Combining Customer Attribute and Social Network Mining for Prepaid Mobile Churn Prediction

Palupi D. Kusuma PKUSUMA@LIACS.NL
Dejan Radosavljevik DRADOSAV@LIACS.NL
Frank W. Takes FTAKES@LIACS.NL
Peter van der Putten PUTTEN@LIACS.NL
LIACS, Leiden University, P.O. Box 9512, 2300 RA Leiden, the Netherlands

Keywords: churn, call graph, social network mining, spreading activation

Abstract

Customer churn, i.e., losing a customer to the competition, is a major problem in mobile telecommunications. This paper investigates the added value of combining regular tabular data mining with social network mining, leveraging the graph formed by communications between customers. We extend classical tabular churn datasets with predictors derived from social network neighborhoods. We also extend traditional social network spreading activation models with information from classical tabular churn models. Experiments show that in the second approach the combination of tabular and social network mining improves results, but overall the traditional tabular churn models score best.

1. Introduction

Churn, which is defined as the loss of customers to another company, is a crucial problem in the telecommunication industry. As the telecom market has matured and opportunities for growth are limited, retaining existing customers has become a higher priority. In order to minimize the churn rate, mobile telecom players have to form defensive strategies to identify and present the appropriate incentive to subscribers with high churn propensity.

The conventional churn models that exploit traditional predictors, such as demographic information (e.g., age, gender or location), contractual details (e.g., package plan type, contract duration or price), usage facts (e.g., voice call duration, the frequency of sending text-messages) and/or other service-related information (e.g., number of interactions with customer service or number of dropped calls), are typically simple and have a good predictive accuracy (Ferreira et al., 2004; Hadden et al., 2006). However, the predictive accuracy of these models cannot be guaranteed if there is few customer data available, namely in the prepaid segment of the telecommunication industry.

This paper investigates the extent to which social network features derived from the graph formed by communications between customers can be exploited to improve churn prediction accuracy in the prepaid segment. Examples of such features include the number of neighbors of a customer and the number of interactions that a customer has with churned neighbors. This research study was conducted at one of the largest telecom providers in the Netherlands, and a dataset containing 700 million call records was used to assess the quality of the various techniques discussed throughout the paper.

We propose two novel models for churn prediction. The first is a hybrid tabular model, which combines both traditional predictors and social network features to predict churn, aiming to gain significant lift. Logistic Regression and the CHAID algorithm are utilized to derive the tabular models. These churn models, however, do not take into account the influential effect of an individual's decision to his/her social network. A recent work by Dasgupta et al. (2008) has been able to address this problem by constructing a churn model based on a traditional social network mining technique, i.e., spreading activation models. The model propagates the negative churn influence from one subscriber to another in a cascade manner. Besides build-

Appearing in *Proceedings of BENELEARN 2013*.

ing hybrid tabular churn models using a combination of the traditional predictors and the social network features, we also propose a second approach, which extends the traditional propagation model to include the output by traditional churn models.

The rest of the paper is organized as follows. Section 2 presents some related work within the field of churn prediction. Section 3 discusses the call graph and proposed algorithms. The research setup and the empirical models are introduced in Section 4. In Section 5, the experimental results and implications of all scenarios are presented. Finally, Section 6 summarizes the paper and presents some suggestions for future work.

2. Related Work

Churn has been widely analyzed not only in the telecommunication industry (Ferreira et al., 2004; Hadden et al., 2006; Radosavljevik et al., 2010), but also, among others, in the online gaming (Kawale et al., 2009) and banking (Prasad & Madhavi, 2012). Many machine learning techniques, such as decision trees, naive bayes, logistic regression, neural networks and genetic algorithms, are often used to build the tabular churn prediction models.

Ferreira et al. (2004) utilize contractual and demographic information of a Brazilian mobile telecommunication provider to build several postpaid churn models using neural networks, decision trees, genetic algorithms and hierarchical neuro-fuzzy systems. Besides evaluating the predictive power, they also assess the profitability value of those models, claiming that even the churn models with the worst performance are still able to save significant cost in the postpaid segment. Hadden et al. (2006) exploit provisions, complaints and repair interaction data to build the churn models. They claim that the regression tree model performs better than one with neural networks or logistic regression. However, there is no further information regarding the performance comparison between the complaints-based model and the benchmark model based on demographic and contractual variables.

