Niko Gamulin, Mitja Štular, Sašo Tomažič

Impact of Social Network to Churn in Mobile Network

As the telecommunications sector has reached its mature stage, maintaining existing users has become crucial for service providers. Analyzing the call data records, it is possible to observe their users in the context of social network and obtain additional insights about the spread of influence among interconnected users, which is relevant to churn. In this paper, we examine the communication patterns of mobile phone users and subscription plan logs. Our goal is to use a simple model to predict which users are most likely to churn, solely by observing each user's social network, which is formed by outgoing calls, and churn among their neighbours. To measure the importance of social network parameters with regard to churn prediction, we compare three models: spatial classification, regression model, and artificial neural networks. For each subscriber, we observe three social network parameters, the number of neighbors that have churned, the number of calls to these neighbors, and the duration of these calls for different time periods. The results indicate that using only one or two of these parameters yields results that are comparable or better than the complex models with large amounts of individual and/or social network input parameters that other researchers have proposed.

Key words: Churn, Social Network Analysis, Machine Learning, Mobile Network

Utjecaj društvene mreže na odljev korisnika u mobilnim mrežama. Kako je telekomunikacijski sektor dosegao zreli stadij, zadržavanje postojećih korisnika od ključne je važnosti za pružatelje telekomunikacijskih usluga. Analizom liste poziva moguće je nadzirati korisnike u kontekstu društvene mreže i dobiti dodatni uvid u širenje utjecaja među povezanim korisnicima, što je relevantno za odljev korisnika. U ovom radu razmatramo obrasce komunikacije korisnika mobilnih mreža i podatke o planu pretplate. Naš cilj je korištenjem jednostavnog modela predvidjeti koji korisnici su najskloniji prijelazu na drugu mrežu, pritom koristeći samo korisnikovu društvenu mrežu koja se formira odlaznim pozivima i prijelazima između mreža njihovih susjeda. S ciljem mjerenja važnosti pojedinog parametra društvene mreže za predikciju prelaska na drugu mrežu uspoređena su tri modela: prostorna klasifikacija, regresijski model i model neuronske mreže. Za svakog pretplatnika razmatramo tri parametra društvene mreže: broj susjeda koji su promijenili mrežu, broj poziva prema njima kao i trajanje spomenutih poziva u različitim vremenskim razdobljima. Rezultati pokazuju kako se korištenjem samo jednog ili dva od navedenih parametara društvene mreže postižu rezultati koji su usporedivi ili bolji od rezultata složenijih modela drugih autora koji koriste veliki broj osobnih parametara i/ili parametara društvene mreže.

Ključne riječi: prijelaz korisnika, analiza društvene mreže, strojno učenje, mobilna mreža

1 INTRODUCTION

To attract new service consumers and retain the existing ones, telecommunications service providers have been constantly forming new subscription plans, adapting the terminal equipment offered according to current trends, and improving the quality of services by upgrading their network equipment. Along with the development and efficiency of machine learning methods, service providers from all industries have become aware of the importance of social network data, which can be used to gain additional insights about their service consumers. Consequently, it is possible to target the important customers, who are more prone to churn, i.e., switch from the current service provider to a competitor.

According to past research, maintaining existing customers is even more important than attracting new ones because the cost of winning a new customer is far greater than the cost of preserving an existing one [1]. Many methods have already been proposed for churn prediction using past data. In some of these, the user is treated as an individual, independent from his acquaintances' data, whereas other methods also consider social network effects.

2 PROBLEM STATEMENT

Although awareness of the importance of social networks has increased significantly with the spread of online social networks such as Facebook, Twitter, Google+ and LinkedIn, some of service providers from non-internet industries have yet to exploit the potential of real social networks of interconnected people influencing each other in the real world. Though several online services and retailers have already developed marketing campaigns based on social networks, some businesses from other industries still try to attract new customers and keep the existing ones by treating each customer as an individual, and mostly investing in broadcast marketing campaigns and forming global-level offers. Some of the service providers and retailers are certainly not able to treat their customers as interconnected peers who are influenced by their friends because the service providers lack the data needed to represent users as nodes, and form edges between them. Conversely, telecommunications service providers have to keep the data of their customers' phone call records for charging purposes. The same data could therefore be used to form a social network of interconnected users and anticipate how they might influence each other. Furthermore, studies have demonstrated that more than 75% of customers will consult a friend before deciding on the purchase of a certain product or adoption of a service [2]. Therefore, social network analysis can be applied to design targeted or viral marketing campaigns.

Our motivation for this research is based on the assumption that users influence each other over phone conversations. The power of influence is conditional on the duration and number of conversations. According to this assumption, a user who has many friends who have already churned is also more likely to churn. Furthermore, we hypothesized that if the observed user is influenced by his peers, he will follow them in a short amount of time.

