Accepted Manuscript

A social model based on customers' profiles for analyzing the churning process in the mobile market of data plans

Marcos Postigo-Boix, José L. Melús-Moreno

Please cite this article as: M. Postigo-Boix, J.L. Melús-Moreno, A social model based on customers' profiles for analyzing the churning process in the mobile market of data plans, *Physica A* (2018), https://doi.org/10.1016/j.physa.2017.12.121

*Highlights (for review)

A social model based on customers' profiles for analyzing the churning process in the mobile market of data plans

Marcos Postigo-Boix[*], José L. Melús-Moreno

Department of Telematics Engineering, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain
marcos.postigo@entel.upc.edu and teljmm@entel.upc.edu
[*]Corresponding author

Highlights

- We use an Agent-Based Model to analyze the churning process.
- The model uses demographic and psychographic features.
- The model uses usage profiles according to the users' social behavior.
- We consider users' profiles and homophily to create social connections.
- We show that customers with greater tendency to churn due to the influence of their social networks can be identified better.

A social model based on customers' profiles for analyzing the churning process in the mobile market of data plans

Marcos Postigo-Boix[*], José L. Melús-Moreno

Department of Telematics Engineering, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain
marcos.postigo@entel.upc.edu and teljmm@entel.upc.edu
[*]Corresponding author

Abstract

Mobile Network Operators (MNOs) present wireless services of the same kind in identical zones, clients select the service taking into account any element they consider relevant. Churning hits on the design of the network and the method to assign prices by MNOs, and of course their earnings. Therefore, MNOs try to reduce churn detecting potential churners before they leave the service. Our approach to churn prediction considers each customer individually. Previous research shows that members of the social circle of a subscriber may influence churn. Thus, many scenarios that describe social relations, and in which churning processes could be expected, set an emerging challenge with practical implications. This paper uses the Agent-Based Modelling (ABM) technique to model customers. The model's parameters include demographic and psychographic features as well as usage profiles according to their social behavior considering their customers' profiles. Our model modifies and extends an existing real social network generator algorithm that considers customer's profiles and homophily considerations to create connections. We show that using our approach, groups of customers with greater tendency to churn due to the influence of their social networks can be identified better.

Keywords: *social network, homophily, customer's profile, churn, customer model, Agent-Based Model*

1 Introduction

Developed and developing countries have evolved substantially over the last years using new wireless technology advances, permitting them a significant diffusion of these technologies and thus a more affordable access to these services [1], [2], [3]. In a market, where a variety of MNOs present wireless services of the same kind in identical zones, clients have the freedom to select the service according to any element they consider relevant. Therefore, MNOs can no longer rely only on offering good QoS (Quality of Service) and novel services, but also have to resort to aggressive customer-centric targeted marketing campaigns, special offers, bundling services, among other strategies to position themselves in the market and make a profit. In addition to this competition

among MNOs, easy and rapid number portability has led to a phenomenon called churning where customers migrate from one service provider to another looking for a service that better satisfies their needs.

Customers churn for a variety of reasons. Competitive pricing, network service quality, discounts, and promotions are some of the primary reasons [4]. Some of the reasons that make customers churn are all those related to campaigns of marketing, availability or quality of the received service, other kinds of complaints, etc.

Churning hits on the income of the MNOs. Consequently, MNOs attempt to diminish this problem. In this sense, MNOs use innovative campaigns to retain customers since it is more costly to attract a new one than to keep existing ones. One of the main factors for the triumph of these campaigns is to identify in advance possible churners and send to them the adequate information to prevent their leaving. As time passes, it is harder for a customer, which has left the service, to want to restart it again.

The approach to churn prediction in this work considers each customer individually, but not isolated from other customers. In this sense, the influence of social networks on leaving an MNO has not been studied when this social network is created under homophily criteria considerations among their users, based on the customers' profiles. Previous work has shown that members of the social circle of a subscriber also influence the subscriber to churn [5]. It is natural to believe that when people leave a service, they also influence the social circle around them with their actions. Many examples can corroborate this idea such as social pressures to adopt new technology may also encourage users to move to an MNO that has the fastest data access or popular handsets.

Some work has already been carried out on identifying churners based on mining existing customer data that MNOs own. These techniques do not always give proper reasons for churning. For this reason, this paper is centered on using Agent-Base Modelling (ABM) that has previously been applied to represent customer and market conduct in [6]. Moreover, some factors that can cause churn are defined with different parameters, depending on the type of data that the MNO has and how would like to use a model that could be able to simulate churn. In this sense, there would be many scenarios that could cause churning, where MNOs could be interested in simulating the effects in advance, such as to change the implementation of a new pricing approach, to present a new plan to customers, to upgrade the facilities offered by the network, etc.

In this paper, we propose a network creation model that incorporates demographic and psychographic characteristics and usage profiles to depict customers, utilizing appropriate individual parameters. The selection of these characteristics is based on findings from research white papers and reports or statistical data from MNO users. Our approach is a constructive and parameterizable model that could be adapted to the circumstances and data of each MNO without losing its general purpose. We propose a model that takes individuals as members of real social networks and considers that social ties between people affect churning. This effect could be more important when the real social network created is based on similarity relations among users (homophily criteria). In [7], the authors create a social graph (underlying social network) that exhibits the features required in real life social networks such as assortative mixing, high clustering, short average path length, broad degree distribution and community existence. Our work

extends and improves that work adding homophily, based on the user and usage profile, to the generated real social network. Homophily establishes that it is more probable for people to associate with similar ones than with dissimilar ones [8]. Proceeding in this way, we make sure not only that members of the same social circle influence each other, but that this influence is stronger. The resultant graph not only represents relations among customers but also the similarities among them, which are represented by the strength of their relationship, and the homophily relations established among users of the same circle. In this sense, our proposal represents an improvement in the existing social network models toward a more real one since it also takes into account customers' behavior and their closest interactions.

The experiments implemented feed the model with data from Eurostat [9]. These real data, without doubt, help to distinguish more accurately customers. The experiments use a hypothetical scenario to show how customers behave according to their environment and the social influence they are subjected. We compare the results of our proposed approach, where homophily in social networks is considered, with results using the original algorithm by [7]. We also consider that this approach could be very useful for MNOs since it may allow them to simulate different scenarios in advance, such as churning with real data.

Our approach could be very useful to MNOs since they can use the already disposable real data from their customers into our proposed simulation model and thus to verify in advance what will be the effect of the implementation of alternative strategies. Additionally, our approach allows getting insights about customers' characteristics that can be relevant later in the decision-making process to elaborate individual plans towards specific users. We also proved that taking into account social aspects helps customers to find their ideal services quicker.

This paper is organized as follows: section 2 mentions some papers and basic concepts related to homophily and social networks, as well as, some recent work on social churn identification. Section 3 explains the proposed approach. Details of implementation and experimentation of the proposed approach are presented in sections 4 and 5, respectively. Finally, section 6 presents the conclusions and future work.

2   State of the art

The proposal developed in this paper includes different topics such as homophily, social networks formation models or social churning prediction techniques. In order to tackle each topic, this section describes and compares how each of these topics is treated in our proposal with respect to other related papers published in the literature about the same topics. Also, it is important to mention that as far as we know, no one of the papers analyzed is similar to the proposal developed in this work. The following subsections describe related papers for each on one of those topics.

2.1   Homophily

Homophily is the principle which establishes that it is more probable for people to

associate with similar people than with dissimilar people [8]. Homophily could be specified or be defined by the number of relationships involved in passing information between two actors. In this sense, social distance could be measured by its network distance. When homophily is presented, the flows of information across the social network are affected, especially by how customers' features influence on social behavior.

In [10], the authors distinguished two types of homophily: status homophily, in which similarity is based on informal, formal or ascribed status and value homophily, which is based on values, attitudes, and beliefs. The main sociodemographic features in the society could be included in gender, age or ethnicity, but there are other characteristics such as education, religion and behavior patterns that are obtained later. This set forms the status homophily, which differs from the value homophily that takes into consideration inner states and their incidence in our future behavior. Based on the same definition of [8], in our paper, homophily is obtained by using the profiles of the customers that form the social network to measure social affinity between customers. This affinity is established through an Euclidean distance in base to the parameters that form the customers' profiles, what clearly differs from the related work analyzed and presented in this subsection.

2.2 Social network formation models

Modeling network formation could be analyzed by a physical/mathematical applied approach that finally allows reproducing the majority of real-world networks or by an economic view that offers a better explanation of why networks emerge. In this subsection, a general and specific review of the literature is included. The first references are related to overviews of the literature or fundamental contributions in this field, and finally, specific references are given, which are close to the work developed in this paper. However, as we presented in the introduction, our paper develops a social network model based on a modification and extension of the proposal in [7]. Thus, our new formation algorithm includes homophily by measuring the affinity between nodes through an Euclidean distance between customers' profiles, what clearly differs from other social network formation models presented in this subsection.

Reference [11] offers an extended overview of the formation of networks. Later on, in [12], the formation of networks is examined from a simple economic model where small-world characteristics appear. In [13], the authors also offer an introduction to establish a model of social and economic network formation that can be applied to many fields.

In many large networks, the number of nodes is continuously in expansion due to the addition of new nodes that at the same time are attached to other already well-connected [14]. Later on, in [15], the authors present advances in complex networks, discussing the principal models and the tools to analyze them.

In [16], models are classified into two categories, first, Network Evolution Models (NEMs) and second, Nodal Attribute Models (NAMs). After some conducted experiments, the NAM model produces assortative networks and community structure but unrealistic degree distributions that are not according to empirical data on large social networks.

NEMs, however, offers better fitting between data and degree distributions and clustering, but assortative networks and community structure do not fit well with data. In [7], an example of NEM is presented that performs well when is compared to the empirical data. It produces very efficiently networks, very similar to real social networks, with assortative degree correlations, high clustering, short average path lengths, broad degree distributions and prominent community structure. The model is based on network growth by two processes: the attachment to random nodes and the attachment to their neighborhood.

In [17], another model of social network formation is proposed. It sets, by means of parameters, the tendency to establish acquaintances that are based on comparative distances in social spaces.

In [18], a social network evolution, which is based on potential connections between the neighbors of a vertex, is analyzed. Using a mechanism described in that paper, previously generated potential edges are converted to edges.

In [19], the authors propose to distribute N-nodes using a homogeneous Poisson process in a social space n-dimensional.

Demographic and network parameters about individuals are described in [20]. Also, they define the formation of friendship, which is organized around different features of the sociality that the individuals present.

The authors of [21] show an approach based on the formation of an induced game that has a threshold equilibrium when the size of outbreed groups falls below a level.

2.3 Social churning prediction

Many research efforts have been aimed at accurately predicting churn in its early stages, involving either intrinsic (customer's profiles) or extrinsic factors (social factors), treating customers as isolated entities. However, friends, friends of friends and others influence customers within their network. Conventional approaches have been focused on intrinsic factors, which treat customers independently, but these approaches do not take into account the role of social ties and interactions between customers, which could affect churn. In our paper, the creation of the social network is based on the customers' profile (nodes) and the affinity that is generated through the Euclidean distance between these profiles. This way of generating links, in a more direct way, between nodes with similar affinities (users that present similar profile) is, in our opinion, a nice and more reasonable way to create a social network to analyze the churning prediction. Of course, our proposal differs substantially from others in this subsection, such as the following papers, related to churning prediction and social churning prediction.

