3-31-2009

Temporal Social Coordination Through Social Networks

Liaquat Hossain
Project Management Graduate Programme, The University of Sydney, NSW 2006, Australia, l.hossain@usyd.edu.au

Sabrina Gao
Project Management Graduate Programme, The University of Sydney, NSW 2006, Australia

TEMPORAL SOCIAL CORDINATION THROUGH SOCIAL NETWORKS

Liaquat Hossain and Sabrina Gao
Project Management Graduate Programme, The University of Sydney
NSW 2006 Australia
Email: l.hossain@usyd.edu.au

Abstract

Temporal communication is mainly associated with the concept of time. The social network derived from temporal environment is constantly changing; a communication link can be connected and disconnected highly frequently. Further with the communication technology such as cell phone, time it self has shifted from an absolute time to a relative time. Mobile communication is closely related with temporal communication, due to its micro coordination property and also the constant establishment of links and breakage of links from time to time. To study the network in the temporal domain, we are constrained by the relative time concept. As communication behaviour is highly dynamic, we expect formation of new ties and breakages of existing ties over time. This is especially different when comparing to social network studies conducted through self report surveys as the network through self report survey remains relatively static for the duration of the survey. In our study, we are interested in how a person would be expanding its network only. Thus we use an accumulated network structure to study the total links a person acquires over time and how such influences the network position.

Keywords: Temporal coordination, social networks, mobile communications, usage behaviour

TEMPORAL SOCIAL CORDINATION THROUGH SOCIAL NETWORKS

Abstract

Temporal communication is mainly associated with the concept of time. The social network derived from temporal environment is constantly changing; a communication link can be connected and disconnected highly frequently. Further with the communication technology such as cell phone, time it self has shifted from an absolute time to a relative time. Mobile communication is closely related with temporal communication, due to its micro coordination property and also the constant establishment of links and breakage of links from time to time. To study the network in the temporal domain, we are constrained by the relative time concept. As communication behaviour is highly dynamic, we expect formation of new ties and breakages of existing ties over time. This is especially different when comparing to social network studies conducted through self report surveys as the network through self report survey remains relatively static for the duration of the survey. In our study, we are interested in how a person would be expanding its network only. Thus we use an accumulated network structure to study the total links a person acquires over time and how such influences the network position.

Keywords: Temporal coordination, social networks, mobile communications, usage behaviour

1.0 Introduction to Mobile Social Networks

In this paper, we assess a selection of factors that could potentially impact the development of the social network within the campus mobile communication domain by drawing an understanding towards how the theory in other domain can be applied to the mobile communication network domain, further, it aims to validate some current debate on the subject on underlying influencing factors on network change over time. We proceed to find and validate the relationship between the proposed influence factors to the actual social network change in the campus mobile communication domain. It should be noted that the proposed factors may not be all the factors that would impact on the change of the social network. Social network analysis is a set of methods that allows exploration into the structure of a network. It is concerned with the relational aspect of the network rather than the attribute level of the network. Social network involves a set of nodes; these nodes are linked to each other by defined relationships [31]. More than two nodes are required for a social network to be constructed. The nodes can be linked by different relationships, and these relationships can have different directions attached and can be of different strength. The data is usually presented using a matrix tabular form or a socio gram which shows the connections among the nodes. The structure characteristic of the network is defined in its social network measures such as power, concentration, degree, centrality etc. A social network's structure would influence the way people behave and interact within a system.

The study begins by gathering background studies and theory on network development process and social network analysis. The findings from the previous researches are collated and discussed, hypothesis are then formed. We obtain data from Massachusetts Institution of Technology (MIT), and proceed to cleanse and perform manipulation on the dataset. Previous studies conducted on the same data set have been analyzed, we identify that although the development of social networks is an important interest from MIT, very small number of studies have been conducted [46]. The data set is provided by the Dynamic group at Massachusetts Institute of Technology Media Lab [46]. The data is collected for the Reality mining experiment, which is the largest mobile phone project in academia, seeking to provide insights into dynamics of individual and group behavior. The data which will be used for this study is collected over the 9 months period on 100 participants from MIT, consisting of a mixture of Media lab students, staff and the business school students. The data is collected through the mobile installed application developed by the Context application from the University of Helsinki on Nokia mobile phone device to capture their communication, proximity, location and activity over the course of 9 months. In addition, participants have been asked to complete a survey regarding their roles and social patterns. The collected data are saved into a Mysql database file. The data represents complex systems and are used for studying in a variety of fields to answer many research questions raised. Various project themes include: complex social systems, sociology, user behavior modeling and prediction, relationship inference, social serendipity, organizational dynamics, epidemiology and information dissemination, Eigen behaviours.

