Social-Network Influence on Telecommunication Customer Attrition

Piotr Wojewnik, Bogumil Kaminski, Mateusz Zawisza, and Marek Antosiewicz

Decision Support and Analysis Division,
Warsaw School of Economics, Al. Niepodleglosci 162, Warsaw, Poland
bkamins@sgh.waw.pl,
{piotr.wojewnik,mat.zawisza,marekantosiewicz}@gmail.com

Abstract. We investigate the network effects of churn in the telecommunication industry. Under calling party pays regime differentiation between on-net and off-net prices implies that customer's calling cost depends on operators chosen by the clients he calls. We assume that clients minimize their expenses. Therefore, after a single person churn we observe churn induced in his social neighborhood. Our aim is to verify, which measures of individual position in a social network are important predictors of induced churn. We control the results for changes in market prices structure, social network structure and number of operators on the market. Using multiagent simulation we show that (a) network structure and number of operators significantly influence induced churn level and (b) weighted prestige is its important predictor.

Keywords: Telecommunication, Customer attrition, Agent-based Models, Social-network analysis.

1 Introduction

One of the key issues driving telecommunication operator's decision making is the ability to forestall losing customers to competition. This is motivated by the fact that acquiring new customers is between 5 to 8 times more expensive than retaining existing ones [10,21]. The level of such voluntary attrition (churn) is estimated as 25-50% of client base per annum, where the particular rate depends mainly on client's segment [4]. The problem of voluntary churn is especially viable for pre-paid customers, where the client might cancel the contract without consequences. The attrition ratio can be reduced by retention campaigns. In order to optimize such actions operators stratify their clients according to their propensity to churn. The process of churn prediction model creation should be automated, because applications reveal the problem of model ageing [3].

Major reasons for voluntary attrition, identified in the literature, are: unsatisfactory service level, unattractive handset portfolio, uncompetitive pricing, poor coverage and bad experiences with call centres [6,11]. This is reflected in standard analytical models forecasting customer churn, where the following predictor domains are used: customer personal socio-economic characteristics, trends

J. O'Shea et al. (Eds.): KES-AMSTA 2011, LNAI 6682, pp. 64–73, 2011.

in number and length of calls over a long period [18], customer's dissatisfaction level measured as number of calls to the mobile carrier's call center [5]. Some modelers employ also minutes of use, balance forward from previous monthly bills, and the tenure of the customer [3]. Generally authors of churn models start with hundreds of variables and choose the most important with the help of some feature selection methods. In recent years clients are analyzed as a social network. Hence, the list of churn predictors includes also some simple network features like connectivity (e.g. number of neighbors) and interconnectivity (e.g. number of churn neighbors connected by an edge) [2].

In our research we focus on customer attrition under *caller party pays* market model. If the charges for the on-net and off-net traffic differ, the optimal choice of operator depends on prices themselves and individual's call structure. Therefore, when a client churns, it might raise the propensity to churn of people calling him, because their call structure changes. We will call such behavior *an induced churn*. Consequently, the problem of identification and monitoring of clients whose churn might lead to high induced churn is important for telecommunication operators.

The objective of this paper is to identify *universal predictors* of the influence of individual customer churn on attrition of other customers. This aim is motivated by the following business scenario. A customer churns due to non-price reasons (for example: poor handset offer, bad experience with customer service). Telecommunication operator wants to estimate how many of his other customers might also churn due to change in their calling costs (induced churn).

We aim to identify measures of customer's position in calls network that are best predictors of induced churn under varying (a) market price structure, (b) call network structure and (c) number of operators. The recognized characteristics will allow telecommunication operators to better identify network users with highest impact on the attrition propensity in their client base. In order to achieve this objective we construct a multiagent model of telecommunication market and simulate churn behavior of customers under different structural assumptions. This differs from standard literature approach, where reported churn prediction models use a single data set, which limits the power of the obtained results to a single market only.

2 Model Architecture

The model consists of two types of agents (1) mobile network subscribers (exactly one mobile phone number) and (2) mobile network operators.

