Measuring the effect of social connections on political activity on Facebook

Olli Parviainen, Petro Poutanen, Salla-Maaria Laaksonen & Mikael Rekola
Communication Studies, Department of Social Research, University of Helsinki
firstname.lastname@helsinki.fi

Abstract

We examined nearly 100 000 Facebook users and their interactions in Finnish presidential elections' campaign pages to understand better the social behavior of political activity through interactions and structures of communication networks. Our attention is focused on the second round of campaigning. The data consists of over 27 000 interactions in the two candidates' Facebook pages during a period of 14 days. Combining statistical and social network analysis we were able to distinguish structural differences in the underlying friendship networks, such as in the interconnectedness of the page users. Our findings imply that activity in political pages is linked to the connectedness of the users generating the content. We also found that content generated by a page admin created most of the activity. The results and future implications of this study could be utilized to understand and possibly predict the patterns of political activity in social media.

1 Introduction

Today the role of social media in political communication is widely discussed both in public media and academia. The recent presidential elections in Finland raised questions concerning the role of social media in mobilizing candidates' support groups and influencing voting behavior. In recent years, also the academic research has recognized the importance of social media in explaining political activity (Castells, 2007; Gibson et al., 2008; Zhang et al., 2010). Traditionally politicians and their interest groups have been enthusiastic to adopt new forms of communication and marketing (Herkman, 2011).

According to the recent release of the Statistics Finland (2011), Internet penetration is now 89% of the population. The same report says that 47% of Finns have registered to a social networking service, and that 49% of Finns searched information on political

parties or candidates online. These numbers explain why politicians' and PR agents' involve so eagerly in the new public sphere provided by online arenas, such as Facebook, Twitter, Google+, blogs and other social media services.
In this paper we are presenting a framework for analyzing candidate supporters' political behavior in social media. The framework is based on Facebook activity measurements and friendship connections of the people involved. The research is also motivated by the interest of using large data sets to better understand social behavior. The behavioral data collected and stored from social media services can possibly be used for analyzing, understanding and predicting candidate supporters' political behavior and mobilization.

Our main interest in this study is to examine the connections between friendship network structures, group-formation and the levels of activity in online environment. In our study we apply social network analysis (Wasserman & Faust, 1994; Monge & Contractor, 2003) and traditional statistical methods. The case under scrutiny is the Finnish presidential elections in 2012. The two candidates in the final round Mr. Sauli Niinistö (National Coalition Party of Finland) and Mr. Pekka Haavisto (The Green League) both gathered over 100 000 users in their Facebook pages. The candidates were using social media excessively for campaigning, and the role of Facebook in this election was constantly an issue brought up by the traditional media.

We collected data on activities and network structures of those who had been active from the both candidates' support pages. The examination period was 14 days before the final Election Day. Our research questions are as follows:

- *What kinds of friendship structures are typical in large support groups (e.g. dyad, triads, bigger cliques, communities)?*
- *What kinds and levels of activity the support pages have (e.g. likes, comments, wall posts)?*
- *How the levels of activity are associated with the friendship structure of the support group?*
- *How the interaction patterns are associated with the friendship structure of the support group?*

1.1 Online social behavior and political activity

Researchers have fostered wide interest to the effects of social ties in political behavior. A classical study by Paul Lazarsfeld, Bernard Berelson, and Hazel Gaudet (1944) introduced the *two-step flow theory* based on the observation that informal and personal contacts were mentioned far more frequently as sources of influence on voting behavior than exposure to radio or newspaper. The recent advances and excessive personal use of social media platforms have raised new questions and new interest on the issue, partly affected by the possibilities of interaction data offered by the platforms.

The views of political participation online seem to float from utopian views to dystopian views. The hopes placed on the political deliberation, increased democracy and political participation offered by interactive online media (e.g. Corrado, 1996; Grossman, 1999; Johnson & Kaye, 2003; Shah, Cho, Eveland, & Kwak, 2005) have been partly undermined as well.

Social networking sites are most often used for the purposes of being entertained and staying connected to other (Boyd, 2008). What is typical to social networking services compared to other types of media and Internet services is that they are structured around existing social networks – around people who already know each other (Boyd, 2008). Concerning political engagement this implies that social networking services can be a powerful tool for amplifying and reinforcing the already existing political views (Baumgartner & Morris, 2009; Prior, 2005). For example, a study conducted with 18- to 24-year-old Americans illustrates that although social networking sites are recognized by youth as a possible source of news and that they act as an important news aggregator for many, the types of news gathered probably do little to inform them or add to democratic discourse, as young adults seem to prefer contacts and information that conforms their preexisting political views (Baumgartner & Morris, 2009). In addition, the same study shows that in spite of the promise social networking holds for increasing political interest and participation, active users are no more inclined to participate in politics than are users of other media (also Davis, 1999; Margolis & Resnick, 2000; Putnam, 2000).

However, some researchers point out that exposure to so called soft news (entertainment-based) can contribute to democratic discourse by imposing individuals not otherwise interested in politics to political information and news, and thus as a by-product increase political awareness or activity (see Baum, 2002; 2003). What is notable, regarding our case, presidential elections – being a person-centered election more than any other – typically create a lot of of soft news type media content and suits well with person-centered social media. Similarly, content in social networking services

can be regarded as soft news (Baumgartner & Morris, 2009), and it can be assumed that usage of social media in general could result in increased political activity and awareness. Some speculated during the presidential elections that the visibility of campaign material in social media provoked people to the extent that they ignored their friends' posts or even changed their voting decisions (e.g. Ilta-Sanomat, 2011).

Recently, advanced studies in political behavior research have started implementing social network analysis as a tool for examining the nature of political interactions online (e.g. Conover, Gonçalves, Flammini, & Menczer, 2012). Social network analysis (SNA) is a set of various methods for describing and studying human interaction, and it is used for studying repeating patterns in connections linking social actors (Wasserman & Faust, 1994).

Using social media as a source for social network analysis is not without some difficulties. For example, services lose popularity (such as MySpace, Friendster etc.) or their functionalities change. This can make comparison of research results difficult, even after a couple of years. The shelf life of social media research is very short indeed.

Network data enable inferences beyond the properties of individual actors. Inferences can be made about the processes involved in interpersonal contacts, such as chaining processes and the characteristics of resulting networks (Weatherford, 1982). Using the data on friendship structures and interactions provided by Facebook allows us to reflect, for example, how existing social networks influence on interactions.

1.2 Case description and Facebook as a platform

In the Finnish presidential elections 2012 the candidate of the Green League (Green party) Pekka Haavisto advanced to the second round with the front runner Sauli Niinistö of the National Coalition party (right) with 18.8 % of the votes even though in the parliamentary elections in 2011 the Green League received 7.3 % of the votes. The term "Haavisto-phenomenon" was coined, and in the mass media the large share of votes was partly credited to social media (Facebook particularly). The second round saw a rapid increase in the volume and intensity of campaigning on Facebook where the Like-counts of both pages were a subject of interest in the main media.

Facebook pages are, according to Facebook *for businesses, organizations and brands to share their stories and connect with people*[1]. The main functionality of the page is the feed or the "wall", where the sharing of information, opinions, promotion, etc. takes place. The page host or admin can publish messages, links, photos etc. (a "wall post")

[1] ("https://www.facebook.com/help/?faq=174987089221178)

on the wall as can individual users. Important factor in the proliferation of the post is the identity of the author. If the page admin publishes a message on the page feed, it reaches far greater mass of the page likers than those posted by an individual user. If an individual user publishes a message on the page feed it proliferates only to the user's friends who also have liked the page. Users can also comment the wall posts, "like" the wall post or the comments it has. Every time a user uses one of these actions her individual Facebook-ID is recorded in the freely retrievable feed of the page. These actions, "wall post", "wall post likes", "comments" and "comment likes", are the only means of interacting within the page and thus are the focus of this research.

2 Page activity

The unit of observation for the analysis was an individual wall post. We collected the time, date, current like-count of the page, type of interaction, the sum of all activities and the number of active users within the post. We defined “activity” as any creation of content (wall posts and comments) or reacting to the content (the “liking” of wall posts and comments). We also distinguished posts made by the admin of the page and posts of the individual users.

We combined the page activity data with the friendship network data. The network metrics were partitioned according to the users' activity in the specific post. Descriptions of all the measured items are on Table 1.

Table 1 measured items

Likes	The number of likes the page has in the time of the post.
Number of wall post likes	The number of likes the wall post has received
Number of comments	The number of comments posted to the wall post
Number of comment likes	The number of likes the comments within the wall post have received
Overall activity	Sum of all activity (wall post likes, comments and comment likes). Measures the response for the post.
Active users	Absolute number of different users activated in the post.
Activity level	The share of the active users of page's all likers in the wall post . Measures the wall post's ability to engage the page likers (audience)
Number of wall post likers	The number of different users liking the wall post
Number of commenters	The number of different users commenting on the wall post
Number of comment likers	The number of different users liking comments
Number of components	Absolute number of friendship components within the post
Friendship network edges	The number of friendship connections within the post
Friendship average component size	Mean of all friendship component sizes within the post
Friend average degree	Mean of number of friends the active users of the post have each other
Friend overall degree	Mean of number of friends the active users of the post have with all the active users in the two week time frame
Friend clustering coefficient	The clustering coefficient of the friendships
Network friends percentage	The percentage of the active users of the post who have at least one friend among the other active users
Poster friend count	Number of friends the author of the post has within all the active users of the page

2.1 Data collection

The data was extracted from Facebook platform via it's FQL[2] interface. The data was collected and analyzed using custom software (C++, Perl, Graph.pm), which collected the needed wallposts, comments, wallpost likes and comment likes from Facebook and stored the data in the database for further processing. Facebook platform offers a way to query if two known users are friends. We queried every possible friendship connection in the two pages active users could have. The number of user pairs that needed to be tested was very large (790130628 for Niinistö and 1802731035 for Haavisto).

