The Impact of Customer Churn on Social Value Dynamics

Przemysław Kazienko
Institute of Computer Science, Wrocław University of Technology
Wyb.Wyspiańskiego 27, 50-370 Wrocław, Poland
BT Innovate, British Telecom Group, Intelligent Systems Research Centre (ISRC), Orion 1/12G, Adastral Park, IP5 3RE Ipswich, UK
Email: przemyslaw.kazienko@pwr.wroc.pl
Dymitr Ruta
BT Innovate, British Telecom Group, Intelligent Systems Research Centre (ISRC), Orion 1/12G, Adastral Park, IP5 3RE Ipswich, UK
Email: dymitr.ruta@bt.com
Piotr Bródka
Institute of Computer Science, Wrocław University of Technology
Wyb.Wyspiańskiego 27, 50-370 Wrocław, Poland
Email: piotr.brodka@pwr.wroc.pl

Abstract: Modern telecommunication service providers implicitly create interactive social networks of individuals that both depend on and influence each other through complex social relationships grown on friendship, shared interests, locality etc. While delivering services on the individual basis, the network effects exerted from customer-to-customer interactions remain virtually unexplored and unexploited. The focus of this paper is on customer churn, where social network effects are widely ignored yet may play a vital role in revenue protection. The key assumption made is that a value loss of a churning customer extends beyond his revenue stream and directly affects interaction within local neighbourhoods. The direction and strength of this impact are evaluated experimentally by direct measurements of the total neighbourhood value of the churning customer along with other standard social network measures taken before and after the churn event.

Keywords: social networks, social network analysis, network dynamics, social position, churn, social value, social neighbourhood

1. Introduction

Social networks are one of possible representation of human communities, in which people interact and get into relationships with one another. These relationships are usually very complex and engage our feelings, emotions, likes and dislikes, etc. Besides, relationships in the social network result from cooperation and work or family dependencies. Simultaneously, social networks evolve: they change their structure; new communities arise while the others disappear; some relationships reinforce while the other abate [18].

In the real world, people depend on each other. Our choices and behaviour also influence behaviour of the others [7]. This is the crucial concept of recommender networks [16, 19] and plays an important role in marketing [17], in which people spread information and opinion about products through mutual, personal contacts.

Ability to predict changes and their consequences is crucial in every business. For that reason, dynamic analysis within the customer network especially in the telecommunication social network is very important. General concept of analysis of dynamic social networks was presented in [1]. In order to forecast such changes and investigate the evolution of social networks even physics and molecular modelling can be utilised [10]. In some other approaches, clustering [5], statistical analyses and visualizations [1] or multi agent systems [2, 21] are used to get an insight into network dynamics. Daspupta *et al.* tried to predict churn based on the analysis of relationship strength in the mobile telecommunication social network [4], whereas Gopal and Meher used typical prediction method – regression to estimate churn time and tenure for the same domain [9].

This paper tries to answer the following question: does our behaviour, as the customers, influence the others and are we able to evaluate this influence based on the available data about mutual contacts or not? In particular, we analyse the influence of churning customers on their neighbourhoods after the churn.

2. Telecommunication Social Network

2.1. Concept of the Telecommunication Social Network

Telecommunication data contains information about some customer activities, namely phone calls and each call can be treated as the evidence of mutual relationship [15].

A telecommunication social network *TSN* is the tuple $TSN=(M,R)$ that consists of the finite set of members (customers, nodes) *M* and the set of relationships *R* that join pairs of distinct members: $R=\{r_{ij}=(x_i,x_j)\colon x_i\in M,\ x_j\in M,\ i\neq j\}$. Note that relationships in *TSN* are directed, i.e. $r_{ij}\neq r_{ji}$.

In practice, a single member corresponds to one phone number, which, in turn, is assigned to one social entity. In case of telecommunication, a social entity can be either residential (a family, single person) or business (a company, department in the organisation, a position or single employee in the organization).

2.2. Strength of Relationship

In order to calculate strength of relationship between two network members, two variants of strength of relationship were utilized. The first one – *RN* takes into consideration number of phone calls, whereas the second one *RD* – their total duration.

The former is calculated in the following way:

$$RN(y\rightarrow x)=\begin{cases}\dfrac{N(y\rightarrow x)}{\sum_{x\in M}N(y\rightarrow x)}, \text{ when} \sum_{x\in M}N(y\rightarrow x)>0 \\ 0, \quad \text{when} \sum_{x\in M}N(y\rightarrow x)=0\end{cases}, \tag{1}$$

where: $N(y\rightarrow x)$ – the number of phone calls that member *y* initialized to member *x*.

Function $RN(y\rightarrow x)$ reflects the commitment of member *x* within *y*'s activities (performed calls). $RN(y\rightarrow x)\in[0;1]$. If member *y* has not made any calls to member *x*, $RN(y\rightarrow x)=0$.

