



**HAL**  
open science

## Influencers and followers on Facebook in Georgia

Zakaria Babutsidze, Tengiz Lomitashvili, Koba Turmanidze

► **To cite this version:**

Zakaria Babutsidze, Tengiz Lomitashvili, Koba Turmanidze. Influencers and followers on Facebook in Georgia. [Research Report] Academic Swiss Caucasus Net. 2012. halshs-01926929

**HAL Id: halshs-01926929**

**<https://shs.hal.science/halshs-01926929>**

Submitted on 19 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Influencers and followers on*  
**Facebook** in Georgia



AN INITIATIVE OF GEBERT RÜF STIFTUNG IN COOPERATION  
WITH THE UNIVERSITY OF FRIBOURG

**Zakaria Babutsidze**

OFCE Sciences Po, 250 Rue Albert Einstein, 06560 Valbonne, France  
SKEMA Business School, 60 Rue Dostoievski, 06902 Sophia Antipolis, France  
Center for Social Sciences, 3 Chavchavadze Avenue, 0179 Tbilisi, Georgia

**Tengiz Lomitashvili**

Center for Social Sciences, 3 Chavchavadze Avenue, 0179 Tbilisi, Georgia

**Koba Turmanidze**

Caucasus Research Resource Center, 16 Zandukeli Street, 0108 Tbilisi, Georgia  
Center for Social Sciences, 3 Chavchavadze Avenue, 0179 Tbilisi, Georgia

November 2012

## Contents

Summary.....	3
Highlights .....	4
Motivation .....	5
Data collection process .....	10
Data description .....	14
Network approach.....	18
Central players in the network.....	22
Influence .....	26
Validity of results.....	32
Conclusions.....	33
List of references.....	35
Appendix .....	37
Acknowledgements .....	41

## Summary

The paper analyzes a unique dataset with information on the demographic characteristics and behavior of Georgian users of Facebook. Based on this data, we construct a network of 2589 Facebook users and 4473 connections among them. This network is representative of the larger network of all Georgian Facebook users. It is the primary object of the analysis. The data also include the behavior of a subset of the respondents with respect to the opportunity to exert influence over their social contacts.

We conclude that the behavior and demographic composition of Georgian Facebook users do not significantly differ from those of users in other countries. Users between 17 and 35 years of age constitute over 80% of the Georgian Facebook population. Women are over-represented in Facebook in Georgia. They account for over 60% of users. Furthermore, young women occupy key positions in the network and possess disproportionately high structural power, which provides them unique opportunity to influence others' behavior and control the flow of information from various parts of the network.

## Highlights

- Many of the Georgian Facebook users might not be aware that by using applications on Facebook, they have been granting commercial institutions access to their private data.
- Users between 17 and 25 years old dominate Facebook in Georgia.
- The average Georgian Facebook user spends a considerable amount of time on Facebook.
- The Georgian Facebook network can handle significant diversity in opinions and behavior and does not require absolute conformity among the users.
- Information can reach most Georgian Facebook users quickly regardless of where in the network it originates.
- The number of friends on Facebook is a good proxy for the extent of direct influence. It is also a good proxy for the degree of helpfulness with respect to the diffusion of others' influence.
- Females influence their social contacts more than do males. Their friends also influence female users more.
- Young Facebook users are more eager to influence others and more reluctant to follow.
- Followers are practically indistinguishable from the general Facebook user population.
- Influencers have a large number of connections and occupy positions that allow them to control the diffusion of information from other nodes.
- Georgian female Facebook users influence their social contacts more intensely relative to the power that they possess.
- Young Georgian Facebook users underutilize their power in the network.

## Motivation

Social interaction among individuals has always been important. It represents the core of social life. Many things in life acquire meaning only in social context; therefore, we pay a great deal of attention to it.

People have been cultivating social networks since the beginning of history. However, as social interactions have typically taken place in the form of face-to-face interaction, little record has been left behind. The recent rise of the Internet and social networking websites presents a unique opportunity to analyze detailed data on social interactions. Activities on social networking sites leave trails that researchers can use to reconstruct subjects' social networking experience. Because many individuals engage in various debates on political and social topics, the analysis of social networks is particularly important.

Political involvement is high in modern electronic social networks. A recent Pew study (Rainie and Smith 2012) finds that 37% of the users of social networking sites express their political views on these platforms. Furthermore, 67% of users receive information about others' political views. The same body of work has attempted to understand how users react to different types of information about the views of their friends on social networking sites. Here, the main interest lies in behavior toward contacts with differing political views. Results have indicated that banning, deleting and ignoring other users are frequent behaviors.

These findings suggest that social networks are constructed based on homophily. Thus, people with similar political views are friends, whereas users with differing political views are not connected. This consideration has raised important concerns about the formation of "echo chambers" online (Manjoo 2008, Sunstein 2009). "Echo chambers" refer to the phenomenon that each person encounters opinions that are in line with his own. Therefore, his opinion is

reinforced and never challenged. This phenomenon is dangerous because it undermines the discussion process among people with diverse beliefs, which is an important building block of modern democracy.

In fact, some scholars argue that “echo chambers” characterize many activities on the Internet, not only social networking. Few authors (Kohut 2004, Von Drehle 2004, Mullainathan and Schleifer 2005, Carr 2006, Friedman 2009) argue that the large amount of news available online makes it easier for people to position themselves to receive information that confirms their views, which is similar to “echo chambers”.

However, a recent large-scale Facebook study (Bakshy et al. 2012) concludes that echo chambers are rare. The study analyzes the behavior of 253 million Facebook users and concludes that even if social networks are based on homophily, diverse ideas can easily penetrate these networks. Few, personally unimportant, links allow information to navigate through social networks easily. This has to do with the redundancy in the possible “chamber” and freshness of the information from the “weak ties”.

This idea dates back to the 1970s, when Mark Granoveter examined the structure of one’s social networks and contributions of different types of contacts toward the value of the information provided to the person. Granoveter (1973) conducted in-depth interviews with workers to investigate how they originally received the information about their jobs. He was primarily interested in whether they received this information from close friends or more distant acquaintances. He called a link to a distant acquaintance a “weak tie” because the frequency with which the information passed through the link was low. Surprisingly, a large share of people named their distant acquaintances as sources of valuable information. By contrast, the share of people who received valuable information from their close friends was small.

An explanation for this finding is that close friends typically share informational space. Therefore, they absorb new information simultaneously. However, acquaintances have distinct informational spaces. Therefore, although information rarely passes through the link, this information has potentially higher value. Social networks that are only built on “strong ties” (i.e., links between close friends) are characterized by “structural holes” (Burt 1992). However, “weak ties” serve as bridges over these structural holes and vastly increase the functionality of the social network. The above-mentioned large-scale Facebook study (Bakshy et al. 2012) found that these “weak ties” were instrumental to eliminating “echo chambers” in social networks. In addition, Gentzkow and Shapiro (2011) find that the Internet is not becoming more segregated over time, which would be the case if “echo chamber”-like behavior existed.

If there are sufficient bridges over structural holes, then networks are not solely composed of likeminded people. In fact, the average Facebook user has over 200 friends, and a significant number of users have a much larger number of contacts. It is difficult to imagine that all of these contacts are likeminded. In addition, a large number of friends implies that it is difficult to pay significant attention to each of them. Therefore, to understand how information flows through the social network and track the possible paths of influence, we must consider the distribution of attention across a user’s friends. It is also important to analyze the variation in attention due to a user’s age and gender.

Another large-scale Facebook study performs this analysis (Backstrom et al. 2011). The authors analyze the data of approximately 16 million heavily active Facebook users to draw conclusions about their attention-paying behavior. A general result that emerges from the study is that attention-paying behavior is stable at the individual level. The authors distinguish between viewing and communication behavior. Viewing behavior can be considered to be passive attention

paying, similar to the consumption of news. By contrast, communication behavior can be viewed as active attention paying, similar to the generation and diffusion of news.

Backstrom et al. (2011) find that older users are more passive. Furthermore, older users disperse their viewing behavior across a greater number of friends than do younger users, who view information from a handful of social contacts. However, in terms of active attention paying, younger users are more proficient. Similar to older users in relation to passive attention, younger viewers disperse their active attention paying across a larger number of social contacts. They communicate information to a larger set of friends. Therefore, younger users are more likely to influence older users. This is somewhat at odds with real-life experience, as older actors are viewed as possessing more wisdom and having greater influence on public opinion. However, this finding might be specific to electronic social networking.

In regard to variation in attention paying due to gender, the study finds that females pay closer attention to a smaller share of friends compared to males. Males disperse their attention, especially active attention paying. Therefore, we might conclude that male users exert an influence on a larger social network, but their efficiency is lower with respect to each social contact that they pay active attention to. By contrast, females influence fewer social contacts, but with higher efficiency.

Interaction across genders differs from that within gender (McPherson et al. 2001). Backstrom et al. (2011) finds that males' attention is balanced across genders. The share of their attention devoted to female social contacts is 53%. However, females devote 60% of attention to other females, and the remaining 40% of female attention is heavily concentrated on a small number of males.

In the current paper, we analyze the social networking activities of Georgian Facebook users. We collected unique data through a

Facebook application that allows us to analyze certain aspects of online social networking behavior, including tie-forming behavior and influence. Unlike Backstrom et al. (2011), who utilize constructed measures of attention, we can directly observe influence in the present data, as users were explicitly motivated to influence each other to install the application.