The purpose of social network study simply stated is to explore the relationship between the group of interacting people [22] within a defined domain, and extrapolate how such a relationship structure impacts on the future interaction and outcome, whether the outcome would be organizational performance or disease dissemination. By using this information, it helps people to achieve the desired outcome by predicting how people would behave and manipulate the network structure. For instance, Goyal has shown that by altering the social network structure, the long run outcome in gaming can be significantly altered [27]. Therefore understanding the development characteristics of a network over time brings benefit in helping to

facilitate network formation and growth in the required manner for a better and more certain outcome.

2.0 Mobile Coordination –Spatial, Time and Micro-coordination

With the advancement of technology and transportation, we will need to coordinate not only the activities but also movements to accommodate for the geographical dispersion [44]. Coordination has thus varied from the simple rules such as to meet at a certain place at a certain time to mediated interaction such as use of mobile phone or email. Communication is regarded as an essential part of group interaction and social activity coordination. [40]. The mobile technology allows people at different geographic locations to contact each other at ease and is beginning to remove the spatial component as an impeding limitation for communication. This deviates with the traditional social network notion that actors bond based on physical proximity as mobile communication allows friends to form outside of their physical proximity more easily. This is supported by Congleton and Nainwal's study [12], it can be seen that amongst the study participants the groups formed by the frequency of calls are not the same as the organizational groups the students are in. In addition to the ability to coordinate and communicate at dispersed geographic locations, one of the most important contribution of mobile phone is mobile telephony allows for more nuanced coordination of everyday life [44]. The concept of time and time keeping are the instrumental integration of society and governs the way we steer through life. It allows activities to be synchronized. With the access to mobile telephony, we move away from the time based coordination to a system that allows more flexibility, where the activities do not need to be synchronized in reference to a secondary system (mechanical timekeeping). The coordination becomes instantaneous rather than scheduled. Alternatively, the concept of time has shifted from the absolute time to a reference time within the social tie itself [55]. Mobile users attribute to the mobile phone communication a strong coordinative value [24]. German survey results show that most of the mobile usages are for making and canceling appointments, verbal communication of good news, verbal reassurance, and in emergency situations. The table below shows that verbal mobile conversations are mainly oriented towards spatial and temporal coordination activities. Most of the phone calls are made to actors that the mobile user already knows, mobile calls is not used as a

primary tool to meet more people [30]. In [44], it was mentioned that "Niche of the mobile telephone is in the realm of small groups for micro-coordination of activity." Thus we expect a network portrayed by mobile communication consists of many small sized cliques. Where the connections between actors are bridged through some key actors within the network and most communications would occur within in group members.

Figure 1: Mobile Usage, Retrieved from [24]

Table 1: Dimensions of verbal mobile communication content. Percents of answers to the open-ended question: About what do you speak at the mobile phone? What are the topics for discussion?

Dimensions		Pct. of responses (Base 125)	Cumulated pct. of responses	Pct. of cases (Base 76)	Cumulated pct. of cases*
Spatial and temporal coordination	Appointments (combined time and place references)	14	50	30	105
	Time clues only	25		52	
	Spatial clues only	11		23	
Task coordination	Problem solving information	25	75	52	157
Sociability	Sociable talk	25	100	52	209

Note: *Sum of pct. over 100 due to the multiple answers

Whilst mobile telephony provides the ability for the users to use 'time' differently in coordinating activities, allow communication between users in two geographically dispersed locations, it does not bring in as much advantage over other forms of coordination (e.g. email) when the coordination or information dissemination is required between a larger group of people. The communication between two people is fast and effective, however with a larger group; a large number of multiple temporal ties need to be established. With any changes to an event, mobile user would require large number of calls to inform the group. Thus mobile coordination is more common in smaller social groups when coordinating activities, and is suitable for micro coordination activities.