Another function of relationship strength $RD(y\rightarrow x)$ from member *y* to *x* is calculated based on a total duration of calls, as follows:

$$RD(y\rightarrow x)=\begin{cases}\dfrac{T(y\rightarrow x)}{\sum_{x\in M}T(y\rightarrow x)}, \text{ when} \sum_{x\in M}T(y\rightarrow x)>0 \\ 0, \quad \text{when} \sum_{x\in M}T(y\rightarrow x)=0\end{cases}, \tag{1}$$

where: $T(y\rightarrow x)$ – the duration of all phone calls that member *y* made to member *x*.

Function $RD(y\rightarrow x)$ also corresponds to the commitment of *y*'s activities towards member *x* but the measure of these activities is duration of calls.

3. Social Value of Customers and Their Neighbourhood

3.1. Social Values of Nodes

Measures (also called metrics) are used in social network analysis (SNA) to describe human profile, specific for the given social network as well as to indicate personal importance of individuals in the community.

There are some common structure measures that reflect centrality rate of the node: centrality degree (*CD*), centrality outdegree (*OD*) and Social Position (*SP*).

Centrality degree is the simplest and the most intuitive measure among all others. It is the number of links that directly connect one node with others [3]. In the undirected graph, it is the number of edges which are connected with the single node. In the directed graph like *TSN*, degree can be divided into indegree (*ID*) – for edges, which are directed to the given node, and outdegree (*OD*) – for edges, which are directed from the given node [22].

The enhanced versions of *CD* and *OD* were used in this paper. Since every single edge has its own weight, function $CD(x)$ is the sum of weights of all edges directly connected to given node x. For strength function $RN(y\rightarrow x)$, centrality degree $CD(x)$ is computed as follow:

$$CD(x) = \sum_{y \in M} RN(x \rightarrow y) + \sum_{y \in M} RN(y \rightarrow x) \,. \tag{3}$$

In fact, the value of $CD(x)$ for strength *RN* equals to the number of all inbound and outbound calls of the member x. Centrality degree for duration is calculated in the similar way: $RD(y\rightarrow x)$ is used instead of $RN(y\rightarrow x)$.

Centrality outdegree $OD(x)$ is the sum of weights of all edges outgoing from node x. For strength function based on duration, $OD(x)$ is evaluated using:

$$OD(x) = \sum_{y \in M} RD(x \rightarrow y) \,. \tag{4}$$

Similarly to $CD(x)$, centrality outdegree $OD(x)$ is simply the total duration of all calls initialized by member x.

Hence, the social value and importance of the member x – $SV(x)$ in the telecommunication social network *TSN* can be derived from the member x's activities, i.e. the number of calls, their duration, frequency, etc., see Eq. (1) and (2). However, it can also reflect the position of other members which are incident to the given one as well as the strength of relationship between member x and his neighbours. In other words, social value of the node may be described by the value of its neighbours and by the fact how important this particular member is for its neighbours [11].

Social position $SP(x)$ is a measure proposed and developed in [11, 13]. It can be used to calculate the importance of every single member of the network in an iterative way:

$$SP^{(n+1)}(x) = (1-\varepsilon) + \varepsilon \cdot \sum_{y \in M} SP^{(n)}(y) \cdot C(y \rightarrow x) \,, \tag{5}$$

where:

$SP^{(n+1)}(x)$, $SP^{(n)}(x)$ – social position of node x after n+1th or nth iteration;

ε – the fixed coefficient from the range (0;1);

$C(y\rightarrow x)$ – the commitment function which expresses the strength of the relation from member y to x.

The constant ε represents the openness of human social position on external influences, in other words high ε means that the social position is highly influenced by others and low ε means that the social position is more static while others' influence is week [12, 13].

The convergence of $SP^{(n)}(x)$ as defined in Eq. (5) was proven regardless of initial values $SP^{(0)}(x)$ in [14].

The importance of a member described by social position $SP(x)$ depends on the social positions of the neighbourhood and the strength of relationship from x's members to member x. More precisely, member x's social position is inherited from x's neighbours which activity is directed to member x and level of inheritance strictly depends on strength of this activity. The activity strength of one member absorbed by another is expressed in Eq. (5) by commitment $C(y \rightarrow x)$. Furthermore, commitment function can be assigned to each relationship based on either duration or frequency (number) of calls. Thus, many different functions can be used as commitment $C(y \rightarrow x)$, including $RN(y \rightarrow x)$, Eq. (1) and $RD(y \rightarrow x)$, Eq. (2). However, one additional restriction has to be applied to them. If member x has no connection to anybody within the network TSN, but there are $k>0$ other members y_i ($i=1, ..., k$) who are connected to member x, the commitment function $C(x \rightarrow y_i)$ should be equally distributed among all members y_i, i.e. $C(x \rightarrow y_i)={}^{1}/_{k}$, for each such member y_i [12]. Note that for all but isolated nodes $\forall\left(x \in M\right)\sum_{y \in M} C(x \rightarrow y) = 1$.