The major result that emerges from the study is that the demographics and behavior of Georgian users of Facebook do not significantly differ from those of other countries. Young female users dominate Georgian Facebook. They occupy key positions in the network, granting them disproportionate power to influence others' behavior. They are also strategically positioned to control the flow of information from different parts of the social network.

In line with Backstrom et al. (2011), we find that younger Facebook users are active attention payers, whereas older users are passive attention payers. Therefore, younger users are more influential and more difficult to influence than older users. Also in line with Backstrom et al. (2011), we find that female users are more efficient influencers.

## Data collection process

To collect the data on the behavior of Georgian Facebook users, we developed a simple Facebook application. Upon the user's installation of the application, we were granted access to the user's profile including the network of friends. We were interested in users' demographic characteristics (age, sex etc.) and social network (number of friends, identity of friends etc.).

Facebook is populated by numerous applications, which many users install. However, our conjecture was that few of these users know that applications allow their developers to access users' private data. Before installing any new application, Facebook warns users about this access. Without accepting these conditions, one cannot install the application. To test the behavior of Georgian users with respect to reading these general conditions, we created a similar warning that we would obtain access to users' private data and required them to accept these conditions again. In addition, we also stated that the data would only be anonymously used for non-commercial research purposes.

To view the welcome page of the application, which repeated the security warning in Georgian, users had to first accept the conditions presented by Facebook and install the application (and by doing this, grant us access to their data). If users are conscious of the applications' access to data when they accept Facebook's conditions, they should accept the same conditions presented by the welcome page of the present application. However, of the 5629 people who installed the application, 3040 (54%) defected after seeing this information provided by the application developer.

This raises a serious concern about Georgian Facebook users' awareness of the explicit security protocols in Facebook and privacy of their data.

Many of the Georgian Facebook users might not be aware of the fact that by using applications on Facebook, they have been granting commercial institutions access to their private data.

Numerous applications developed by commercial parties are popular in Georgia. The current exercise shows that many of the Georgian Facebook users might not be aware of the fact that they have been granting commercial institutions access to their private data. This behavior might be influenced by the English proficiency of Georgian Facebook users, as Facebook's original warning is in English, whereas our warning was in Georgian. However, this does not change the fact that Georgian Facebook users are unaware of the consequences of their actions on electronic social networking sites.

As stated above, 5629 people installed the application and granted us access to their data. However, once they were warned again about the fact that we could access their profiles, 54% of them defected and refused to continue using the application. Therefore, we discarded the data of these users from the analysis. We only analyzed the data collected from the profiles of 2589 Georgian Facebook users who accepted the application conditions twice.

Upon approval to collect a user's private data, we presented each user with a simple questionnaire. We posed the following four multiple-choice questions:

1. How did you learn about the application?
2. How many times do you go on Facebook per week?
3. How many minutes do you spend on Facebook during one visit?
4. Which age group do you belong to?

Once users submitted the answers to these questions, they were provided the opportunity to invite their friends to install the application. They were presented with a list of their Facebook friends and could choose an unlimited number of people to invite to install the application. People who were invited by their friends to install the application had the option to report this in the first question of the questionnaire. If they chose this option, they were presented with a list of their Facebook friends and asked to select the person who invited them. This gave the selected friend one credit. To encourage

the invitation behavior, we offered valuable gifts to the top three referrers (1st place: iPad, 2nd place: iPod Nano, 3rd place: iPod shuffle). This information was clearly communicated to the users on the final screen of the application.

The invitation mechanism served a double purpose. First, it helped to diffuse the application and collect more data. Second, and more important, it served as the influence exercise. The user only received the credit if her friend explicitly chose her as the referrer. As all of the users sent invitations to a large number of friends and response from friends was limited, each response can be interpreted as the exercise of influence. If user A did not respond to the invitation to the application from user B (given that user A is not yet a user of the application), we can infer that B does not have sufficient influence over A. By contrast, if A has responded (and explicitly chose user B as the source), we can infer that B has influence over A. Influence in this context must be interpreted as the power of persuasion.

As previously mentioned, one major channel through which the application has spread among Georgian Facebook users is explicit invitations from friends who have installed the application. However, if this was the only channel used, the representativeness of the sample would be compromised. We would collect a growing interconnected network that originated from few nodes. Many Georgian Facebook users would not receive the information about the application, and as a consequence, we would not obtain information about them. Therefore, we devised a paid Facebook advertisement campaign that targets all Georgian adult Facebook users. This campaign reached various Georgian Facebook networks and, therefore, significantly improved the representativeness of the sample.

Beyond these two major channels, the application was also diffused by users receiving information on their walls that their friends have installed this application. This channel also significantly contributed to the data collection effort.

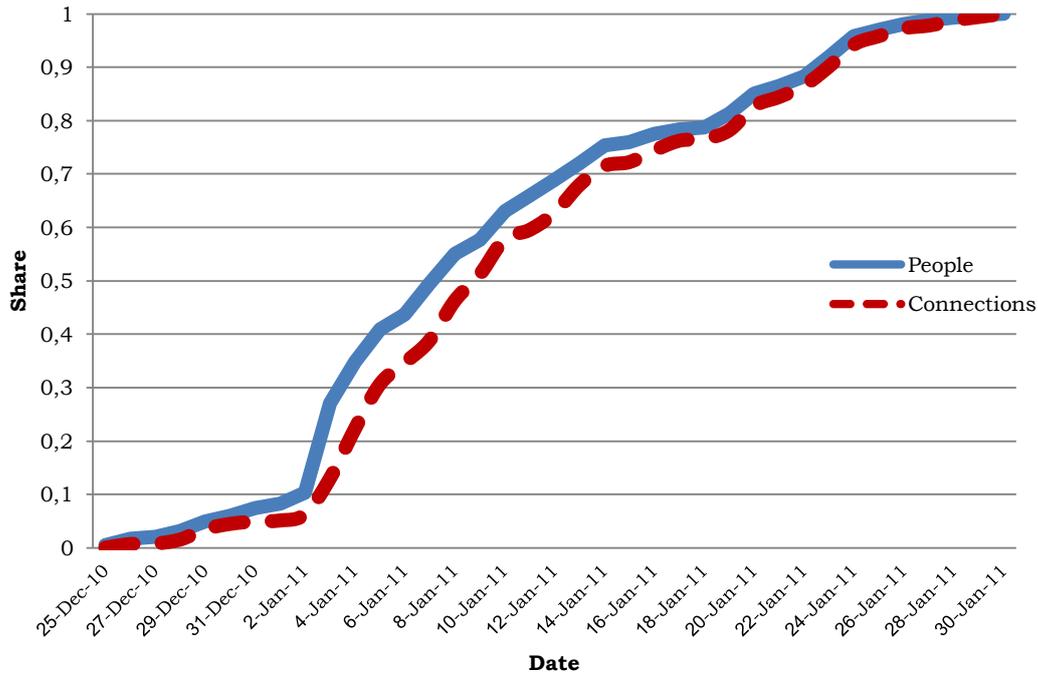


Figure 1: Dynamics of application installation.

The application was launched on December 25, 2010 and ran through January 31, 2011. To ensure that we have not prematurely terminated the data collection process, we plot the dynamics of application installation during this period in figure 1. The figure presents the cumulative number of installations as of a given date as the share of total installations by January 31, 2011. The solid line represents the users, and the dashed line represents the connections among these users. The large jump on January 2, 2011 coincides with the start of the advertising campaign. As shown, the application installation process has the expected S shape, which saturates close to the end of the data collection process. In late January, the growth rate of the number of users and links between them neared zero, indicating that the application installations have saturated. The illustration of network forming over the data collection period is presented in appendix A.

## Data description

During the data collection period, 5629 users installed the application, but only 2589 users responded to the questionnaire. The data of the 3040 users who declined to accept the conditions the second time are discarded. We detect 4473 friendship links between the 2589 users in the database. Furthermore, we detect 1552 acts of influence.

The general Georgian Facebook user's profile is similar to that of the general user in developed countries. For example, female users comprise 62% and male users 38% of the Georgian Facebook users. This result is in line with the findings from data on US Facebook users.<sup>1</sup> Furthermore, the distribution of user age profiles, presented in the right panel of figure 2, closely follows the user age distribution in the United States.<sup>2</sup>

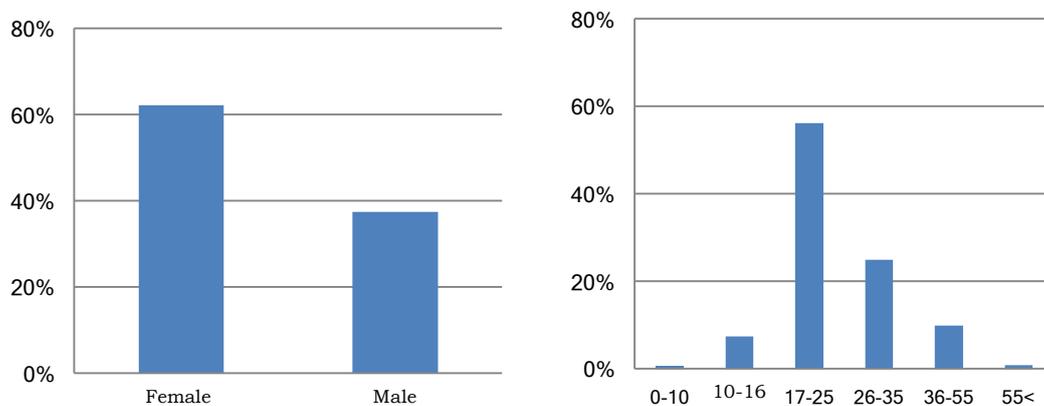


Figure 2: Gender (left) and age (right) breakdown of Georgian Facebook users.