If an individual's churn decision is influenced by prior churn among his neighbors, it should be possible to confirm the assumption by analyzing only phone conversation data and subscription plan logs. With such an approach, it might be possible to isolate the importance of peers' influence from other possible factors relating to individual properties or tendencies, such as demographic parameters, bill payments, or complaints about services, to name a few.

Our goal was also to use as simple a model as possible, i.e., a model that required a minimal number of input parameters and that had minimal complexity, so it could be implemented in real time in the large networks of telecommunications service providers. A churn detection algorithm could be applied in such a way that the provider could target the users which could potentially leave with specific actions, such as promotions or discount.

In the following chapter, we provide an overview of some of the previously proposed methods for churn prediction, in which users were observed either as individuals or in the context of a social network.

3 PREVIOUS RESEARCH

A number of studies have been conducted on the churn problem within the telecommunications sector and other fields that are relevant to churn phenomena in a broader context, such as social networks, diffusion, probabilistic models and machine learning.

When considering the churn problem independently from the broader context, one must first be aware of the possible significant difference between prepaid and postpaid users. The former users are not bound by contracts, while some service providers require the latter to sign a service contract for a specific period of time. In the case of prepaid users, it is easy to observe them in the context of a social network and extract the rules for churn diffusion because these users are free to make decisions about changing their service plan at any time.

The model in which prepaid users were observed in the context of a social network is presented in [3]. Conversely, though it is easier to observe prepaid users in the context of a social network, it is not easy to determine a user's churn status because such users do not explicitly cancel their subscription plan. In [4], a model for prepaid user labelling was proposed along with a churn prediction technique in which users were observed as individuals, without considering the influence between interconnected users. In addition to the service plan (prepaid or postpaid), if model complexity and data availability are not an issue, it is possible to build a prediction model using a number of other variables related to the observed users.

Other studies[5], [6], [7], [8] , [9], [10], [11], [12], [13] proposed different approaches for churn prediction in which users were considered individually.

In [5], Datta et al. described an automated system for modeling mobile subscriber churn that predicted which customers would discontinue their cellular phone service. The authors described a feature selection process, in which they selected the best 30 to 50 individual parameters. These were then used to build cascade neural networks and predict which users were most likely to discontinue the service.

In [6], Guojie et al. proposed a Mixed Neural Network model, intended for churn prediction. The model was based on demographic parameters and information about telephone usage.

In [7], Rupesh et al. discussed a method based on ordinal regression to predict churn time or tenure of the observed users.

In [8], the authors proposed two genetic algorithm based neural network models to predict churn among mobile phone users.

In [9], Wei & Chiu proposed and evaluated a churn prediction model that predicted churn from subscriber contractual information and call pattern changes extracted from call details.

In [10], Hung et al. compared various prediction models based on user demographics, billing information, contract/service status, call detail records, and service change logs.

In [11], Mozer et al. experimented with various prediction models, such as logistic regression, decision trees, and neural networks to predict churn among mobile phone users.

In [12], Au et al. proposed a data mining algorithm, called data mining by evolutionary learning (DMEL), which estimated the churn likelihood of the observed users.

In [13], the authors proposed a support vector machine model to predict churn, and compared it with artificial neural networks, decision trees, logistic regression, and naive Bayesian classifiers.

In comparison to the models noted above, which are based solely on individual properties, the following include either both network and individual, or solely network properties.

Using Markov Logic Networks (MLNs), Dierkes et al. [14] examined whether the user churn decisions of individuals in previous time periods impacted on other users with whom the target customer interacted via voice call, short message service (SMS), or multimedia message service (MMS). To design the prediction model, the authors selected a data sample consisting of 2,654 users, of which 645 had churned and 2,000 had not. The non-churners were who had churners as neighbors in the call graph were selected. In addition to the call and message logs, the authors included an additional 32 attributes in the model, such as handset equipment, contract duration, interaction with service center, age and gender. To evaluate the performance of the MLN model they compared it with a logistic regression model which considered the same combination of individual and network attributes and benchmarked it against the logistic regression model which considered only individual attributes.

In [15], Kiss and Bichler evaluated the relevance of several centrality measures to user's importance in both simulated and real-world networks extracted from call detail records (CDRs). The authors observed phone call connections in the context of a social network and discussed the relevance of several network centrality measures to estimating an individual's influence on his community.

Xiaohang et al. [16] investigated the effects of network attributes on the accuracy of predicting churn. The authors used logistic regression, decision trees, and neural network models built in SAS Enterprise Miner. For the research, the authors analyzed several users' personal characteristics and network attributes. The authors concluded that as the number of churn neighbors increases, the probability of individual churns also increases.

Richter et al. [17] proposed an approach for churn prediction for groups of interconnected users by only analyzing call data without demographic information. The authors first employed a novel, information theoretic based measure to quantify the connectivity between each pair of subscribers in the network. Having established the connectivity measure, they partitioned the network into a small disjointed, densely connected group of users by keeping only the strongest connections. Finally, the authors queried the statistical model to establish a churn risk score for each group and its individual members.