There are n mobile subscribers, who are characterized by the following factors: the operator they are using, a list of unique clients they are calling along with calls' intensity during a certain period of time (e.g. a month). Based on this information, the subscribers are placed in a *calls network* - a weighted directed graph, where the clients are represented by nodes and the intensity of their calls are weights of edges.

Each subscriber i has the ability to calculate his total calling cost for his current operator as well as his potential total calling cost for all other operators. The total calling cost takes into account call frequency (represented as matrix $D = [d_{ij}] \in \Re_{+}^{n \times n}$) and prices of operator k ($c_{k,1}$ for on-net and $c_{k,2}$ for off-net prices). The cost assuming client i chooses operator k is the sum of his on-net and off-net calls volume, multiplied by on-net and off-net prices:

$$cost_i(k) = \sum_{\{j:operator(j)=k\}} d_{ij} c_{k,1} + \sum_{\{j:operator(j) \neq k\}} d_{ij} c_{k,2}$$

where function $operator(j)$ returns the operator number of j^{th} subscriber. Subscribers are aware of network operators being used by their peers at any moment.

The assumption, that the subscribers know the network operators used by their peers, seem to be strong, though e.g. in Polish it is plausible. In particular, in 2010 the four main Polish mobile operators possessed over 95% of the market, and their prefixes for new clients are generally known. These clients might churn carrying their phone number with, whereas the level of such migration is only of 2-3% in a year, so the prefix is very good predictor as to the operator used. Moreover, the subscribers might recognize the operators of their peers basing on the billing list (in postpaid market) or single call cost (in prepaid market).

In each time period of simulation, subscribers are activated in a random order. Upon activation, subscribers optimize the choice of their operator. In order to do this, he calculates his total calling cost for each operator and chooses an operator with the lowest total cost.

The *operator* agent is characterized by his calls tariff: the pair of prices for calling to subscribers in the same network (on-net) and in another network (off-net). Since we consider only the short term market effects, we assume operators keep their tariffs constant during a single simulation. Operators do not perform any optimization of their tariffs, so the on-net and off-net prices are modeled as external parameters.

The *calls network* is modeled as a Small World Network, because such topology approximates call graphs for the whole populations well, see case of Finland [8,14,15]. It is generated by the procedure proposed by [20], which is parametrized by *rewiring probability* μ and *radius* r. Firstly, we create one-dimensional lattice network with radius r. On the lattice network, nodes are equally distributed on a circumcircle, and each node is connected to his nearest neighborhood within radii of r. Hence, each node has $2r$ connections (neighbors on the left and right), which is average degree of Small World Network [16]. Lattice network features high level of clustering, which is typical for telecom networks. Next this lattice network is altered randomly. For each node we divide the links into two groups: left and right. Next we replace the right links with new remote connections with rewiring probability at level μ. Higher value of parameter μ makes the resulting social network more similar to random network than lattice network. In extreme case the calls graph has random network property - the existence of each link is equally probable. Hence, the rewired network is always between lattice network

($\mu = 0$) and random network ($\mu = 1$). Weights of network links are drawn from the exponential distribution, which is parametrized by average call frequency $\lambda = 100$.

Network parameters μ and r are chosen to fit real data taken from anonymous telecommunication operator. Using theoretical probability distribution according to [1], we optimize μ and r to minimize sum square difference between theoretical and empirical probability distribution of neighborhood size. The following objective function is taken:

$$objective(\mu, r) = \sum_{i=0}^{n} (\widehat{f_{\mu,r}}(i) - f_{emp}(i))^2 \rightarrow \min$$

where $\widehat{f_{\mu,r}}(i)$ is the theoretical frequency of clients, having exactly i neighbors, for the network with μ and r parameters. This is compared to empirical frequency $f_{emp}(i)$ from the real data. The estimated parameter values resulting from this procedure are $\hat{\mu} = 1.00$ and $\hat{r} = 3$.

Moreover, we depict the frequency of neighborhood size on Figure 1, as well as, present the descriptive statistics in Table 1. We can observe that the average number of neighborhoods is 7.475, which corresponds to radius r of 3 or 4, since mean neighborhood size in Small World Network is $2r$.