This distribution implies that users between 17 and 25 years old dominate Facebook in Georgia. This age group is followed by users between 26 and 35 years old. Users who are younger than 17 years and older than 35 compose less than 19% of the user population.

Users between 17 and 25 years old dominate Facebook in Georgia.

<sup>1</sup> Source: [www.insidefacebook.com](http://www.insidefacebook.com)

<sup>2</sup> Ibid.

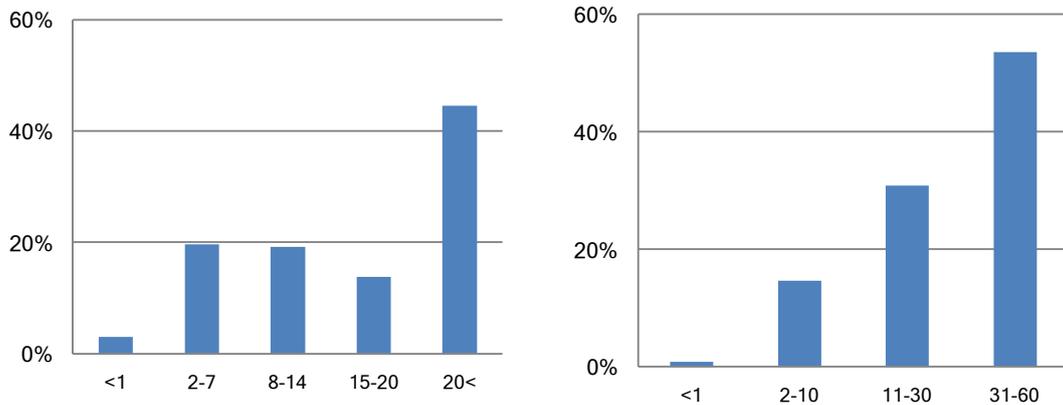


Figure 3: Facebook usage frequency (left) and intensity (right).

We also examine the frequency and intensity of Facebook usage in Georgia. The distributions of frequency of visits (per week) and duration of stay during each visit (in minutes) are presented in figure 3. The data suggest that the average Georgian Facebook user spends a considerable amount of time on Facebook.

The average Georgian Facebook user spends a considerable amount of time on Facebook.

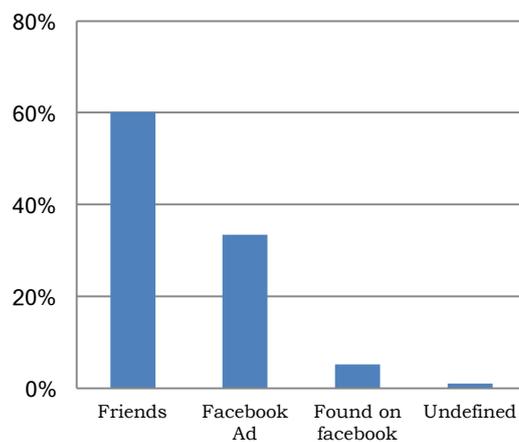


Figure 4: Source of information about the application.

Figure 4 presents the frequency distribution of referrals to the application. As shown, 60% of respondents were referred to the application by their friends. Of the users, 33% found the application through the advertising campaign that did not target any particular demographic group. Therefore, the sample is representative of Georgian Facebook users.

## *Network architecture*

The current paper investigates Georgian users of Facebook. These users constitute a sub-network of the much larger network that includes all Facebook users. It is interesting to observe how the sample relates to its umbrella network. Of note, we have limited information about the subjects. For example, we detect only the links between the two users who are in the dataset.<sup>3</sup> These connections comprise a very small portion of the subjects' connections. However, we have information on the number of friends that each subject has on Facebook. This directly corresponds to the number of connections in the global Facebook network. Technically, the number of connections each person (node in network) has is referred to as the degree of a node.

The overwhelming evidence suggests that many real-life networks (natural and manmade) have particular structures. If the connections through nodes were formed randomly, the degree distribution of nodes would be normal (Erdos and Renyi 1960). Therefore, the degree of the large majority of the nodes would be very close to the mean of the degree distribution, and there would be few extremes (i.e., nodes with very low and very high degrees). However, empirical evidence suggests that networks are characterized by a scale-free nature (Barabasi and Albert 1999, Barabasi 2003). Nodes with an extreme degree occur at a much higher frequency than is predicted by the normal distribution. The distribution that best approximates the degree distribution in social networks is the log-normal distribution (Ebel et al. 2002), whereas other types of real-life networks are best described by the power law distribution (Border et al. 2000). Various studies have found that the global Facebook network conforms to the log-normal degree distribution (Backstrom et al. 2011, Bakhty et al. 2012).

---

<sup>3</sup> In principle, we also detect links to individuals outside the dataset. However, as we do not have access to the profiles of those individuals, we cannot understand the representativeness of the detected links. Furthermore, many of the individuals outside the dataset are foreigners. Therefore, we discard this information.

The left panel of figure 5 presents the distribution of the number of friends of the subjects. Of note, the number of friends in the dataset corresponds to the degree of the user in the global Facebook network. Each marker on the graph shows how many people have the given number of connections. For example, if the marker corresponds to  $X$  on the horizontal axes and  $Y$  on the vertical axes, then  $Y$  people have  $X$  connections in the global Facebook network. The right panel of figure 5 presents the cumulative distribution of the number of friends. This plot shows the share of users with a degree that is equal to or greater than the user under consideration. In this case, the marker corresponding to  $X$  and  $Y$  indicates that  $Y$  share of people have a degree that is greater than or equal to  $X$ . Plots are on a log-log scale to easily identify the log-normal distributions. Both the density and cumulative density distributions of the log-normal distribution have a linear right tail on the log-log scale.

Figure 5 confirms the representativeness of the data with respect to the global Facebook network. The right tail of both distributions is approximated by the straight line. This implies that users with very few and very many friends are frequent.

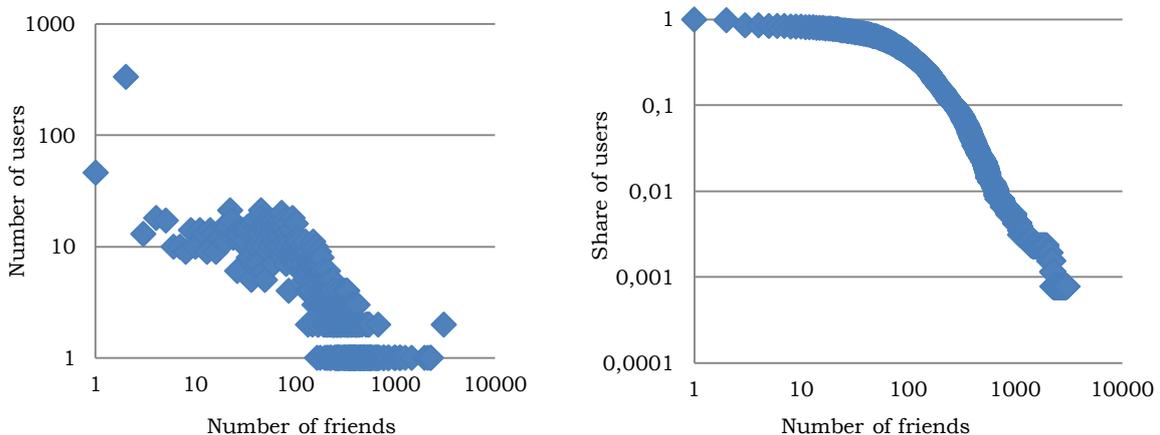


Figure 5: Degree (left) and cumulative degree (right) distributions in the global Facebook network.

## Network approach

In this section, we present the characterization of the Georgian Facebook users' network that has been constructed from the data of 2589 respondents. In the previous section, we presented the degree distribution of the sample in the global Facebook network. Figure 6 presents the degree distribution of each node (user) in the local Georgian Facebook network. Of note, the local Facebook network is in principle only a sub-network of all Georgian users. This network comprises only 2589 nodes. The network is illustrated in the lower right panel of figure 12 in appendix A.

As shown in the figure, the degree distribution is positively skewed and possesses a heavy right tail. The average degree in the network is 3.45, and the mode is 1. The skewedness is due to 8 users in the data with a degree above 30.