Umayaparvathi and Iyakutti [22] review the existing works on churn prediction based on datasets, methods, and metrics, and they argue that using these three perspectives is crucial for developing more efficient churn prediction models. References [23], [24] and [25] also survey churn prediction techniques, but primarily focus on different modeling techniques and the accuracy of churn prediction.

Karnstedt et al. [26] argue that churn is not only a phenomenon due to individual decisions and profiles. Rather, it is influenced by external events and, more importantly,

by community effects. In this sense, they analyze social roles of the individuals and their influence on the community.

A proposal that considers different kinds of factors that could influence the churning process in individuals is introduced in [27]. The authors use the idea of Collective Classification (CC) to examine local characteristics and interdependencies between individuals in a group.

A churn prediction algorithm that quantifies the strength of social ties, which are defined by means of several parameters, is presented in [28]. Later on, over this scenario, the authors apply a diffusion model to assess the influence of churners. By means of machine-learning methods and by mixing that influence with other social factors, it is demonstrated the more accuracy obtained with this procedure.

In [29], a method to recognize "churn influencers" is suggested. In this paper, the authors score the subscribers' influence level in a similar way as in other churn models, using techniques to extract the most powerful influencers.

The effect of peer influence on churn is addressed in [30]. The empirical analysis is conducted from a data set from a wireless carrier. The peer influence is analyzed by applying survival models. It was exhibited that the tendency to churn rises more when strong friends are churning. This result also suggested that survival models overestimate the effect of influence between peers.

In [31], a data set is used to implement the graphical representation of nodes and edges. Later on, a graph of callers to whom is associated an activity vector that records its utilization statistics is developed. Proceeding in this way permits to classify callers mixing naive Bayes methods and those based on decision trees.

In Abd-Allah et al. [32], they constructed an undirected call graph and evaluated the social tie strength using new interactional attributes, derived from basic call attributes, to ensure the validity of the calculated tie strength. They proposed an influence propagation model, where the strongest ties are exploited for churn influence transfers from one node to another.

In [33], the authors present a framework to deduce how many potential users would be able to change to another service operator by means of extracting and selecting some features that later are classified with the idea of learning and detecting them. In this sense, they utilize explicit and implicit features that allow classify users and finally to detect them.

In [34], it is presented a method, the "Group-First Churn Prediction", which recognizes social leaders in each identified group. Using KPIs, this method predicts the churn of the members of the groups.

In [5], the authors present a prediction technique of potential churners by considering not only the actual set of churners but also their subjacent social network. They obtained the call graph from Call Detail Record data. Relationships between people are based on the duration of voice calls, call frequency etc. that are exchanged during a certain period. They found that this spreading activation-based technique performs better than a decision-tree technique to predict churners. They also found that connectivity and interconnectivity attributes improve the performance of the decision-tree technique.

3 Proposed approach

The aim of this work is to analyze the effect of homophilic social networks on the churning process. Figure 1 shows the different parts of our proposal to analyze the churning process. These parts are simulated using an Agent-based Model structure:

1) *Population generation*: The model generates a population of customers that interact with MNOs. The use of statistical sources allows us to generate a population with age/gender characteristics similar to real ones.

2) *Customer's profile determination*: The statistical sources, allow us to define customers by their personal profile or characteristics and their usage profile [35]. In subsection 3.1, we describe the main parameters of the customer's profile used in [43].

3) *Social Network creation*: The creation of the social network takes into account nodal attributes, which are encoded in the customer's profile, and considers that homophily (people are more likely to associate with similar ones than with dissimilar) plays an important role in establishing connections in the social network created. Thus, our proposal includes the formation of real-life social networks, based on the algorithm proposed by Toivonen et al. [7]. This formation algorithm is extended in two ways: we accommodate customer's profile information in each node and, second, we establish a measure (based on homophily metrics) to form ties between them. Section 3.2 describes those extensions.

4) *Customer-MNOs interaction*: The customers' behavior determines the interaction between customers and MNOs in a mobile market. In section 3.3, it is described the customer's purchase decision process and how they evaluate their subscribed data plan and MNO.

5) *Churning Process Analysis*: Finally, we analyze the churning process considering the usage profiles of customers. In section 5, we present the results of this analysis for an example implementation that is presented in section 4.
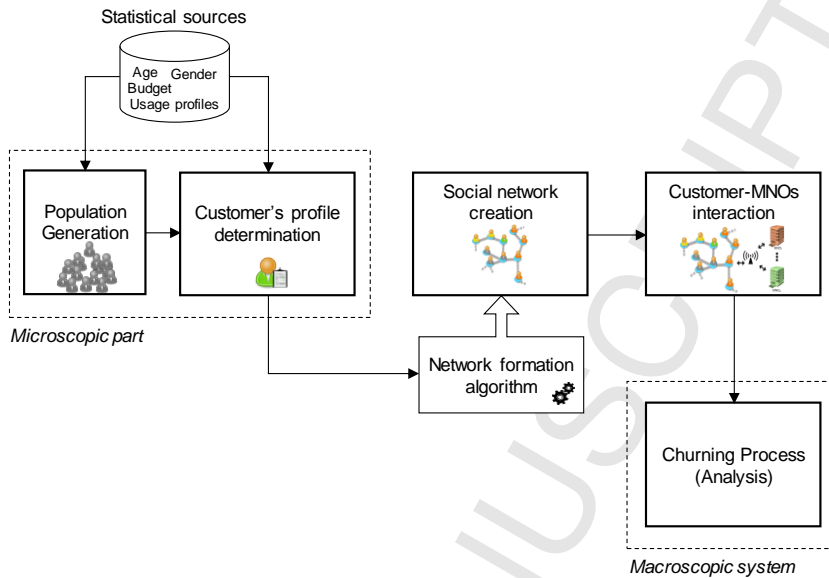
Figure 1: Churning process analysis.

3.1 Customer's profile

In Section 3.2.1 of [35] was described how customers' profiles are generated, including the characteristics and attributes selected as well as the relationships between them. Although, we use the same customer's profiles that in [35], in this section, we briefly describe the customer's profile to make this paper more self-contained. Figure 2 shows the customer's profile we consider in this work.
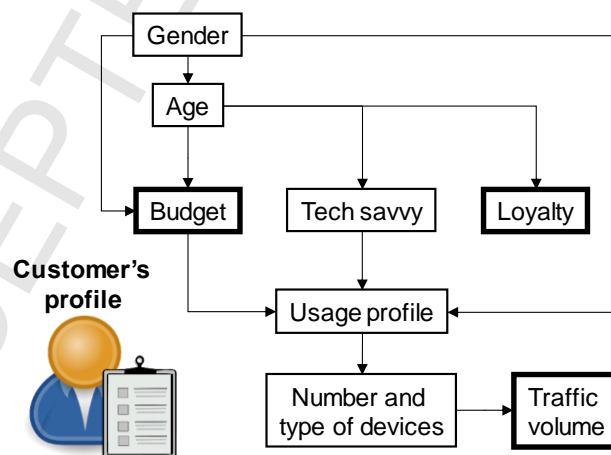

Figure 2: Customer's profile.

The customer's profile is defined by demographic and psychographic characteristics. The three main characteristics of customer's profile are budget, traffic volume and loyalty.

The budget of customers is crucial in their purchase behavior since they only can subscribed affordable data plans. As it is shown in Figure 2, the budget depends on two

more customer characteristics: age and sex. This is because of the statistical sources (in this work, we used Eurostat [9] statistics) we use to complete the costumer's profile. Eurostat gives information about income according to age and sex, and therefore, the budget depends on them.

The traffic volume is necessary to evaluate the conditions of the network that can be related to the level of service the user receives. In addition, it is necessary to evaluate the subscribed data plan, since the traffic volume is decisive to determine if the data cap is enough or not. Customers use these evaluations in the post-purchase stage to decide if they change to another data plan generating churning. The traffic volume depends on the number and types of devices that costumers use and their usage profiles. The usage profile describes the applications that a customer uses and the mean use of each application. In section 5, we assume six usage profiles considering the main use of the MNO's network: moderate use customers, online gamers, customers that use the service for work-related activities, online social networks users, on-demand music users and on-demand video users. The usage profile depends on three characteristics: gender, budget and tech savviness. Findings in [36] [37] [38] [39] showed that different genders tend to use different types of applications. The budget also determines the traffic the user can generate to the network for a particular data plan, and therefore, the usage profile depends on budget. The tech savviness (i.e., affinity for technology) also influences the usage profile of customers [40].

The loyalty is the consumer's preference for a particular brand and it depends on age as in [41]. Loyalty affects directly to customers' migration from one MNO to another, which is decided in the post-purchase behavior.

3.2   Social network

As already mentioned, it has been shown in preliminary studies, such as the one by [5], that members of the social circle of a customer also influence the customer to churn. It is natural to believe that when people leave a service, they also affect the social circle around them through their actions. In [26] is argued that churn is an event due to personal decisions and outlines, and of course, it is affected by external occurrences and, among them, by the impacts of the community where this individual belongs. In [28], the authors present an approach that takes into account intrinsic and extrinsic factors, using the idea of Collective Classification (CC). With this in mind, we decided it was important that the proposed model took into account social behavior.

3.2.1   Network creation

In our proposal, it is fundamental that the social networks resemble those in real life. A general review of the social and economic structure of networks can be found in [42]. Therefore, we looked for a model that produced a social network close to reality. In [7], as it was already explained in section 1, an example of NEM that produces very efficiently networks resembling real social networks is presented. In this work, we extend that NEM by means of accommodating customer's profile information in each node that can be used as a measure (based on homophily metrics) to form ties between nodes.

Therefore, the network formation algorithm proposed by Toivonen et al. [7] to create an underlying social network structure has the required properties for real social networks such as assortative degree correlations, high clustering, short average path lengths, broad degree distributions and prominent community structure. The algorithm is based on network growth by two processes: attachment to random nodes and attachment to their neighborhood. We extend this formation algorithm in two ways:

- First, we accommodate customer's profile information in each node based on statistical sources.
- Second, we establish a mechanism (based on homophily metrics) to form ties between them. Thus, ties are more probable between similar people than between dissimilar ones. Homophily has important implications on how information flows through the social network and for our purposes of detecting churning among customers it is a key factor to include due to its incidence.

Other models of network formation including homophily exist in the literature such as [43] but it totally differs from our approach in how the nodal attributes are defined and are encoded and later, in how the established mechanisms are to form ties between nodes.