Radosavljevik et al. (2010) investigate the extent to which Customer Experience Management (CEM) data could improve prepaid churn prediction. Several Key Performance Indicators (KPI) of service quality combined with other subscriber data are used to train the decision tree models. Since the CEM data is always available, the constraint on lacking demographic information on the prepaid subscribers could be eliminated. Although the CEM data is predictive, the empirical study shows that there is insignificant gain on this model performance compared to the benchmark.

Several social network studies have been conducted by utilizing mobile call graph data to examine the structure and evolution of social networks (Backstrom et al., 2006; Seshadri et al., 2008), the human mobility patterns (Gyan et al., 2012) and their social interactions (Dasgupta et al., 2008). Dasgupta et al. (2008) analyze the influential impact of the churned neighbors to their social circle by applying a spreading activation-based technique similar to trust metric computations (Ziegler & Lausen, 2004). Using call graph data, they are able to show that churn can be propagated through a social network. Although the study is limited to use social ties information only, reasonable predictive accuracy could still be achieved. The analysis identifies that the churn propensity of a subscriber correlates positively with the number of churned neighbors.

Kawale et al. (2009) conducted a similar study using social network data from a popular online gaming community. They propose a new twist to the existing churn propagation model proposed by Dasgupta et al. (2008) by combining the social influence and user engagement in the game. The user engagement property, which refers to the length of the playing session during the observation period, can be classified as an intrinsic variable. The research shows that the models trained using a combination of social factors and this user engagement property perform better than traditional propagation models. Using collective classification techniques, Oentaryo et al. (2012) are also able to demonstrate that the churn prediction accuracy could substantially be improved by utilizing the combination of traditional user profile and social features.

We apply similar ideas from the above mentioned works. A customer's decision to churn might not only depend on the social influences but also on how they perceive the products and services. On our initial observation, we found that the ratio of the immediate churned neighbors to the number of adjacent neighbors (degree) positively correlates to the churn behavior. When half of the neighbors have churned, the probability of a subscriber to churn is 2 times higher than the baseline churn rate. It implies to some extent that social behavior might have an impact on the subscribers' churning decision. It could be that the hybrid models, which exploit both traditional predictors and social relationships, could outperform the simple social network and the tabular churn model built exclusively using traditional predictors. However, the question is also whether it adds actionable value over existing data. We suspected there may have been some element of publication bias: positive results get published more often, thus easier to find than non-significant or negative results, at least for trending topics. Hence, we decided to evaluate the business value experimentally.

3. Methodology

A call graph can be derived from raw data of communications between customers. This graph, further discussed in Section 3.1, is essentially a social network which can be leveraged in two ways. Classical 'tabular' models are built on rectangular data sets, one row per customer with subscriber level information. This can be simply extended with attributes (columns) that contain information derived from the social network, as we will outline in Section 3.2. Likewise, a traditional approach to modeling social network dynamics is the spreading activation model, which can be used to model how customer behavior such as churn spreads over the network. Insights from traditional tabular models, more specifically churn scores, can be used to improve these classical social network models, a technique on which we will elaborate in Section 3.3.

3.1. The Call Graph

The *call graph* can be constructed from the Call Detail Records (CDRs) provided by the telecom provider. These CDRs contain detailed facts about mobile interactions, such as source phone number, destination phone number, the type of mobile communication, duration and a timestamp. This information is mapped to a directed social graph $G = (V, E)$ as illustrated in the Figure 1.

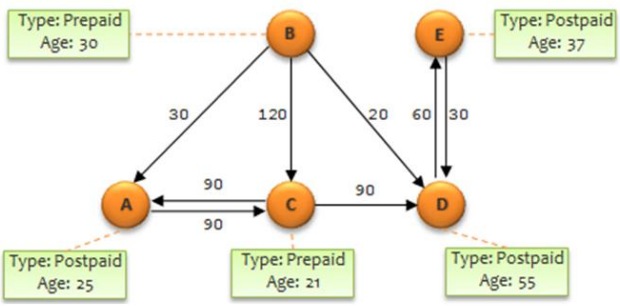

Figure 1. Telecom call graph.