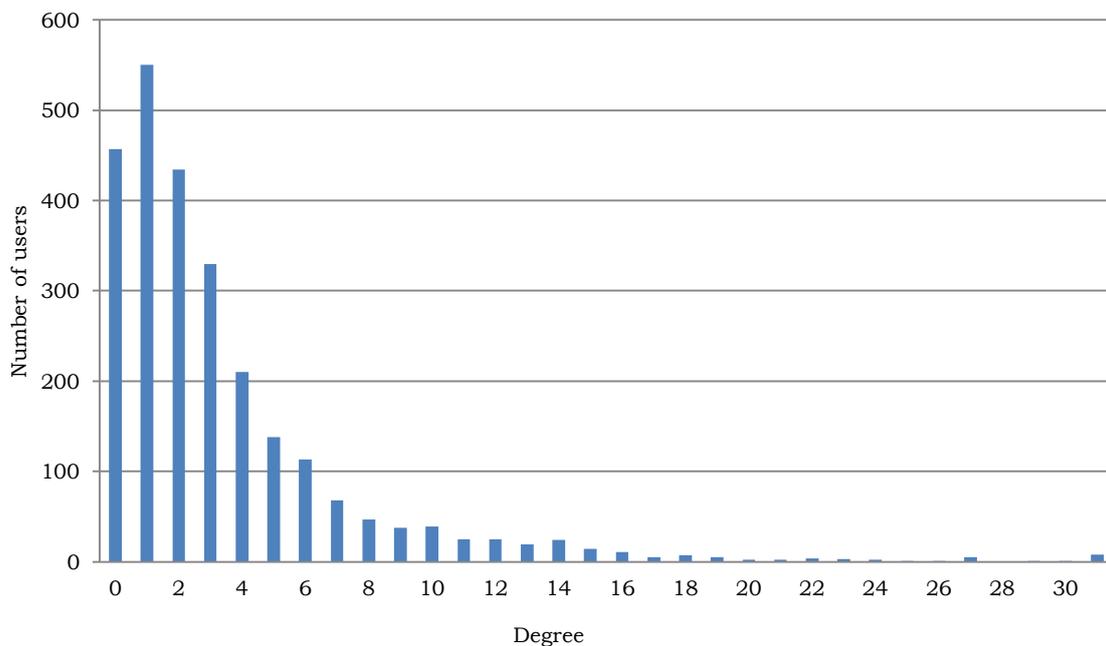


Figure 6: Degree distribution in the Facebook network.

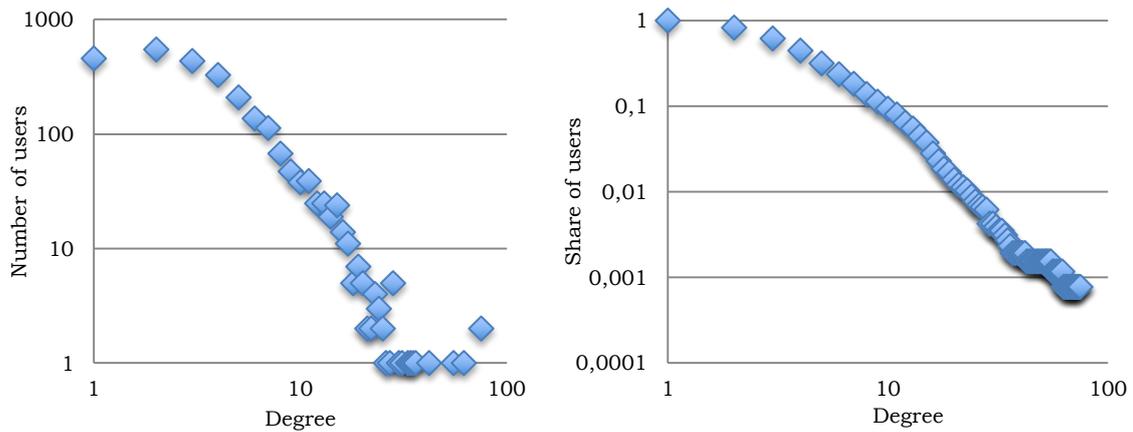


Figure 7: Degree (left) and cumulative degree (right) distributions on a log-log scale.

To determine how the local network relates to the anticipated social network structures, we plot the same distribution on a log-log scale, as presented in the left panel of figure 7. We plot the cumulative distribution in the right panel of the same figure, which is also on a log-log scale. As demonstrated, the local Facebook distribution conforms to the log-normal distribution.

An important characteristic of interaction in the social networks is the distance between two nodes (people). If two people are connected to each other, they can easily affect each other's choices. However, if two people are not directly connected, they can have some influence over each other's opinions and choices through common friends. In this case, a person can affect a friend of her friend. This generalizes over the social distance that is measured by the minimum number of interactions that one needs to reach node B from node A in the network. It is intuitive that the potential influence is inversely related to the social distance, i.e., the larger the distance is, the smaller the influence is. However, as long as there exists at least one route to reach node B from node A, A can potentially affect the opinion of B.

In this respect, two network characteristics are important. One such characteristic is the *diameter of the network*. The diameter of the network is the length of the longest route between two nodes in the network, where route is defined as the shortest possible path between a given pair of nodes. As the network contains some isolated nodes

(nodes that are not linked to any other node), this distance cannot be calculated. In other words, the diameter of the network is infinite. Therefore, information or influence cannot originate from any node and reach all the other nodes of the network. However, there is heterogeneity in the importance of the nodes. Typically, the nodes that are isolated are socially less active and cannot assert their influence. They are members of virtually any network that is created from the data with some degree of random sampling. In this case, we concentrate on the so-called giant component of the network. This approach disregards isolated nodes and small isolated interconnected groups of nodes. The giant component consists of nearly 73% of the respondents in the network. It consists of 1889 nodes and 4145 links. The average degree in this sub-network is 4.39. As this is a connected network, the diameter is finite by definition. In this case, the diameter is 18. Therefore, the longest path from one node to another in the network is through 17 other nodes. This number is quite high for such a network and implies that the Georgian Facebook network can handle significant diversity in opinions and behavior and does not require absolute conformity among the users.

The Georgian Facebook network can handle significant diversity in opinions and behavior and does not require absolute conformity among the users.

The other important characteristic that measures the reach of influence in the network is the *average shortest path length* between the nodes. This measure, similar to the network diameter, concerns the distance between the nodes. However, in this case, it is the average distance between two users on Georgian Facebook. The average shortest path length measures the distance between two randomly chosen users on Georgian Facebook. This measure was pioneered by Stanley Milgram, who observed that randomly selected pairs of people were surprisingly close to each other in social space (Travers and Milgram 1961). He termed this phenomenon as the Small World. This phenomenon has been confirmed to characterize both social and physical networks (Wagner and Fell 2001) and comprehensive models have been created to

explain its emergence (Watts and Strogatz 1998). The most well-known incarnation of the Small World phenomenon is six degrees of separation (Watts 2003). Starting with Milgram's experiment and ending with The Kevin Bacon Game,<sup>4</sup> it has been observed that the average shortest path length in many diverse social and physical networks is about 6. Surprisingly, this does not depend on the size of the network. In fact, a recent study of 721 million Facebook users finds that the distance between two random users has been decreasing and currently stands at 4.74 (Backstrom et al. 2012). The current network conforms to six degrees of separation. The average path length in the Georgian Facebook network is 6.01, indicating that information (and influence) can reach most Georgian Facebook users quickly regardless of where in the network it originates.

Information can reach most Georgian Facebook users quickly regardless of where in the network it originates.

---

<sup>4</sup> <http://www.thekevinbacongame.com>

## Central players in the network

In the previous section, we characterized the network of Georgian Facebook users. In this section, we take a micro approach and attempt to characterize the individuals' power and position in the network. As the study is anonymous, we cannot name the individuals who occupy special positions; however, many interesting conclusions can be drawn.

There are three important node centrality measures in networks, degree centrality, closeness centrality and betweenness centrality. Before discussing the analysis of the data, we outline each of the three measures.

Degree centrality is measured as a given user's number of friends. As the number of friends increases, the number of people the user can directly influence increases. Therefore, we expect that nodes with a higher degree centrality are powerful players in the network. In the data, we observe two such measures. One measure is the degree in the global network, which is measured as the number of friends on Facebook. The other is the degree in the local network, which is measured as the number of friends in the dataset.

Closeness centrality measures how close the node is to all other nodes in the social network. This is measured in terms of social distance (similar to shortest path length and diameter, which were discussed in the previous section). As closeness to all other users increases, the assertion of influence, directly and indirectly, becomes easier.

Betweenness centrality measures how important the node is in terms of connecting other nodes to each other. In other words, it relates to the number of shortest paths between all of the pairs of users that go through this particular node. For example, if the node lies on the only route between the two large components in the network, its betweenness centrality is high due to its crucial importance for passing information from one part of the network to another. This idea

relates to the structural holes and weak ties (Granoveter 1973). Nodes with high betweenness centrality bridge the holes that would constitute the structural divide between substantial parts of the network. Therefore, they carry the important auxiliary function of aiding in the spread of information and influence of other people.