Many different measures known from social network analysis and describing single node in the network can be used to reflect social value of member x – $SV(x)$. Three of them were utilized in this paper, namely $CD(x)$, $OD(x)$ and $SP(x)$. First two, CD and OD, were chosen for experiments because they are the most efficient to calculate. Besides, outdegree centrality OD directly reflects the income related to the member. However, CD and OD take into consideration only the first level neighbours, whereas SP depends, in the indirect way, on the positions of all other members. For that reason, SP was used in experiments. Each of these measures can use different strength functions, e.g. based on a number of calls – $RN(y \rightarrow x)$ or their duration – $RD(y \rightarrow x)$.

3.2. Neighbourhoods

The neighbourhood of member x is the set $N(x)$ of members y_i which are directly connected to member x. In other words, it is a set of members with which x maintains the closest relationships. We can expect that the members from $N(x)$ have the biggest influence on member x and member x has big influence on them.

Social value of the neighbourhood of member x – $SVN(x)$ is the sum of social values $SV(y)$ of all x's neighbours y:

$$SVN(x) = \sum_{y \in N(x)} SV(y). \tag{6}$$

Note that the neighbourhood does not include member x. Furthermore, also other churning nodes y were excluded from set $N(x)$, see Fig. 4.

In the telecommunication business, members and separately their neighbours can belong to various classes. Two of them are usually distinguished: residential (individuals and their families, acquaintances, friends, etc.) or business (a company, units in the organisation and their collaborators).

4. Network Effect of Customer Churn

4.1. Churn

Today's global telecommunication market environment can be characterized by the strong competition among different telecoms and a decline in growth rate due to maturity of the market. Furthermore, there is a huge pressure on those companies to make healthy profits and increase their market shares. Most telecom companies are in fact customer-centric service providers and offer to its customers a variety of subscription services. One of the major issues in such environment is customer churn known as a process by which a company loses a customer to a competitor. Recent estimates suggest that churn rates in the telecom industry could be anything between 25% to 50% [8]. Moreover on average it costs around $400 to acquire a new customer which takes years to recoup [8]. These huge acquisition costs are estimated to be between 5 to 8 times higher than it is to retain the existing customer by offering him some incentives [23]. In this competitive and volatile environment, it makes therefore every economic sense to have a strategy to retain customers which is only possible if the customer intention to churn is detected early enough [20].

There are many different reasons for customers to churn, some of them, like moving home, unstoppable, other, like sudden death, undetectable. Any churn prevention systems should therefore focus on detecting those churners that are deliberately moving to a competitor as these customers are most likely to leave data traces of their intent prior to churn and can be potentially persuaded to stay [20].

This work, however, is not concerned with churn prediction (like for example in [4], [9] and [20]) nor the effectiveness of actual actions preventing churn. What it is trying to deliver is the preliminary intelligence about the impact of customer churn on the dynamics of a service value within a social neighbourhood of the churning customer, such that a decision to retain or rescue a churning customer can be better aligned to the potential value impact. It is reasonable to assume that a churn of an active network member may have some impact on his direct and indirect network neighbours that could range from fading, redirected or reinvigorated activity up to the follow-up churn in extreme cases. From the business point of view the importance of this change lies in individual changes of the neighbours' value streams and can be considered within a generic context of social value and its dynamics

4.2. Dynamics of Social Values

Let us consider a social network of customers interacting by means of telephone calls provided by a telecom service provider. Each such customer established his local social network consisted of customers whom he called or who called him at least once during his lifetime. We would refer to such customers the first-level neighbours with respect to the customer in question as depicted in Fig 1. Since the acquisition each customer generates a dynamic value stream consisting of a value of his outbound calls as well as value added network component stemming from the fact that his presence drives inbound calls from his neighbours. We refer to such network value added component as social value of a customer.

While a dynamics of customer value stream is explicitly evident in his outbound calls that translate into telephone bills, the social value remains implicit and is hidden from direct observations. One naïve way of estimating the social value of a customer is to periodically measure the value of his inbound calls. The problem with this method is that it is unclear to which degree the calls made by customer neighbours are driven by the customer presence, or in other words it is not clear if under customer absence his neighbours would call less, redirect calls from the absent customer to other customers or perhaps even get stimulated to grow their neighbourhood and call more.

Customer churn gives a realistic opportunity to evaluate the social value of a customer and explore its dynamics over time. By comparing the value of customer neighbourhood before and after the churn one can truly estimate the impact of churning customer on the change in neighbourhood value which is equivalent o the social value of churning customer. Such experimental scenario is illustrated in Fig. 1 and Fig. 2.