Table 1. Descriptive statistics of neighborhood size.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max	St. Dev.
1	3	5	7.475	9	69	7.83

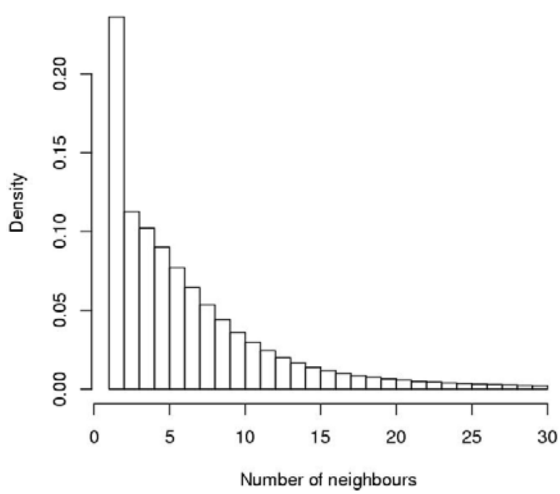

Fig. 1. Histogram of neighborhood size

Following the objectives of the paper we do not want to stick to single network parametrization and we use a range of μ and r parameters values. Such approach allows us to obtain generalizable results and check their robustness.

3 Simulation Results

3.1 Experiment Setup

The parameters of the simulation model are given in Table 2. In our experiment we explore 54 different parameter configurations. For each parameter setting we run the simulation $T = 4$ times. The parameters that are constant for each simulation are: the number of subscriber agents ($n = 1500$), and the parameter for the exponential distribution of the intensity of calls ($\lambda = 100$). The number of operators can be equal to 3 or 5. The price tariffs are defined as a pair of an on-net price and off-net price levels, which is the same for every operator. Normalizing the on-net price to unity, the off-net price can be equal to 1.1, 2 and 3. The experiments are also carried out for different parameter values for the procedure generating the call network. The rewiring probabilities used are 50%, 75% and 100%, and the radius takes the values of 3, 4 and 5. For each simulation a new call network is generated.

The first step of the simulation is to generate a Small World Network for 1500 agents according to the method described in Section 2. Next, the set of measures of customer position in the call network is calculated. The list is given in Table 3. The first three are simple measures of direct links between the subscriber i and its peers. *Degree* is the size of the neighborhood of subscriber i, where the neighborhood is defined as the set of peers who both call and are called by subscriber i. *Prestige* is the number of other subscribers who call subscriber i,

Table 2. Model parameters and their values

Parameter	Simulation values	Meaning
	Telecommunication Market	
n	1500	The number of subscribers in the system
k	$\{3, 5\}$	The number of operators
c	$\forall_i c_{i,1} = 1$ $\forall_i c_{i,2} \in \{1.1, 2, 3\}$	Price tariffs c. c is a $k \times 2$ matrix, where $c_{i,1}$ is on-net price and $c_{i,2}$ is off-net price of operator i
	Telecommunication Network	
μ	$\{50\%, 75\%, 100\%\}$	Rewiring probability
r	$\{3, 4, 5\}$	Radius
λ	100	Average call frequency
D	$\Re_+^{n \times n}$	Distance matrix of call network, containing the information of call time. d_{ij} is call time the caller i talks with the receiver j.
	Simulation Technicality	
T	4	Number of times whole simulation is run for single parametrization

Table 3. Measures of subscriber's position in the call network

Measure	Symbol	Meaning
Degree	DEG	Number of other subscribers - called by and calling the subscriber [12]
Prestige	PRE	Number of other subscribers who call the subscriber [12]
Weighted prestige	WPR	Sum of intensities of other subscribers' calls to this subscriber[12]
Local clustering coefficient	LCC	Degree to which callers of the client consist of a clique [20]
Eigenvector centrality	ECT	Importance of subscriber[9,19]

and *Weighted Prestige* is the total sum of call intensities from other subscribers to subscriber i. The next two measures take into account the whole structure of the call graph and not just direct links between pairs of subscribers. *Local Clustering Coefficient* is the number of links between the subscribers who belong to the neighborhood of subscriber i divided by the total number of possible links, which is $s(s-1)$, where s is the size of the neighborhood. *Eigenvector Centrality* of subscriber i is the i-th coordinate of the eigenvector corresponding to the greatest eigenvalue of the call network matrix. It is the measure of subscriber's centrality within a network.