In this call graph, *nodes* denote subscribers and an *edge* represents a mobile interaction between two subscribers. The *edge weight* can be calculated from one variable or a combination of interaction variables, e.g., voice call duration or SMS frequency. It could indicate the interaction intensity or the relationship strength between two nodes. As several interactions could exist between the same pair of nodes, we treat duplicate edges between two nodes as a single edge, by aggregating the weight values. The aggregation method applied in this research is explained in Section 3.3.

3.2. Extended Tabular Churn Models

Many tabular churn models generally exploit either subscriber profile information or social network statistics separately. The predictive power of churn models based merely on the traditional predictors might be reduced in case of many missing values. In our prepaid churn study, we only have an access to limited demographic data because prepaid subscribers are not required to fill in their (accurate) personal information. On the other hand, the social network features might not be predictive enough to influence the churn decision. Neither the traditional models nor the models based exclusively on social networks can cover all aspects of churn on their own. Therefore, we propose to combine both elements to predict churn, adding the features listed in Table 1.

Table 1. Social network features used in the extended tabular churn models.

CATEGORY	VARIABLE
CONNECTIVITY	Count of in/out-degree
	Sum & average of in-/out-weight
	Count & average of voice, SMS & voice+SMS to/from neighbors
	Total and average of edge weight*
	Total interaction frequency with neighbors*
	Total and average frequency with neighbors for voice & SMS separately*
	Degree, 2nd degree & 3rd degree count*
CHURNER CONNECTIVITY	Count of in/out-degree churners
	Sum & average of in/out-weight with churners
	Count & average of voice, SMS & voice+SMS to/from churners
	Total & average edge weight with churners*
	Total interaction frequency with churners*
	Ratio of in/out-degree churners to the total in/out-degree
	Ratio of in/out-weight churners to the total in/out-weight
	Ratio of in/out voice, SMS & voice+SMS frequency with churners to the total in/out-weight
	Ratio of churner weight to the total weight*
	Ratio of interaction frequency with churners to the total interaction frequency*
	Churner degree, 2nd & 3rd degree count*
	Ratio of churner degree to the total degree*
	Ratio of 2nd churner degree to the total 2nd degree*
	Ratio of 3rd churner degree to the total 3rd degree*

*direction is not taken into account

When creating the extended tabular churn models we started with a model based on traditional predictors and added connectivity features from the social network call graph: the in-degree and out-degree, the number of second degree neighbors, sum and average of in-weight and out-weight calculated from duration of voice conversations, SMS and a combination thereof. We also added churn connectivity variables (in-degree and out-degree with churners, etc.), as well as the ratios of the total connectivity measures vs. the churners connectivity measures. A detailed overview of the added social network graph features is presented in Table 1. For a more detailed feature analysis, we refer the reader to Kusuma (2013).

3.3. Extended Social Propagation Models

In this subsection, we discuss an extension of the spreading activation model to measure how churn is diffused around telecom social network (Dasgupta et al., 2008). The churn propagation process begins by initialization of all nodes. In this study, we set the energy of *non-churners* using two different values (see Figure 2). For the *simple propagation approach*, the initial energy of non-churners is set to 0; for the hybrid *extended approach*, it is set to the churn score returned from the regular tabular models.

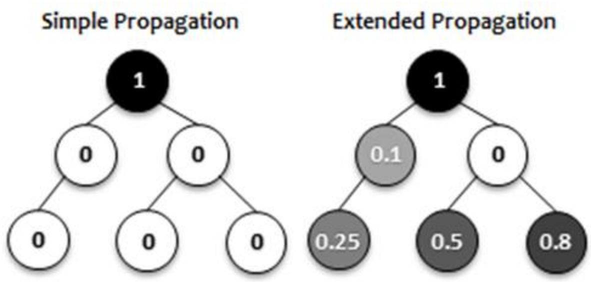

Figure 2. Initial energy of the simple and extended propagation technique.

In the propagation process, for a node $x \in V$, the value of $En(x)$ represents the current amount of energy of a node, and the $En(x, i)$ represents the amount of energy or social influence transmitted to the node x via one or more of its neighbors at stage i (Dasgupta et al., 2008). After energy initialization, a set of previous churners (seeds) is activated. In stage 0, the current energy of the seeds $En(x)$ is used as initial spreading value. Therefore, the current energy value $En(x)$ becomes 0 and amount of energy in a node x at step 0 or $En(x, 0)$ becomes equal to 1.