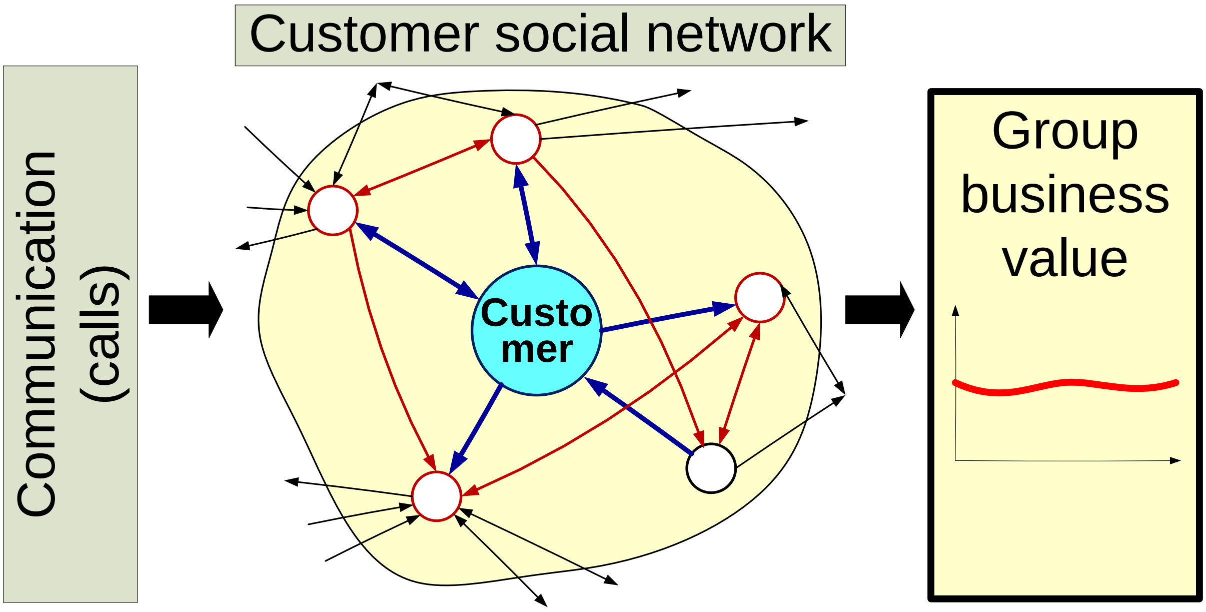

Fig. 1. A customer and their first-level (nearest) neighbours

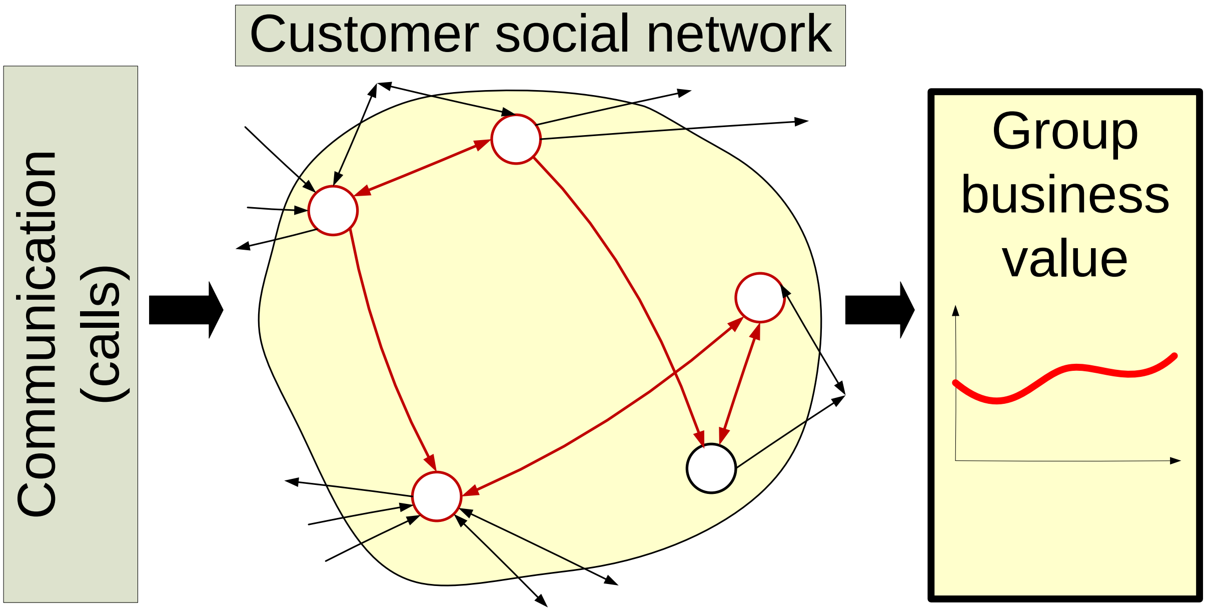

Fig. 2. The neighbourhood of the churning customer

It is important to remember that the neighbourhood value may continue to change at different rates well after the churn event until the new equilibrium is achieved. Moreover, it might be very difficult to extract a direct impact of particular customer's churn on his neighbourhood value as there may be many other concurrent drivers of value dynamics like other customers' churn, acquisition, customer moves and other significant network events. From the global perspective, all these additional processes impacting social value dynamics happen continuously anyway and are part of the ongoing network value fluctuations; hence their impact should be statistically similar before and after the churn event.