The measure of the homophily in our approach is evaluated as the Euclidean distance between two nodes, according to their nodal attributes, that in this case are the customer's profile characteristics (gender, age, budget, tech-savvy, usage profile and traffic volume) as described in 3.1. Each attribute could be measured in a different dimension and so the resultant distance between two nodes would be the Euclidean distance in an n-dimensional space (1), n being the number of attributes a node has.

$$distance_{pq} = \sqrt{w_1 \; p_1 - q_1 \;^2 + w_2 \; p_2 - q_2 \;^2 + \ldots + w_n \; p_n - q_n \;^2} \qquad (1)$$

where p and q are nodes, $1,2,\ldots,n$ are the number of attributes and w is a weight given to a dimension. w can be used to normalize attributes or give more or less importance to a given attribute. In this work, we normalize the maximum distance value to 1, and therefore, we can define affinity as in (2), which is 0 if the two nodes have the maximum distance and 1 if they have distance 0.

$$affinity = 1 - distance_{pq} \qquad (2)$$

Thus, our proposal is to extend the algorithm proposed by [7]. In particular, we add steps 1, 2 and 6, and we modify step 3. The algorithm works as follows:

1. Select the size of the desired social network (i.e., the number of customers) and for each customer determine the customer's profile.
2. Obtain the affinity matrix for each pair of customers taking into account their customers' profiles.
3. Start with a seed network of N_0 nodes. In contrast to the random selection in [7], we select the nodes with the highest affinity according to the affinity matrix.
4. Pick on average $m_r \geq 1$ random nodes as initial contacts.
5. Pick on average $m_s \geq 0$ neighbors of each initial contact as secondary contacts.
6. Choose the customer with the highest affinity for the selected initial and secondary contacts.

7. Connect the new vertex to the initial and secondary contacts.
8. Repeat steps 3–7 until the network has grown to the desired size.

3.2.2 Influence on friends

Although people can have a number of friends, not all of them share the same bond as the others. Many measurement indexes could be proposed to establish an influence value. In this work, we evaluate this index by measuring the affinity, that is, to use the converse value of the distance between two customers as an influence value, meaning that similar customers share a stronger bond. As the minimum value of the distance is 0, meaning that both customers have the exact same profile, the influence (affinity) between two customers with the exact same profile would be 1, which is the maximum value of influence. The influence on friends will be used in the Customers-MNO interaction, in particular, it will affect the customer's purchase decision process as it will be shown in the following section.

3.3 Customer's behavior

In Section 3.2.2 of [35], it is described the purchase decision process as well as the evaluation of the MNO to which customers are subscribed. In this section, we briefly describe the customer's behavior to make this paper more self-contained. Figure 3 shows the purchase decision process we consider in this work, and how social influence affects this process.
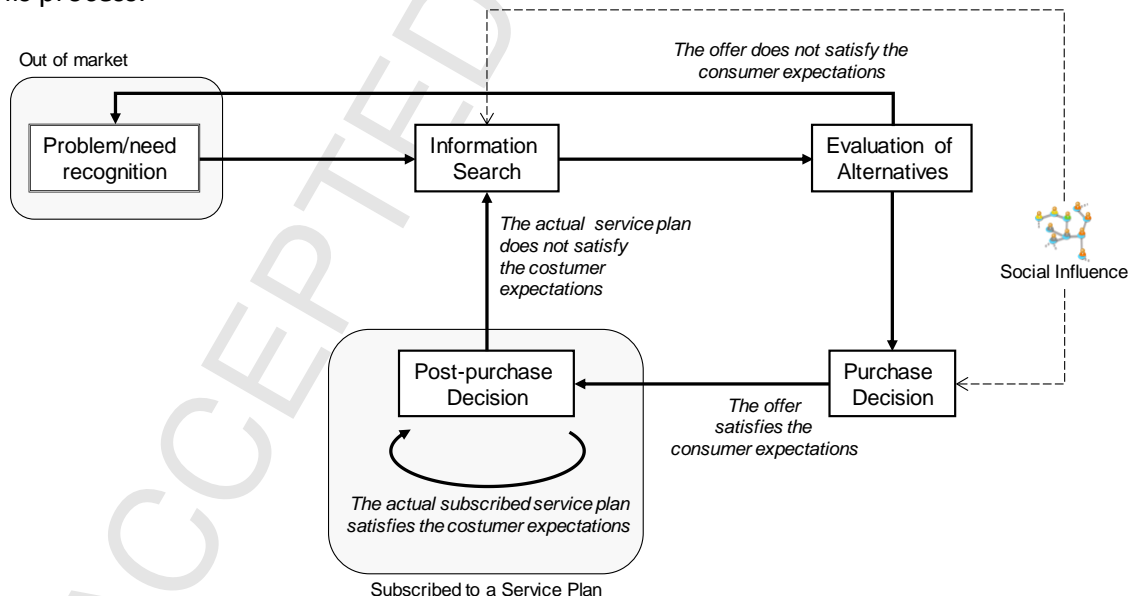

Figure 3: Purchase decision process.

The purchase decision process consists of five stages:
1. *Problem/need recognition*: The customer is not subscribed to a data plan. The offer of data plans does not satisfy its expectations.

2. *Information search*: The customer obtains information about the MNOs offering the service and their data plans. This information includes the price, the overage charge and the data cap of each data plan. The social network of the customer may influence this information since the customers will consider first the information about the data plans their friends have subscribed.

3. *Evaluation of alternatives*: Customers evaluate the different alternatives they have found in the Information search stage. The price of the data plan, its data cap and the evaluation of the quality of service are the main elements to decide about the convenience of a data plan. Therefore, budget and traffic volume are decisive costumer's profile characteristics to make this evaluation.

4. *Purchase decision*: Costumers will decide a data plan alternative based on the information shared by their friends. The affinity between friends will determine the influence of the friend on the purchase decision. Therefore, customers will have to take into account the actual evaluation of the alternatives and the influence of their friends to decide the data they subscribe.

5. *Post-purchase decision*: In this stage, customers will evaluate the service plan they are subscribed to and the MNO that is providing the service. If the evaluation does not satisfy the customer expectations, the customer will go the information search stage to decide to change to a different service plan within the same MNO or churn. Customers will evaluate service plans based on price, a plan's data cap and the perceived QoS. Customers will also evaluate the MNO they are subscribed to, based on the Customer Experience model proposed in [44]. The model uses a utility function that is the weighted sum of each category. Table 1 shows, in a 0-100 scale, the categories and scores used in this work. Customers share the evaluations of the service plans and MNOs with their friends.

Category	Weighting	Description
Cost	18.18	Cost competitiveness and plan usage
Handset	9.09	Handset repairs and known issues
Coverage	23.38	QoS and coverage
Customer services	18.18	Complaint volume and repetition
Offers and promotions	18.18	Savings from offers and promotions
Billing	12.99	Billing complaints

Table 1: Customer Experience categories and weights.

4 Implementation

The preferred framework to support the implementation of our proposed approach was OMNeT++ simulator by OpenSim Ltd [45] since it provides adequate infrastructure,

tools for writing discrete-event simulations and offers a generic architecture that can model and simulate any system that can be mapped into entities communicating by exchanging messages. Models are made up of reusable components called modules. Modules can be combined to form compound modules. Modules may have parameters that can be used to customize module behavior and/or to parameterize the model's topology. Modules at the lowest level of the module hierarchy are called simple modules, and they encapsulate model behavior. Simple modules are programmed in C++ and make use of the simulation library. Although OMNeT++ is primarily used to simulate communication networks with several available external frameworks, such as INET, we use this framework as a generic modular simulation framework.

Three types of modules were created as depicted in Figure 4: a module to create the social network, a module to define the customers, and a module to define MNOs.
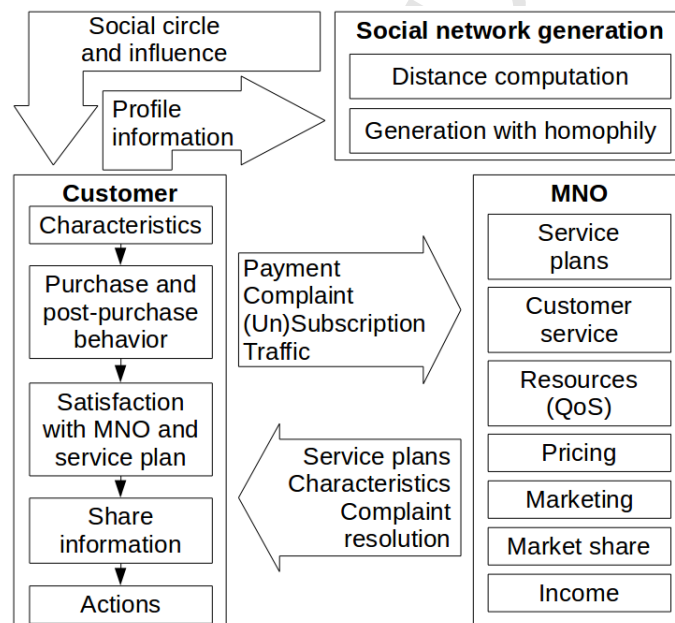

Figure 4: General model.

In the simulation initialization, the customer and MNO characteristics are assigned according to distributions or scalar values. In the starting point, customers are not subscribed to any specific MNO, but the can be assigned to an MNO according to the MNO market share. When the simulation runs, in the customer-MNO interaction stage (Figure 1), each customer will interact with the MNO that better fits the purchase decision process described in section 3.3. In this work, the MNOs' behavior is not the objective, although could be implemented following several criteria.

4.1 Social network creation module

We designed a module to create the social network using the proposed network creation algorithm as explained in section 3.2.1. The tool selected to handle and perform all network operations was the Stanford Network Analysis Platform (SNAP) [46]. SNAP is a

general-purpose software in C++ for network analysis with a graph-mining library. This software packet easily scales large networks with millions of nodes, manipulating and calculating structural properties of massive graphs.

Customers' profile characteristics such as gender, age, budget, tech savviness and mean traffic volume were chosen as the characteristics for calculation of the measure of homophily (affinity). In order to include the usage profiles as a dimension in the homophily space, the importance that usage profiles give to QoS, and the rate at which usage profiles increment their dislike towards a plan depending on billing, QoS and data cap, were used. The minimum and maximum values of each dimension were used to normalize each dimension. Initially, all users are created individually and then they send their profiles to the network generator. The social network creation module calculates the social distances and influences (affinities) and then uses the proposed network creation algorithm to create the social network. Once the social network is created, customers are informed of their social circles and their influence among friends. This information exchange is done through appropriate messages that the OMNeT++ tool generates as it can be seen in Figure 4.

4.2 Customer module

The customer model is implemented as a simple module that encloses the particularization of the customers' profiles and the implementation of their behaviors. This model includes the likelihood of customers complaining when an incident happens regarding billing or QoS, which are defined as parameters. For simplicity, this work assumes that customers consider that the QoS is their average bandwidth available, which depends on the MNO's available bandwidth, the subscribed customers, and their amount of traffic generated and the time they use the bandwidth.

4.3 MNO module

The MNO module includes some common features such as the pricing scheme, the amount of resources (bandwidth), the service plans presented and other special features such as customer services or marketing parameters.

A set of plans is defined by parameter price, overage charges, and data cap. General parameters of the MNO include the available bandwidth, cost competitiveness and plan usage, handset repairs and known issues, and savings from offers and promotions. The probability of resolving complaints is also a parameter.