The experiment is performed in two phases (a) *burn in* period and (b) *analysis* period. In the burn in phase each subscriber is assigned randomly to a network operator and subscribers optimize their call costs by changing the network operators according to the optimization rule described in Section 2. The optimization phase is simulated for 1000 iterations, which is always enough for the operator choice of each agent to become stable.

Next, the analysis period is performed as follows. Starting from the network structure (the choice of operators) obtained after the burn in phase, the following procedure is applied for each subscriber. We switch the operator of a chosen subscriber to some other random operator and force him to remain with that operator. We will call this operation *forced churn* and interpret it as originated by reasons other than calling cost minimization (for instance customer's dissatisfaction with operators offer of handsets).

Afterwards we run the simulation for another 50 iterations with a probability $p = 0.1$ for each agent to optimize their operator choice in a single period. The customers that switch operators during this period are induced churners. We calculate two indices of the level of induced churn. Firstly, we calculate the total number of subscribers who have chosen a different network operator than the one chosen after the burn in period. This index takes into account the number of churners from the operator of the initial forced churner as well as churners from other operators. Therefore we calculate a second measure of induced churn level, which is the number of churners from the initial operator of the forced churner. After this, agents return to their operators chosen after the burn-in period and this step is repeated for next subscriber.

Using the above simulation procedure, for each customer, we obtain his metrics of social position in the call graph and two measures of his induced churn level. For a single parametrization the procedure is repeated T times. Simulations are conducted using MASON Multiagent Simulation Toolkit [13], as well as JUNG and JAMA packages of Java programming language. As a result we obtain a data set for each parametrization of the simulation ($3 * 3 * 3 * 2 = 54$ datasets containing $4 * 1500 = 6000$ observations).

To identify churn predictors we binarize induced churn values. Since for 6000 clients only a low proportion of induced churn measures is greater than 1, we create a binary variable $target = 0$ for *zero churn* and $target = 1$ for *non-zero churn*. Such formulation enables us to employ logistic regression to identify the factors influencing the churn intensity. We also tested an alternative – linear probability model [7] – to check the robustness of the logistic model. The results do not differ significantly.

The following network characteristics have correlation coefficient above 0.7: *prestige*, *weighted prestige* and *degree*. This fact makes its coefficient estimates, in logistic regression, negatively correlated. As a consequence, we abandon both *prestige* and *degree* variables, because their information is nearly contained in *weighted prestige*. In the logistic regression model we estimate the probability that he will induce churn from the following formula:

$$P(target = 1) = \frac{exp(\beta_0 + \beta_1 \mathrm{WPR} + \beta_2 \mathrm{LCC} + \beta_3 \mathrm{ECT})}{1 + exp(\beta_0 + \beta_1 \mathrm{WPR} + \beta_2 \mathrm{LCC} + \beta_3 \mathrm{ECT})},$$

where coefficients are computed from data by Maximum Likelihood Estimator [7]. We work on the standardized data, so that coefficients are comparable with each other and denote the strength of churn intensity of respective variables. Computational routines are conducted in GNU R [17].

3.2 Social Factors Influencing Customer Attrition

The results of performed simulations are given in Table 4 and Table 5.

What seems non-obvious, is the impact of average radius, i.e. average neighborhood size, on the average churning. As we can see in Table 4, the higher radius r, the smaller churning. This is explained by the fact, that in small neighborhoods, each neighbor has greater relative impact on a given client, so the induced churn occurs more often. On the other hand, in larger neighborhoods, the switching of single neighbor might be not sufficient to convince others to change an operator.

The existence of difference between on-net and off-net prices are necessary for churn to take place. Additional simulations were conducted where $c_{i,1} = c_{i,2}$ to confirm that result. The size of this price difference plays role only for the small neighborhood of radius $r = 3$. In Table 4, we see that larger difference coincide with higher churn for $r = 3$. However, this relationship does not hold for larger neighborhood size of $r = 4$ or $r = 5$.