In [18], the authors clearly proved the importance of social aspects by employing a proportional hazard model to estimate the hazard of churn for mobile service users. To build and test the proposed model, the authors used social variables, satisfaction-related variables, and economic incentive. The social variables included the number of neighbors who had already churned, edge strength, homophily, and the number of neighbors the focal user had. The edge strength between two users was determined by the relative communication volume among them, and homophily was determined by their similarity in terms of gender, age, segment, and socioeconomic status. The satisfaction-related variables were determined by change in the level of usage and the number of service records. Economic incentive was determined by the relative duration of calls within a service provider network. In addition to these variables, the authors included control variables in the hazard model. These included usage, tenure, age, gender, and sociodemographic segment.

Another example of research, from a non-telecom industry, worth mentioning as it emphasized the importance of social networks is [19]. The authors analyzed churn among participants of Massively Multiplayer Online Role Playing Games (MMORPGs). Similar to mobile phone users, MMORPGs players can be represented as nodes, and edges between two participants if they participate in the same game. Each participant forms his own opinion, which is based on experience and might be influenced by other participants. Based on participants' engagement, the authors proposed a diffusion model, which took into account the participant's game engagement and social influence from neighbors in influence propagation.

4   ANALYSIS OF SOCIAL IMPACT ON CHURN PROBABILITY

The aim of this research was to demonstrate that the user's network attributes extracted from CDR data can be used to detect potential churners, without additional demographic information or any variables related solely to observed individual users. By isolating social network attributes and achieving a prediction accuracy that was comparable to existing, more complex models based on additional personal attributes, the importance of the social network was unambiguously visible. We thus analyzed three models that only used one to three social network attributes as input parameters.

Although many churn prediction models have been proposed to date, our aim was to prove it is possible to reduce model complexity and predict postpaid user churn by observing only a small number of network attributes.

The previously proposed models for churn prediction were either designed for different purposes or considered different user attributes. The model proposed by Richter et al. [17] is similar to our research from the context standpoint and the observed attributes. However, the quantification of social relatedness between two subscribers is significantly more computationally expensive; specifically, the relatedness of two users is calculated by examining the number of calls to common neighbors.

Another previous study [15] is also similar to our research in that it discusses the relevance of several network centrality measures. However, whereas we observed how an individual is influenced by his peers, this study discussed how the observed individual influenced his peers.

In [14], the authors compared the performance of MLN and logistic regression models that considered both network and user description attributes. The benchmark against the logistic regression model which considered only individual attributes proved that the MLN model did not perform better than the benchmark. Conversely, the logistic regression model with additional network attributes performed significantly better than the benchmark. Their obtained results were comparable to ours; however, their model's complexity was much greater, and it is not possible to distinguish the relevance of network and personal attributes to the model's accuracy.

The comparison of the models presented in [18], clearly showed that the social model, which included both social and individual variables, outperformed the model which included only individual variables. In terms of the importance of social variables, the presented findings were consistent with our results. However, the presented model was not comparable to ours as it was based on both individual and social variables. Conversely, our aim was to isolate social network effects and therefore our model did not include individual variables.

When considering large datasets, achieving good or better results using simple models, rather than complex ones, means that it is possible to reduce the required computer resources without prediction quality degradation. In [14], the authors used a dataset sample consisting of 2,645 customers, containing 645 churners and 2,000 non-churners. In the paper, they stated it was not possible to run the model with a larger dataset. The ratio of churners to non-churners was considerably larger and model complexity was greater in comparison with the same ratio in our dataset and our model complexity.

Similarly, in [16], the authors proposed a complex model that estimated the churn probability based on a large number of personal and network attributes. They estimated model performance using a dataset with a higher percentage of churners; however, their results are still comparable to those obtained in our analysis with simple models using only one or two network attributes on a dataset with a low percentage of churners.

In this paper we present the results of three different models used to predict churn solely on the social network obtained from the CDRs of mobile operator.

4.1   Dataset

We used anonymized call logs from CDRs for a two-month period, for September and October 2010. Approximately 42,000,000 aggregated daily call connections of postpaid users were included in the analysis. The call connection record contained the number of calls between the caller and called person and the sum of call durations for the observed day.

In addition to the call connections data, the churn log for the time period 2005 and 2011 was available. Using these data, it was possible to label each observed user either as a churner or a non-churner, depending on subscription plan changes made during the two-month observed CDR period. Accordingly, in these two months 0.67% of all observed users were labeled as churners.

To perform the analysis, the original dataset was reshaped in the following ways:

A list of all active postpaid users was made from the aggregated daily call records.

For each user, the sum of all outgoing calls was calculated.

Three different period lengths, L: 60 days, 30 days and 15 days were defined to observe which neighbors had churned.

For each user and each period length L, we counted the number of neighbors who churned at most L days prior to

the observed date, the number of calls toward them, and summed their duration.