4.4 Customer-MNO message exchange

MNO and customer modules, interchange data through adequate messages as it is shown in Figure 4, such as subscription or not subscription of the customer, inquiries, complaints, payment procedure, etc. An example of an exchange of messages between a customer and its MNO is when a customer sends a complaint to the MNO, and then it resolves its question. Additionally, customers communicate among themselves sharing

their MNO's evaluations with their social circles.

5 Evaluation and results

In order to evaluate the proposed model, a hypothetical scenario was devised. In this scenario, depicted in Figure 5, 2000 customers interact with two MNOs, MNO 1 and MNO 2, with similar features such as customer service level, billing complaint, handset availability and marketing aggressiveness indexes. The value for these indexes is established to 90% except for the billing complaint index, which is set to 10%. Initially, both MNOs supply similar resources to customers, which are established on 608 Kb/s. Later each MNO offers one service plan with equivalent features: 2 GB of data cap for 10 currency units, each additional GB costs 5 currency units more (overage charge) and the period of billing and subscription is 30 days (1 month). While simulation is running, MNO 2 increases its available resources (bandwidth) to 768 kb/s and it informs of this change to customers. In this scenario, the MNO 2 operator offers an increase in bandwidth for the users. This fact will cause customers sensitive to this change to churn and influence similar customers.
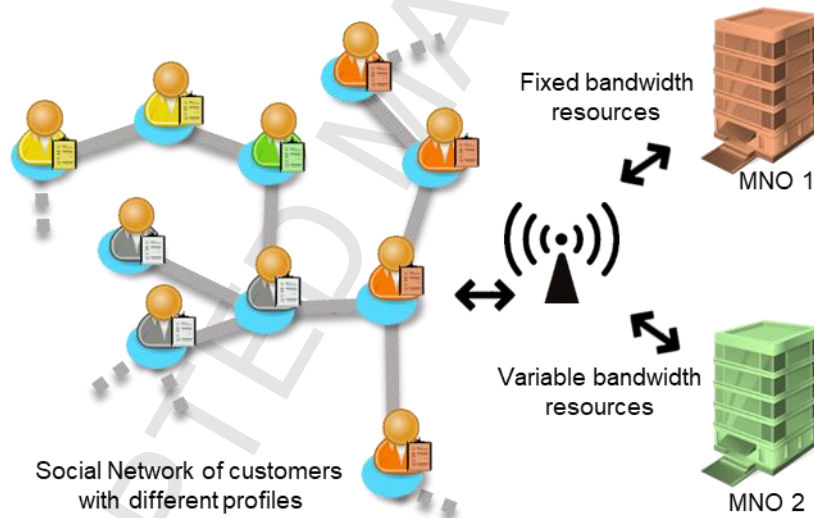

Figure 5: Simulated scenario.

For these experiments, the total simulation time considered is 4 years (1440 days if we consider months of 30 days). Operator MNO 2 changes its available bandwidth at day 720. This makes current MNO 2 customers to experience an increase in the QoS. As input for the customer module, Eurostat data from 2014 [9] are applied in reference to gender, age and income distribution. When data from Eurostat could not be used as input for other features and parameters, we utilized other distributions and values that we consider adequate as mentioned in 3.1. As an example and for simplicity, the general and related parameters used in these experiments are the same as in Section 4 of [35] and we show them in Appendix A.

The network generator uses the same parameters as Toivonen et al. [7]. The initial network size is set to 10% of the number of customers in the experiments and it is

generated by adding pairs of nodes with the lowest distance as presented in section 3.2. The probability defining the number of initial connections was set to 0.95 for one initial contact and 0.05 for two initial contacts. The number of secondary connections is selected from $U[0,3]$. In order to better differentiate profiles, the distances between profiles are normalized using the minimum and maximum values found in the experiment, 0.1 and 0.9 respectively, but if lower or higher values are found, these should be used instead.

For the case implemented in the simulator, the results obtained in the evaluation of the proposed model are compared with those obtained by using Toivonen et al. [7] and are referred to two categories: the circle of friends on same MNO ratio (section 5.1) and the value of the churn for different usage profiles (section 5.2). Section 5.3 presents the topologies of the social network graph generated by using the original and our proposal for the network creation algorithm.

5.1 Circle of friends on same MNO ratio

The circle of friends on same MNO ratio is a representative way to show the benefits of taking into account homophily in the process of the social network generation. Customers with similar profiles should behave similarly and make similar decisions. This would have the effect of a majority of customers in a social circle being subscribed to the same MNO. Moreover, due to customers having social ties with other customers that have higher influence from those with similar behavior, the shared experience in the social group will help these customers find quicker, according to their profile, their "ideal" MNO and service plan. Consider $\mathcal{N}_i \in \mathcal{N} = \{ \mathcal{N}_1, \mathcal{N}_2 \}$ as the set of subscribers of MNO i. We define the number of subscribers in the set $\mathcal{N}_i$ as $|\mathcal{N}_i|$. The ratio of friends on same MNO of a customer $c \in \mathcal{N}_i$ ($FR_{i,c}$) is obtained counting individually the number of friends in $\mathcal{N}_i$ ($F_{i,c}$) divided by the number of friends subscribed to any MNO, the same or other. Unsubscribed friends of c ($F_{0,c}$) are not considered.

$$FR_{i,c} = \frac{F_{i,c}}{\sum_{j \neq 0} F_{j,c}} \tag{3}$$

Now, consider the set of subscribers with a particular usage profile $\mathcal{P}_j \in \mathcal{P} = \{ \mathcal{P}_{music}, \mathcal{P}_{gamer}, \mathcal{P}_{moderate}, \mathcal{P}_{video}, \mathcal{P}_{social}, \mathcal{P}_{worker} \}$. The mean ratio of friends on same MNO of customers $c \in \mathcal{N}_i$ with usage profile $\mathcal{P}_j \in \mathcal{P}$ ($c \in \mathcal{N}_i \cap \mathcal{P}_j$) is defined as in (4), where $|\mathcal{N}_i \cap \mathcal{P}|$ denotes the number of customers in the set $\mathcal{N}_i \cap \mathcal{P}_j$.

$$\overline{FR_i^j} = \frac{1}{|\mathcal{N}_i \cap \mathcal{P}_j|} \sum_{c \in \mathcal{N}_i \cap \mathcal{P}_j} FR_{i,c} \tag{4}$$

Figures 6, 7, 8, 9, 10 and 11 show the customers' circle of friends on same MNO mean ratio ($\overline{FR_i^j}$) for different usage profiles: music customers in Figure 6, gamer customers in Figure 7, moderate customers in Figure 8, video customers in Figure 9, social customers in Figure 10, and work customers in Figure 11. In these figures, the original

algorithm in [7] (random algorithm) and the proposed social network creation algorithm are compared. Each figure shows the ratio for MNO 1 and MNO 2 for each usage profile. It can be seen that for the case of MNO 2, the average circle of friends on same MNO ratio for profiles with a higher sensitivity to QoS (gamer, worker, music and video), according to Table A.9, is higher than that obtained with the original algorithm. This is more significant after the increment in the available bandwidth of MNO 2, which means that friends with a higher sensitivity to QoS join the MNO 2 that offers the higher bandwidth. In the same way, customers subscribed to MNO 1 with profiles with lower sensitivity to QoS (moderate and social) have an average circle of friends on same MNO ratio higher than that obtained with the random social network generator. This means that friends with a lower sensitivity to QoS group in the MNO 1 that is offering the lowest bandwidth.
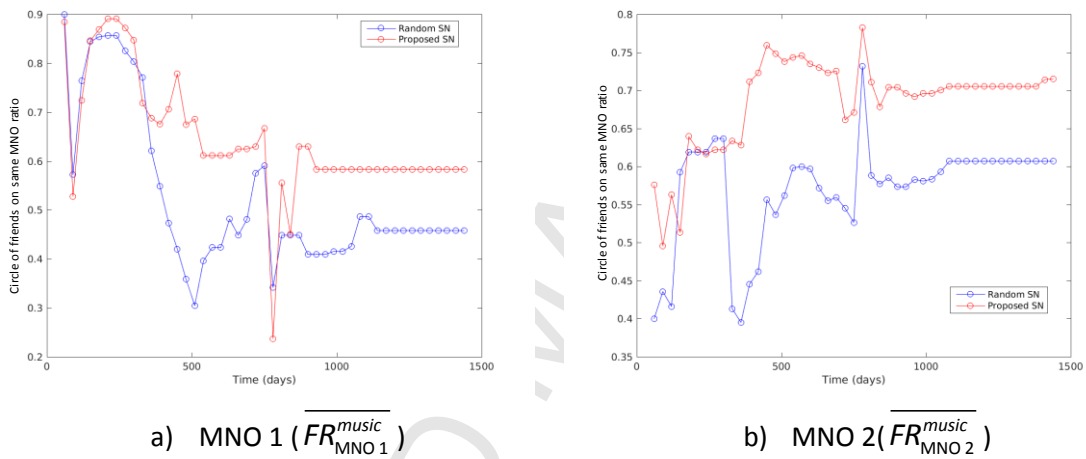

a)   MNO 1 ($\overline{FR_{\text{MNO 1}}^{music}}$)   b)   MNO 2($\overline{FR_{\text{MNO 2}}^{music}}$)

Figure 6: Mean ratio of friends on same MNO ratio for music customers.

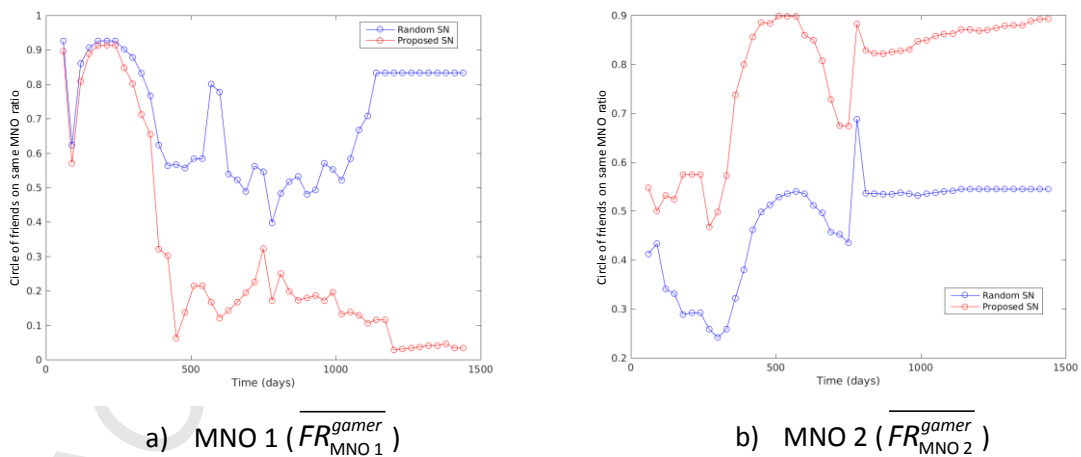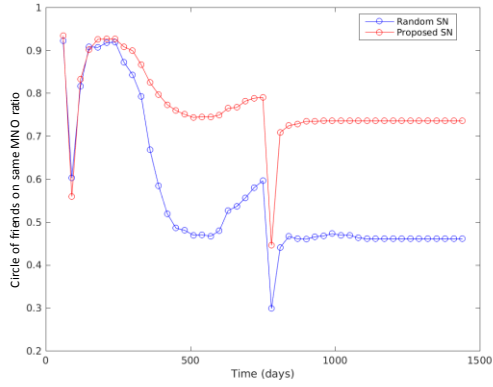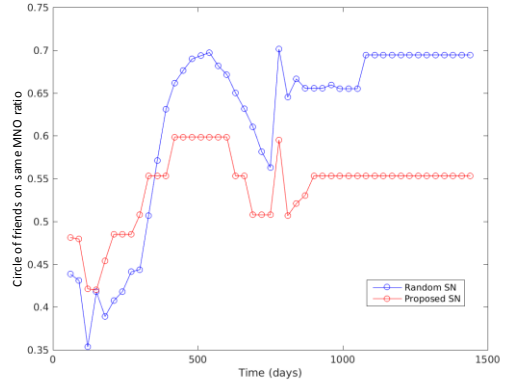a)   MNO 1 ($\overline{FR_{\text{MNO 1}}^{gamer}}$)   b)   MNO 2 ($\overline{FR_{\text{MNO 2}}^{gamer}}$)