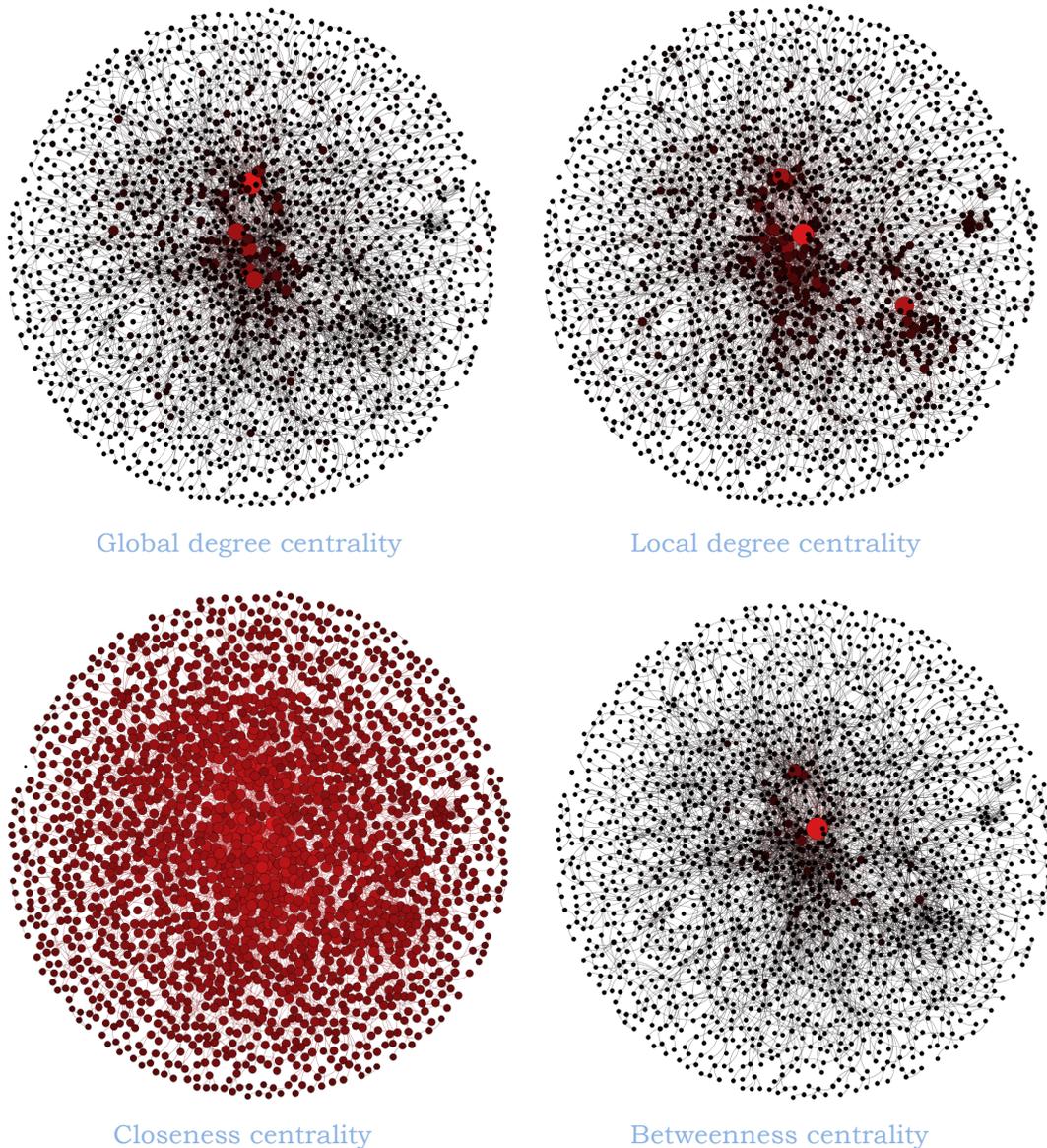


Figure 8: Illustration of node importance with respect to the centrality measures.

Figure 8 illustrates the user importance distributions across several measurements in the giant component of the network. In each panel, the color and size of the node corresponds to the centrality of the node with respect to the appropriate indicator. The upper panels illustrate the degree centrality distributions. Degree centrality is calculated with

respect to two networks. The left panel concerns the global Facebook network, for which degree centrality is measured as the number of friends the user has on Facebook. The right panel concerns the local network, for which degree centrality measures the number of friends each user has in the sample. The lower panels illustrate the importance distributions with respect to closeness and betweenness centralities.

	Number of friends	Degree centrality	Closeness centrality	Betweenness centrality
Number of friends	1	0.521	0.042	0.458
Degree centrality	-	1	0.126	0.758
Closeness centrality	-	-	1	0.012
Betweenness centrality	-	-	-	1

Table 1: Correlations across variables in data.

Each of these measures characterizes a node’s importance for the functioning of the network. Figure 8 shows that there is some overlap between the measurements of importance. However, there is a discrepancy in calculation difficulty across the measures. For example, degree centrality is easily calculable based on a portion of the graph, whereas closeness and betweenness centralities require much more detailed information and computational sophistication. Therefore, to characterize a node’s importance based only on degree centrality, we calculate correlations across these measures. The correlations are presented in table 1. As shown in the table, number of friends (a readily visible node characteristic on Facebook) correlates with degree centrality and betweenness centrality. Its correlation with closeness centrality is weak. This leads us to conclude that the number of friends on Facebook is a good proxy for the extent of direct influence. It is also a good proxy for the degree of helpfulness with respect to the diffusion of others’ influence. However, it does not quantify well the

The number of friends on Facebook is a good proxy for the extent of direct influence. It is also a good proxy for the degree of helpfulness with respect to the diffusion of others’ influence.

power of the node with respect to indirect influence. Thus, indirect influence does not seem to be related to the number of friends.

## Influence

In this section, we focus on the analysis of the data generated by the influence exercise. Of note, the application allowed users to send invitations to their Facebook contacts. It also allowed newcomers to state which friend referred them to the application. We interpret each disclosed referral as the influence. We state that user A has exercised influence over user B if B has listed A as the referrer to the application.

All of the exercises of influence are recorded in the data; therefore, we can clearly identify people who have influenced others (A-type nodes) and those who have been influenced (B-type nodes). We label the former nodes as influencers and the latter nodes as followers. Of note, one node can be both an influencer and a follower. Furthermore, a node can be neither an influencer nor a follower.

We detect 932 influencers and 1552 followers. They constitute a network with 2296 nodes and 1552 connections. The characterization of the network creation dynamics is presented in figure 13 in appendix B. These dynamics are illustrated in appendix C. As demonstrated in the figures, the influencer-follower network has been created through temporal dynamics that are similar to the original network creation. Thus, the influence exercise has not been temporarily clustered and has been stretched across the entire data collection period.

Figure 14 in appendix B presents the degree and cumulative degree distributions that characterize the influencer-follower network. Once again, the degree distribution seems to be well characterized by the log-normal distribution. This is interesting because the current network is not a typical social network, as it represents only a sub-network of the underlying network. Yet, it conforms to the general regularities and is characterized by a large number of nodes with a small or large number of connections.

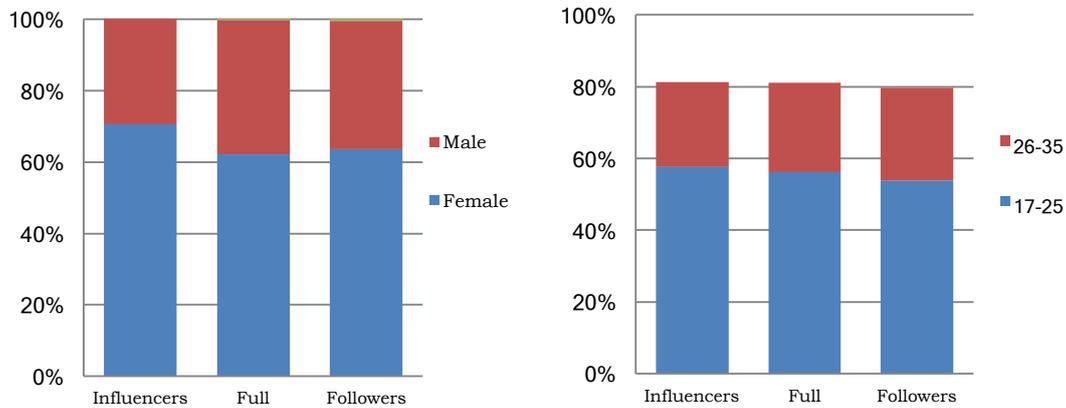


Figure 9: Comparison of gender (left) and age (right) profiles.

The difference between the two populations in the network is quite interesting. Figure 9 contrasts the gender and age profiles of influencers, followers and the general Georgian Facebook user population. A complete analysis of this contrast with respect to these and Facebook usage intensity variables is presented in appendix D. An important conclusion from this analysis is that females are engaged in the exercise more actively. Females influence their social contacts more than do males. Their friends also influence female users more; however, this difference is small.

Females influence their social contacts more than do males. Their friends also influence female users more.

There are no stark differences by age. However, 17-25 year olds seem to be overrepresented in the population of influencers and underrepresented in the population of followers. Thus, young Facebook users are more eager (or able) to influence others and more reluctant to follow.

Young Facebook users are more eager to influence others and more reluctant to follow.

To characterize influencers and followers more accurately, we perform two-sided t-tests on the mean difference between influencers and the full population and between followers and the full population. Table 2 presents p-values for these tests. As shown, followers are practically indistinguishable from the general Facebook user population. However, influencers are statistically different from the average Georgian Facebook user. Influencers have a large number of

Followers are practically indistinguishable from the general Facebook user population.

connections and occupy positions that allow them to control the diffusion of information from other nodes.

	Number of friends	Degree centrality	Closeness centrality	Betweenness centrality
Full vs. Influencers	0.03	0.00	0.06	0.00
Full vs. Followers	0.86	0.00	0.24	0.85

Table 2: Test of mean difference in key characteristics. The values presented are the p-values of two-tailed t-test on mean difference with unrestricted variance.

Furthermore, figure 10 presents the dynamics of the arrival of influencers and followers during the data collection process. It shows that followers are randomly scattered across the data-gathering period. However, most of the influencers have arrived during the early stages of data collection. Therefore, it is possible that influencers are more curious about the novelties that appear on Facebook and easily attracted to the application. Alternatively, the realization of influence may require time. In other words, invitation efforts (that could have been repeated) require time to trigger the potential follower to follow. In this case, we would see a disproportionate share of followers to the early entrants. Consequently, a disproportionately high number of early entrants are qualified as influencers. Both of these possible causes explain the patterns in the data presented in figure 10. Unfortunately, we cannot determine which of the two possible causes affected the data.