Note that the process of analyzing the impact of churn on social value dynamics is directly reverse to the process of analyzing the impact of customer acquisition hence the two can be analyzed together possibly even sharing similar observations and conclusions. A diagram illustrating such process is shown in Fig. 3.

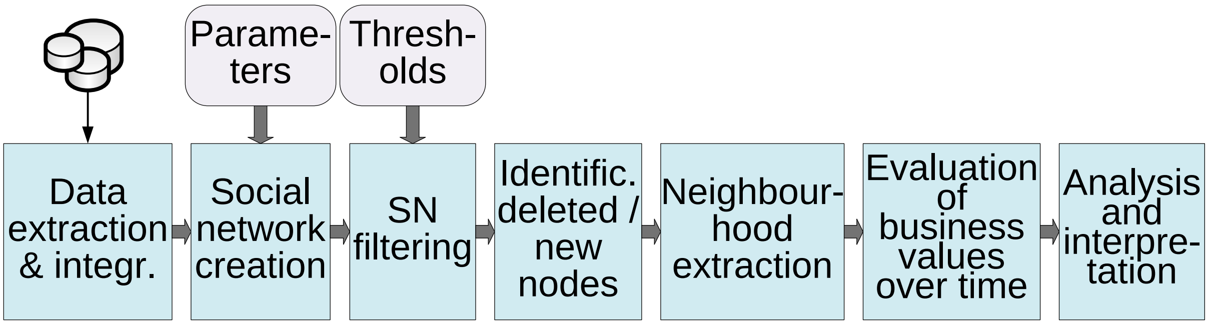

Fig. 3. Process of analysis of social neighbourhood for churning and acquired customers

The first part of this process is the identification of relationships in social network which allows establishing the neighbourhood of any particular customer. The next step is finding customers for which we want to analyse the social value dynamics, and those would be the customers who churn or are acquired preferable during the middle part of the period the analysis is conducted for. Then the key part involves establishing the neighbourhoods of such customers (prior- for churn and post- for acquisition) and measure the time series of their total business values from before the event until the point after the event for which the neighbourhood value time series attains again stationarity. Note that as illustrated in Fig. 4, the value of churned or acquired customer is excluded from the neighbourhood value both before and after the churn/acquisition event and thus social value of a customer is a measure of customers ability to drive business value from the rest of the social network

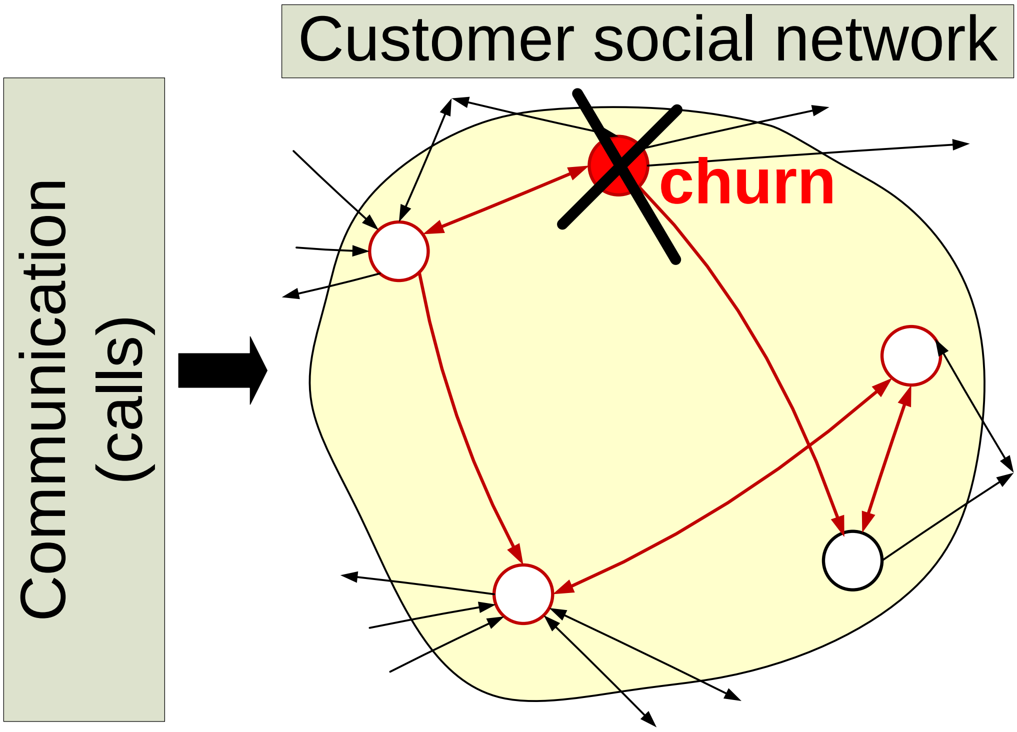

Fig. 4. Process of analysis of social neighbourhood for churning and acquired customers

5. Experiments

The experiments were performed for the real telecommunication social network *TSN* with several dozens of thousand of regular members and a few hundred of churning members in line with the process shown in Fig. 3. The data came from 1-month period that has been split into three 10-day slots. The second slot was used to identify churning members, the first one to extract their neighbourhoods and calculate neighbourhood values “Before”, whereas the last slot was utilised to evaluate values “After” and relative change: “After” compared to “Before”.