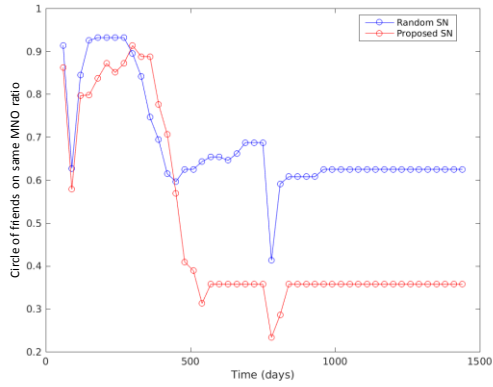Figure 7: Mean ratio of friends on same MNO ratio for gamer customers.

a) MNO 1 ($\overline{FR_{MNO\,1}^{moderate}}$)
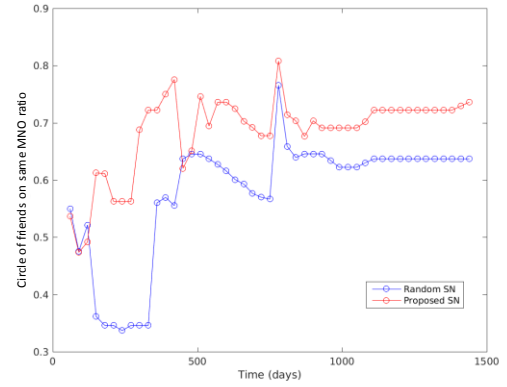
b) MNO 2 ($\overline{FR_{MNO\,2}^{moderate}}$)

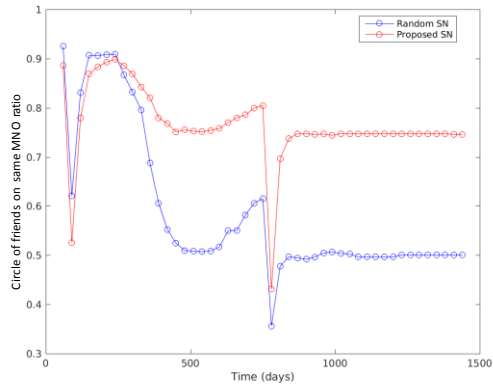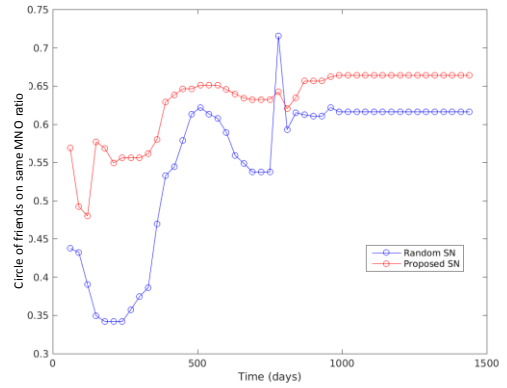Figure 8: Mean ratio of friends on same MNO ratio for moderate customers.

a) MNO 1 ($\overline{FR_{MNO\,1}^{video}}$)

b) MNO 2($\overline{FR_{MNO\,2}^{video}}$)

Figure 9: Mean ratio of friends on same MNO ratio for video customers.

a) MNO 1 ($\overline{FR_{MNO\,1}^{social}}$)

b) MNO 2 ($\overline{FR_{MNO\,2}^{social}}$)

Figure 10: Mean ratio of friends on same MNO ratio for social customers.

a) MNO 1 ($\overline{FR_{MNO\,1}^{worker}}$)   b) MNO 2 ($\overline{FR_{MNO\,2}^{worker}}$)
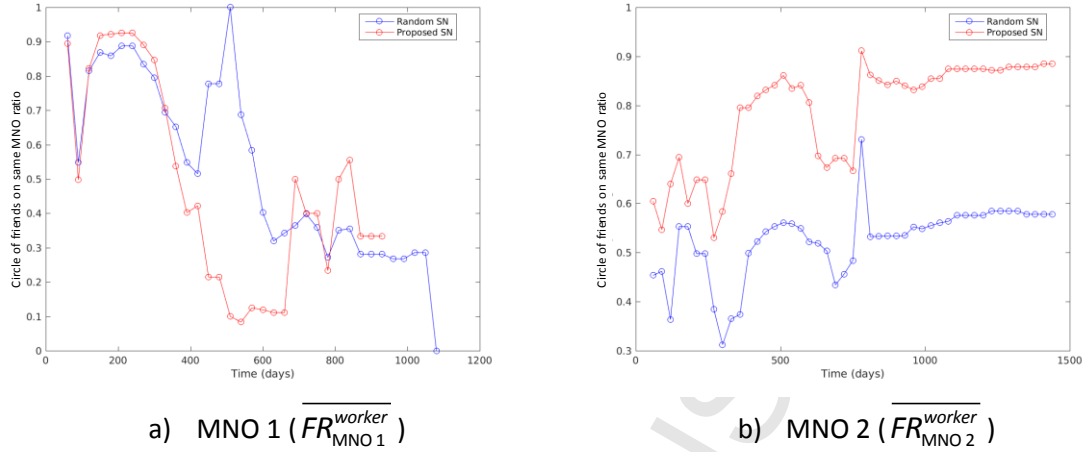
Figure 11: Mean ratio of friends on same MNO ratio for worker customers.
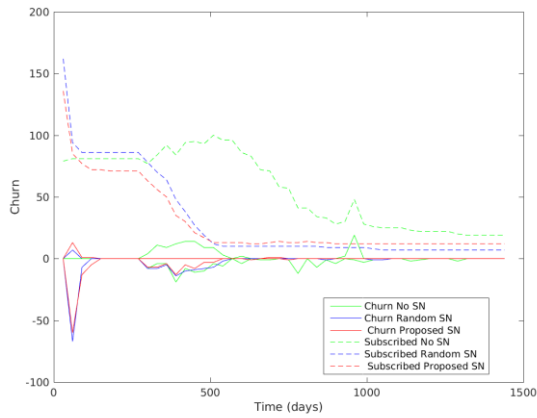
5.2 Churn and subscriptions

Figures 12, 13, 14, 15, 16 and 17 show the amount of churn and the number of subscriptions to MNO 1 and 2, respectively, for different usage profiles: music customers in Figure 12, gamer customers in Figure 13, moderate customers in Figure 14, video customers in Figure 15, social customers in Figure 16, and work customers in Figure 17.

Let $\mathcal{N}_i \in \mathcal{N} = \{\, \mathcal{N}_1, \mathcal{N}_2 \,\}$ be the set of subscribers of MNO i. Consider that we measure $S_i\,n = |\mathcal{N}_i\,n|$ that is the number of subscribers of MNO i at discrete time n. The churn is the variation of S_i due to customers that leave to other MNOs, $C_i^+\,n = \sum_{j\neq i} C_{ij}$. We call to this measure, positive churning, to differentiate from what we call negative churning that is the variation of S_i due to customers that come from other MNOs, $C_i^-\,n = \sum_{j\neq i} C_{ji}$. In addition, S_i changes due to subscribers of MNO i that leave the market, $Out_i\,n$, and newcomer customers that subscribe to MNO i, $In_i\,n$.

We can express $S_i\,n$ as:

$$S_i\,n = S_i\,n-1 + C_i^+\,n - C_i^-\,n + Out_i\,n + In_i\,n \tag{5}$$

Figures 12-17 show, for different usage profiles, the values of $S_i\,n$ (number of subscribers), $C_i^+\,n$ (positive churn) and $-C_i^-\,n$ (negative churn). It can be seen that for both cases where there is no social network, values for both churn and subscriptions are similar. However, it can be noted that when a social network is present, both indicators reach their final values more quickly than in the case where there is no social network. This can be attributed to the information that is shared between similar customers that get to their "ideal" option faster than in the case where there is no information sharing.
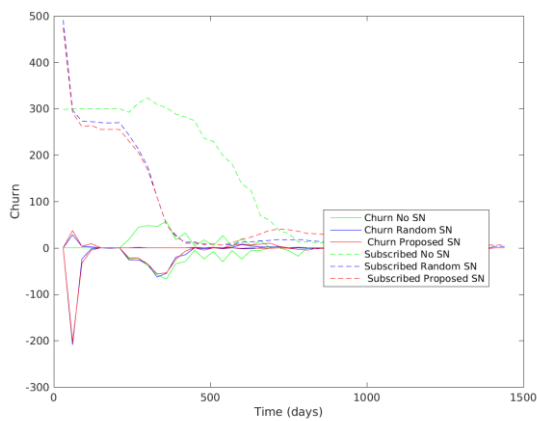
a) MNO 1
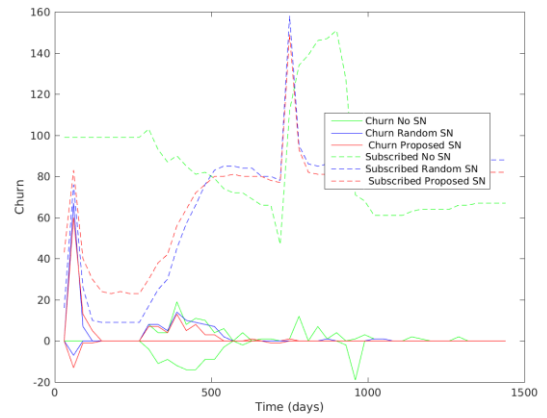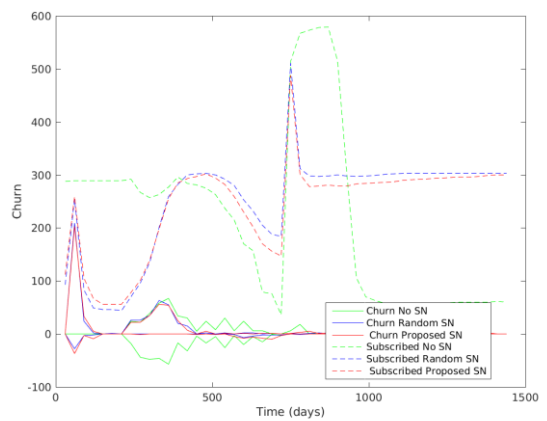
b) MNO 2