2.3.3 Mobile Phone Usage

In the paper 'The rise of the mobile communication', the German scholars conducted an extensive study concerning the usage of mobile phones through a detailed survey on German mobile phone users, including the functionality being used, the usage patterns in relation to a range of demographics – age, sex, education, the relationship between mobile use and traveling, the usage affected by one's experience with technology. Results have shown a these factors affects the usage of mobile phone differently. Mobile phone users tend to have a larger group

of friends than non user pointing to the extension of close relationships. The use of mobile phone communication helps in connecting people, but just connects rather than building relationships. Thus it seems that mobile telephony usage may have a complex effect on the construction of social relations [8]. The majority of mobile research has focused on individual usage; other studies have shown that remote people in the user's social network dramatically affect device usage and user behavior [23], however little work has been done to derive how such a social network can change over time.

Social networks are highly dynamic; they change over time through establishment of new links and additions of new edges. In the mobile context, cell phone is perceived as a peripheral tool, that is, its use is to strengthen relationships rather than building new relationships. Studies in the mobile domain generally indicate that the call frequency is related to the actor's maintenance of relationships. However, in the real world domain, social network is formed with respect to one's proximity, position, and interest group. As mobile devices are one of the main communication means used for coordination, the mobile phone would therefore reflect the established immediate links and provide a picture of the actors' communication network.

A key focus has been on the link prediction problem. Where social interaction pattern and social network can be drawn based on different parameters. A study has been conducted [54] to explore how the network changes and how actors interact with each other in playing games. We want conduct a similar study in the mobile community, more specifically to seek some answers to the questions below:

- Does mobile network reflect the off line social networks change?
- What specific social factors affect the network development over time?
- The extent of the different factors' change on mobile network
- What relationship is there between the demographic of the group and the change in their social network?
- Is the initial social network the only explanatory variable that determines the network at a later time interval?

2.4.1 Model

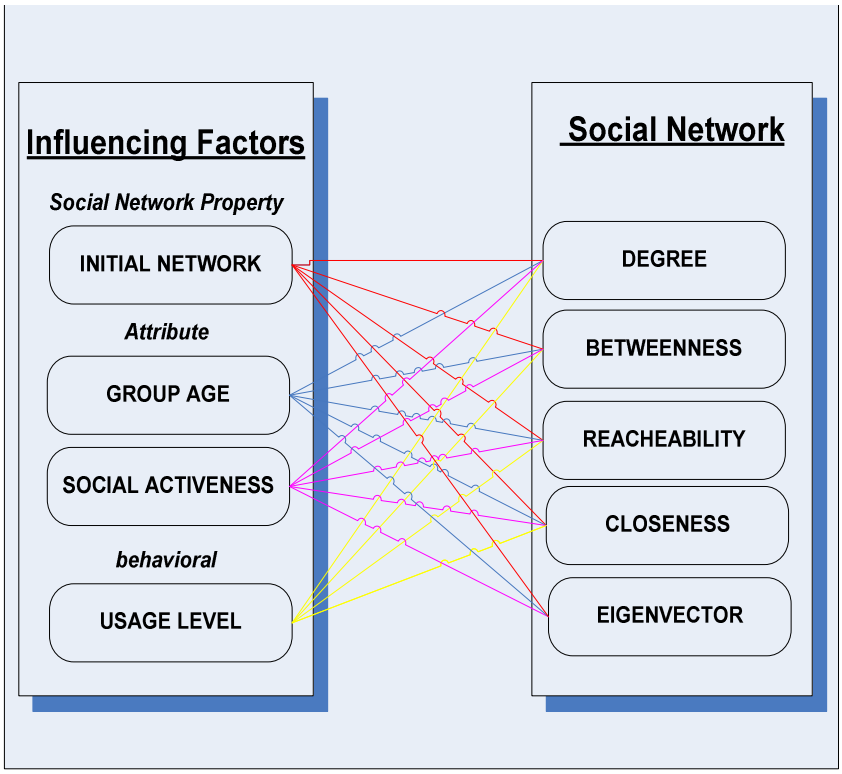
In the paper 'exploring temporal communication in a social network', the social network structure, is related to the communication behavior within the network [10]. In addition, the social network is affected by the environmental, individual and network attribute and other factors. As mobile technology is used as a peripherals tool to support the maintenance of social relations and development of new connections, it is essential to consider how the real world factors in conjunction with the properties of mobile communication would impact on the network development process.

Based on the previous sections, a social network develops based on actor attributes (actor's personality and attitude), network structure (initial network), network properties (network density and network duration), technology domain, behavior traits (communication) and other external factors (calendar).