In each consecutive stage i, the activated nodes transfer a portion of their energy to their neighbors and retain certain portion for themselves. The spreading factor $\delta \in [0, 1]$ controls the proportion of the transmitted energy, denoted by $\delta * En(x, i)$ and the amount of retained energy $(1 - \delta) * En(x, i)$. A spreading factor value of $\delta = 0.8$ means that 80% of the energy is transferred to the neighboring nodes and 20% of the activated energy is retained by the node. This factor value could also be seen as a decay measure because the transferred energy will decline as it gets further away from the source. It implies that the direct neighbors will receive more influence than second degree neighbor and so on. The trust propagation study of Ziegler and Lausen (2004) has shown that people tend to trust individuals trusted by own friends more than individuals trusted only by friends of friends.

Since nodes can have multiple neighbors, the amount of the distributed energy from an active node to each neighbor depends on the tie strengths between the node pair. In Figure 3, for example, the amount of energy transferred from node 1 to node 2 might not be the same as the amount transferred from node 1 to node 3, because the edge weights are not equal.

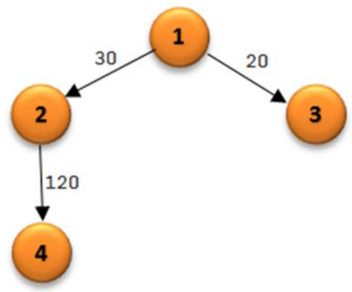

Figure 3. Spreading activation in a weighted graph.

Let y be a neighboring node of an active node x (with $x, y \in V$). We denote the amount of energy transferred from node x to node y in the i-th stage with $En(x, y, i)$. This amount depends on the relative edge weight of the paired nodes. This is determined by a transfer function $f(x, y)$, described in Equation 3 below. The amount of energy transferred is then:

$$En(x, y, i) = \delta * En(x, i) * f(x, y) \tag{1}$$

The amount of energy of node x after the spreading computation is as follows:

$$En(x) = En(x) + (1 - \delta) * En(x, i) \tag{2}$$

There are multiple functions to determine the relative weight between two nodes. The simplest method is using linear edge weight normalization function (Ziegler & Lausen, 2004).

$$f(x, y) = w(x, y) / \sum_{(x,z)} w(x, z) \tag{3}$$

Here, $f(x, y)$ denotes the relative weight of the edge between x and y, $w(x, y)$ represents the weight of that corresponding edge, and $\sum_{(x,z)} w(x, z)$ represents the total weight of all edges connecting node x to its adjacent nodes.

We propagated the churn energy through both a directed and an undirected version of the graph. In the directed graph, energy is propagated only to outgoing edges, and in the undirected graph, both outgoing and incoming edges are used. For churn propagation, the remaining energy after termination ultimately determines the probability of a network member to churn. These churn probability scores are then distributed into score intervals. The upper interval groups contain more subscribers with high churn propensity behavior compared to the lower interval groups. Using the threshold score-based technique, the subscribers/groups with churn scores above a predefined threshold score can each be labeled as a 'churner', and otherwise as a 'non-churners'. As an alternative, a cut-off point can also be determined by specifying the target group size.

4. Experimental Setup

This section describes different specific techniques and assumptions with respect to telecommunications data in Section 4.1, after which the dataset and weighting technique is discussed in Section 4.2. We then give an overview of the seven different scenarios that were used to construct the churn models, outlining our experimental setup in Section 4.3.

For our experiments, we use Chordiant Predictive Analytics Director software to automate variable discretization, variable selection and grouping, to train the scoring models and also to compare the models performance. The default evaluation statistic that is used to measure the performance of the predictors and models is Coefficient of Concordance (CoC). CoC measures the area under the Lorenz curve formed by the percentage of cases with positive behavior against the percentage of cases with negative behavior for each unique score (Harell, 2001).

4.1. Operational Definition of Churn

We constructed models for the prepaid and postpaid telecom segments. Although the definition of churn is different for each segment, we will only discuss the prepaid results because both studies have come to the similar conclusion. Unlike postpaid subscribers, prepaid subscribers are not bound by a contract, which makes it easier for them to churn. Prepaid subscribers need to purchase a credit voucher before using any telecom service. If they do not have sufficient voucher credit, they could not initiate any calls, send SMS/MMS or connect to internet. They could re-enable the service by recharging or topping-up their voucher credit.