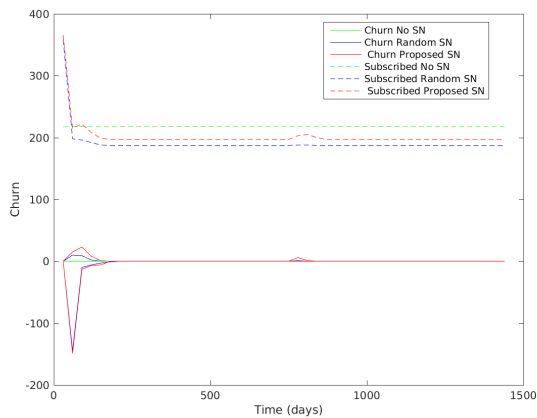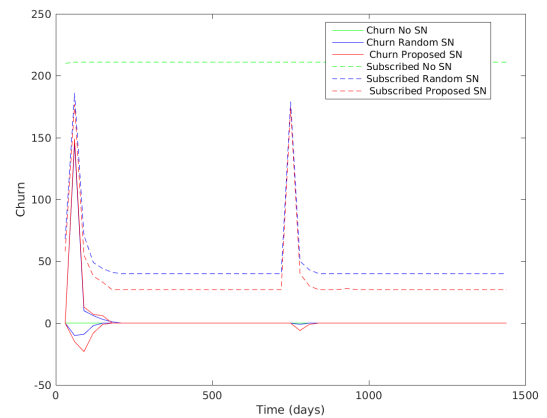Figure 12: Music customers' churn and subscriptions.

a) MNO 1

b) MNO 2

Figure 13: Gamer customers' churn and subscriptions.

c) MNO 1

d) MNO 2

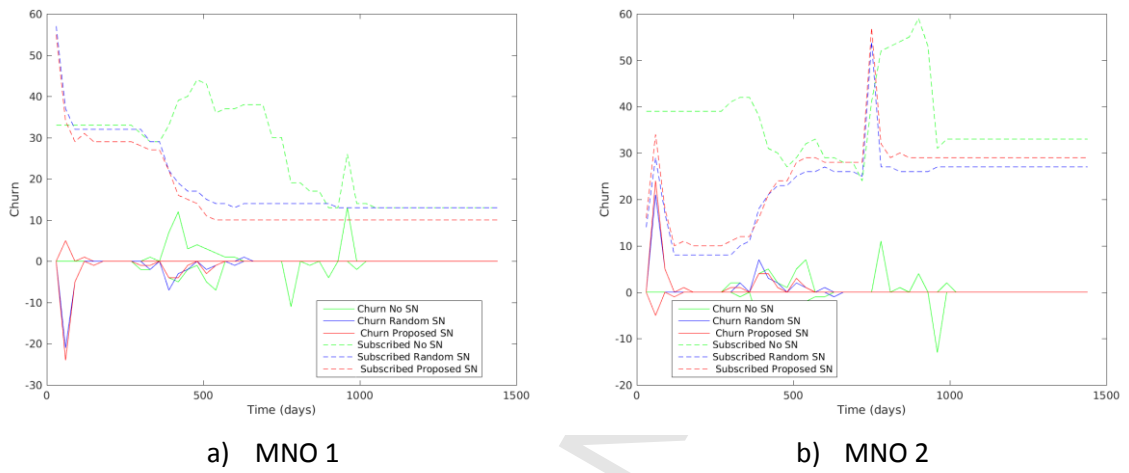Figure 14: Moderate customers' churn and subscriptions.

a) MNO 1   b) MNO 2

Figure 15: Video customers' churn and subscriptions.

a) MNO 1   b) MNO 2

Figure 16: Social customers' churn and subscriptions.
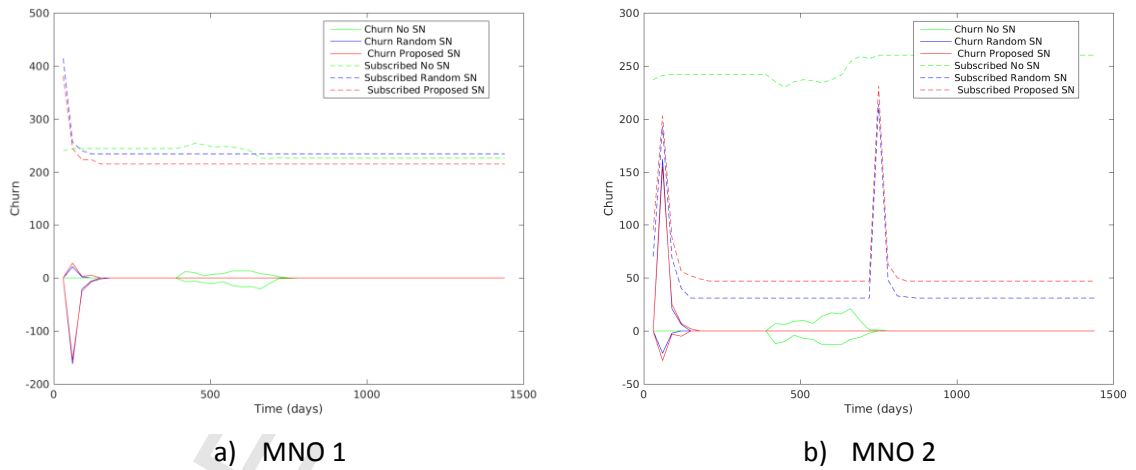
<div align="center">a)   MNO 1   b)   MNO 2</div>
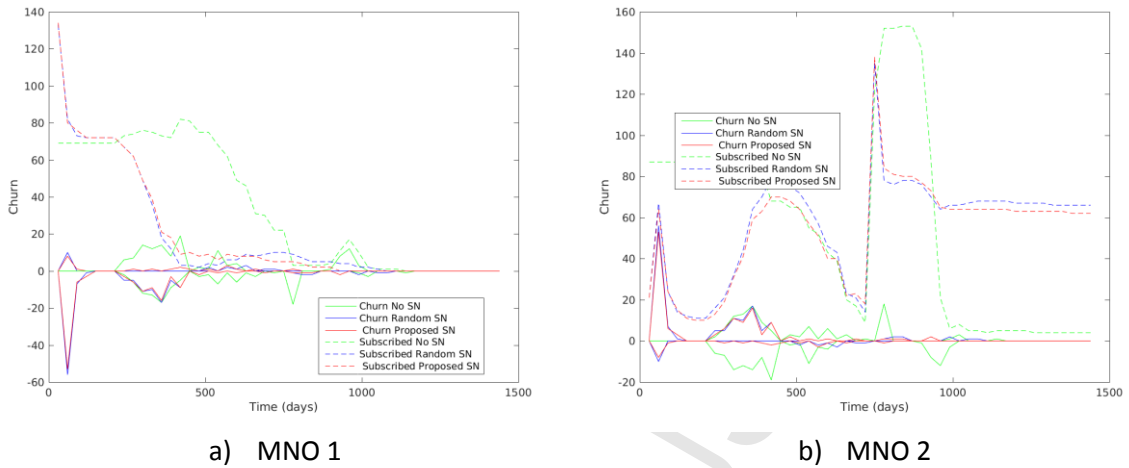
<div align="center">Figure 17: Work customers' churn and subscriptions.</div>

5.3 Social network graph

Figures 18 and 19 show the biggest components of the social network graph obtained with the original algorithm and with the proposed algorithm, respectively.

Each node represents a customer and each edge represents a bidirectional social tie. Nodes are colored according to their usage profile. Blue nodes represent moderate customers, green nodes represent gamer customers, cyan nodes represent online social customers, pink nodes represent music customers, red nodes represent worker customers and yellow nodes represent movie customers.

Comparing the two figures, it can be seen that by adding the homophily measure, that is, a higher probability that customers with similar profiles have social ties, into the algorithm proposed by Toivonen et al. [7], our goal is accomplished. In this case, it is represented by the usage profiles with different colors. This is also proven using a measure of heterogeneity proposed in [47]. This measure assesses the heterogeneity of a customers' network with respect to the value of an attribute. The formula used to calculate customer-network heterogeneity with respect to attribute A for customer i is:

$$HETEROGENEITY_{iA} = 1 - \left[\sum_{1}^{n} \left(\frac{A_k}{en} \right)^2 \right] \tag{6}$$

where A is the attribute, in this case, the usage profile. A_k the number of friends with usage profile k in the customer network. en is the number of friends in the ego network with valid data on A, where n is the total number of usage profiles of A represented in the customer network. Lower values mean higher numbers of friends with the same usage profile, and higher values mean a more heterogeneous customer network. Figure 20 shows boxplots of the heterogeneity in the network created by the original algorithm and the proposed algorithm. On the boxplot, the central mark is the median value of the heterogeneities of all the nodes in the social network. The bottom edge of the box shows

the 25th percentile and the top edge the 75th percentile. The whiskers range to the most extreme heterogeneities without considering 0 and 1 values (outlier values). As it can be seen in Figure 20, the values of heterogeneity for our proposed algorithm are lower than for the original one and therefore, it considers the homophily effect of the social network.
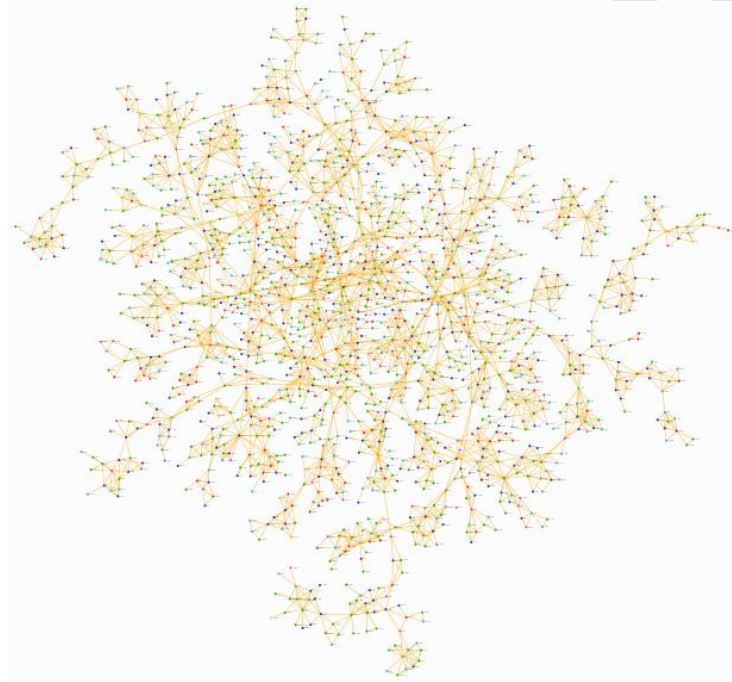

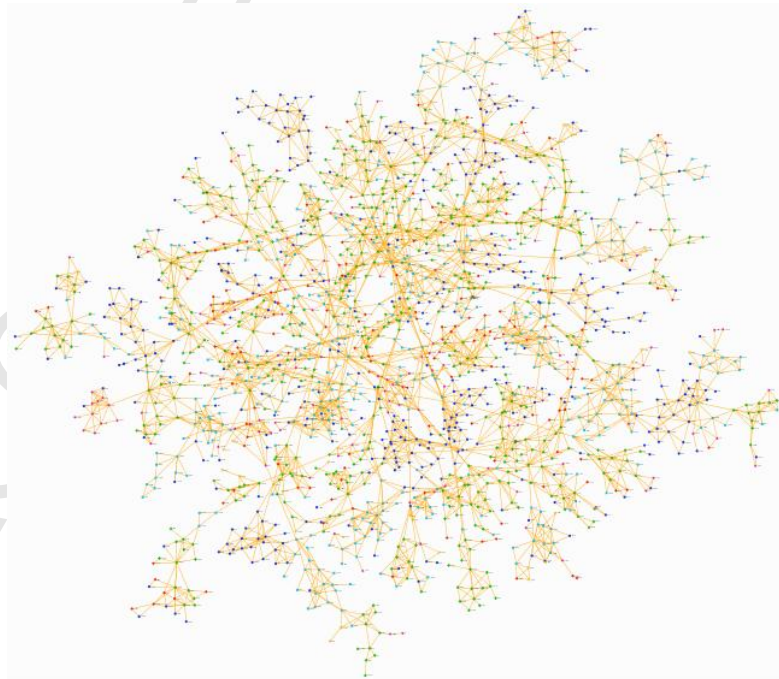Figure 18: Network created with random social network generator.