The friendship network was collected several months after the election and predictably, some of the user accounts that participated in the election discussion, were no longer active at that time and such a user would appear to have no friends. Furthermore, it's probable that some of the friendship connections have changed since the elections, and the current friendship network does not fully correspond to the network that existed during the elections.

We used metrics such as component size, degree centrality, betweenness and clustering coefficient to describe the individual actors in the network, whereas average degree centrality, average path length, network diameter, and average clustering coefficient were used to describe the overall structure of the network (Wasserman & Faust, 1994).

[2] FQL is a query language based on SQL, which returns the result data as JSON.

3 Results

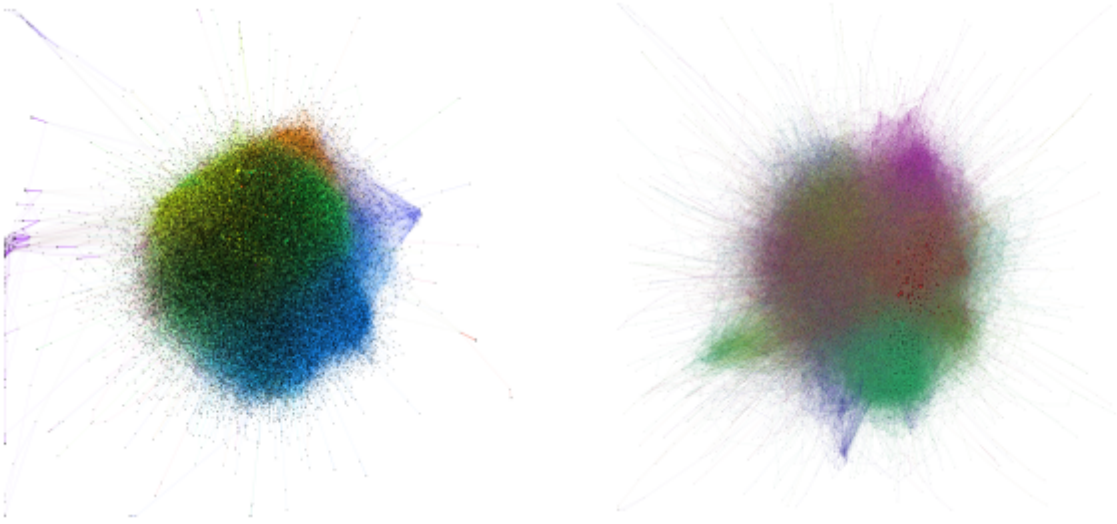
3.1 Descriptives

The friendship networks of both pages consist of large main component and several smaller components. As the main component in both cases contains over 90 % (Niinistö 90 %, Haavisto 95 %) of all active users, our main focus on the network metric comparisons is on the main component. The users that are a part of the smaller components are incorporated into the analysis of interactions. When comparing the share of common users between different pages (Table 2) and component types, we see that the users in the Haavisto main component make up almost one in ten of active users in the Niinistö (Fig. 1) main component. This represents the overlap between the two user communities. It is also interesting that between the components outside of the main components there is no overlap. The Niinistö non-main component(s) seems to be isolated as it lacks any overlapping with the Haavisto page. Overall there were combined 96947 individual users in the two pages within the two week period.

Table 2: Component overlap

	Haavisto main	Haavisto non main	Niinisto main	Niinisto non main
Share of Haavisto main	0%	0%	6%	0%
Share of Haavisto non main	0%	0%	7%	0%
Share of Niinisto main	9%	1%	0%	0%
Share of Niinisto non main	0%	0%	0%	0%

Fig 1.: Friendship network main component. Haavisto on the left, Niinistö on the right



Diameter in the Niinistö page is significantly larger than in the Haavisto page (18 steps vs. 13 steps) and the radius of the network is 9 steps for the Niinistö page (Table 3). However, the average path length in the Niinistö page is 4,70 steps and 6,80 steps in Haavisto page. Explanation for this can be larger "cliquishness" of the Niinistö page (Cowan & Jonard, 2004, 1560). This suggests that the Niinistö page shows more 'small world phenomenon' (Watts & Strogatz, 1998) than the Haavisto page.

Table 3: Network topology

Descriptive statistics

	Nodes	Edges	Diameter	Radius	Average path length
Niinistö	35372	189673	18	9	4.699
Haavisto	57696	461744	13	7	6.804

Most notable difference can be found in degree centrality: members of the Haavisto page have in average five friends more than the members of the Niinistö page have (Table 4). The Niinistö page degree centrality also varies more as the most "popular" user in the Niinistö page has over 200 more friends than the most popular user in the Haavisto page (1043 friends over 825 friends). According to Two-Sample Kolmogorov-Smirnov Test, all of the centrality metrics from different pages (degree, closeness, betweenness, eigenvector centrality and clustering coefficient) differ statistically significantly (p<0,000) from each other.

TABLE 4: Network descriptive (*All figures are significant (p<0,001))

.

Descriptives

						95% Confidence Interval for Mean			
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Degree	Niinistö	35372	10,72	27,915	,148	10,43	11,02	1	1043
	Haavisto	57696	16,01	26,116	,109	15,79	16,22	1	825
		93068	14,00	26,936	,088	13,83	14,17	1	1043
Closeness Centrality	Niinistö	35372	4,6994	,69134	,00368	4,6922	4,7066	2,83	12,59
	Haavisto	57696	4,3895	,53965	,00225	4,3851	4,3939	2,92	8,60
		93068	4,5073	,62035	,00203	4,5033	4,5113	2,83	12,59
Betweenness Centrality	Niinistö	35372	65426,4611	411644,43946	2188,73128	61136,4812	69716,4410	0,00	30983647,17
	Haavisto	57696	97777,7065	411343,85471	1712,50623	94421,1866	101134,2264	0,00	27120682,46
		93068	85482,0920	411755,46863	1349,70602	82836,6832	88127,5008	0,00	30983647,17
Clustering Coefficient	Niinistö	35372	,1374	,20445	,00109	,1353	,1395	0,00	1,00
	Haavisto	57696	,1129	,15641	,00065	,1117	,1142	0,00	1,00
		93068	,1222	,17661	,00058	,1211	,1234	0,00	1,00
Eigenvector Centrality	Niinistö	35372	,0080	,03857	,00021	,0076	,0084	,00	1,00
	Haavisto	57696	,0088	,02994	,00012	,0086	,0091	,00	1,00
		93068	,0085	,03348	,00011	,0083	,0087	,00	1,00

Betweenness centrality of the Haavisto page is much larger than in the Niinistö page. The higher clustering coefficient in the Niinistö page is also visible when comparing the

network visualizations; the overall structure in the Haavisto page is more uniform and fewer distinct groups emerge. The network visualizations were made with Gephi and Force directed algorithm was used to calculate the layout of the network graph.

Differences in activity are visible. In the case of Haavisto a greater part of the overall activity count and the levels of activity were initiated by the page users, whereas in the Niinistö case the activity was more admin-originated (Table 5). Appendix 1 contains detailed tables.

Table 5: Activity comparison

Descriptives								
	Niinistö admin		Niinistö users***		Haavisto admin		Haavisto users***	
Activity	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Activity level (%)	1,725	2,045	0,011	0,012	1,468	1,594	0,017	0,018
Overall activity count	1512,702	1519,108	12,984	17,691	1531,52	1454,071	20,742	24,898
Number of active users	1357,024	1439,509	9,632	10,018	1367,205	1370,87	15,924	15,922
Number of wall post likes	1297,643	1338,515	7,292	8,706	1303,047	1324,11	11,773	14,222
Number of wall post likers	1297,643	1338,515	7,292	8,706	1303,047	1324,11	11,773	14,222
Number of comments	98,071	202,832	1,841	4,201	68,591	93,631	3,177	5,481
Number of comment likes ^	116,988	112,111	2,851	9,553	159,882	195,69	4,792	13,704
Number of comment likers ***	40,25	31,745	1,469	3,759	76,236	77,271	3,057	6,348
Number of commenters in the post	27,952	25,269	0,923	2,086	26,291	25,971	1,444	2,614
Structure								
Average friendship component size of active users of the post	1,273	0,232	1,404	0,697	1,332	0,327	1,104	0,238
Median friendship component size of active users of the post	1	0	1,129	0,611	1	0	1,01	0,155
Average number of friends the users of the post have within the post	0,477	0,311	0,485	0,527	0,521	0,435	0,161	0,279
Average number of friends the users of the post have within the all active users of the page	10,04	1,521	30,994	34,694	15,413	2,137	21,803	17,344
Percentages of active users within the post having at least one friend within the post	25,445	12,289	30,09	29,239	27,413	15,063	11,339	16,556
Number of components within the post	951,798	768,489	7,24	7,743	895,969	653,452	14,486	13,777
Number of friendship connections the users have within the post	511,31	911,195	3,019	5,609	616,205	1047,849	1,673	5,429
Number of friends the poster has within ALL the active users **			22,255	71,066			23,962	44,452