A prepaid subscriber is disconnected from the network and he/she is marked as a churner after six consecutive months of inactivity. A prepaid activity could be translated to an outbound voice call, an inbound voice call, an outbound SMS, a data usage or a commercial voucher recharge, also known as top-up. As churn should be detected as early as possible, the disconnection date might not be the appropriate churn date measure (Kraljevic & Gotovac, 2010). The prepaid subscribers might be long gone before they are actually disconnected from the network. Therefore, we define churn as two consecutive months of inactivity. This definition is aligned with many internal studies that are conducted within the company.

4.2. Dataset

We use the CDRs from the whole month of February 2012, which is roughly about 700 million records, to construct the social graph. We include subscribers who have at least one call in February and we base our social network graph on the interactions that occurred in that month. The end goal is to use the traditional predictors as well as the social network information obtained in February, March and April 2012 to predict churn in June 2012. We assume that churn is also a social networking phenomenon, thus subscribers that communicate with people that have churned are more likely to churn themselves. Therefore, we label the nodes/subscribers that churned in the period before May 1, 2012 ('observation 1' in Figure 4) as seeds/churners of the propagation graph explained in Section 4.3. The churn we are trying to predict occurs between May and June 2012 ('observation 2' in Figure 4).

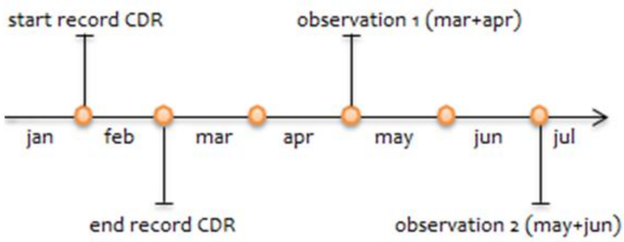

Figure 4. Call Graph Details.

In this research study, we only consider the duration of voice calls in minutes and the count of text messages. We could not explore mobile interactions utilizing the data connection, i.e., using over the top (OTT) services[1], due to legal issues. Within the company, the postpaid cost of making one minute of a voice call is the same as one SMS. In the prepaid segment, SMS is typically charged roughly half of a minute of voice call. Therefore, we furthermore assume that a text message is equivalent to a voice call of 30 seconds. Hence, we could generalize the edge weight $w(x, y, t)$ between a pair of nodes x and y at time t to include both types of mobile communication, voice calls and SMS and all interactions could all be measured uniformly in seconds. The identifier t represents the hourly timestamp at which the interaction starts, and is ranged from 1 until 29 February 2012.

$$w(x, y, t)' = w(x, y, t) * \begin{cases} 1, & \text{if voice call} \\ 30, & \text{if SMS} \end{cases} \tag{4}$$

Interactions that occurred outside working hours are assigned twice the weight to emphasize their importance. The underlying assumption here is that interactions within working hours mostly indicate communication of professional nature, whereas interactions outside working hours may involve communication of more personal nature (e.g., friends, family), which could have higher influence on the decision to churn. Motahari et al. (2012) shows that members of a family/friends social network are more likely to call each other on the weekend and the engagement ratio value within the family/friends network is at least twice as much compared to the rest of the population. Therefore, we introduce a weight scale $\rho(t)$, which is defined as follows:

$$\rho(t) = \begin{cases} 1, & \text{if t=weekdays (8-17)} \\ 2, & \text{otherwise} \end{cases} \tag{5}$$

$$w(x, y, t)'' = \rho(t) * w(x, y, t)' \tag{6}$$

We also assume that a recent interaction should carry more weight than older ones. Therefore, the daily decay rate $\alpha = 0.2$ is manually selected. The weight value of an edge that is measured on a certain day exponentially decayed according to a predefined rate as follows:

$$w(x, y, t)''' = w(x, y, t)'' * e^{-\alpha * d} \tag{7}$$

Here, the symbol d corresponds to the gap measured in days between the interaction timestamp and the end of the observation period. In our case, d is equal to 28 measured from 1 until 29 February 2012. In the end of the observation period, the weight values are aggregated. As a result, each node pair could only have maximum one edge in each direction, so two edges in total. The equation below formulates the aggregation process of the weight values.

$$w(x, y) = \sum w(x, y, t)''' \tag{8}$$

For an undirected graph, we could simply add up the weights for both directions together as follows:

$$w(x, y) = \sum w(x, y, t)''' + \sum w(y, x, t)''' \tag{9}$$

4.3. Churn Predictive Models

To investigate to which extent social network data could be used to predict churn and possibly could improve churn prediction performance, we trained three tabular data mining models using scoring algorithms and four social network models using a spreading activation algorithm (see Figure 5).

