to predict potential churners than a random selection.



SNA— Social network analysis

Figure 4.
Lift plot for churn prediction based on accumulated influence at nodes alone.

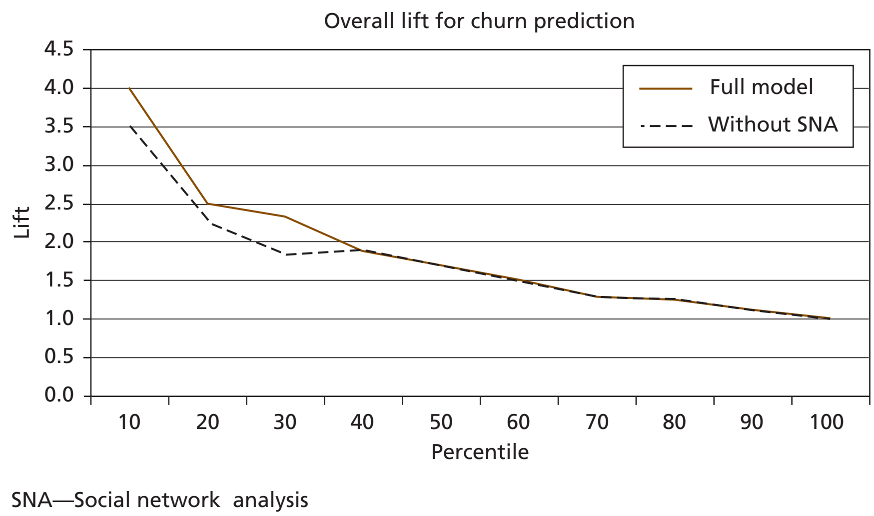


Figure 5.
Overall lift comparison with and without using SNA predictors.

Table I. Important predictor variables.

Variable name	Description
ContractEndIndicator	Indicator representing if a contract has ended, will end, or if there was no contract, in terms of number of days (negative if already ended)
DaysConnected	Number of days since the start of service
Influence	Accumulated influence from churners as outcome of influence diffusion model
NumCallsToChurners	Total number of calls to churners made over past three months
CallsRange	Range of calls made in the past three months
BillchargeAmtRange	Range of bill/charge amount over the past three months
MouPeakMeanOff	Minutes of use (MOU) during peak period that were made to off-net callers
OneMinMeanOff	Mean number of one minute calls made that were made to off-net callers
MouOnPct	Percentage of minutes of use that were made to on-net callers
OneMinRange	Range of calls under one minute over thee months

on our tie-strength algorithm is able to identify many more churners than a random selection, suggesting that the accumulated churn influence is a good predictor of the likelihood of the person churning. **Figure 4** shows a lift curve for churn prediction using accumulated SNA influence alone over the first decile of the population.

Lastly, we used a boosting model [15] that combined the various predictors, both social as well as usage-based, to create a churn prediction model that scores the users. Given the limited period for which CDRs were available, we used the same data set for training as well as scoring. The churn rate in the sample was very low (<1 percent), so various methods such as upsampling and appropriate weight selection were applied to overcome the issue of a limited and small target set. As a result, we observed a large variation in lift performance based on how the data was partitioned for training, testing, and validation. The model was run for numerous data partitions and the average lift for the first decile was computed. Using the full model which included all the predictors, we observed an average first decile lift of 4. **Figure 5** shows a comparison of the overall lift for the model run with and without using the SNA predictor variables. Clearly, in predicting churners in the top decile, the full model, which incorporates the SNA predictors fares much better than one which does not take social predictors into account. This illustrates the benefit of using social influence analysis in churn prediction. Attributes of the top predictive variables are listed in **Table I**. We observed that the contract end-date remains one of the top variables for predicting churn. Given that this data is for post-paid customers, who are typically locked in a contract, it is no surprise. However, many of the socially relevant predictors such as accumulated influence and number of calls to churners are also some of the top variables, which suggests that social predictors play an important role in churn prediction.