We also identify the impact of the number of operators: the higher the number of operators, the more induced churn we observe. This result should encourage

Table 4. Average number of churned clients induced by subscriber's switching. For each parametrization we present: the overall churning and a ratio of churners for the initially churned customer's operator.

r	μ	$k = 3$			$k = 5$		
		$c_{i,2} = 1.1$	$c_{i,2} = 2.0$	$c_{i,2} = 3.0$	$c_{i,2} = 1.1$	$c_{i,2} = 2.0$	$c_{i,2} = 3.0$
3	50%	3.03(85%)	3.12(87%)	3.24(87%)	3.42(81%)	3.20(82%)	3.54(80%)
	75%	2.43(87%)	2.22(88%)	2.88(86%)	3.88(77%)	3.55(77%)	2.77(79%)
	100%	1.99(85%)	2.71(85%)	3.83(83%)	2.31(79%)	2.31(78%)	2.55(81%)
4	50%	1.10(86%)	1.22(81%)	0.17(100%)	1.11(73%)	1.02(79%)	0.20(100%)
	75%	0.18(100%)	0.16(100%)	0.18(100%)	0.17(100%)	0.16(100%)	0.16(100%)
	100%	0.18(99%)	0.16(100%)	0.19(100%)	0.17(100%)	0.17(100%)	0.15(100%)
5	50%	0.06(100%)	0.72(99%)	0.06(100%)	0.07(100%)	0.06(100%)	0.06(100%)
	75%	0.06(100%)	0.06(100%)	0.07(100%)	0.06(100%)	0.06(100%)	0.06(100%)
	100%	0.07(100%)	0.06(100%)	0.08(100%)	0.07(100%)	0.07(100%)	0.06(100%)

Table 5. Importance of subscribers' network characteristics for different parametrizations. Each of 18 estimated logistic regressions contains coefficients for: weighted prestige, local clustering coef. and eigenvector centrality. The absolute value of each coefficient divided by the largest coefficient in the whole sample gives the normalized measure of variable importance.

r	μ	$c_{i,2}$	$k = 3$			$k = 5$		
			WPR	LCC	ECR	WPR	LCC	ECR
3	0.50	1.1	0.81	0.04	0.12	0.98	0.02	0.02
		2.0	0.84	0.03	0.06	0.94	0.01	0.02
		3.0	0.80	0.01	0.03	0.90	0.02	0.05
	0.75	1.1	0.75	0.01	0.10	1.00	0.01	0.02
		2.0	0.69	0.01	0.17	0.95	0.03	0.03
		3.0	0.73	0.01	0.06	0.74	0.05	0.00
	1.00	1.1	0.66	0.03	0.09	0.66	0.01	0.01
		2.0	0.71	0.01	0.38	0.64	0.01	0.04
		3.0	0.96	0.01	0.05	0.75	0.02	0.14

telecommunication regulators to enable new operators to enter the market and boost competition.

Table 4 also shows that the lower rewiring probability μ, the higher average churning rate. Low μ makes more cliques to occur in the network, which could contribute to the domino effect, i.e. one churned client will result in a few more churn events. Finally we note that the majority of induced churn events happen for the initially churned customer's operator.

Next let us analyse the role of client's social network position on induced churn, see Table 5. Variable *weighted prestige* has the greatest impact on induced churn. This holds for all network parametrizations and is statistically significant at the level of $\alpha = 1\%$. However, we observe higher importance of this variable for small clustered neighborhood, i.e. with $r = 3$ and $\mu \in \{0.5, 0.75\}$. *Eigenvector centrality* is found important just occasionally. We found no significant impact of *local clustering coefficient*. The same calculations were performed for $r = \{4, 5\}$, resulting in similar conclusions.

4 Conclusions

The problem of monitoring induced churn is crucial for the maintenance of profitable customer retention policy by telecommunication operators. Next to the typical factors like: customer service satisfaction, prices or device availability, the customers choosing the operator consider also their individual call structure. Therefore its changes might increase churn level.

We have shown that churn induced by switching of a single customer is the most significant issue for markets with large number of operators and call networks with low radius. At individual customer level we have found that weighted prestige is a good predictor of induced churn independent of simulation parametrization.