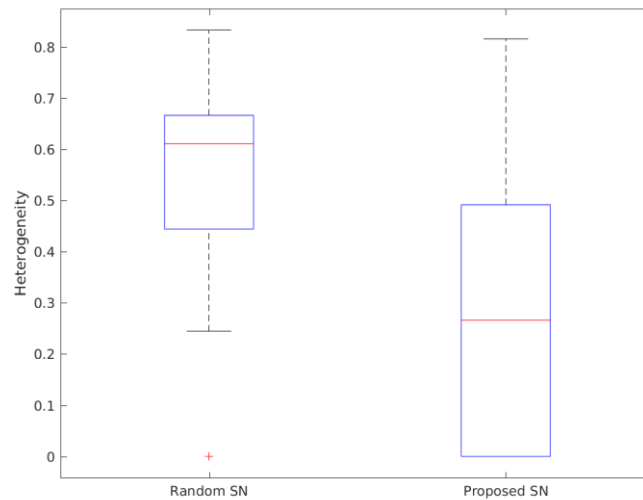Figure 19: Network created with proposed social network generator.

Figure 20: Heterogeneity.

6 Conclusions

MNOs compete against each other in the telecommunications market with the aim to get a better status among customers. In this sense, churning is an issue that MNOs must face reducing its impact by means of customer retention techniques. In order to facilitate this retention, potential churners should be identified in advance.

The mainstream approach to churn prediction considers each customer individually. Preliminary studies have shown, however, that members of the social circle of a subscriber also influence the subscriber to churn [5]. Developing churn prediction systems that consider social aspects poses an emerging theoretical challenge with potentially great practical implications.

Classic proposals use data mining strategies to identify potential churners, but usually, they do not facilitate clear reasons in the churning process. Using ABM models allow exploring the impact that different features of customers and MNOs have in analyzing and predicting churning processes.

Our model incorporates demographic and psychographic features and interdependencies among them. These features show customer behavior in the mobile market. This model also describes a well-defined purchase decision process in deciding if to continue or to leave an MNO. Moreover, we take into account the social aspects that have to do with the flow of information and their effect on customers' decisions. We use the work by Toivonen et al. [7] to create a social graph that exhibits the features present in real-life social networks and we propose extending the algorithm used there in order to take into account homophily between customers (based on their customers' profiles).

The incorporation of usage profiles is very significant since they necessarily implied to establish the quantity of data, QoS sensibility or even price/data rate, what allow us to differentiate among usage profiles in their purchase and post-purchase behavior.

We showed with our experiments that when taking into account social influence, the users find their ideal service quicker than when there is no flow of information between customers. Furthermore, we showed that taking into account homophily helps customers to find this ideal option more quickly. Using homophily in the algorithm proposed in [7] effectively groups similar customers together.

Our approach is potentially useful to MNOs since it can admit any type of related data set. In this sense, MNOs have already available data from their customers and over them; different actions could be adopted whether to change of scenarios or to test new pricing or marketing strategies. This facility will permit MNOs to assess new strategies in advance before their implementation. Our approach is also constructive and parametrizable, and in this sense enables the values of the included parameters to be changed and new ones to be included.

Another possible advantage in using this approach to MNOs is that will let them use the customer features incorporated into the model for segmentation intention. This will allow MNOs center on specific subsets of customers to get insights into such segments with the aim to be later used for decision-making purposes. Thus, an MNO could get benefit from these insights to create marketing and sales strategies addressed to concrete customers. Moreover, MNOs could target or identify groups of customers with a tendency to churn due to social influence and target these social circles in order to keep these customers.

As future work, we contemplate experimentation on more complex scenarios. For example, multiple MNOs, each offering multiple nonequivalent service plans. Another opportunity for future work could be letting customers change and adjust their profiles over time, so profiles are non-static.

Acknowledgements

References

[1] Chau, A., "Asia/Pacific (Excluding Japan) Mobile Services Year-End Review, 2009," vol. 1, May 2009.

[2] Strategy Analytics, "Wireless Operator Performance Benchmarking Q4 2009," Strategy Analytics, Santa Fe, 2010.

[3] ITU, "The world in 2013: ICT facts and figures," ITU, 2013.

[4] D. Lonergan, "Improve Communication to Minimize Churn," Yankee Group, 2012.

[5] K. Dasgupta, R. Singh, B. Viswanathan, D. Chakrabo, S. Mukherjea, A. A. Nanavati and A. Joshi, "Social ties and their relevance to churn in mobile

telecom networks," in *Proceedings of the 11th International Conference on Extending Database Technology: Advances in Database Technology, EDBT '08*, New York, USA, 2008.

[6] P. Twomey and R. Cardman, "Agent-based modelling of customer behaviour in the telecoms and media markets," *info,* vol. 4, no. 1, pp. 56-63, 2002.

[7] R. Toivonen, J. Onnela, J. Saramäki, J. Hyvönen and K. Kaski, "A model for social networks," *Physica A: Statistical Mechanics and its Applications,* vol. 371, no. 2, pp. 851-860, 15 Nov. 2006.

[8] M. McPherson, L. Smith-Lovin and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology,* vol. 27, pp. 415-444, August 2001.

[9] Eurostat, "Statistics by theme," 8 May 2014. [Online]. Available: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes. [Accessed 8 May 2014].

[10] P. F. Lazarsfeld and R. K. Merton, "Friendship as a social process: A substantive and methodological analysis," *Freedom and control in modern society,* vol. 18, pp. 18-66, 1954.

[11] M. O. Jackson, "A Survey of Models of Network Formation: Stability and Efficiency," California Institute of Technology, 2003.

[12] M. O. Jackson and B. W. Rogers, "The Economics of Small Worlds," *Journal of the European Economic Association,* vol. 3, no. 2/3, pp. 617-627, May 2005.

[13] M. O. Jackson, Social and Economic Networks, Princeton University Press, 2010, p. 520.

[14] A. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science,* vol. 286, no. 5439, pp. 509-512, 15 October 1999.

[15] R. Albert and A. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics,* vol. 74, no. 1, pp. 47-97, 30 January 2002.

[16] R. Toivonen, L. Kovanen, M. Kivelä, J. Onnela, J. Saramäki and K. Kaski, "A comparative study of social network models: Network evolution models and nodal attribute models," *Social Networks,* vol. 31, no. 4, pp. 240-254, October 2009.

[17] M. Boguña, R. Pastor-Satorras, A. Diaz-Guilera and A. Arenas, "Emergence of clustering, correlations, and communities in a social network model," *eprint arXiv:cond-mat/0309263,* September 2003.

[18] A. Vázquez, "Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations," *Physical Review E,* vol. 67, no. 5, 07 May 2003.

[19] L. H. Wong, P. Pattison and G. Robins, "A spatial model for social networks," *Physica A: Statistical Mechanics and its Applications,* vol. 360, no. 1, pp. 99-120, 15 January 2006.

[20] S. M. Goodreau, J. A. Kitts and M. Morris, "Birds of a feather, or friend of a

friend? using exponential random graph models to investigate adolescent social networks," *Demography,* vol. 46, no. 1, pp. 103-125, 2009.

[21] S. Currarini, J. Matheson and F. Vega-Redondo, "A simple model of homophily in social networks," *European Economic Review,* vol. 90, pp. 18-39, November 2016.

[22] V. Umayaparvathi and K. Iyakutti, "A survey on customer churn prediction in Telecom Industry: Datasets, Methods and Metrics," *International Research Journal of Engineering and Technology,* vol. 3, no. 1, pp. 1065-1070, April 2016.

[23] S. Liao, P. Chu and P. Hsiao, "Data mining techniques and applications – A decade review from 2000 to 2011," *Expert Systems with Applications,* pp. 11303-11311, 15 September 2012.

[24] N. Kamalraj and A. Malathi, "A Survey on Churn Prediction Techniques in Communication Sector," *International Journal of Computer Applications,* vol. 64, no. 5, pp. 39-42, February 2013.

[25] N. Hashmi, N. Buttand and M. Iqbal, "Customer Churn Prediction in Telecommunication. A Decade Review and Classification," *IJCSI International Journal of Computer Science Issues,* vol. 10, no. 5, pp. 271-282, September 2013.

[26] M. Karnstedt, T. Hennessy, J. Chan, P. Basuchowdhuri, C. Hayes and T. Strufe, "Churn in Social Networks," in *Handbook of Social Network Technologies and Applications*, B. Furht, Ed., Boston, MA, Springer US, 2010, pp. 185-220.

[27] R. J. Oentaryo, E. P. Lim, D. Lo, F. Zhu and P. K. Prasetyo, "Collective Churn Prediction in Social Network," in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2012.

[28] C. Phadke, H. Uzunalioglu, V. B. K. D. Mendiratta and D. Doran, "Prediction of subscriber churn using social network analysis," *Bell Labs Technical Journal,* vol. 17, no. 4, pp. 63-76, March 2013.

[29] S. Motahari, T. Jung, H. Zang, K. Janakiraman, X. Y. Li and K. S. Hoo, "Predicting the influencers on wireless subscriber churn," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, 2014.

[30] Q. Han and P. Ferreira, "The Role of Peer Influence in Churn in Wireless Networks," in *Proceedings of the 2014 International Conference on Social Computing*, Beijing, China, 2014.

[31] J. A. Deri and J. M. F. Moura, "Churn detection in large user networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.

[32] M. N. Abd-Allah, A. Salah and S. R. El-Beltagy, "Enhanced Customer Churn Prediction Using Social Network Analysis," in *Proceedings of the 3rd Workshop on Data-Driven User Behavioral Modeling and Mining from Social Media*, Shanghai, China, 2014.