Influencers have a large number of connections and occupy positions that allow them to control the diffusion of information from other nodes.

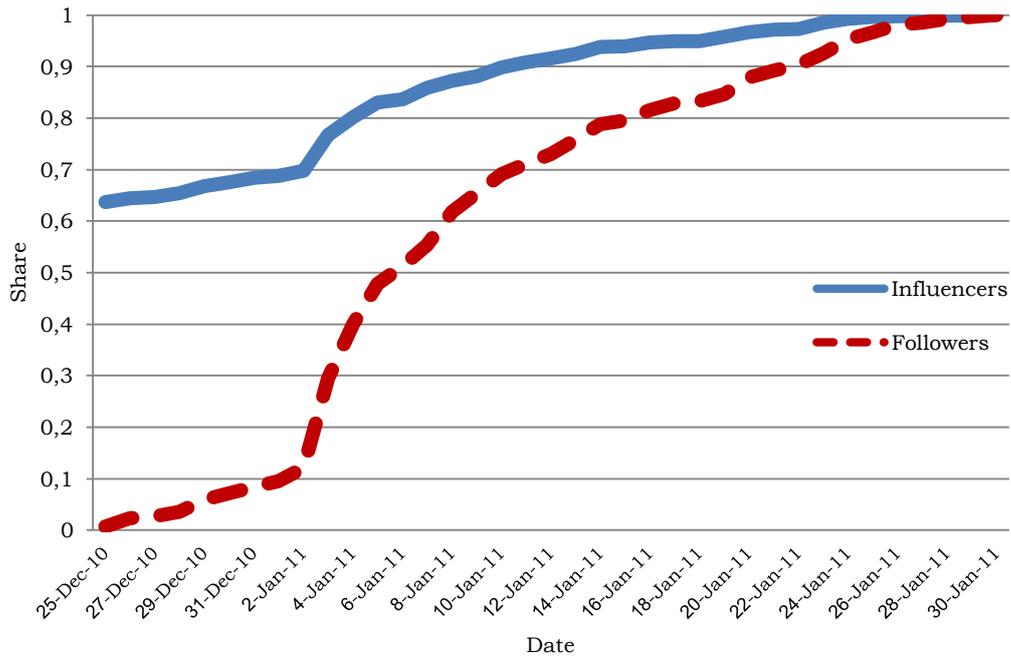


Figure 10: Dynamics of influencer and follower arrival.

### Closer look at influencers

In previous section, we identified the central players in Georgian Facebook network. These players can potentially be powerful in influencing the public opinion held by the social network users. In the first part of this section, we identify the users who have influenced the actions of other users. Below, we attempt to understand how these two groups of users are related to each other. In other words, we examine whether users who are potentially powerful have used their power during the exercise.

As demonstrated above, the number of connections that a user has is a good predictor of the user's general power in the network. In this respect, we identify two types of powerful users. The user is *globally powerful* if he commands a high rank in terms of total number of Facebook friends. By contrast, the user is *locally powerful* if he commands a high rank in terms of connections in the sample. A user can be both globally and locally powerful. To contrast the groups of influencers with the groups of locally and globally powerful users, we must truncate the power ranking. To avoid problems due to the

difference in sample size, we limit the size of the powerful user population to the size of the population of influencers. This method solves the truncation issue.

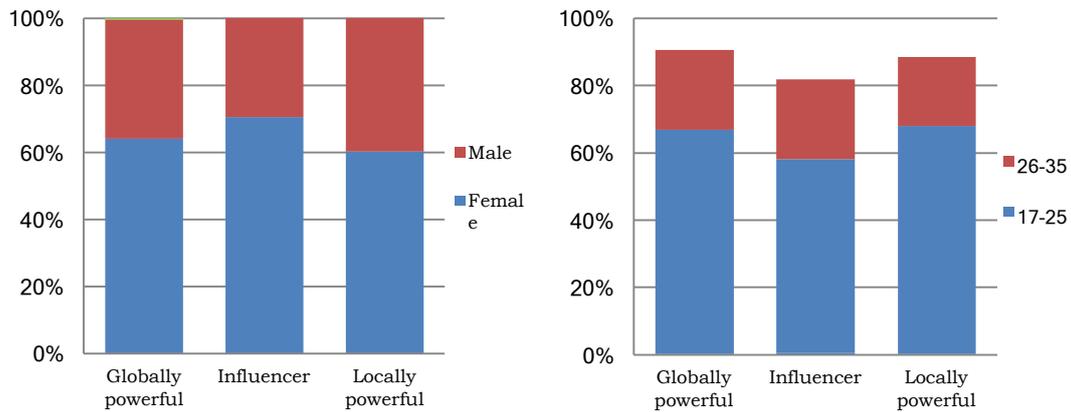


Figure 11: Comparison of gender (left) and age (right) profiles of powerful users with those of influencers.

Figure 11 presents the gender and age profiles of the influencers and the globally and locally powerful users. Females are overrepresented in sample of influencers; thus, they perform better relative to the power that they possess. By contrast, 17-25 year olds are underrepresented in the sample of influencers, suggesting that young Facebook users underutilize their power in the network.

Georgian female Facebook users influence their social contacts more intensely relative to the power that they possess.

	Number of Friends	Degree centrality	Closeness centrality	Betweenness centrality
Influencers vs. globally powerful	0.00	0.00	0.28	0.01
Influencers vs. locally powerful	0.00	0.00	0.04	0.00

Table 3: Test of mean difference between actual and potential influencers.

The data permits us to go one step further and test the mean difference between the basic network characteristics. Table 3 presents the p-values of these t-tests. A somewhat surprising result is that nearly all influencers' characteristics statistically differ from those of both globally and locally powerful users. Thus, in the current exercise, the choice to influence social contacts has been determined by

Young Georgian Facebook users underutilize their power in the network.

characteristics other than the power of the user. This is understandable because the true reason behind the influence exercise was not significant from the individual's prospective. Furthermore, it was deliberately not communicated to the respondents to promote behavior in the environment that is similar to that in the natural environment.

## Validity of results

The discussion in the previous section highlights important limitations in the present research. We have previously mentioned one shortcoming in the influence exercise. Due to the motivation problem, the analysis in this section is somewhat speculative. We cannot determine whether powerful people will use their power to influence the public opinion in the social network when there are more important questions at stake. In presence of a motivation that will trigger the effort of key players in the network, the outcome may be quite different from that observed in the present study. More precisely, we expect that p-values in the tests reported in table 3 would show no statistically significant difference between influencer and powerful populations.

Another important shortcoming that affects the strength of the results is shared among all studies that analyze only a portion of a network. The current network is only a sub-network of the full Georgian Facebook network. Therefore, despite the representativeness of the sample, the user characteristics are definitively miscalculated. However, the important issue is whether this miscalculation is systematic and, therefore, compromises the results. We believe that the miscalculation is not systematic for two reasons. First, numerous tests throughout the paper demonstrate that the sample is representative of the general Georgian Facebook user population in many ways. Most important, it is representative in terms of network structure. We can expect to find the log-normal degree distribution in the full Georgian Facebook network, and we find this distribution in the sample. Second, this particular structure implies the network characteristic of scale-freeness, which emphasizes the importance of the rankings with respect to certain network characteristics rather than the values of the node-specific characteristics.

## Conclusions

In this paper, we have analyzed data collected from Georgian users of Facebook. Facebook is the most popular social networking site. In light of the digitalization of the country, it is important to study the characteristics of such socialization platforms, especially when they are heavily used for sharing users' political views. The data were collected in December 2010 and January 2011 and contain information on 2589 Georgian Facebook users. The data were collected using a Facebook application that was created solely for this purpose. The dataset includes information on the gender, age and Facebook usage intensity of the respondents. Furthermore, we have information on their friend lists. We know both the number and identity of the subjects' contacts on Facebook. This allows us to create a social network of 2589 users. This network, which includes 2589 nodes and 4473 connections between the pairs of nodes, is the main object of the study.

The application also allowed us to formulate an exercise in which users were invited to influence their social contacts. Users were ignorant about the true aim of this exercise and were motivated to exert an influence over their friends by the possibility of winning a valuable prize. We recorded and analyzed 1552 cases of exercised influence.

Several important conclusions result from the analysis. First, drawing on the experience of the application usage, we find that many of the Georgian Facebook users might not be aware that they grant application owners the access to their private data when using applications on Facebook. Facebook warns users before installing every new application; however, this message does not get delivered to Georgian Facebook users (potentially due to a language barrier or carelessness).

Second, the representativeness of the sample allows us to conclude that the demographic characteristics of Facebook users in Georgia are similar to those of users in developed countries. Young female users dominate Georgian Facebook. We also conclude that the social network in Georgia conforms to the well-known phenomenon of six degrees of separation. The average shortest distance between two nodes in the network is 6.01.

Third, the number of friends a user has on Facebook is a good measure of the given user's power in the local network. This concerns the power to influence the public opinion on Facebook and to control the flow of information from other users.

Finally, we find that young female users are more active on Facebook. Young females who occupy strategic positions on Georgian Facebook possess disproportionately large power. We also find that young people have underutilized their power. Conversely, females have "over-reacted" compared to the average Facebook user.

Overall, we can conclude that user demographics and characteristics in Georgian do not systematically differ from those in more developed, populous and researched countries. Therefore, policy conclusions based on the analysis of other representative Facebook populations can be readily applied to the Georgian case. In light of this, the general conclusion that young females should be major targets of efforts to influence opinion is relevant to the Georgian case.