The following social values were used during experiments: social position *SP* for number of calls, centrality degree *CD* for both number of calls and duration of calls and centrality outdegree *OD* also for number of calls and duration of calls.

The average social position of the churning nodes turned out to be about 40% lower than the average social position of all other network members (Fig. 5.). It may indicate that churning customers lower their activities within the network just before they actually leave the network. Additionally, weaker social position of the churning customer also affects his neighbourhood although the average neighbours’ social position is only 2.3% smaller than the average neighbour of the non-churning customer. However, if we look into the dynamics of social position over the three time slots, the impact of churn on the entire neighbourhood is becoming more apparent.

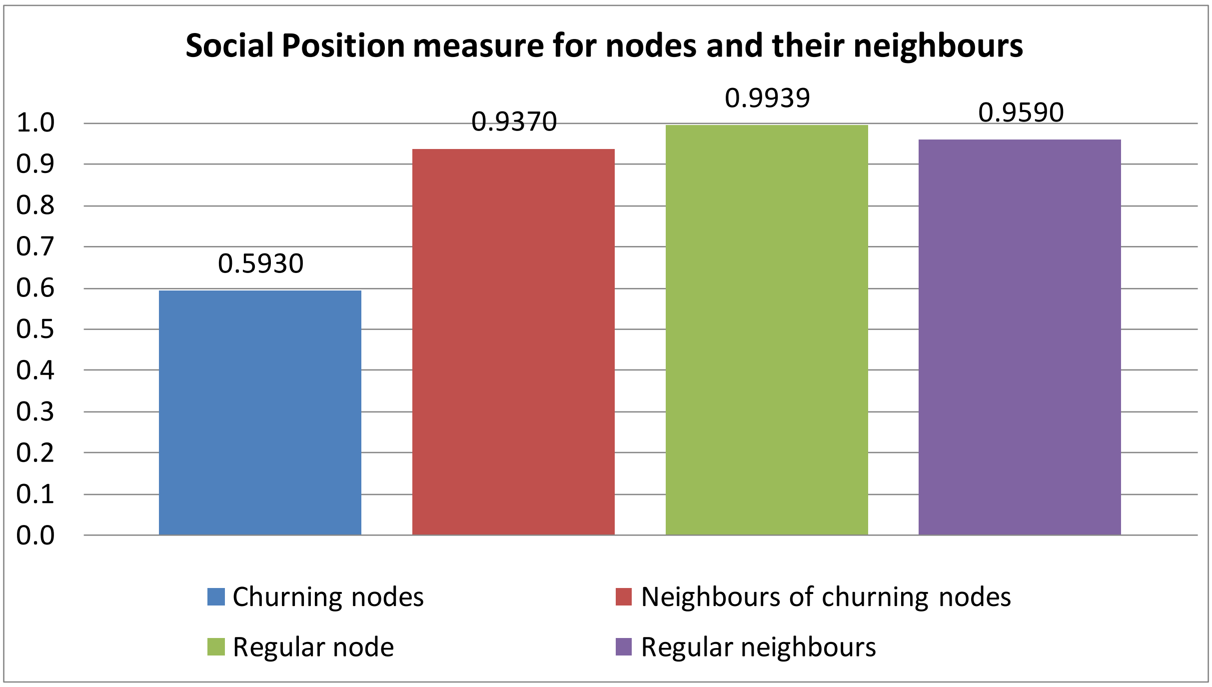

Fig. 5. Average social position value for members in the first time slot

The average social position of churner’s neighbourhood fell by 23% after the churn event compared to regular members as shown in Fig. 6. This reveals how much individual customer churn affects his neighbourhood. Due to customer churn, the network not only loses the member but also the member’s neighbours activities deteriorate within the network. Before these results were obtained, the trend, reflecting general change of customer activities between the third and the first period, was removed.