^ Significant at level p<0,10

*** Significant at level p<0,001

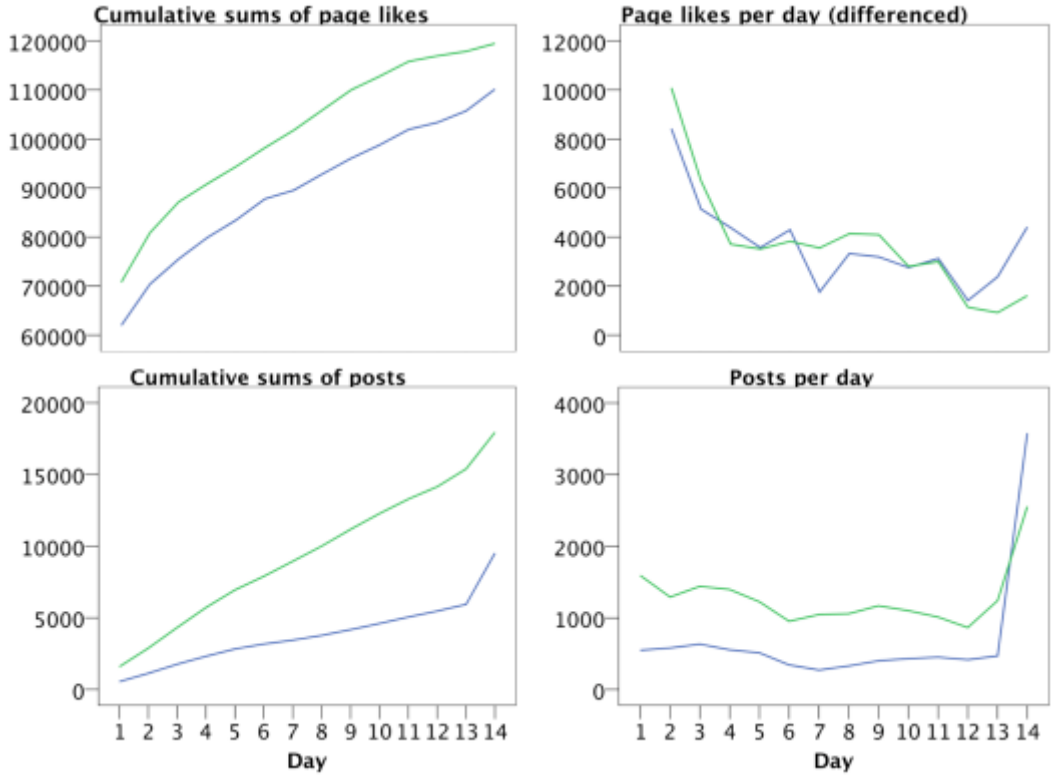
3.2 Analyzing activity over time

In order to understand how the two support pages distinguished with relation to time, time series were sampled from the data. The time series comprise 14 days, each day containing 8 time slices (3 hours sampling sequence). Time sampling was possible, since each post contained a time code. Means, sums, or other statistics were used to aggregate values into time slices. Additionally, daily sums and means of values were used when appropriate.

Page likes and posts

The Figure 2 shows the daily sums of page likes and posts. It can be seen that in the both groups the typical amount of posts remained quite stable during the whole period (daily median in Haavisto page is 1197 and Niinistö page 460), apart from the last day when there were more posts in both pages.

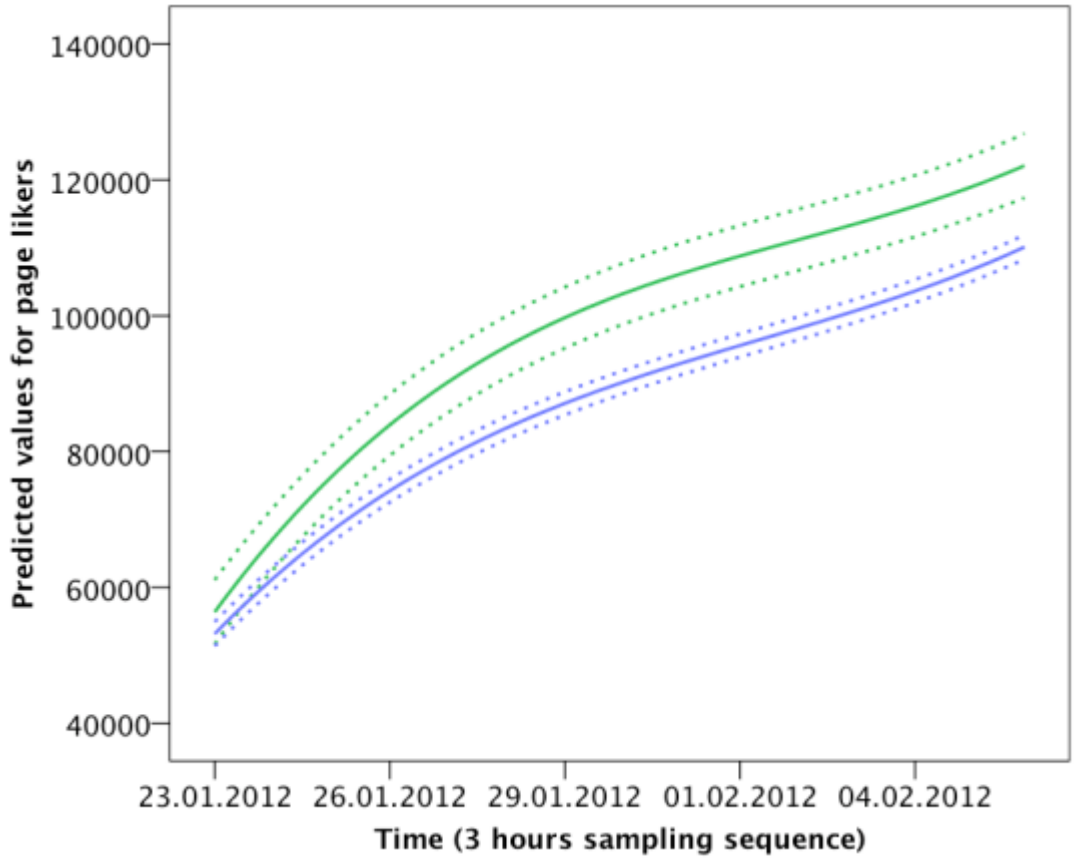
Fig.2: Page likes and posts per day in Haavisto (green line) and Niinistö (blue line) pages.



The cumulative growth of new likers is quite stable for the both pages. The slope of the linear fitted line was for Haavisto 519,625 (t=35,247, p<0,000) and for Niinistö 457,541

(t=52,525, p<0,000), corresponding the amount of new likers per 3 hours. Because of the slightly bended natures of the cumulative growth curves, other than linear types of models were also tested. The best fit was found with cubic model (R squares: Haavisto=0,984, Niinistö=0,997) defined by the equation Y = b0 + (b1 * t) + (b2 * t**2) + (b3 * t**3). F test showed better fit for cubic model with Niinistö page (F=11788,118, p<0,000) than with Haavisto page (F=2239,773, p<0,000). This is most likely due to harder sprint of Niinistö page's new likers towards the final election day (see the differenced values above). In the figure 3 below are depicted the predicted values of page likers for Haavisto and Niinistö pages with 95 % confidence intervals according to the cubic model.

Fig. 3: Predicted values of cumulative page likes for Haavisto (green line) and Niinistö (blue line) pages by cubic model and with 95 % confidence intervals.