For each user we calculated the relative number of neighbor churners, the relative duration of calls to these churners and the relative number of calls by dividing the absolute values by the number of all neighbors, the duration of all calls and the number of all calls for each user respectively.

In order to train, evaluate, and compare the performance of different models, we split the reshaped data into training and test datasets. The training dataset consisted of roughly 1% of all users, among which half were churners and half non-churners. To form the test dataset, we selected the remaining churners and the appropriate number of non-churners, in order to have the churn ratio comparable to the original dataset.

The further discussion is limited to a 60-day period observation length, where all of the models presented achieved the best results.

Vector representation   Each record from the reshaped data consisted of three input variables. In our case the input variables were the relative number of neighbors who had already churned, the relative total duration of calls to churners, and the relative number of calls to churners. To shorten the notation, the input variables have been labeled with the letters x, y, and z respectively, where x is the relative number of neighbors that have churned, y is the relative durations of calls to neighbors that have churned, and z is the relative number of calls to neighbors that have churned, and the input vector X containing one, two or all three input variables has been formed.

The vector notation for all three input variables is shown in (1).

$$X \equiv \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad (1)$$

4.2   Churn prediction models

To evaluate the impact of social network to churn probability we used three different churn prediction models, namely, spatial, regression and an artificial neural network model. We also selected the best model among them with regard to its gain and its simplicity, i.e., computational efficiency.

All models take vector X as their input and return classification function $h(X)$, which is proportional to the predicted churn probability. All models were tested with all possible combinations of input variables x, y, and z in input vector X.

First we calculated the classification function $h(X^{(i)})$ for each user, where $X^{(i)}$ denoted vector X
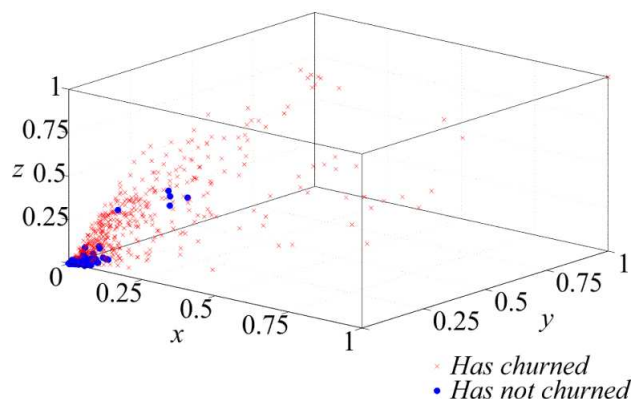

Fig. 1. Sample set of users from the observed dataset, represented as points in 3D space

of the i-th user in the test dataset. Then we sorted users by descending value of $h(X^{(i)})$ and captured the first N_{cu} users, i.e., the users with the highest $h(X^{(i)})$.

Given the actual churn facts for the users from the test dataset, the performance of all models was evaluated with recall $R(N_{cu})$ and gain $G(N_{cu})$ as defined below.

Recall is defined as the fraction of churners among captured users:

$$R(N_{cu}) = \frac{N_{cc}}{N_{ac}}, \qquad (2)$$

where N_{cc} is the percentage of churners among the percentage of captured users N_{cu}, whereas gain is the ratio:

$$G(N_{cu}) = \frac{N_{ru}}{N_{cu}}, \qquad (3)$$

where N_{ru} is the percentage of randomly captured users which would yield the same percentage of captured churners N_{cc} as capturing N_{cu} users using the churn prediction model. Gain thus directly reflects the savings of telecommunications service providers when they target a certain limited number of users to prevent their churn.

4.3   Spatial model

To intuitively understand the relevance of each observed input variable, we first represented the data visually, by representing each observed user as a colored point in 3D space: Red for churners and blue for non-churners.

Based on the visual representation shown in Figure 1, it was possible to draw the following conclusions: If many of the observed user's neighbors churned and if one spent a relatively long time talking with churners, the probability that the observed user would churn was higher than in the case when the observed users did not have many neighbors who churned.
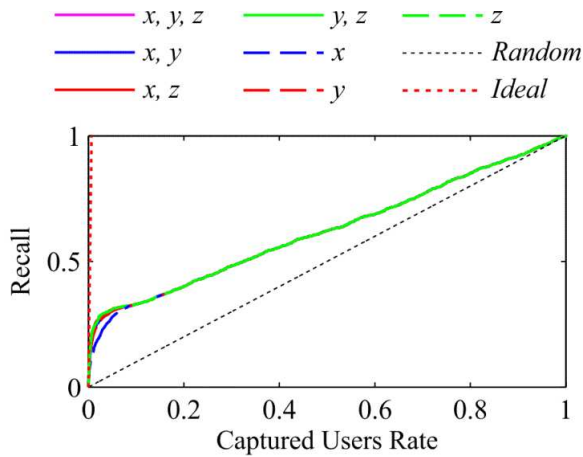
Fig. 2. *Spatial model recall curves for different combinations of input variables for the 60-day time period*

Fig. 3. *Linear regression recall curves for different combinations of input variables for the 60-day time period*

Accordingly we defined the classification function $h(X)$ of the spatial model as the length of vector X:

$$h(X) = \|X\| \qquad (4)$$

which is equal to the Euclidian distance of the user from the coordinate system origin.