[33] T. H. Hsu, C. C. Chen, M. F. Chiang, K. W. Hsu and W. C. Peng, "Inferring potential users in mobile social networks," in *2014 International Conference on Data Science and Advanced Analytics (DSAA)*, 2014.

[34] 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*, 2010.

[35] M. Flores-Méndez, M. Postigo-Boix, J. L. Melus-Moreno and B. Stiller, "A model for the mobile market based on customers profile to analyze the churning process," *Wireless Networks,* pp. 1-14, 4 August 2016.

[36] U. Kumar and A. Helmy, "Extract: Mining Social Features from WLAN Traces -- a Gender-based Case Study," in *Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, Bodrum, Turkey, 2010.

[37] E. Papaioannou, C. Georgiadis, P. Kourouthanassis and G. M. Giaglis, "Profiling the mobile phone users and their relationship to the internet services and portals," in *2011 10th International Conference on Mobile Business*, 2011.

[38] A. Peslak, L. J. Shannon and W. Ceccucci, "An empirical study of cell phone and smartphone usage," *Issues in Information Systems,* vol. 12, no. 1, p. 407–417, 2011.

[39] C. K. Shi, X. Hao and R. S. Sharma, "A Cross-Disciplinary Study of What Determines the Adoption of MDS," in *2010 43rd Hawaii International Conference on System Sciences*, 2010.

[40] A. Sell, J. Mezei and P. Walden, "An attitude-based latent class segmentation analysis of mobile phone users," *Telematics and Informatics,* pp. 209-219, May 2014.

[41] A. Kumar and H. Lim, "Age differences in mobile service perceptions: comparison of Generation Y and baby boomers," *Journal of Services Marketing,* vol. 22, no. 7, pp. 568-577, 10 October 2008.

[42] M. O. Jackson, "The Economics of Social Networks," in *Advances in Economics & Econometrics, Theory & Applications: Ninth World Congress of the Econometric Society*, 2006.

[43] D. J. Watts, P. S. Dodds and M. E. Newman, "Identity and search in social networks," *Science,* vol. 296, no. 5571, pp. 1302-1305, 17 May 2002.

[44] M. Anaman and M. Lycett, "Toward a Model of Customer Experience - An Action Research Study within a Mobile Telecommunications Company," in *Proceedings of the 12th International Conference on Enterprise Information Systems*, Funchal, Madeira, Portugal, 2010.

[45] OpenSim Ltd., "OMNeT++ Discrete Event Simulator - Home," 10 March 2017. [Online]. Available: https://omnetpp.org/.

[46] J. Leskovec, "SNAP: Standford Network Analysis Platform," 10 March 2016. [Online]. Available: https://snap.stanford.edu/snap/.

[47] K. M. Harris, C. T. Halpern, E. Whitsel, J. Hussey, J. Tabor, P. Entzel and J. R. Udry, "The National Longitudinal Study of Adolescent to Adult Health: Research Design," 2009. [Online]. Available: http://www.cpc.unc.edu/projects/addhealth/design.

[48] MathWorks, "MATLAB," 2017. [Online]. Available: http://www.mathworks.com/products/matlab. [Accessed 2017].

[49] A. C. Bemmaor and N. Glady, "Modeling Purchasing Behavior with Sudden "Death": A Flexible Customer Lifetime Model," *Management Science,* vol. 58, no. 5, pp. 1012-1021, 2 December 2011.

[50] Hutchinson 3G UK Ltd., "How much data do I need? | Data Calculator," 2017. [Online]. Available: http://www.three.co.uk/datacalculator. [Accessed 14 November 2017].

[51] Verizon Wireless, "Data Calculator," 2017. [Online]. Available: https://www.verizonwireless.com/data-calculator/. [Accessed 14 November 2017].

Appendix A. Parameters of the simulations

This appendix lists the parameters used in the simulations to model customers and MNOs that were taken from [35].

MNO-related parameters can be seen in Table A.1 and Table A.2, showing the parameters of the plan offered by both MNOs.

Minimum bandwidth	384 kb/s
Maximum bandwidth	768 kb/s
Increment/decrement bandwidth step	16 kb/s
Marketing aggressiveness index	90%
Customer service level	90%
Billing complaint index	10%
Handset availability index	90%

Table A.1: MNO General Parameters.

Data cap	2 GB
Price	10 currency units
Overage charges	5 currency units per GB
Billing time	30 days
Subscription time	30 days

Table A.2: MNOs offered data plan.

Data from Eurostat [9], whenever possible, was used as input for the customer module. For other characteristics and parameters where data from Eurostat could not be used as input, distributions and values we deemed appropriate according to findings already mentioned in [35] were used instead. Eurostat data from 2014 was used for gender, age and income distribution. For loyalty, a function that increments with age was used. For tech savviness, a function that decreases with age was used.

The particular distributions, probabilities and functions to assign the characteristics of customer's profile, as already explained in 3.1, will now be mentioned. Eurostat data from 2014 was used as input of the MathWorks MATLAB [48] curve-fitting tool to create equations (A.1), (A.2) and (A.3). The following are the applied distributions and functions:

- *Gender and age*: Discrete distributions from Eurostat's data [9] for gender share and age distribution according to gender.
- *Budget*: The Gamma-Gompertz function, reported in [49] was chosen as wealth distribution because it was the distribution that better fitted maximum monthly income percentile data. Parameters for this function are shown in Table A.3.

Wealth distribution	
s	0.2782
b	24.56
β	52.72

Table A.3: Wealth distribution function parameters.

The function meanbudget (A.1) was obtained from Eurostat's data [9] regarding mean income for age intervals according to sex from. Parameters for this function can be seen in Table A.4.

$$\text{meanbudget}(age) = a \cdot e^{b \cdot age} + c \cdot e^{d \cdot age} \tag{A.1}$$

Mean budget function		
	Males	**Females**
a	-0.2411	-0.9996
b	0.07355	0.06179
c	61.79	56.94
d	0.01057	0.01395

Table A.4: Mean budget function parameters.

- *Tech savvy*: Data for computer and Internet usage in different age intervals was used to obtain the function probabilityTechSavvy (A.2). Table A.5 shows the values for the parameters.

$$\text{probabilityTechSavvy}(age) = (p1 \cdot age^2 + p2 \cdot age + p3) / (age + q1) \tag{A.2}$$

Tech-savvy probability	
$p1$	-0.02287
$p2$	2.061
$p3$	6.4x10-3
$q1$	14.88

Table A.5: Tech-savvy probability function parameters.

- *Loyalty*: The loyalty index (A.3) was used following the findings in [41]. The values of these parameters are shown in Table A.6.

$$\text{loyalty } age = \frac{age - age_{min}}{age_{max} - age_{min}} \cdot (loyalty_{max} - loyalty_{min}) + loyalty_{min} \tag{A.3}$$

	Loyalty
age_{min}	14
age_{max}	74
$loyalty_{min}$	0.8
$loyalty_{max}$	0.2

Table A.6: Loyalty function parameters.

- *Usage profiles*: We consider six usage profiles, but others could be chosen. It seems quite reasonable to consider: moderate use customers, customers that play online games, customers that use the service for work-related activities, customers that are very active in social networks, customers who mainly listen to music and customers who mainly watch videos. Table A.7 presents the number of events that each profile generate in a month.

	Moderate	Gamer	Social	Music	Worker	Video
Email	30	150	150	150	2400	150
Music stream (min)	0	0	240	1200	0	240
Music download (song)	5	20	30	180	10	30
Video stream (min)	12	120	120	600	0	1800
Video call (mins)	0	0	20	0	240	0
Audio call (mins)	0	0	120	0	480	0
Surf web (pages)	150	500	1500	600	600	600
Social media (posts w/photo)	600	1500	4500	1500	1500	1500
App/game download	5	50	20	20	5	20
Online gaming (min)	0	3600	0	450	0	450
Instant messages	600	600	12000	1500	1500	1500
File download	5	5	5	5	5	5

Table A.7: Applications and usage profiles.

- *Usage profile probabilities*: Probabilities for usage profiles according to gender and tech savviness are shown in Table A.8. Females were given a higher probability to be social, and a lower probability to be gamers. Males were given a higher probability to be gamers, and a lower probability to be social. Tech-savvy customers were given a higher probability to be gamers. Nontech-savvy customers were given a higher probability to be moderate.

	Moderate	Gamer	Social	Music	Worker	Video
Female	0.16	0.08	0.25	0.16	0.16	0.16
Male	0.16	0.25	0.08	0.16	0.16	0.16
Tech savvy	0.07	0.21	0.21	0.21	0.14	0.14
Non tech savvy	0.30	0.10	0.10	0.10	0.20	0.20

Table A.8: Usage profile probabilities for sex and tech savviness.

- *QoS importance level*: For each usage profile we assigned different QoS importance levels. The share of each usage profile in the population and the importance each usage profile gives to QoS are shown in Table A.9.

	Share	QoS importance
Moderate	26%	70%
Gamer	34%	95%
Social	21%	80%
Music	8%	90%
Worker	7%	95%
Video	4%	90%

Table A.9: Usage profile share and QoS importance.

- *Dissatisfaction level*: Different factors for evaluating dissatisfaction levels with the service plan are assigned to each usage profile. These dissatisfaction levels, shown in Table A.10, define how fast customers with different usage profiles start considering changing their current plan. Each period, in this case, a month, customers evaluate their plans summing these dissatisfaction levels for each factor they are not satisfied with. When customers reach a dissatisfaction level of 100, they will start looking for other plans.

	Billing	QoS	Data cap
Moderate	40	10	10
Gamer	30	40	30
Social	30	20	30
Music	20	30	20
Worker	10	40	30
Video	20	30	20

Table A.10: Dissatisfaction step by usage profiles.

- *Number of devices*: As already mentioned, customers can have more than one type of mobile device. Device probabilities for each usage profile are shown in Table A.11. Moderate, music and social customers were given a higher probability to have a smartphone. Gamers and video customers were given a higher probability of having a tablet. Workers were given a higher probability of

having a mobile computer.

	Smartphone	Tablet	Mobile computer
Moderate	0.80	0.10	0.10
Gamer	0.10	0.80	0.10
Social	0.25	0.25	0.50
Music	0.50	0.40	0.10
Worker	0.80	0.10	0.10
Video	0.10	0.80	0.10

Table A.11: Device probabilities for usage profiles.

- *Traffic volume*: Data rates for each type of application's event according to the type of device are shown in Table A.12. This table was made with data from MNO websites from different countries [50] [51].

	Smartphone	Tablet	Mobile computer
Email (no attach 75%, w/attach 25%)	20KB/300KB	20KB/300KB	20KB/300KB
Music stream (min)	1MB	1MB	1MB
Music download (song)	7MB	7MB	7MB
Video stream (min)	5.1MB	5.1MB	15MB
Video call (mins)	12MB	12MB	12MB
Audio call (mins)	2MB	2MB	2MB
Surf web (pages)	1MB	1MB	2MB
Social media (posts w/photo)	350KB	350KB	500KB
App/game download	4MB	5MB	30MB
Online gaming (min)	85KB	85KB	85KB
Instant messages	15KB	15KB	15KB
File download	4MB	4MB	30MB

Table A.12: Applications traffic and devices.