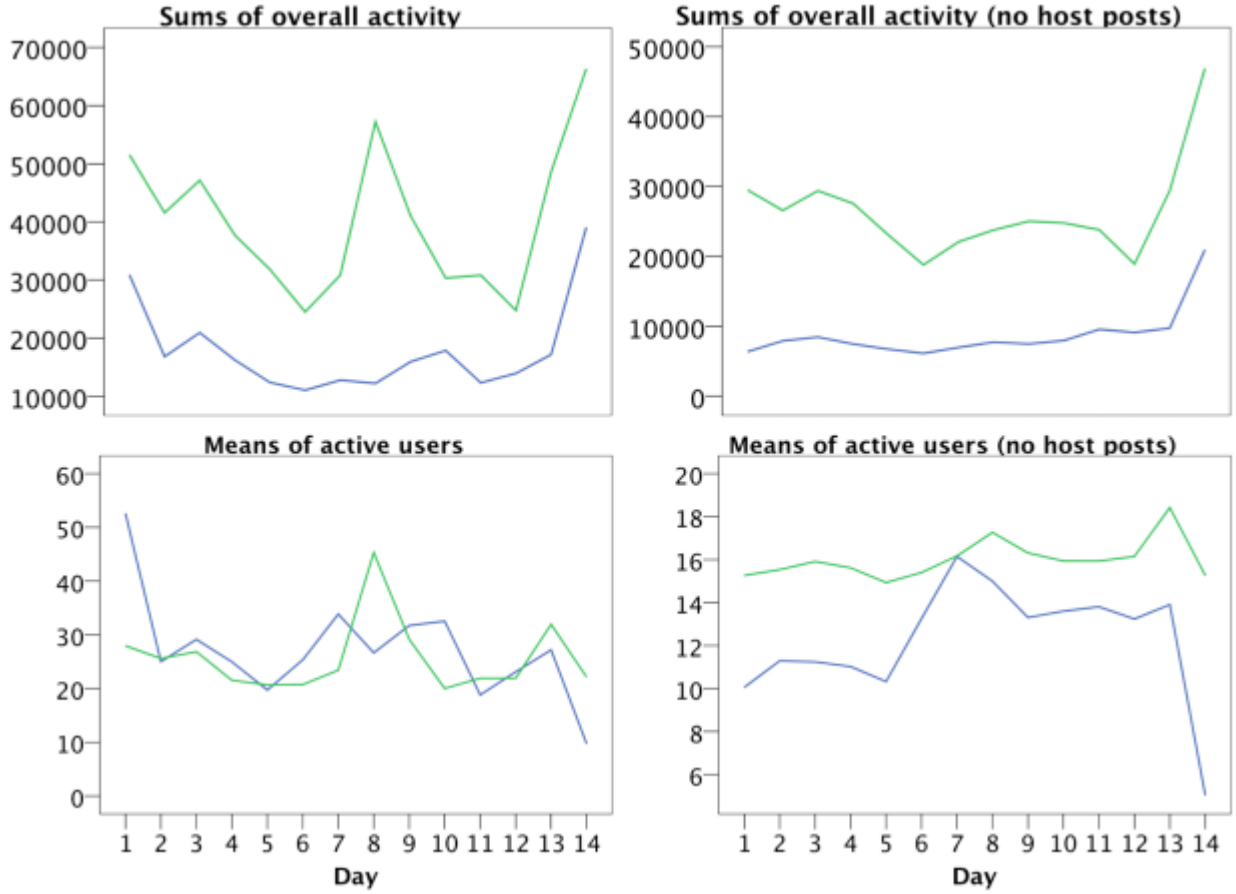


The cubic model illustrates how the growth is steeper on the beginning of the second round and then calms down moderately after five days or so, before rising again slightly towards the final Election Day at Feb 5. The same trend fits to the both pages.

Activity levels

Figure 4 shows the levels of overall activity and active users for both pages (see appendix 2 for the illustrations of the each activity component separately). The plots on the right hand side show the activity levels without the contributions accounted for the page admin's posts. It is easy to see that admin posting accounts a lot of the overall activity – in Haavisto group about 34% of all activities, and in Niinistö group almost 51%. In addition, the overall activity of Haavisto page is in absolute numbers at much higher level than Niinistö page's.

Fig 4. Overall activities and active users in Haavisto (green line) and Niinistö (blue line) pages. Posts made by the page admin excluded from the plots on the right side.



Interactions

The network metrics used in this study open up an interesting window for comparing the structural relations of the pages' members and different types of activities. In order to describe the relative levels of activities on the pages we calculated two measures for each posts: the ratio of comment likes and comment likers and the ratio of comments

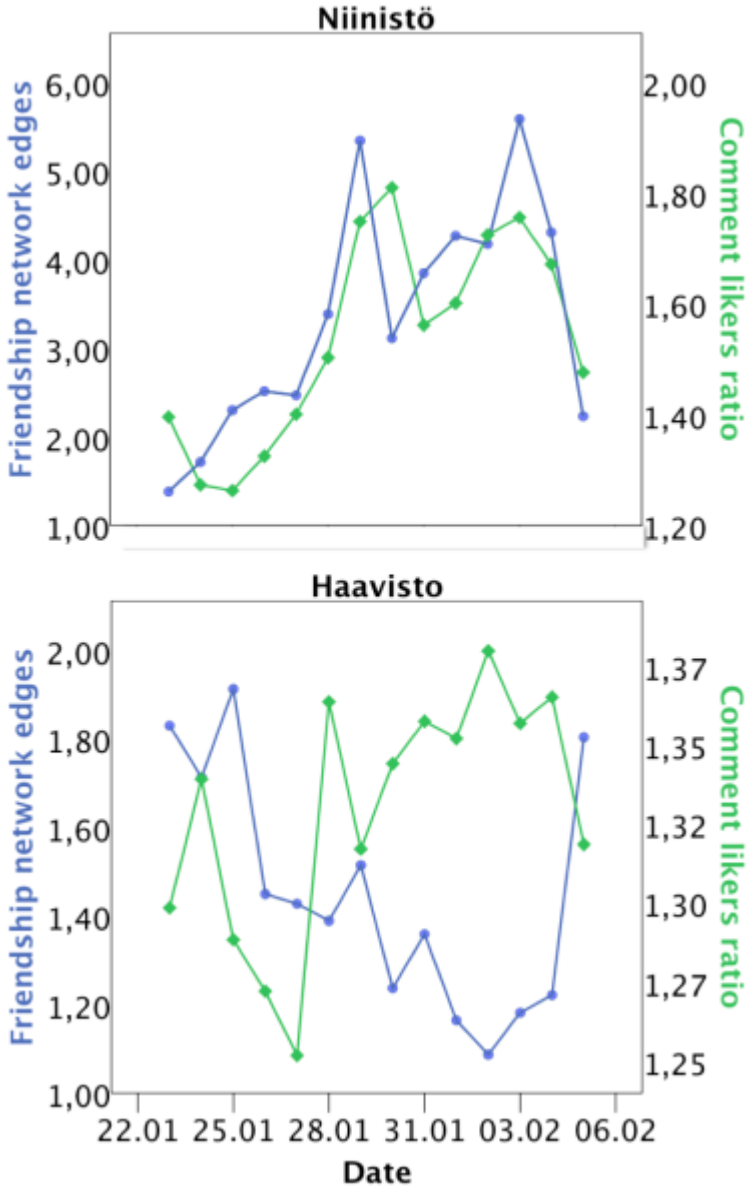
and commenters. These relative measures of activities describe the levels of participation in each post. If the value is close to 1 (minimum), it means that there are as many contributors as contributions, whereas the bigger values describe the situation in which there are fewer contributors than contributions. For example, if there are 15 comment likes within the conversation thread of one post and 3 likers, the ratio is 5 indicating that few likers have liked many of the comments. Similarly, if there are 5 comment likes and only one liker, the ratio is 5. In contrast, if there are as many likers as likes, the ratio approaches minimum value. Thus, the ratio may distinguish between balanced (many contributors) and unbalanced (few active contributors) activities.

The support pages differ from each other in terms of the both ratios. Let us, for example, consider the mean values of 3 hours time slices, excluding the posts made by the page admin. The mean value of comments–commenters ratio is 2,28 (std. = 0,35) in Haavisto page and 1,84 (std. = 0,4) in Niinistö page, indicating that the commenting activity was accounted, in average, by fewer people in Haavisto page than in Niinistö page. In contrast, comment likes–comment likers ratio is 1,33 (std. = 0,01) in Haavisto page and 1,54 (std. = 0,3) in Niinistö page. This means that in Niinistö page more same people like the comments within a post than in Haavisto page (see appendix 3 for the illustrations of the time series).

We used cross correlations to explore the associations of ratio measures with the amount of friendship network edges within a post, i.e. the amount of friends the contributors has within the post. We found positive correlation between comment likes–comment likers ratio and the amount of friendship network edges in Niinistö group (0,429 in lag 0 and 0,418 in lag -1). It is possible that the amount of friendship connections within the post predicts fewer comment likes because the commenting happens between friends and does not encourage participating. However, this possible link will have to be further investigated.

The following plots (5 and 6) show the relationships between the likes–likers ratio and the amount network edges for both pages. The trend lines in Haavisto page seem to follow each other more tightly than in Niinistö page. It seems to be also that in Haavisto page the activity is much more maintained by non-friends, whereas in Niinistö page friendship connections are much frequent in each post. The higher level of friendship connections may be associated with the correlating values, as suggested above.