The recall and gain were for different combinations of input variables x, y, and z as a function of the number of captured users, as represented in Figure 2.

It is worth noting that the model using the single input variable z, i.e., the relative number of calls to neighbors that churned, performed as well as models using two or all three of the variables. When only z was used as the input variable, the classification function $h(X)$ was simplified to:

$$h(X) = z, \qquad (5)$$

thus we concluded that the choice of the single input variable z was optimal for the spatial model.

4.4 Regression model

In the regression model we used a linear classification function:

$$h_\theta(\theta_0, \theta, X) = \theta_0 + \theta^T \cdot X, \qquad (6)$$

where θ was a weight vector of the same dimension as the input vector X. The use of higher order functions did not bring any improvement.

To determine the values of θ and θ_0 we used the training set. We assigned churn variable c, to each user in the training set, with value 1 for churners and value 0 for non-churners.
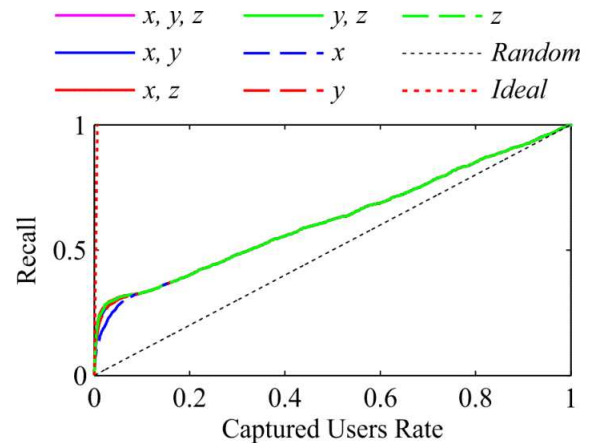
The sum of the square values of the differences between the classification function $h_\theta(X)$ and churn variable c of all users in the test set was used as a cost function:

$$J(\theta_0, \theta) = \sum_{i=1}^{M} \left(h_\theta(\theta_0, \theta, X^{(i)}) - c^{(i)} \right)^2, \qquad (7)$$

where M was the number of users in the training set and $c^{(i)}$ was the churn variable of the i-th user. The cost function had minimal value when its derivatives equaled zero:

$$\frac{\partial J(\theta_0, \theta)}{\partial \theta_0} = 0 \qquad \frac{\partial J(\theta_0, \theta)}{\partial \theta} = 0, \qquad (8)$$

where 0 was a zero vector of the same dimension as θ. To obtain optimal values $\theta_{0,opt}$ and θ_{opt} we searched for the values θ_0 and θ which satisfied (8) for data in the training set for each combination of input variables x, y, and z. The classification function of the regression model was then set to:

$$h(X) = h_\theta(\theta_{0,opt}, \theta_{opt}, X) \qquad (9)$$

The recall and gain curves for the regression models with different combinations of two and three input variables are presented in Figure 3.

All combinations of input variables yielded nearly the same results, except when only using x, where the results were worse. The combination of x and z yielded slightly better results than other combinations, thus we asserted this combination as the best for the regression model.

4.5 Artificial Neural Network Model

For different combinations of input variables, we trained an Artificial Neural Network (ANN) with one hidden layer, which consisted of a different number of neurons using a training dataset. In the case of two input variables, the best performance was achieved with three neurons, and in case of three input variables, the network with
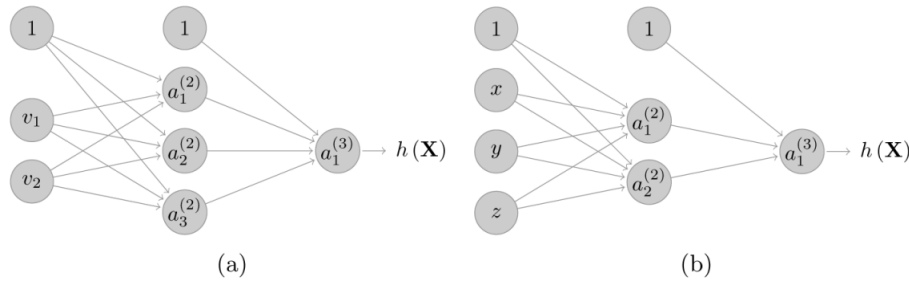
Fig. 4. ANN models for the combination of two input variables (a) and three input variables (b)

two neurons in the hidden layer demonstrated the best performance. Because the ANN for a single input variable is trivial, it was not considered. The ANN configurations for two and three variables are shown in Figure 4.