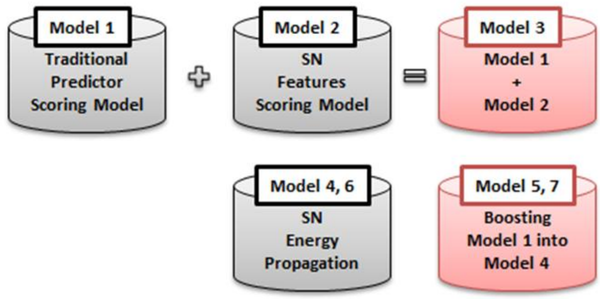

Figure 5. Implementation scenarios.

4.3.1. Scoring Models

We apply a logistic regression and a CHAID decision tree algorithm to train our three scoring models:

- Model 1: simple scoring model
- Model 2: social network (SN) scoring model
- Model 3: extended scoring model

Model 1, a simple scoring model, is trained using the traditional churn predictors, using features such as demographic, contractual, handset and usage information. We employ this model as the benchmark model. Model 2 is a social network scoring model, which focuses solely on the social network attributes extracted from call graph, such as the number of incoming and outgoing ties of the first and second degree neighbors. The extended scoring model or the second hybrid model, Model 3, combines the dataset of the first

[1] An over the top service is utilizing the telecom network to perform. However, it does not require any explicit affiliation with the network provider. Examples of over the top application are WhatsApp, Skype or Viber application.

and the second model. This last model is learned from both social network features as well as the traditional churn variables.

4.3.2. Propagation models

The remaining four models are trained using energy propagation techniques based on the previously discussed spreading activation algorithm:

- Model 4: simple propagation model
- Model 5: extended propagation model
- Model 6: simple propagation model undirected
- Model 7: extended propagation model undirected

March and April's churners are used as the source of the energy propagation. Each churned node is given an initial energy of 1. Model 4, which is a simple propagation model, sets the initial energy of non-churners to 0. Model 5 is actually boosting of Model 1 into Model 4. It indirectly incorporates subscribers' intrinsic information into the propagation model. Instead of setting the energy of non-churners to 0, this model assigns the churn score obtained from Model 1 as the initial energy of the non-churner nodes. The intuition behind this idea is that a subscriber might already have a certain tendency to churn due to his/her experience with the provided service. Model 6 and Model 7 are similar to Model 4 and Model 5, except that those models are trained using an undirected instead of a directed graph.

The total energy value that remains after termination is assumed to be the probability of a network member to churn. To study the influential effect of churned neighbors in the social network, we then compare the propensity values of non-churners to the actual known churn class.

5. Results

In this section, the empirical result for each of the implementation scenarios is discussed (see Table 2 and Figure 6). We present and discuss only the scoring models based on decision trees, because these models have a slightly better predictive performance compared to the ones built using logistic regression. Moreover, we only include propagation models with the spreading factor that yield the best prediction results.

Model 3, which is the hybrid model that combines tabular churn predictors and social network variables derived from social network graph, has the highest CoC score on the test set (64.98%). Since it only slightly outperforms Model 1 (64.88%), we can conclude that adding social network features on top of the traditional churn predictors does not appear to provide any substantial improvement for our scoring model. Model 2 built solely using social network predictors has the lowest predictive accuracy compared to the rest of the scoring models (56.57%). By targeting the top 30% of the subscribers, Model 2 can find only 37% of the

Table 2. Coefficient of Concordance of the scoring and propagation models.