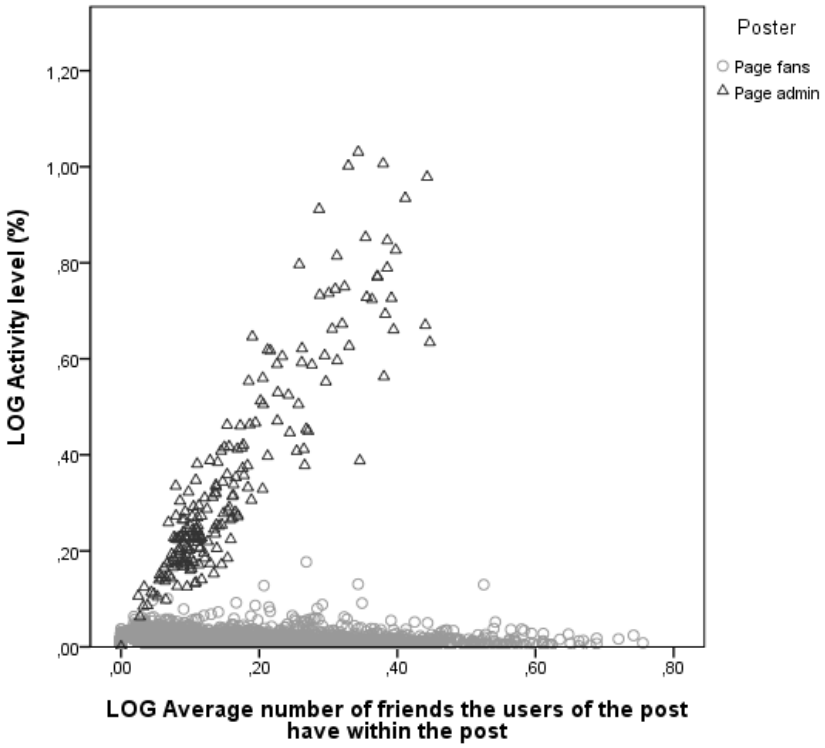
Figures 5 and 6: Niinistö's and Haavisto's pages in dual mean plots of friendship network edges (left vertical axis) and comment likes–comment likers ratio (right vertical axis). Admin posting is excluded.



3.3 Explaining activity

Regarding the posts made by users all measures of *activity* are greater for the Haavisto page (all p>0,001). However, the opposite is true for Niinistö page where most of the variables indicating friendship *structure* are higher (see Appendix 4 Table 4B). The mean of activity level for posts made by the users is greater on Haavisto page (p<0,001), but for admin-initiated posts the Activity level is greater for Niinistö page (p=ns.). ANOVA indicated that for admin-initiated posts comment liking are more prominent activity for the Haavisto page than for the Niinistö page, but otherwise no significant differences were found. Thus, it seems that for the admin-initiated posts the activity levels and friendship structure variables all in all are rather similar, but within the user-initiated posts Haavisto page creates more activity between users that are not connected to each other. This observation is also visible in the time series analysis above.

Fig. 7: Activity level plotted against number of user friends



All in all, the effects of friendship structure to the activity levels are more clearly shown in the activity levels of the posts made by the page admin. Figure 7 shows how considering the posts made by page likers (fans) the average number of friends does not correlate with the activity level (R^2 Linear = 0.014). However, for posts made by the page admin the more friends users have on average within other users in the same post, the higher is the level of overall activity (R^2 Linear 0.828). Similarly, as shown in Figure 8, when looking at the posts made by the page admin the total activity level correlates with the number of friendship connections the users have within the post, and with the average number of friends the users of the post have within the post (see Figure 9).

Fig 8: Posts by admin activity level – average number of friendship connections within the post plot

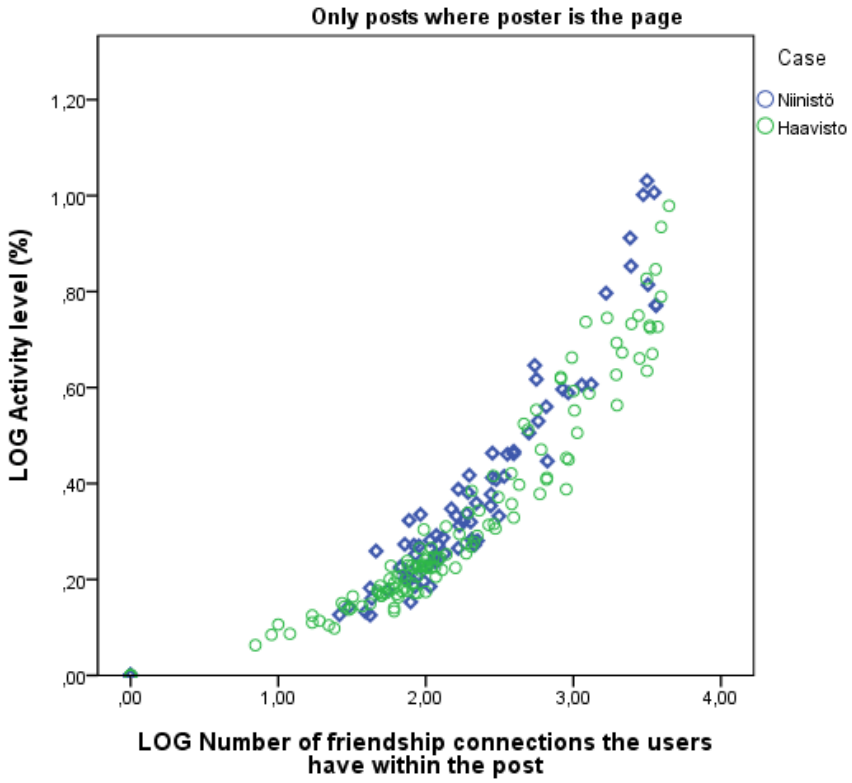
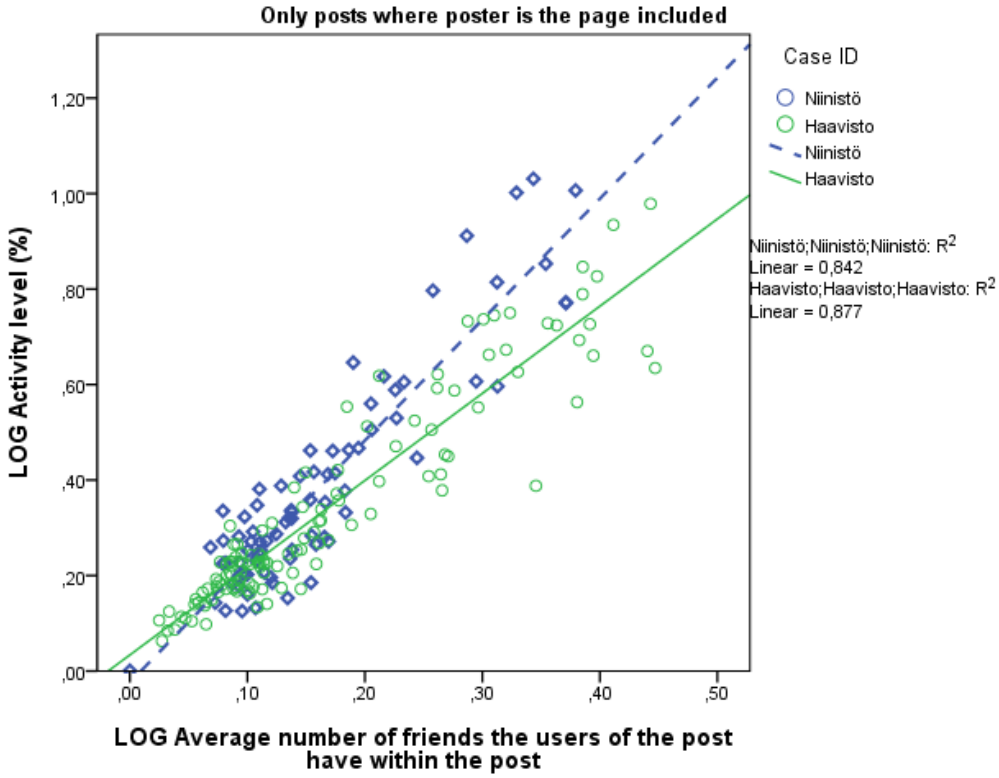


Fig 9: Posts by admin activity level – friendship average degree within the post plot



3.4 Regressions

In the regression analysis a stepwise method was used to enter the variables. The dependent variable used was Overall Activity count (see Table 1) (after logarithmic transformation) instead of Overall Activity level, which was highly skewed despite the log transform. Examination of the tolerance and VIF scores testing multicollinearity revealed that there were no too high correlations among the independent variables.

The regression models accounted for about 81 to 99 percent of the variance, which is a very high accuracy for a model and probably partly relates to the big amount of observations in the data.

Regression model for the admin-initiated posts

In the first regression, number of components within the post ($\beta = 1.202$, $p<0.001$ for Niinistö; $\beta = 1.281$, $p<0.001$ for Haavisto) surfaced as a significant positive predictor of overall activity count of a post (Table 6 and 7). The addition of the variable average

number of friend the users of the post have within the post (β = 0.511, p<0.001 for Niinistö, β = 0.8, p<0.001 for Haavisto) slightly improved the explanatory power of the overall model (R2 = .992 for Niinistö, R2 =.991).