The presented findings are important for shaping the telecommunication operators marketing policies. Using weighted prestige measure they can identify the clients whose churn can cause most induced churn. In such situation operator will be willing to invest into ensuring that such customers do not switch their provider.

In further research it is interesting to test a broader range of customer's network position characteristics, such as: betweenness or closeness and check robustness of our results through implementing different subscribers' operator choice rules.

Acknowledgments

This research was supported by Warsaw School of Economics Research Grant number 03/S/0025/10.

References

1. Barrat, A., Weigt, M.: On the properties of small-world network models. The European Physical Journal B - Condensed Matter and Complex Systems 13(3), 547–560 (2000)
2. Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S., Nanavati, A.A., Joshi, A.: Social ties and their relevance to churn in mobile telecom networks. In: EDBT 2008: Proceedings of the 11th International Conference on Extending Database Technology, New York, NY, USA, pp. 668–677 (2008)
3. Datta, P., Masand, B., Mani, D.R., Li, B.: Automated Cellular Modeling and Prediction on a Large Scale. Artificial Intelligence Review 14, 485–502 (2000)
4. Furnas, G.: Framing the wireless market. The Future of Wireless, WSA News:Bytes 17(11), 4–6 (2003)
5. Hadden, J., Tiwari, A., Roy, R., Ruta, D.: Churn prediction using complaints data: Proc. of world academy of science. Engineering, and Technology 13, 158–163 (2006)

6. Hadden, J., Tiwari, A., Roy, R., Ruta, D.: Computer assisted customer churn management: State-of-the-art and future trends. Computers & Operations Research 32, 2902–2917 (2007)
7. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Data Ming Inference and Prediction. Springer, Heidelberg (2008)
8. Hidalgo, C., Rodriguez-Sickert, C.: Persistence, Topology and Sociodemographics of a Mobile Phone Network. Technical report, Center for Complex Network Research, Department of Physics, University of Notre Dame (2007)
9. Katz, L.: A new status index derived from sociometric analysis. Psychometrika 18, 39–43 (1953)
10. Kazienko, P., Bródka, P., Ruta, D.: The Influence of Customer Churn and Acquisition on Value Dynamics of Social Neighbourhoods. In: Lytras, M.D., Damiani, E., Carroll, J.M., Tennyson, R.D., Avison, D., Naeve, A., Dale, A., Lefrere, P., Tan, F., Sipior, J., Vossen, G. (eds.) WSKS 2009. LNCS, vol. 5736, pp. 491–500. Springer, Heidelberg (2009)
11. Kim, H., Yoon, C.: Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. Telecommunications Policy 28, 751–765 (2004)
12. Knoke, D.: Social Network Analysis in Quantitative Applications in the Social Sciences. Sage Publications, Inc., Thousand Oaks (2007)
13. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G.: MASON: A Multiagent Simulation Environment. Simulation 81(7), 517–527 (2005)
14. Newman, M.: The Structure and Function of Complex Networks. SIAM Review 45, 167–256 (2003)
15. Onnela, J., Saramaki, J., Hyvonen, J., Szabo, G., Lazer, D., Kaski, K., Kertesz, J., Barabasi, A.: Structure and Tie Strengths in Mobile Communication Networks. Proc. Natl. Acad. Sci. USA 104, 7332–7336 (2007)
16. Vega-Redondo, F.: Complex Social Networks. Cambridge University Press, Cambridge (2007)
17. R Development Core Team: R: A Language and Environment for Statistical Computing. Foundation for Statistical Computing, Vienna, Austria (2010)
18. Richter, Y., Yom-Tov, E., Slonim, N.: Predicting customer churn in mobile networks through analysis of social groups. In: SIAM Int. Conf. on Data Mining (2010)
19. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge (1994)
20. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature 393(6684), 409–410 (1998)
21. Yan, L., Miller, D.J., Mozer, M.C., Wolniewicz, R.: Improving prediction of customer behaviour in non-stationary evnironments. In: Proc. of International Joint Conference on Neural Networks, IJCNN, vol. 3, pp. 2258–2263 (2001)