For the activation functions ($a_1^{(2)}$, $a_2^{(2)}$, $a_3^{(2)}$, $a_1^{(3)}$), we selected sigmoid functions and used a back propagation algorithm [20] with the cost function:

$$
\begin{aligned}
J\left(\theta_1, \theta_2\right) \\
= \sum_{i=1}^{M}\left(c^{(i)} \log \left(h_\theta\left(\theta_1, \theta_2, X^{(i)}\right)\right)\right) \\
+ \sum_{i=1}^{M}\left(\left(1-c^{(i)}\right) \log \left(1-h_\theta\left(\theta_1, \theta_2, X^{(i)}\right)\right)\right),
\end{aligned}
\tag{10}
$$

where θ_1 was the weight matrix of the 1st layer, θ_2 was the weight vector of the 2nd layer, and $h_\theta\left(\theta_1, \theta_2, X\right)$ was the classification function used in the back propagation algorithm to obtain optimal values $\theta_{1,opt}$ and $\theta_{2,opt}$ and the ANN classification function:

$$
h_\theta\left(X\right) = h_\theta\left(\theta_{1,opt}, \theta_{2,opt}, X\right). \tag{11}
$$

We determined optimal weights for three input parameters and all combinations of two input parameters. The obtained recalls and gains as functions of captured numbers of users are presented in Figure 5.

The best result were obtained with the combination of the two input variables x and z.

4.6   Comparison of models

In this section, we present a comparison of the three models. For each model we selected the combination of variables which achieved the best performance. The comparison of recall and gain curves is presented in Figure 6.

By comparing the best combinations of input variables of the models discussed, we can see that the performance of the spatial model using a single variable is approximately equal to the performance of more complex models using more variables. Thus, the spatial model with single input variable z (the relative number of calls to neighbors that already churned) was the best choice due to its simplicity.
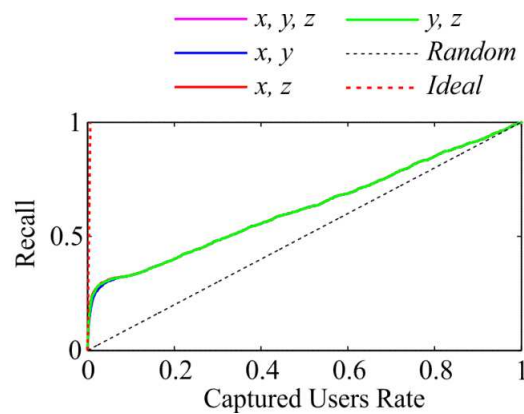

Fig. 5. ANN recall curves for different combinations of input variables for the 60-day time period

Table 1. Spatial model, ANN, and regression model results for the 60-day time period, using the best combinations of input variables, where x represents the relative number of neighbors that have already churned, and z the relative number of calls to neighbors that have already churned

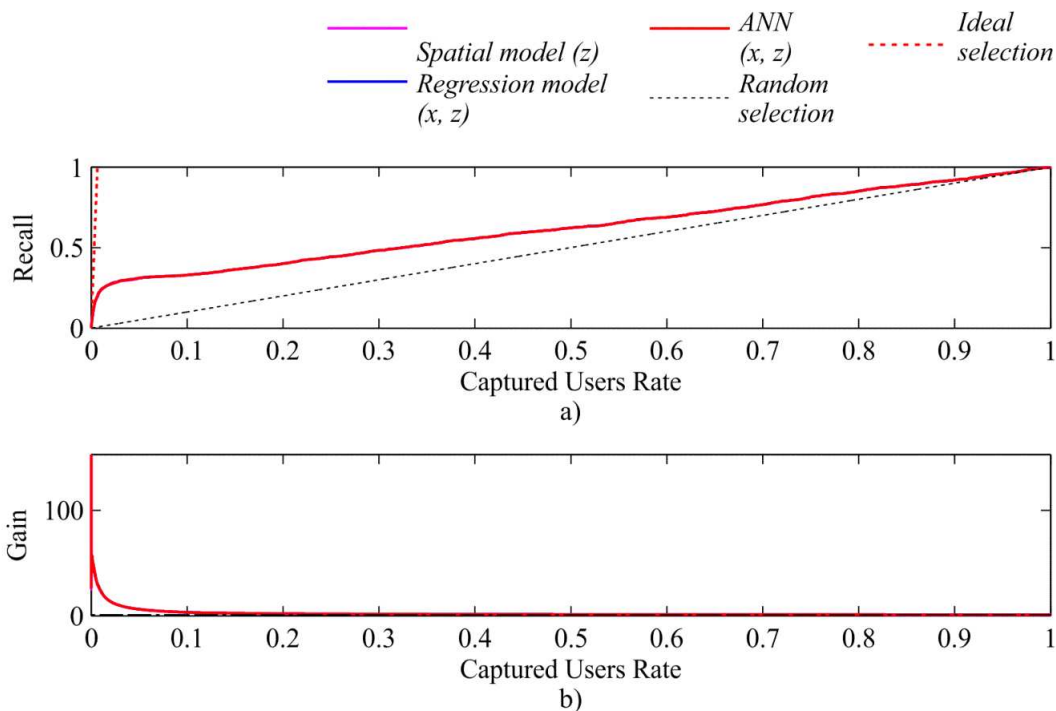| Model	Input variables	Recall		Gain	
		5% users	*10% users*	*5% users*	*10% users*
Spatial	z	0.310	0.330	6.194	3.298
Regression	x, z	0.312	0.330	6.230	3.298
ANN	x, z	0.312	0.330	6.230	3.298