Table 6A: Explaining LOG Overall Activity Count for user-initiated posts in case Niinistö

Variable	Model 1	Model 2
Constant	-0.437***	-0.149*
LOG Number of components within the post	1.202***	1.075***
LOG Average number of friends the users of the post have within the post		0.511***
R^2	0.988	0.992

Table 7B: Explaining LOG Overall Activity Count for user-initiated posts in case Haavisto

Variable	Model 1	Model 2
Constant	-0.638***	0.039
LOG Number of components within the post	1.281***	0.999***
LOG Average number of friends the users of the post have within the post		0.800***
R^2	0.976	0.991

Regression model for user-initiated posts

The overall model explained about 91,5 percent of the variance in the dependent measure of overall activity count. The first regression yielded the number of components within the post (β = 1.080, p<0.001 for Niinistö, β = 1.068, p<0.001 for Haavisto) as a significant positive predictor for overall activity (Table 8 and 9). The addition of the attitudinal variables related to the friendship structure within the post further improved the regression model; number of friendship connections the users have within the post (β = 0.331, p<0.001 for Niinistö, β = 0.159, p<0.001 for Haavisto) and the number of friends the poster has within ALL the active users (β = 0.010, p<0.001 for Niinistö, β = 0.290, p<0.001 for Haavisto) positively predicted overall activity.

However, average number of friends the users of the post have within the *all* active users of the page (not just within the active users within the post) interestingly was a negative predictor of overall activity for Niinistö (β = -0.022, p<0.001 in the final model), but a slightly positive predictor of overall activity for Haavisto (β = 0.025, p<0.001 in the final model).

Table 8 Explaining LOG Overall Activity Count for admin-initiated posts in case Niinistö

Variable	Model 1	Model 2	Model 3	Model 4
Constant	0.110***	0.115***	0.141***	0.135***
LOG Number of components within the post	1.080***	0.895***	0.894***	0.898***
LOG Number of friendship connections the users have within the post		0.331***	0.340***	0.335***
LOG Average number of friends the users of the post have within the all active users of the page			-0.022***	-0.024***
LOG Number of friends the poster has within ALL the active users				0.010***
R^2	0,817	0,914	0.915	0,915

Table 9 Explaining LOG Overall Activity Count for admin-initiated posts in case Haavisto

Variable	Model 1	Model 2	Model 3	Model 4
Constant	0.037***	0.071***	0.038***	0.036***
LOG Number of components within the post	1.068***	0.998***	0.999***	1.003***
LOG Number of friendship connections the users have within the post		0.159***	0.154***	0.144***
LOG Average number of friends the users of the post have within the all active users of the page			0.026***	0.025***
LOG Number of friends the poster has within ALL the active users				0.290***
R2	0.914	0.931	0.932	0,932

Thus, for both admin-initiated posts and user-initiated post the best predictor for overall activity was the Number of components within the post, i.e. how many different groups or cliques of people were activated. In addition, for admin posts the *Average number of friends the users of the post have within the post* was a positive predictor, whereas for user-initiated posts the additional predictors were *Number of friendship connections the users have within the post*, the *Average number of friends the users of the post have within the ALL active users of the page* and the *Number of friends the poster has within ALL the active users*.

This implies that for the user-initiated posts more activity is gained when a poster or the commenters were well connected to the other users of the page. For the admin-initiated posts the inter-connectedness within the post, i.e. within the comments elicits more activity. Naturally this effect partly originates to the technical allowances of Facebook as the user-initiated posts by users who have more connections are also more likely shown in the newsfeed for those connections, as well as the admin-initiated posts are shown to the commenters' friends.

4 Discussion and conclusions

The underlying friendship networks of the two campaign pages differed in many aspects. Most notable difference was the greater degree centrality of the Haavisto page and the larger clustering coefficient of the Niinistö page. Haavisto's campaing was regarded more "Facebook-campaign" in the main stream media. The results of friendship network support this assumption as the users of the Haavisto page were in overall more interconnected and suggested a more "social media savvy" users compared to the users of the Niinistö page. Interesting future research would also be the change of the friendship network structure *during* the campaign as our analysis was conducted post hoc. The scalability of the friendship network analysis is limited as the number of possible connections grows exponentially.

More user-initiated / user-caused activity on Haavisto page, also between users that are not connected to (friends with) each other.

Analyzing activity over time revealed several interesting points. First, the evolution of the both groups in terms of new page likers and the amount of posts was quite steady, although in Haavisto's group, the cumulative growth of new page likers was slightly faster, as was the daily amount of posts much bigger. The cubic function model was fitted with the cumulative growth of new page likers, indicating the bended nature of the curve and faster growth rate in the beginning and towards the end. Secondly, admin posting - posts made by the page admin - were found to contribute nearly half of the all activities in the page in Niinistö's page. When comparing the levels of activities without admin posting, both pages show more similar activity levels. Thirdly, associations between the amount of friendship connections within the post and different "activity ratios" were explored. Speculations about how the amount of friends being active within the posts influence on the other members' participation were presented.

Combining the time series analysis with a timeline of external events (news stories etc.) could provide more insight on the dynamics of the online interaction. Future research on this matter is suggested.

The regression models indicated that different measures for the inter-connectedness of the active users positively predict the overall activity. Most definitive indicator was the number of components, thus the more different, distinct groups of people a post activates the more active the specific post will be. Regarding the user-initiated posts the numbers indicating the popularity of the users and the poster herself within all the active users of the page also predicted the activity increase. This implies that activating the most connected users of a page will yield the page more activity. Again, this is partly due to the technical ways of presenting the content by Facebook, and is thus applicable to all Facebook pages. For admin-initiated posts the effect of total connectedness is not that significant.

References

Baum, M. A. (2002). Sex, lies, and war: How soft news brings foreign policy to the inattentive public. American Political Science Review, 96, 91-109.

Baum, M. A. (2003). Soft news and political knowledge: Evidence of absence or absence of evidence? Political Communication, 20, 173-190.

Boyd, D. (2008). Can social networking sites enable political action. In A. Fine, M. Sifry, A. Raseij, & J. Levi (Eds.), Rebooting democracy. New York: Personal Democracy.

Castells, M. (2007). Communication, power and counter-power in the network society. International journal of communication, 1(1), 238–266.

Conover, M.D.; Gonçalves, B., Flammini, A., & Menczer, F. 2012. Partisan Asymmetries in Online Political Activity. arXiv:1205.1010. Available at: http://arxiv.org/abs/1205.1010 [Accessed June 20, 2012].

Corrado, A. (1996). Elections in cyberspace: Prospects and problems. In A. Corrado, & C. M. Firestone (Eds.), Elections in cyberspace: Toward a new era in American politics. Washington, DC: Aspen Institute.

Davis, R. (2005). Politics online: Blogs, chatrooms, and discussion groups in American democracy. New York: Routledge.

Gibson, R. K., Lusoli, W., & Ward, S. (2008). Nationalizing and normalizing the local? A comparative analysis of online candidate campaigning in Australia and Britain. Journal of Information Technology & Politics, 4(4), 15–30.

Grossman, L. K. (1999). The electronic republic. In R. E. Heibert (Ed.), Impact of mass media: Current issues. New York: Longman.

Herkman, J. (2011). Politiikka ja mediajulkisuus. Tampere: Vastapaino.

Presidentinvaalien herutus rikkoo Facebook-kaveruuksia. (n.d.).Ilta-Sanomat. Retrieved August 11, 2012, from http://www.iltasanomat.fi/presidentinvaalit/art-1288444875102.html

Johnson, T. J., & Kaye, B. K. (2003). A boost or bust for democracy? How the web influenced political attitudes and behaviors in the 1996 and 2000 presidential elections. Harvard International Journal of Press/Politics, 8, 9-34.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). The People's Choice. How the Voter Makes up His Mind in a Presidential Campaign. New York: Columbia University Press.

Margolis,M., & Resnick, D. (2000). Politics as usual: The cyberspace ‘‘revolution’’. Thousand Oaks, CA: SAGE.

Monge, P. & Contractor, N. (2003). Theories of Communication Networks. New York: Oxford University Press.

Official Statistics Finland (2011). Tieto- ja viestintätekniikan käyttö 2011. Retrieved from http://www.stat.fi/til/sutivi/2011/sutivi_2011_2011-11-02_fi.pdf (Accessed 27 February 2012)

Prior, M. (2005). News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. American Journal of Political Science, 49, 577-592.

Putnam, R. D. (2000). Bowling alone: The collapse and revival of American community. New York: Simon & Schuster.

Shah, D. V., Cho, J., Eveland, W. P., Jr., & Kwak, N. (2005). Information and expression in a digital age: Modeling Internet effects on civic participation. Communication Research, 32, 531-565.

Sometrik (2012). About Sometrik.com. Retrieved from http://sometrik.com/perl/fb/about (Accessed 14 March 2012).

Wasserman, S. & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge University Press, New York

Watts, D.J.; Strogatz, S.H. (1998). "Collective dynamics of 'small-world' networks". Nature 393 (6684): 440–442.

Weatherford, M.S., 1982. Interpersonal Networks and Political Behavior. American Journal of Political Science, 26(1), pp.117–143.

Zhang, W., Johnson, T. J., Seltzer, T., & Bichard, S. L. (2010). The revolution will be networked. Social Science Computer Review, 28(1), 75–92.

Appendix 1. Detailed descriptive tables.