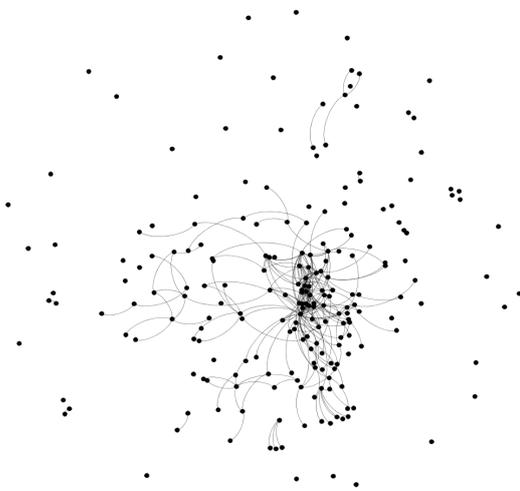
## List of references

1. Backstrom, L., E. Bakshy, J. Kleinberg, T. Lento and I. Rosenn (2011) Center of Attention: How Facebook users allocate attention across friends. In Proceedings of 5th International Conference on Weblogs and Social Media.
2. Backstrom, L., P. Boldi, M. Rosa, J. Ugander and S. Vigna (2012) Four Degrees of Separation. arXiv:1111.4570v3.
3. Bakshy, E., I. Rosenn, C. Marlow and L. Adamic (2012) The role of social networks in information diffusion. WWW2012, Lyon, France.
4. Barabasi, A-L. (2003) *Linked: How everything is connected to everything else and what it means for business, science, and everyday life*. Plume, Penguin Group.
5. Barabasi, A-L. and R. Albert (1999) Emergence of scaling in random networks. *Science* 286:509–512.
6. Border, A., R. Kumar, F. Maghoul, P. Raghavan, S. Sajagopalan, R. Stata, A. Tomkins and J. Wiener (2000) Graph structure in the web. *Computer Networks* 33:309–320.
7. Burt, R. (1992) *Structural Holes: The Social Structure of Competition*. Harvard University Press.
8. Carr, N. (2008) *The Big Switch: Rewiring the World, from Edison to Google*. New York: Norton.
9. Ebel, H., L-I. Mielsch and S. Bornholdt (2002) The scale free topology of e-mail networks. *Physical Review E* 66:035103.
10. Erdos, P. and A. Rényi (1960) On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5:17–61.
11. Friedman, T. (2009) “Advice from Grandma,” *New York Times*, November 21.
12. Gentzkow, M. and J. Shapiro (2011) Ideological segregation online and offline. *Quarterly Journal of Economics* 127:1799–1839.

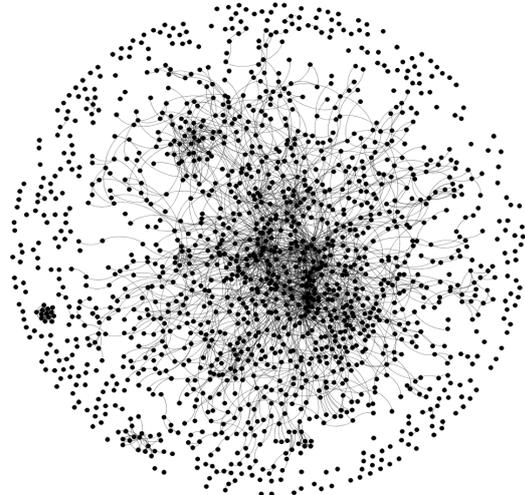
13. Granovetter, M. 1973. The strength of weak ties. *American Journal of Sociology* 78:1360–1380.
14. Holme, P., C. Edling and F. Liljeros (2004) Structure and time evolution of Internet dating community. *Social Networks* 26:155–174.
15. Kohut, A. (2004) “The Nation: Media Myopia; More News Is not Necessarily Good News,” *New York Times*, July 11.
16. Manjoo, F. (2008) *True Enough*. Wiley, Hoboken NJ, United States.
17. McPherson, M., L. Smith-Lovin and M. Brashears (2006) Social isolation in America: Changes in core discussion networks over two decades. *American Sociological Review* 71:353–375.
18. Mullainathan, S. and A. Shleifer (2005) “The Market for News,” *American Economics Review* 95:1031–1053.
19. Newman, M., A-L. Barabasi and D. Watts (2006) *The Structure and Dynamics of Networks*. Princeton University Press, Princeton NJ.
20. Rainie, L. and A. Smith (2012) *Social networking sites and politics*. Pew Internet & American Life project.
21. Sunstein, C. (2009) *Republic.com*. Princeton University press, Princeton NJ, United States.
22. Travers, J. and M. Milgram (1961) An experimental study of the small world problem. *Sociometry* 32:425–443.
23. Von Drehle, D. (2004) “Political Split Is Pervasive: Clash of Cultures Is Driven by Targeted Appeals and Reinforced by Geography,” *Washington Post*, April 25.
24. Wagner, A. and D. Fell (2001) The small world inside large metabolic networks. *Proceedings of the Royal Society of London B* 268:1803–1810.
25. Watts, D. (2003) *Six Degrees: The Science of a Connected Age*. Norton, New York NY.
26. Watts, D. and S. Strogatz (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393:440–442.

## Appendix

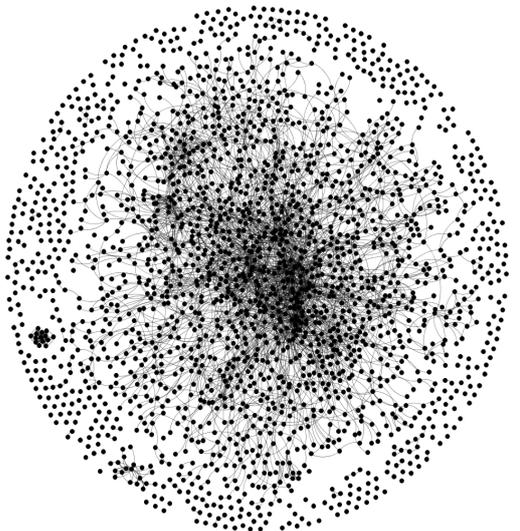
*Appendix A:* Dynamics of the data collection illustrated through the network



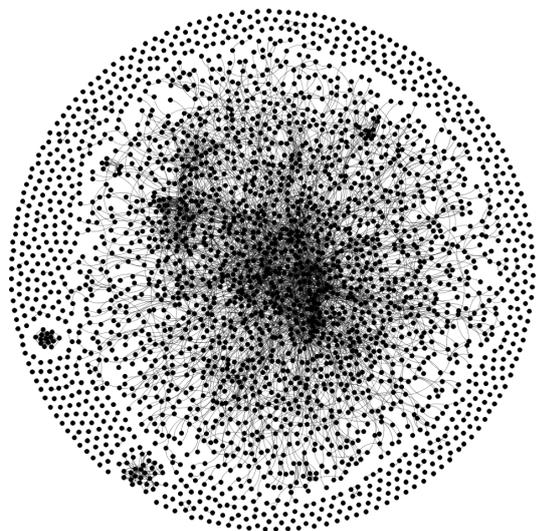
Network on January 2, 2011



Network on January 10, 2011



Network on January 20, 2011



Network on January 31, 2011

Figure 12: Illustration of the data collection dynamics.

Appendix B: Characterization of influencer-follower network.

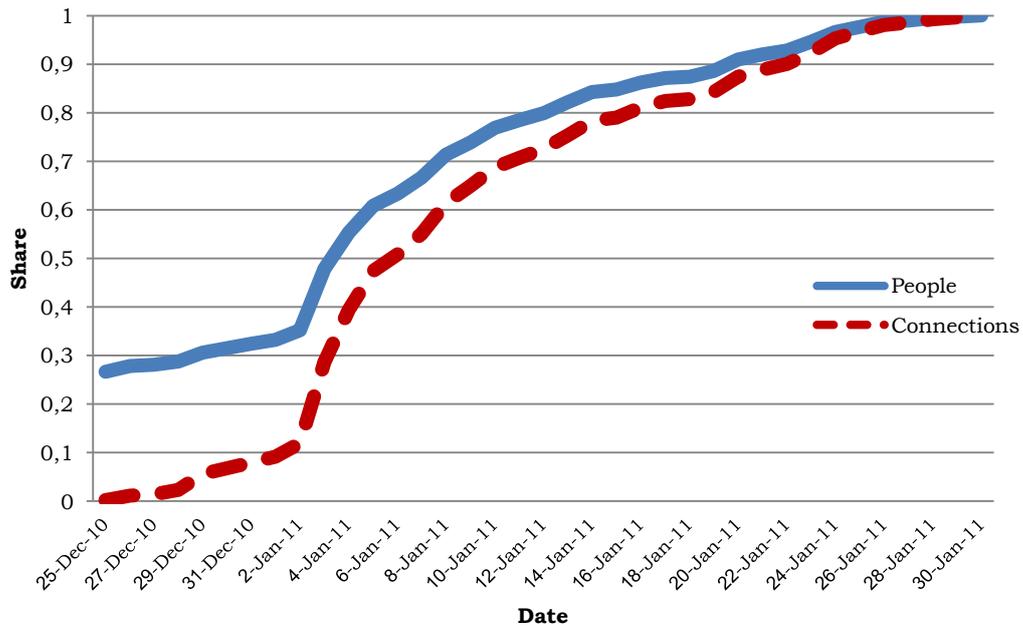


Figure 13: Dynamics of sub-network creation.