Conclusions and Future Work

This paper presented a methodology to predict customer churn in telecommunication services that integrates SNA concepts with traditional churn prediction methods. We demonstrated that the tie-strength and influence propagation algorithm we developed can be integrated into machine learning-based churn prediction models to improve their accuracy. Our approach is generic, and is applicable to any phenomenon that has influence diffusion, such as up-sell and cross-sell of services and application downloads. Future research seeks to enhance the tie-strength and information diffusion model to detect social influencers for better targeting of new services and applications. Various extensions can also be made to the information diffusion principles—such as decay of influence with the number of hops, and with time and consideration of the directionality of information propagation with asymmetrical tie-strengths. Finally, social media analytics that involve text mining of social web sites can be employed to learn about users' sentiment toward their telecommunication service provider. Churn prediction may be further enhanced by linking user identity in the social media to subscriber identity in the telecom domain.

Acknowledgements

The authors would like thank anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

*Trademarks

Facebook is a trademark of Facebook, Inc.
iPhone is a registered trademark of Apple Inc.
TreeNet is a registered trademark of Salford Systems.
Twitter is a registered trademark of Twitter, Inc.

References

[1] R. A. Becker, R. Cáceres, K. Hanson, J. M. Loh, S. Urbanek, A. Varshavsky, and C. Volinsky, "Using Cellular Network Data for Urban Planning," Proc. 2nd Conf. on the Anal. of Mobile Phone Datasets and Networks (NetMob '11) (Cambridge, MA, 2011), pp. 50–52.

[2] A. Chau, Asia/Pacific (Excluding Japan) Mobile Services Year-End Review, 2009, IDC, Doc. No. AP3053401S, vol. 1, May 2010.

[3] K. Dasgupta, R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjea, A. A. Nanavati, and A. Joshi, "Social Ties and Their Relevance to Churn in Mobile Telecom Networks," Proc. 11th Internat. Conf. on Extending Database Technol. (EDBT '08) (Nantes, Fra., 2008), pp. 668–677.

[4] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, 2nd ed., John Wiley and Sons, New York, 2001.

[5] J. H. Friedman, "Stochastic Gradient Boosting," J. Comput. Statistics and Data Anal., 38:4 (2002), 367–378.

[6] R. K. Gopal and S. K Meher, "Customer Churn Time Prediction in Mobile Telecommunication Industry Using Ordinal Regression," Proc. 12th Pacific-Asia Conf. on Advances in Knowledge

Discovery and Data Mining (PAKDD ’08) (Osaka, Jpn., 2008), LNCS vol. 5012, pp. 884–889.

[7] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, “Churn Prediction Using Complaints Data,” World Acad. Sci., Eng., Technol., 13:2006 (2006), 158–163.

[8] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, New York, 2009.

[9] S. Isaacman, R. Becker, R. Cáceres, S. Kobourov, M. Martonosi, J. Rowland, and A. Varshavsky, “Identifying Important Places in People’s Lives from Cellular Network Data,” Proc. 2nd Conf. on the Anal. of Mobile Phone Datasets and Networks (NetMob ’11) (Cambridge, MA, 2011), pp. 19–21.

[10] D. Lonergan, Improve Communication to Minimize Churn, Yankee Group, Jan. 2012.

[11] A. Marin and B. Wellman, “Social Network Analysis: An Introduction,” The SAGE Handbook of Social Network Analysis (J. Scott and P. J. Carrington, eds.), Sage, London, 2011, pp. 11–25.

[12] M. C. Mozer, R. Wolniewicz, D. B. Grimes, E. Johnson, and H. Kaushansky, “Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunications Industry,” IEEE Trans. Neural Networks, 11:3 (2000), 690–696.

[13] Orange, “KDD Cup 2009,” <http://www.kddcup-orange.com/>.

[14] Y. Richter, E. Yom-Tov, and N. Slonim, “Predicting Customer Churn in Mobile Networks Through Analysis of Social Groups,” Proc. SIAM Internat. Conf. on Data Mining (SDM ’10) (Columbus, OH, 2010), pp. 732–741.

[15] G. Ridgeway, “Generalized Boosted Models: A Guide to the gbm Package,” Aug. 3, 2007, <http://cran.r-project.org/web/packages/gbm/vignettes/gbm.pdf>.

[16] Salford Systems, “TreeNet,” <http://www.salford-systems.com/>.

[17] Strategy Analytics, Wireless Operator Performance Benchmarking Q3 2009, Dec. 2009.