Fig. 6. Recall (a) and gain (b) values of the spatial model using single variable z, and the regression model and ANN using variables x and z
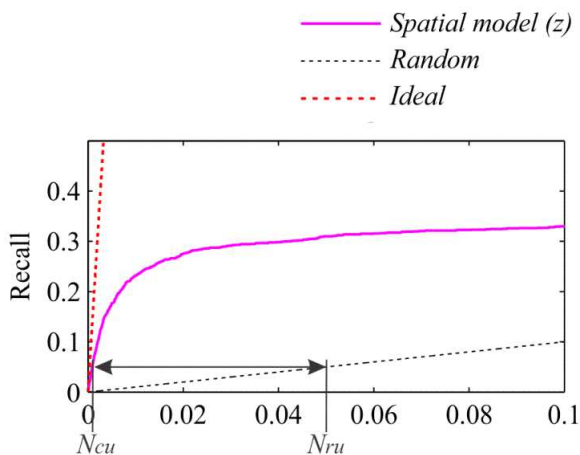

Fig. 7. Spatial model gain

5 CONCLUSION

In areas where the majority of the developed population has already adopted mobile services, it is critical to implement churn prediction methods to retain market share. In addition to observing user behaviour as an individual, it is crucial to discover the patterns and rules that hold for networks of interconnected users.

In this research, we observed a user's behavior in terms of churn and proved that this depends on his neighbors' prior behavior. To evaluate the relevance of available input variables, we evaluated three different classification models that used combinations of a single, two, and three variables.

By comparing the performance of a model using a single variable with the performance of the model using two or three variables, we have shown that the performance when using a single variable was approximately the same as in cases using two or three variables. Furthermore, prediction accuracy was not improved by using a regression model or ANN, which are more complex than the spatial model.

The performance of the models was comparable to existing models, which are more complex than the simple spatial model where predictions are based on a single input variable. Due to its low complexity, the model was capable of calculating churn probability for all users in real time.

To describe the model's performance in practice, we might consider a case in which the mobile provider has 2,000,000 subscribers and wants to target 5% of churners. If the provider does not use any classification model, it has to select 100,000 random users, whereas if employing the proposed model, 5% of all churners are captured in a segment of 1,780 users. The size of targeted segment is therefore reduced by gain value, in this case, factor 56.

In further studies we are planning to extend the proposed model by including individual properties for the observed users. By including these parameters we expect to improve its predictive ability, especially with regard to predictions for the segment of users who do not have a great deal of contact with prior churners. Additionally we are planning to develop a model to identify influential users whose behavior, in terms of churn, could potentially spread among their neighbors.

ACKNOWLEDGMENT

REFERENCES

[1] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, "Computer assisted customer churn management: State-of-the-art and future trends," Computers & Operations Research, vol. 34, pp. 2902-2917, 2007.

[2] S. Doyle, "The role of social networks in marketing," Journal of Database Marketing & Customer Strategy Management, vol. 15, pp. 60-64, 2007.

[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," presented at the Proceedings of the 11th International Conference on Extending Database Technology: Advances in Database Technology, Nantes, France, 2008.

[4] L. Alberts, I. R. L. M. Peeters, R. Braekers, and C. Meijer, "Churn Prediction in the Mobile Telecommunications Industry," Master Thesis, Department of General Sciences-Maastricht University, Maastricht, Netherlands, 2006.

[5] P. Datta, B. Masand, D. R. Mani, and B. Li, "Automated Cellular Modeling and Prediction on a Large Scale," Artificial Intelligence Review, vol. 14, pp. 485-502, 2000.

[6] S. Guojie, Y. Dongqing, W. Ling, W. Tengjiao, and T. Shiwei, "A Mixed Process Neural Network and its Application to Churn Prediction in Mobile Communications," in Data Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE International Conference on, 2006, pp. 798-802.

[7] T. Washio, E. Suzuki, K. Ting, A. Inokuchi, R. Gopal, and S. Meher, "Customer Churn Time Prediction in Mobile Telecommunication Industry Using Ordinal Regression," in Advances in Knowledge Discovery and Data Mining. vol. 5012, ed: Springer Berlin Heidelberg, 2008, pp. 884-889.