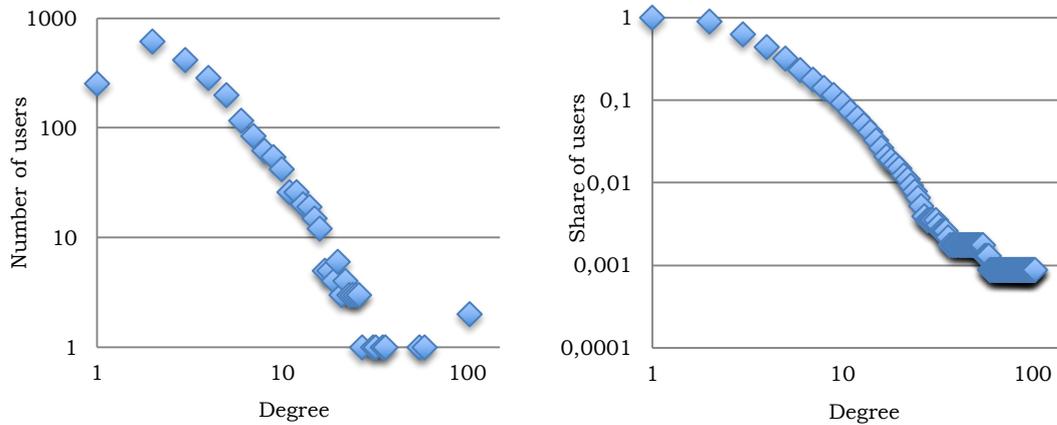
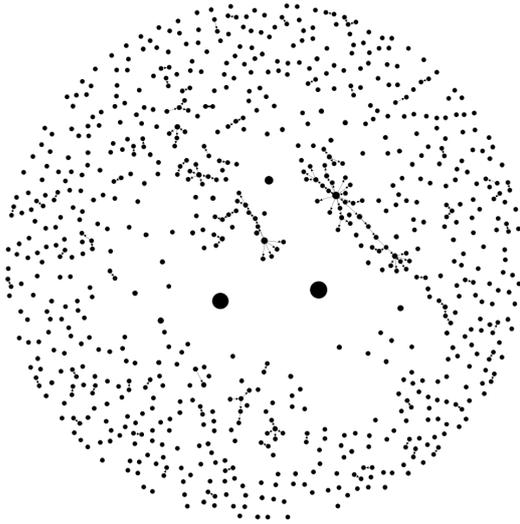
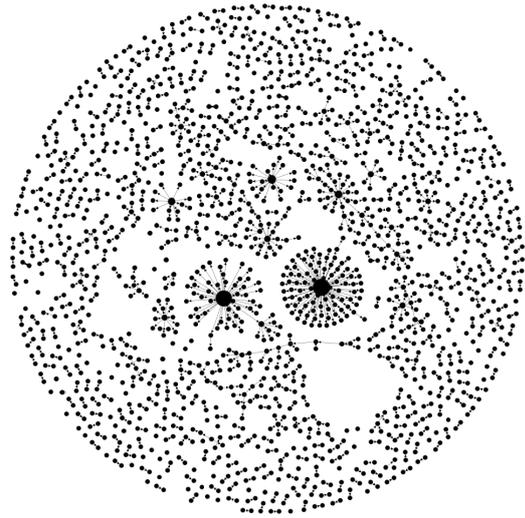


Figure 14: Degree and cumulative degree distributions in the sub-network.

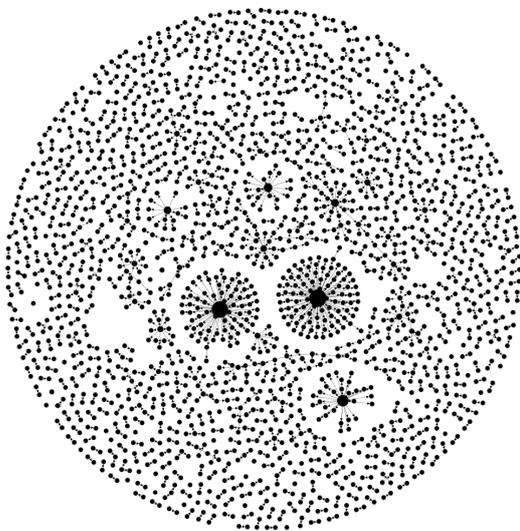
*Appendix C: Illustration of the influencer-follower network creation dynamics.*



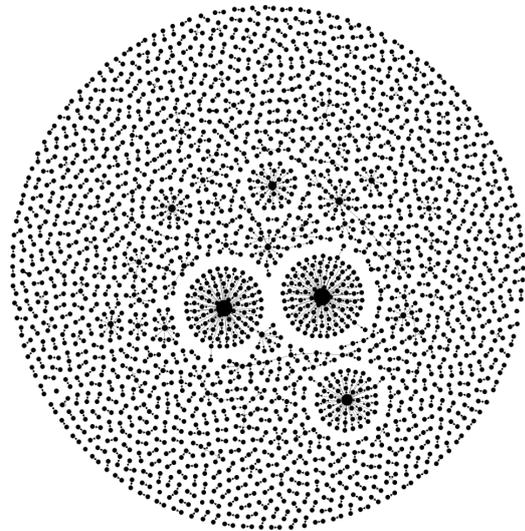
Network on January 2, 2011



Network on January 10, 2011



Network on January 20, 2011



Network on January 31, 2011

Figure 15: Illustration of the influencer-follower network creation.

*Appendix D: Comparison of influencer and follower profiles with the total population of Facebook users in Georgia.*

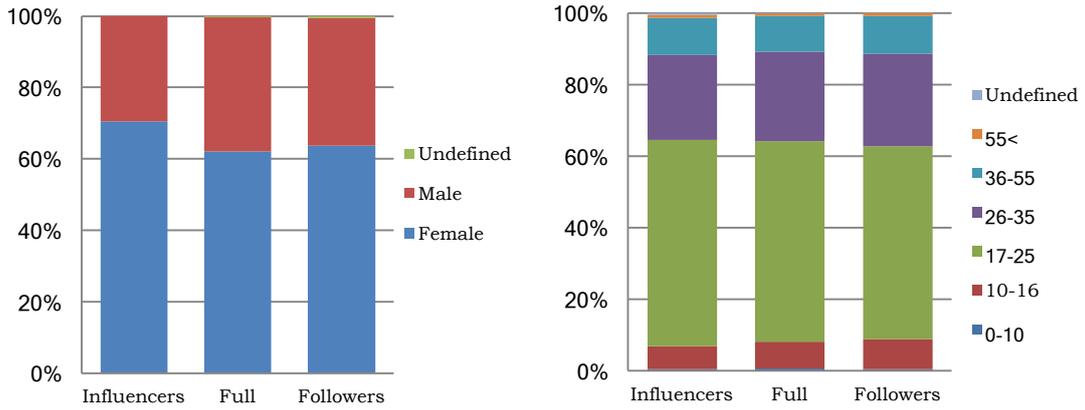


Figure 16: Comparison of full gender and age profiles.

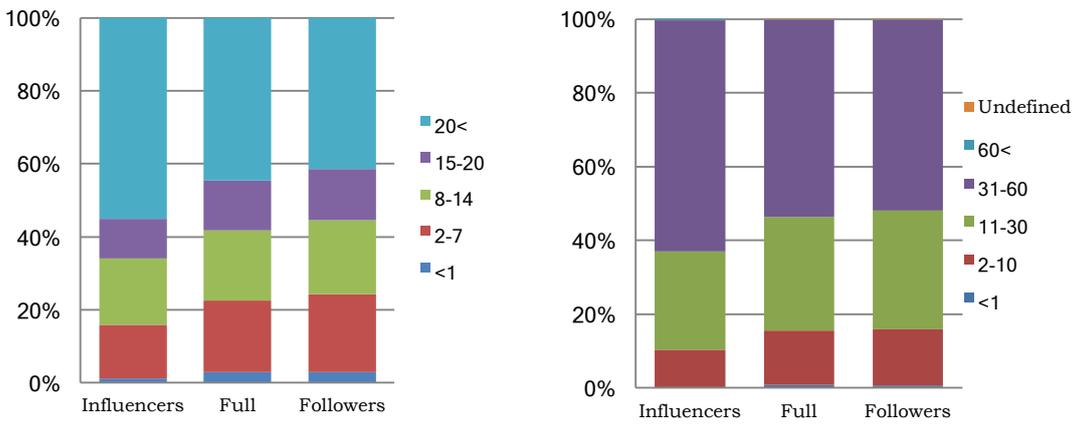


Figure 17: Comparison of Facebook usage profiles.

## Acknowledgements

The present study was conducted with the support of the Academic Swiss Caucasus Net (ASCN). ASCN is a program that promotes the social sciences and humanities in the South Caucasus (primarily Georgia and Armenia). Its activities foster the emergence of a new generation of talented scholars. Promising junior researchers receive support through research projects, capacity-building trainings and scholarships. The program emphasizes the advancement of individuals who, thanks to their ASCN experience, become better integrated in international academic networks. The ASCN program is coordinated and operated by the Interfaculty Institute for Central and Eastern Europe (IICEE) at the University of Fribourg (Switzerland). It is initiated and supported by Gebert RUF Stiftung.