[18] J. Yang, X. He, and H. Lee, “Social Reference Group Influence on Mobile Phone Purchasing Behaviour: A Cross-Nation Comparative Study,” Internat. J. Mobile Commun., 5:3 (2007), 319–338.

(Manuscript approved September 2012)

CHITRA PHADKE is a member of the Network Performance and Reliability Department at Bell Labs in Murray Hill, New Jersey. Her interests include application enablement, mobile applications and services, and data analytics. She was a finalist in the Americas Entrepreneurial Boot Camp. She has led the development of the Geographic Messaging Service Platform (GMSP), a location based application platform that transitioned from a Bell Labs Venture to a product of the Alcatel-Lucent Applications business group. She has also been a key team member in the development of several network planning and optimization tools. She is currently a member of the Alcatel-Lucent Technical Academy and recipient of the two Bell Labs President’s Gold Awards. She holds a M.Sc. in physics from the Indian Institute of Technology and an M.S. in electrical and computer engineering from the University of Texas at Austin.

HUSEYIN UZUNALIOGLU is the director of the Bell Labs Network Performance and Reliability Department in Murray Hill, New Jersey. In his 14 year tenure with the company, he has contributed to a wide variety of projects in diverse subject areas including packet network performance modeling, Internet Protocol (IP) quality of service (QoS), call admission control, overload control, network design and routing, modeling of telecom infrastructure resiliency, and wireless network capacity planning. His work in the call admission control area influenced the development of the resource and admission control function (RACF) in the ITU. Currently, he is leading a team that is developing predictive analytics algorithms for customer experience solutions and proactive performance management. He received a B.S. in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, and a Ph.D. in electrical and computer engineering from Georgia Institute of Technology, Atlanta. He is a senior member of the IEEE and a recipient of the Bell Labs President’s Gold Award. He is also a co-author of a paper that received the Best Paper Award in the 5th International Workshop on Design of Reliable Communication Networks (DRCN), 2005.

VEENA B. MENDIRATTA leads the next-generation

solutions, services and systems reliability work in the Bell Labs Network Performance and Reliability Department in Naperville, Illinois. She began her career at AT&T Bell Labs over 25 years ago. Her work has focused on reliability and performance analysis for telecommunications systems products, networks, and services to guide system architecture solutions. Her technical interests include architecture, system and network dependability analysis, software reliability engineering, and telecom analytics. Current work is focused on predictive analytics for customer experience solutions, Long Term Evolution (LTE) solution reliability engineering, and service reliability modeling for mission critical networks for various domains. She holds a B.Tech in engineering from the Indian Institute of Technology, New Delhi, India, and a Ph.D. in operations research from Northwestern University, Evanston, Illinois. Dr. Mendiratta is a senior member of IEEE; a member of INFORMS; a member of the Alcatel-Lucent Technical Academy; and a recipient of the Bell Labs President's Award.

DAN KUSHNIR is a researcher in the Bell Labs Network

Performance and Reliability Department in Murray Hill, New Jersey. He received his B.Sc. degree in computer science from Hebrew University, Jerusalem, and his the M.Sc. and Ph.D. degrees both from the Department of Computer Science and Applied Mathematics at the Weizmann Institute of Science, Rehovot, Israel. Prior to joining Bell Labs, he held the J. W. Gibbs Assistant Professor position in the Mathematics Department at Yale University, New Haven, Connecticut. His current research interests include data mining, machine learning, applied harmonic analysis, numerical analysis, and signal processing. He developed efficient spectral algorithms for data analysis and focused on the development of fast methods for massive numerical computations in machine learning and data analysis.

DEREK DORAN is a Ph.D. candidate in the Department

of Computer Science and Engineering at the University of Connecticut. His research interests include web mining, social network analysis, and the performance and reliability modeling of distributed systems. He was a summer intern in the Bell Labs Network Performance and Reliability Department, investigating social-based churn prediction techniques across mobile calling networks, when the research detailed in this paper was conducted. Prior to his internship at Bell Labs, Mr. Doran spent a summer at National Taiwan University as an NSF EAPSI Fellow where he developed performance models for novel Long Term Evolution (LTE)-femtocell interference avoidance schemes. ◆