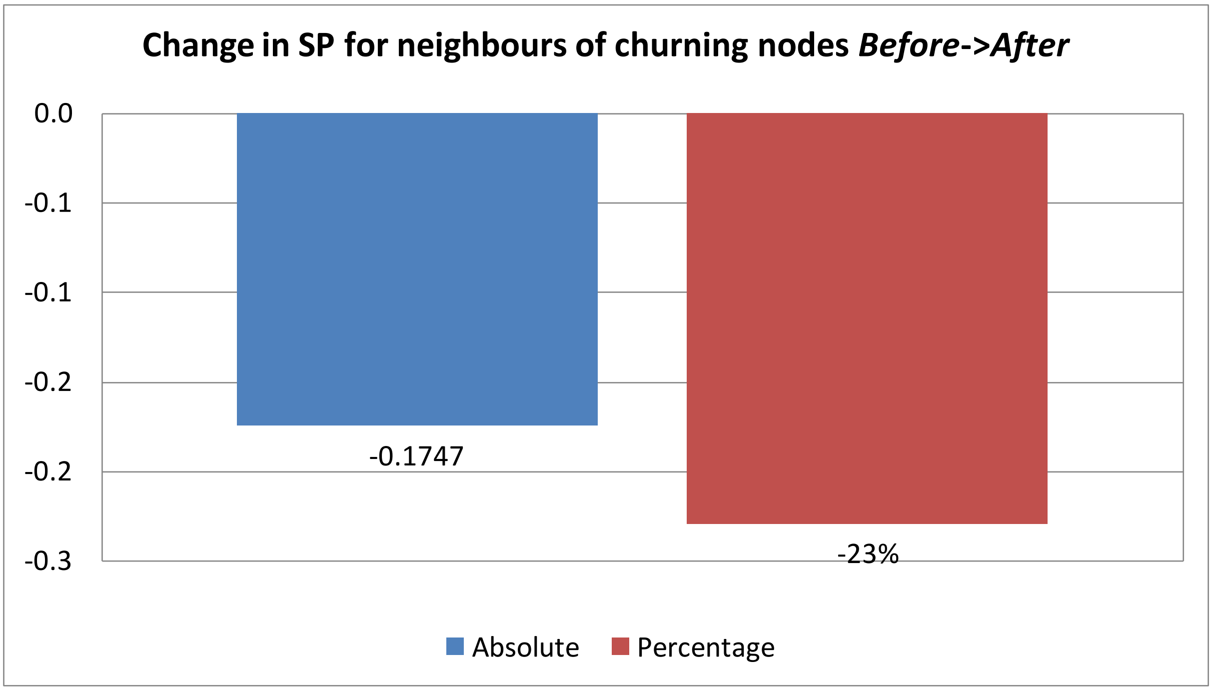

Fig. 6. Average change of neighbourhood social position value for churning members compared to regular members; the third time slot compared to the first one

Similar studies were carried out for centrality degree *CD* and centrality outdegree *OD*, as illustrated in Fig. 7. Neighbours of churning members were compared to neighbours of regular nodes, before and after the churn, for the third and the first time slot. The trend from Fig. 6 was confirmed by *CD* and *OD* based on the number of calls, while the duration of calls increased after the churn.

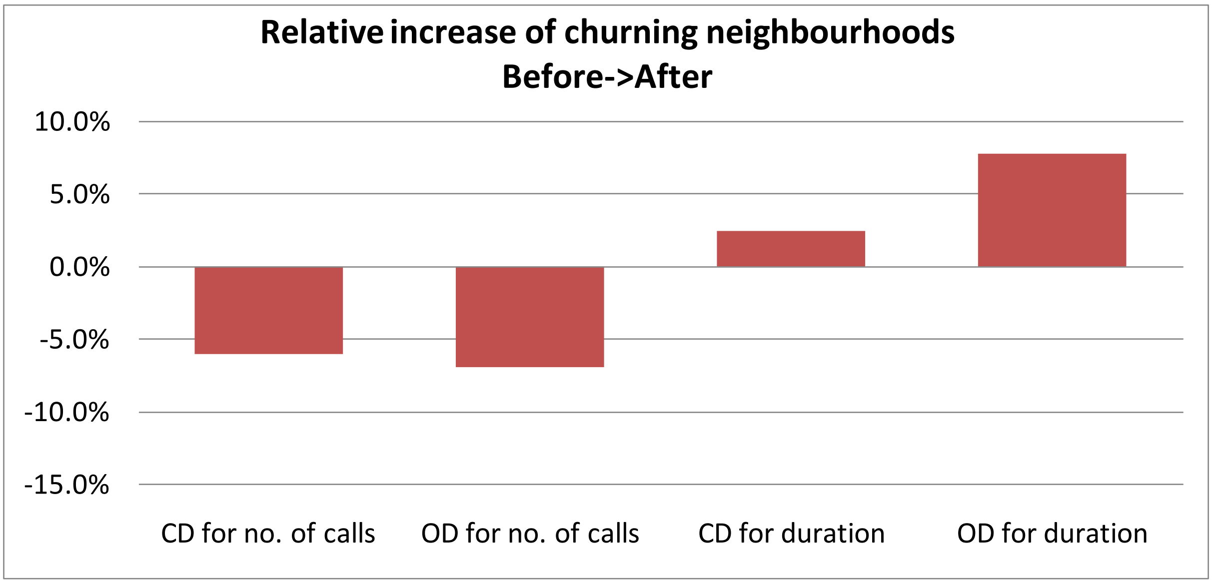

Fig. 7. Average change of neighbourhood value for churning members compared to regular members; the third time slot compared to the first one

6. Conclusion and further study

People influence one another and this principle can be used to analyse and understand customer behaviour especially in retail companies. This impact can be observed by means of social network analysis and changes in social value of nearest neighbours.

The preliminary experiments presented in the paper revealed that the churning customers influence their neighbourhoods. In particular, social position of the neighbours drops significantly after the churn. Smaller values of social position also point to the churning customers before they churn.