[8] P. C. Pendharkar, "Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services," Expert Systems with Applications, vol. 36, pp. 6714-6720, 2009.

[9] C.-P. Wei and I. T. Chiu, "Turning telecommunications call details to churn prediction: a data mining approach," Expert Systems with Applications, vol. 23, pp. 103-112, 2002.

[10] S.-Y. Hung, D. C. Yen, and H.-Y. Wang, "Applying data mining to telecom churn management," Expert Systems with Applications, vol. 31, pp. 515-524, 2006.

[11] M. C. Mozer, R. Wolniewicz, D. B. Grimes, E. Johnson, and H. Kaushansky, "Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry," Neural Networks, IEEE Transactions on, vol. 11, pp. 690-696, 2000.

[12] W. H. Au, K. C. C. Chan, and Y. Xin, "A novel evolutionary data mining algorithm with applications to churn prediction," Evolutionary Computation, IEEE Transactions on, vol. 7, pp. 532-545, 2003.

[13] G.-e. Xia and W.-d. Jin, "Model of Customer Churn Prediction on Support Vector Machine," Systems Engineering - Theory & Practice, vol. 28, pp. 71-77, 2008.

[14] T. Dierkes, M. Bichler, and R. Krishnan, "Estimating the effect of word of mouth on churn and cross-buying in the mobile phone market with Markov logic networks," Decision Support Systems, vol. 51, pp. 361-371, 2011.

[15] C. Kiss and M. Bichler, "Identification of influencers – Measuring influence in customer networks," Decision Support Systems, vol. 46, pp. 233-253, 2008.

[16] Z. Xiaohang, L. Zhiyu, X. Yang, S. Wenhua, and W. Qi, "Predicting customer churn by integrating the effect of the customer contact network," in Service Operations and Logistics and Informatics (SOLI), 2010 IEEE International Conference on, 2010, pp. 392-397.

[17] Y. Richter, E. Yom-Tov, and N. Slonim, "Predicting customer churn in mobile networks through analysis of social groups," in Proceedings of the 2010 SIAM International Conference on Data Mining (SDM 2010), 2010.

[18] I. Nitzan and B. Libai, "Social effects on customer retention," Journal of Marketing, vol. 75, pp. 24-38, 2011.

[19] J. Kawale, A. Pal, and J. Srivastava, "Churn prediction in MMORPGs: A social influence based approach," in Computational Science and Engineering, 2009. CSE'09. International Conference on, 2009, pp. 423-428.

[20] J. Han and C. Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning," in From Natural to Artificial Neural Computation. vol. 930, J. Mira and F. Sandoval, Eds., ed: Springer Berlin Heidelberg, 1995, pp. 195-201.
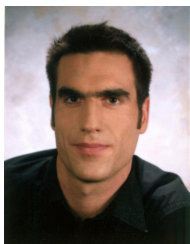
Niko Gamulin received B.Sc. degree in Electrical Engineering from the Faculty of Electrical Engineering, University of Ljubljana, Slovenia in 2009. In 2015 he received a Ph.D. in Electrical Engineering from University of Ljubljana. He is employed as a researcher at Telekom Slovenije. His main research interests cover the area of machine learning, big data, and churn modelling in telecommunications industry. His research is carried out in collaboration with the Laboratory of Information Technologies at Faculty of Electrical Engineering.

Mitja Štular graduated from the Faculty of Electrical Engineering and Computer Science of the University of Ljubljana, Slovenia in 1994. He successfully defended his master thesis in 1997 and Ph.D. thesis in 2000. In 2001 he began a career at Mobitel as a consultant to the CEO and also as the UMTS project director and been involved in a variety of research and development projects. In 2006 he took on the position of CTO responsible for networks, services and IT, and in 2009 a position of a Senior Strategy Adviser to the CEO, CSO.

Sašo Tomažič received his Ph.D. in Electrical Engineering from University of Ljubljana, Slovenia, in 1991. He is a full professor on the Faculty of Electrical Engineering, Ljubljana. He is the head of the Laboratory of Communication Devices and the head of the Department of Telecommunications. He was an advisor for networking, teleworking, and telematics at the Ministry of Education of Slovenia, an advisor for Telecommunications at the Ministry of Defense and the National Coordinator for telecommunications research of the Republic of Slovenia.

AUTHORS' ADDRESSES

Niko Gamulin, Ph.D.
Mitja Štular, Ph.D.
Telekom Slovenije,
Cigaletova 15, 1000 Ljubljana, Slovenia
email: {niko.gamulin, mitja.stular}@telekom.si

Prof. Sašo Tomažič, Ph.D.
Department of Telecommunications,
Faculty of Electrical Engineering,
University of Ljubljana,
Tržaška 25, 1000 Ljubljana, Slovenia
email: saso.tomazic@fe.uni-lj.si