	Performance on		
	Train set	Validation set	Test set
Model1	65.48	64.47	64.88
Model2	57.93	56.72	56.57
Model3	65.65	64.45	64.98
Model4	53.34	53.43	53.04
Model5	55.26	54.58	55.24
Model6	52.07	52.15	52.26
Model7	58.39	57.66	58.30

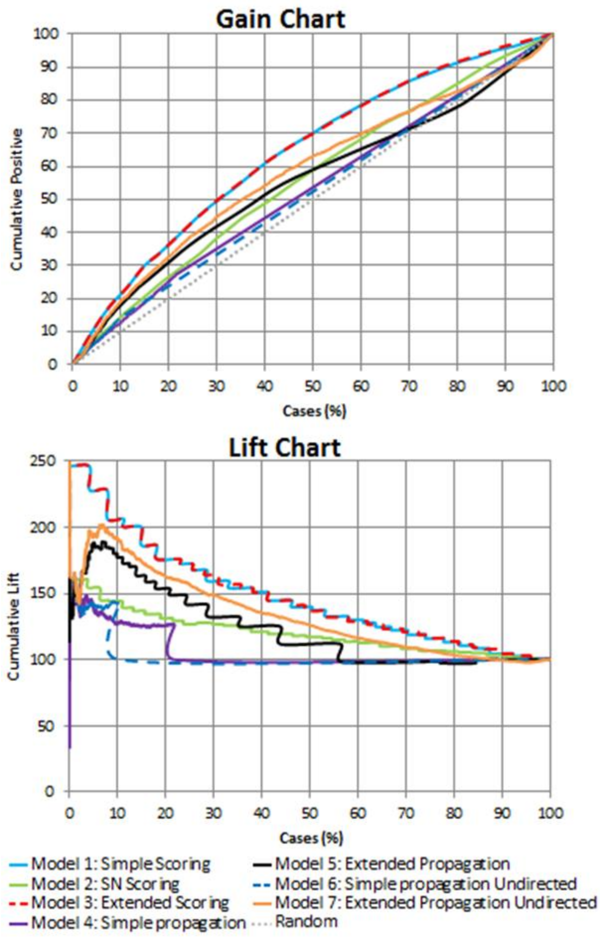

Figure 6. Gain and Lift chart of all models.

churners, while Model 1 and Model 3 are able to return about 50% of the churners. The lift chart shows that in the top 30% of the cases Model 2 has cumulative lift of 130%, whereas other scoring models have cumulative lift of 160%. In other words, the information derived from the social network is weakly predictive by itself and it fails to outperform the predictive power of the traditional predictors.

As expected, the extended propagation models (Model 5 and Model 7), which incorporate churn scores of the simple scoring model as the initial energy value in the propagation process, outperform the traditional social network propagation models (Model 4 and Model 6). These extended or hybrid models provide better predictive accuracy than the simple propagation models for the directed and the undirected graph. By targeting 30% subscribers, Model 7 is able to correctly predict about 45% churners. It returns 5% less than the tabular churn models, Model 1 and Model 3. Although Model 7 incorporated the traditional predictor elements in the propagation process, the predictive power is still lower than that of the traditional tabular churn scoring models.

The simple propagation models that incorporate only the social neighborhood information, Model 4 and Model 6, have even lower performance compared to Model 2. Unlike Model 2, the simple propagation model uses only the previous churner information within the social network without considering the individual churn propensity. This leads us to believe that the churning behavior of neighbors does not have enough influential effect on other members within a prepaid telecom subscriber social network. Traditional churn predictors apparently have a stronger influence on churn compared to social relationships.

6. Conclusions and Future work

Throughout this paper we have investigated the extent to which social network information can be used to predict telecom churn, and how this information could potentially improve the predictive performance of the conventional churn prediction method. We have assessed the performance of models constructed using the classical tabular data mining, the social network mining and the combination of both mining techniques. The first hybrid model is built by extending the traditional tabular churn predictors with social network variables extracted from the social graph. The second hybrid model is obtained by incorporating results of the traditional tabular churn models to the social propagation graph.

The performance of our models was verified using a large dataset of 700 million call data records. Our initial observation shows that the churn probability is positively aligned with the number of churned neighbors. The regular tabular churn models and the traditional social network models constructed exclusively using social network information score the least. This indicates that social network information alone is not sufficient to predict churn. Overall, the traditional tabular churn models have the best predictive accuracy. The added value of the social network variables to the tabular churn models is rather minimal. Although the second hybrid models are able to outperform the regular propagation models, it still could not beat the performance of the traditional tabular churn models. The contribution of traditional predictors to churn prediction is substantially higher than that of the social network behavior. Moreover, the performance gain of both hybrid models is not substantial enough to justify the computational costs.