However, to indicate why it happens and to verify the extent of the presented phenomena, some additional studies on larger data sets stretching along longer periods are necessary. Once these observations are validated and formally described the next step could be to try to predict the magnitude of the change in network's activity based on the local properties of the node that triggers the change.

7. References

[1] Berger-Wolf T. Y., Saia J.: *A Framework for Analysis of Dynamic Social Networks*, Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, August 20-23, 2006, Philadelphia, PA, USA

[2] Bocalettia S., et.al.: *Complex networks: Structure and dynamics*, Physics Reports 424 (2006) 175 – 308.

[3] Carrington P., Scott J.: Wasserman S.: *Models and methods in Social Network Analysis*, Cambridge University Press, 2005.

[4] Dasgupta K., Singh R., Viswanathan B., Chakraborty D., Mukherjea S., Nanavati A.A.: *Social ties and their relevance to churn in mobile telecom networks*. Proc. of the 11th International Conference on Extending Database Technology: Advances in Database Technology, EDBI'08, March 25-30, 2008, Nantes, France, ACM Press, 2008, 668-677.

[5] Ebel H., Davidsen J., Bornholdt S.: *Dynamics of social networks*. Complexity 8 (2), 2002, 24-27.

[6] Falkowski T., Bartelheimer J., Spiliopoulou M., *Mining and Visualizing the Evolution of Subgroups in Social Networks*, Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, p.52-58, December 18-22, 2006

[7] Fowler J. H., Christakis N. A.: *Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study*, BMJ 2008; 337:a2338.

[8] Furnas, G.: *Framing the wireless market. The Future of Wireless*, WSA News:Bytes 17(11), 2003, 4-6.

[9] Gopal R. K., Meher S. K.: *Customer Churn Time Prediction in Mobile Telecommunication Industry Using Ordinal Regression*. 12th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, PAKDD 2008, Osaka, Japan, May 20-23, 2008, LNCS 5012, Springer, 2008, 884-889.

[10] Juszczyszyn K., Musiał A., Musiał K., Bródka P.: *Molecular Dynamics Modelling of the Temporal Changes in Complex Networks*, , submitted to CEC'09

[11] Kazienko P., Musiał K., Zgrzywa A.: *Evaluation of Node Position Based on Email Communication*, Control and Cybernetics, 2009, in press.

[12] Kazienko P., Musiał K.: *Assessment of Personal Importance Based on Social Networks*, The 6th Mexican International Conference on Artificial Intelligence, November 4-10, 2007, Aguascalientes, Mexico, Springer Verlag, LNAI 4827, 529-539.

[13] Kazienko P., Musiał K.: *On Utilising Social Networks to Discover Representatives of Human Communities*, International Journal of Intelligent Information and Database Systems, Special Issue on Knowledge Dynamics in Semantic Web and Social Networks, Vol. 1, Nos. 3/4, 2007, pp. 293-310

[14] Kazienko P.: *Associations: Discovery, Analysis and Applications*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2008.

[15] Kazienko P.: *Expansion of Telecommunication Social Networks*, The Fourth International Conference on Cooperative Design, Visualization and Engineering, CDVE 2007, September 16-20, 2007, Shanghai, China, Springer Verlag, LNCS 4674, 2007, 404-412.

[16] Kempe D., Kleinberg J. M., Tardos E.: *Maximizing the spread of influence through a social network*, The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2003, Washington, DC, USA, August 24 - 27, 2003, ACM Press, 2003, 137-146

[17] Leskovec J., Adamic L. A., Huberman B. A.: *The dynamics of viral marketing*, ACM Transactions on the Web 1(1), 2007.

[18] Leskovec J., Backstrom L., Kumar R., Tomkins A.: *Microscopic evolution of social networks*, Proc. of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008, ACM Press, 2008, 462-470.

[19] Leskovec J., Singh A., Kleinberg J. M.: *Patterns of Influence in a Recommendation Network*, Advances in Knowledge Discovery and Data Mining, 10th Pacific-Asia Conference, PAKDD 2006, Singapore, April 9-12, 2006, Lecture Notes in Computer Science LNCS 3918, Springer, 2006, 380-389.

[20] Ruta D., Adl C., Nauck D.: *Data Mining Strategies for Churn Prediction in Telecom Industry*, in Hsiao-Fan Wang (eds) Intelligent Data analysis: Developing New Methodologies Through Pattern Discovery and Recovery. IGI Global, New York, 2008, pp 218-235.

[21] Schweitzer F.: Brownian *Agents and Active Particles – Collective Dynamics in the Natural and Social Sciences*, Sporinger Series in Synergetics, New York, 2007.

[22] Scott J.: *Social network analysis: A handbook (2nd ed.)*, London: SAGE Publications Ltd, 2000.

[23] Yan L., Miller D.J., Mozer M.C., Wolniewicz R.: *Improving prediction of customer behaviour in non-stationary environments*. Proc. of International Joint Conference on Neural Networks, IJCNN’01, Vol.3, 2001, 2258-2263.