Table 1A: Comparing means, posts made by page **admin** only

	Niinistö					**Haavisto**					**All**				
	Mean	N	SD	Min	Max	Mean	N	SD	Min	Max	Mean	N	SD	Min	Max
Activity															
Activity level (%)	1,725	84	2,045	,000	9,739	1,468	127	1,594	,000	8,524	1,570	211	1,787	,000	9,739
Overall activity count	1512,702	84	1519,108	0	6253	1531,520	127	1454,071	0	5661,000	1524,028	211	1476,741	0	6253
Number of active users	1357,024	84	1439,509	0	6073	1367,205	127	1370,870	0	5497,000	1363,152	211	1395,207	0	6073
Number of wall post likes	1297,643	84	1338,515	0	4965	1303,047	127	1324,110	0	4971,000	1300,896	211	1326,682	0	4971
Number of wall post likers	1297,643	84	1338,515	,00	4965,00	1303,047	127	1324,110	,00	4971,000	1300,896	211	1326,682	,00	4971,00
Number of comments	98,071	84	202,832	0	1277	68,591	127	93,631	0	661,000	80,327	211	147,410	0	1277
Number of comment likes ^	116,988	84	112,111	0	503	159,882	127	195,690	0	965,000	142,806	211	168,486	0	965
Number of comment likers ***	40,250	84	31,745	,00	138,00	76,236	127	77,271	,00	380,000	61,910	211	65,518	,00	380,00
Number of commenters in the post	27,952	84	25,269	,00	112,00	26,291	127	25,971	,00	154,000	26,953	211	25,646	,00	154,00
Structure															
Average friendship component size of active users of the post	1,273	78	0,232	1,000	2,044	1,332	123	0,327	1,030	2,376	1,309	201	0,295	1,000	2,376
Median friendship component size of active users of the post	1,000	78	0,000	1,00	1,00	1,000	123	0,000	1,00	1,000	1,000	201	0,000	1,00	1,00
Average number of friends the users of the post have within the post	0,477	78	0,311	,000	1,395	0,521	123	0,435	,059	1,799	0,504	201	0,392	,000	1,799
Average number of friends the users of the post have within the all active users of the page	10,040	78	1,521	6,656	14,556	15,413	123	2,137	10,919	21,873	13,328	201	3,250	6,656	21,873
Percentages of active users within the post having at least one friend within the post	25,445	78	12,289	,000	57,432	27,413	123	15,063	5,212	62,355	26,649	201	14,052	,000	62,355
Number of components within the post	951,798	84	768,489	,00	3503,00	895,969	127	653,452	,00	3122,000	918,194	211	700,263	,00	3503,00
Number of friendship connections the users have within the post	511,310	84	911,195	,00	3627,00	616,205	127	1047,849	,00	4445,000	574,445	211	994,785	,00	4445,00

^ Significant at level p<0,10

*** Significant at level p<0,001

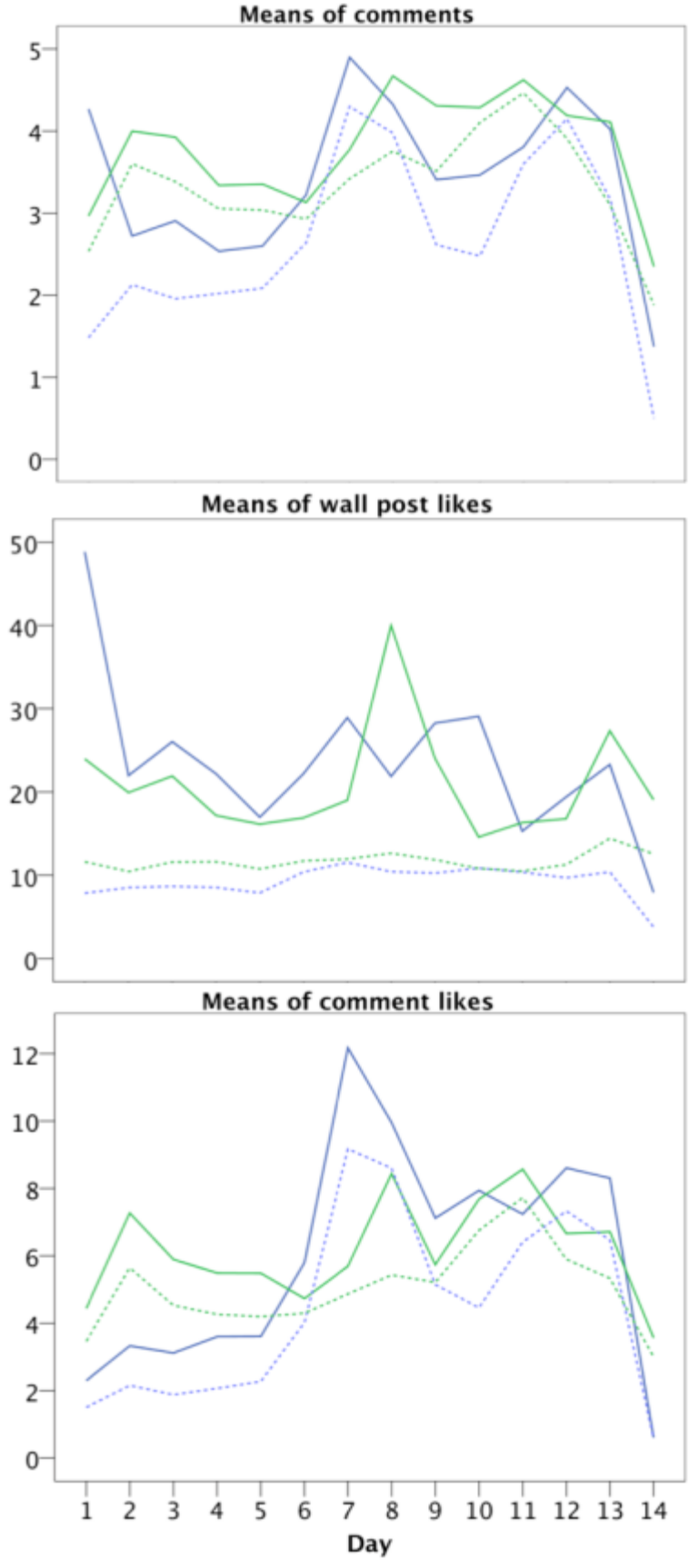
Table 1B: Comparing means, posts made by **users** only

	Niinistö					**Haavisto**					**All**				
	Mean	N	SD	Min	Max	Mean	N	SD	Min	Max	Mean	N	SD	Min	Max
Activity															
Activity level (%) ***	0,011	9435	0,012	,001	,261	0,017	17818	0,018	,001	,503	0,015	27253	0,016	,001	,503
Overall activity count ***	12,984	9435	17,691	1	305	20,742	17818	24,898	1	695	18,056	27253	22,962	1	695
Number of active users ***	9,632	9435	10,018	1	287	15,924	17818	15,922	1	356	13,746	27253	14,472	1	356
Number of wall post likes ***	7,292	9435	8,706	0	280	11,773	17818	14,222	0	296	10,221	27253	12,768	0	296
Number of wallpost likers ***	7,292	9435	8,706	,00	280,00	11,773	17818	14,222	,00	296,00	10,221	27253	12,768	,00	296,00
Number of comments ***	1,841	9435	4,201	0	106	3,177	17818	5,481	0	124	2,715	27253	5,114	0	124
Number of comment likes ***	2,851	9435	9,553	0	170	4,792	17818	13,704	0	499	4,120	27253	12,459	0	499
Number of comment likers ***	1,469	9435	3,759	,00	58,00	3,057	17818	6,348	,00	126,00	2,507	27253	5,640	,00	126,00
Number of commenters in the post ***	0,923	9435	2,086	,00	26,00	1,444	17818	2,614	,00	50,00	1,264	27253	2,457	,00	50,00
Structure															
Average friendship component size of active users of the post ***	1,404	9435	0,697	1,000	17,000	1,104	17818	0,238	1,000	8,000	1,208	27253	0,475	1,000	17,000
Median friendship component size of active users of the post ***	1,129	8998	0,611	1,00	17,00	1,010	17721	0,155	1,00	8,00	1,050	26719	0,381	1,00	17,00
Average number of friends the users of the post have within the post ***	0,485	9435	0,527	,000	4,250	0,161	17818	0,279	,000	4,696	0,273	27253	0,413	,000	4,696
Average number of friends the users of the post have within the all active users of the page ***	30,994	9435	34,694	,000	440,000	21,803	17818	17,344	,000	560,000	24,985	27253	25,149	,000	560,000
Percentages of active users within the post having at least one friend within the post ***	30,090	9435	29,239	,000	100,000	11,339	17818	16,556	,000	100,000	17,831	27253	23,553	,000	100,000
Number of components within the post ***	7,240	9435	7,743	1,00	263,00	14,486	17818	13,777	1,00	237,00	11,978	27253	12,520	1,00	263,00
Number of friendship connections the users have within the post ***	3,019	9435	5,609	,00	119,00	1,673	17818	5,429	,00	356,00	2,139	27253	5,529	,00	356,00
Number of friends the poster has within ALL the active users **	22,255	9435	71,066	,00	818,00	23,962	17818	44,452	,00	764,00	23,371	27253	55,144	,00	818,00