The current research study only explores the negative influential effect of previous churners within the social network. Future research could potentially be focused on removing this limitation. The influences from both churners and non-churners could be taken into account, as subscribers might spread messages based on how they perceive the product/service quality. Assuming bad news can have a stronger influential effect than good news, positive influence from non-churners to stay within the network might not be as strong as negative influence from churners. Since our energy propagation model is purely derived from node and neighborhood-based relationships, the spreading activation computations are done locally and subscribers do not have knowledge beyond their direct neighbors. Other algorithms, for example from the field of community detection, are capable to identify the role of subscribers within the social network, such as influencer or adopter. Rather than targeting all future churners, we can minimize our resources by focusing only on churners with high influential power.

References

Backstrom, L., Huttenlocher, D., Kleinberg, J., & Lan, X. (2006). Group formation in large social networks: membership, growth, and evolution. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 44–54.

Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S., Nanavati, A. A., & Joshi, A. (2008). Social ties and their relevance to churn in mobile telecom networks. *Proceedings* *of the 11th international conference on Extending database technology*, 668–677.

Ferreira, J. B., Vellasco, M., Pacheco, M. A., & Barbosa, C. H. (2004). Data mining techniques on the evaluation of wireless churn. *Proceedings of the European Symposium on Artificial Neural Networks*, 483–488.

Gyan, R., Hui, Z., Zhi-Li, Z., & Jean, B. (2012). Are call detail records biased for sampling human mobility?. *ACM SIGMOBILE Mobile Computing and Communications Review*, *16*, 33–44.

Hadden, J., Tiwari, A., Roy, R., & Ruta., D. (2006). Churn prediction using complaints data. *International Journal of Intelligent Technology*, *13*, 158–163.

Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California, Riverside. `http://faculty.ucr.edu/~hanneman/`.

Harell, F. E. J. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis.* New York: Springer.

Kawale, J., Pal, A., & Srivastava, J. (2009). Churn prediction in mmorpgs: A social influence based approach. *Proceedings of the 2009 International Conference on Computational Science and Engineering*, *4*, 423–428.

Kraljevic, G., & Gotovac, S. (2010). Modeling data mining applications for prediction of prepaid churn in telecommunication services. *Automatika*, *51*, 375–383.

Kusuma, P. D. (2013). Extending traditional telecom churn prediction using social network data. Master's thesis, Leiden University. `http://www.liacs.nl/assets/2013-03PalupiKusuma.pdf`.

Motahari, S., Mengshoel, O. J., Reuther, P., Appala, S., Zoia, L., & Shah, J. (2012). The impact of social affinity on phone calling patterns: Categorizing social ties from call data records. *In Proceedings of the 6th SNA KDD Workshop*, 9–17.

Neo4j. (2012). Neo4j: Community edition (version 1.8.m05) [software]. `http://Neo4j.org/`.

Oentaryo, R. J., Lim, E. P., Lo, D., Zhu, F. D., & Prasetyo, P. (2012). Collective churn prediction in social networks. *In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 210–214.

Pegasystems. (2008). Predictive analytics director (version cdm 6.3) [software]. `http://www.pega.com/products/decision-management`.

Prasad, U. D., & Madhavi, S. (2012). Prediction of churn behavior of bank customers using data mining tools. *Business Intelligence Journal*, *5*, 96–101.

Radosavljevik, D., van der Putten, P., & Larsen, K. (2010). The impact of experimental setup in prepaid churn prediction for mobile telecommunications: What to predict, for whom and does the customer experience matter?. *Transactions on Machine Learning and Data Mining*, *3*, 80–99.

Seshadri, M., Machiraju, S., Sridharan, A., Bolot, J., Faloutsos, C., & Leskovec, J. (2008). Mobile call graphs: beyond power-law and lognormal distributions. *In Proceedings of the 14th ACM SIGKDD*, 596–604.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques.* San Francisco: Morgan Kaufmann. 3rd edition.

Ziegler, C. N., & Lausen, G. (2004). Spreading activation models for trust propagation. *In Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service*, 83–97.