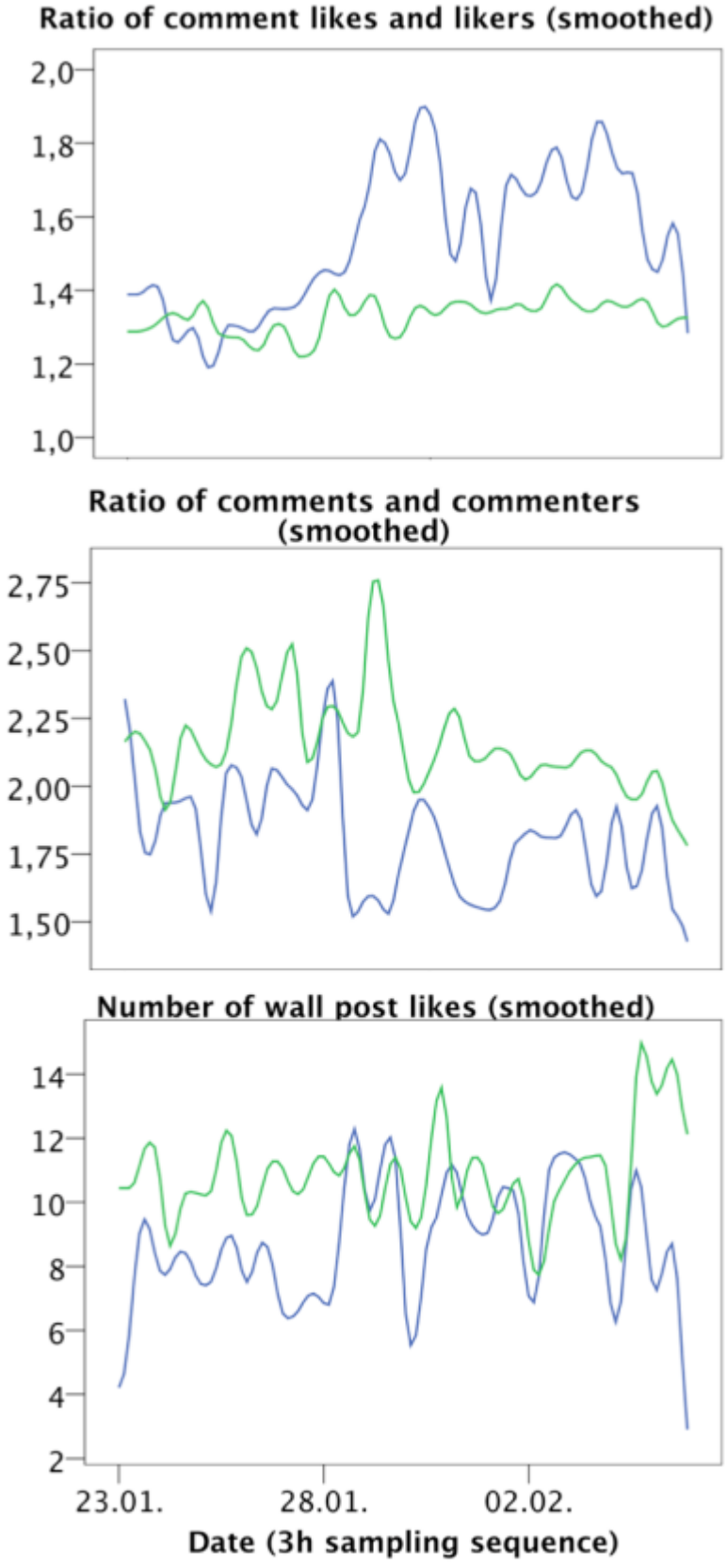
*** Significant at level p<0.001

** Significant at level p<0.01

Appendix 2. Means of comments, wall post likes and comment likes for Haavisto's (green line) and Niinistö's (blue line) pages. Dotted lines show the same measures without the contributions accounted for the posts made by the page admin.



Appendix 3. The ratios (first two plots) are calculated for the each post separately. Average values of ratios for each 3 hours time slice are used. The third plot shows the average number of wall post likes per three hours. (Haavisto = green line, Niinistö = blue line, admin posting excluded)



Appendix 4 Correlations

Table 4A Correlations - Spearman's rho, Posts made by the users

Page ID			LOG Activity level (%)	LOG Overall activity (count or all act, log trans.)
Niinistö	LOG Activity level (%)	Correlation Coefficient	1,000	,957**
		Sig. (2-tailed)	.	,000
		N	9435	9435
	LOG Number of components within the post	Correlation Coefficient	,947**	,931**
		Sig. (2-tailed)	,000	,000
		N	9435	9435
	LOG Number of friends the poster has within ALL the active users	Correlation Coefficient	,181**	,178**
		Sig. (2-tailed)	,000	,000
		N	9435	9435
	LOG Number of friendship connections the users have within the post	Correlation Coefficient	,698**	,731**
		Sig. (2-tailed)	,000	,000
		N	9435	9435
	LOG Average number of friends the users of the post have within the post	Correlation Coefficient	,411**	,452**
		Sig. (2-tailed)	,000	,000
		N	9435	9435
	LOG Average number of friends the users of the post have within the all active users of the page	Correlation Coefficient	,429**	,420**
		Sig. (2-tailed)	,000	,000
		N	9435	9435
	LOG Probability of finding completed triad within the post	Correlation Coefficient	,410**	,447**
		Sig. (2-tailed)	,000	,000
		N	9435	9435

Page ID			LOG Activity level (%)	LOG Overall activity (count or all act, log trans.)
Haavisto	LOG Activity level (%)	Correlation Coefficient	1,000	,949**
		Sig. (2-tailed)	.	,000
		N	17818	17818
	LOG Number of components within the post	Correlation Coefficient	,958**	,965**
		Sig. (2-tailed)	,000	,000
		N	17818	17818
	LOG Number of friends the poster has within ALL the active users	Correlation Coefficient	,133**	,123**
		Sig. (2-tailed)	,000	,000
		N	17818	17818
	LOG Number of friendship connections the users have within the post	Correlation Coefficient	,559**	,536**
		Sig. (2-tailed)	,000	,000
		N	17818	17818
	LOG Average number of friends the users of the post have within the post	Correlation Coefficient	,392**	,366**
		Sig. (2-tailed)	,000	,000
		N	17818	17818
	LOG Average number of friends the users of the post have within the all active users of the page	Correlation Coefficient	,225**	,183**
		Sig. (2-tailed)	,000	,000
		N	17818	17818
	LOG Probability of finding completed triad within the post	Correlation Coefficient	,249**	,236**
		Sig. (2-tailed)	,000	,000
		N	17818	17818

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

[Table 4BS here]

Table 4B: Correlations - Spearman's rho, Posts made by the page admin

Page ID			LOG Activity level (%)	LOG Overall activity (count or all act, log trans.)
Niinistö	LOG Activity level (%)	Correlation Coefficient	1,000	,961**
		Sig. (2-tailed)	.	,000
		N	84	84
	LOG Overall activity (count or all act, log trans.)	Correlation Coefficient	,961**	1,000
		Sig. (2-tailed)	,000	.
		N	84	84
	LOG Number of components within the post	Correlation Coefficient	,968**	,993**
		Sig. (2-tailed)	,000	,000
		N	84	84
	LOG Number of friends the poster has within ALL the active users	Correlation Coefficient	.	.
		Sig. (2-tailed)	.	.
		N	84	84
	LOG Number of friendship connections the users have within the post	Correlation Coefficient	,948**	,971**
		Sig. (2-tailed)	,000	,000
		N	84	84
	LOG Average number of friends the users of the post have within the post	Correlation Coefficient	,834**	,853**
		Sig. (2-tailed)	,000	,000
		N	78	78
	LOG Average number of friends the users of the post have within the all active users of the page	Correlation Coefficient	-,428**	-,456**
		Sig. (2-tailed)	,000	,000
		N	78	78
	LOG Probability of finding completed triad within the post	Correlation Coefficient	,627**	,656**
		Sig. (2-tailed)	,000	,000
		N	78	78

Page ID			LOG Activity level (%)	LOG Overall activity (count or all act, log trans.)
Haavisto	LOG Activity level (%)	Correlation Coefficient	1,000	,976**
		Sig. (2-tailed)	.	,000
		N	127	127
	LOG Overall activity (count or all act, log trans.)	Correlation Coefficient	,976**	1,000
		Sig. (2-tailed)	,000	.
		N	127	127
	LOG Number of components within the post	Correlation Coefficient	,976**	,991**
		Sig. (2-tailed)	,000	,000
		N	127	127
	LOG Number of friends the poster has within ALL the active users	Correlation Coefficient	.	.
		Sig. (2-tailed)	.	.
		N	127	127
	LOG Number of friendship connections the users have within the post	Correlation Coefficient	,969**	,970**
		Sig. (2-tailed)	,000	,000
		N	127	127
	LOG Average number of friends the users of the post have within the post	Correlation Coefficient	,909**	,900**
		Sig. (2-tailed)	,000	,000
		N	123	123
	LOG Average number of friends the users of the post have within the all active users of the page	Correlation Coefficient	,352**	,314**
		Sig. (2-tailed)	,000	,000
		N	123	123
	LOG Probability of finding completed triad within the post	Correlation Coefficient	,794**	,787**
		Sig. (2-tailed)	,000	,000
		N	123	123

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).