A selection of measure has been proposed in this model to be related to the standard measure which would define the general structure of a social network.

The main measure for social network are defined below in the diagram, these are regarded as the dependant variable for this study. The factors which would cause changes to the network are regarded as the independent variables.

Figure 2: Model



2.4.2 Hypothesis

Hypothesis one: The social network development correlates with its initial network structure.

Hypothesis two: The social network development correlates with the established group age..

Hypothesis three: The social network development correlates with the Socialness factor of the actors.

Hypothesis four: The social network development correlates with the amount of calls of the actors.

3. Social Network Analysis Methods

Social network are concerned with the network rather than the node attribute level data. In order to study the social network, methods need to concentrate on drawing the network structural picture as a whole network, and perform analysis on the relational aspects of the network rather than treating the nodes independently [66]. To see the overall network structure, the linkages between actors need to be mapped against their relationship. A common way to do so is by computing this information into a relational matrix, where it indicates the tie relationships among the actors within the network. Such mapping can be binary or multiple; binary simply shows the existence of connection, multiple shows the types of the relations that are existent between the actors, and the strength of the connections. The same matrix can be manipulated in software programs to provide graphical representations of the network connections, called a socio gram [31]. Depending on the input data, the diagram shows the links between actors, and provides visual indication on the density of the network and the structures of the network. It aids in social network analyst in looking for patterns based on preliminary observations presented in graphical form.

Network analysis can be performed at two levels. At the whole network level, or at the ego level. The network level analysis requires a thorough identification of the complete network, including all interacting actors [31]. Such boundaries are hard to be defined, as it is highly likely that the actors know someone outside of the existing defined boundary. Further, the small world phenomenon indicates that anyone in the world can be connected in just a few connections [1]. The population for a complete network analysis is thus rarely possible. Ego level

analysis looks at the subjects and their corresponding position and role within their immediate network. Boundaries are defined much easier. Majority of studies are conducted at the Ego level.

A popular software used to analyze social network is UCINET. The program automatically computes the social network structure values based on a relational matrix input. UCINET calculates the ego network or the whole network information. The reliability of the result depends on the validity of the dataset.

Social networks can be studied with many different types of variables, at different levels of analysis. In this paper, we are analyzing the resultant social network and the node positions in conjunction with attributes and initial network structure to gain a better understanding of why and how the social network is formed in such a way. We are only concerned if there is a connection established between the actors such that the actors have more acquaintance in the network. At this point, we are not concerned with the strength of the relationships or the breakage of the relationship. Therefore, the network adjacency matrix which looks at binary connections will be used and imported into UCINET for calculating the network centrality measures.

The Reality Mining Dataset [46]

The reality mining data set is collected by the MIT research lab for the purpose of studying human behavior. The dataset has been made available to the general academic community for research purposes. The data is collected by using 100 Nokia 6600 smart phones preinstalled with several pieces of software developed for MIT as well as a version of the Context application from the University of Helsinki. The participants were given the phones which would automatically log their daily activities. Including Bluetooth devices in proximity, call logs, cell tower encountered, application usage. The data collected represents 500000 hours of data on the user's behavior over the 9 months period. Along with the automatically logged details, participants were also asked to provide a voluntary survey regarding their positions, and their social behaviors. The 100 users consist of 25 incoming students at the MIT Sloan business school adjacent to the laboratory and 75 are either students or faculty in the MIT media Laboratory. Of

these 75 users, 20 are incoming masters students and 5 are incoming MIT freshman.

The reality mining data set is a Mysql data dump file containing of eight tables. They include details on the location of the users, the call activity of the users, the functions used by the users, and a general voluntary survey on the users. The locations of the users are indicated in the Cellspan, Cellname, and Coverspan table. The device and device span table indicates the usage of the phone functions. The Activityspan and Callspan table indicates the calling activities of the user, and the person table contains the self reported survey information on the participants.

4. Stability of the Network and Actor Behaviour

If we assume that a network is in a perfect stable state, we would expect to see no change in the underlying social network factors, such that no new connections will be formed, and the node position would remain the same. However, the measure over the nine month period clearly indicates that the actors in the social network accumulated more ties.

By observing the change of the social network over the period, the results indicate that the frequency of contacts and network structure depends very much on the behavior patterns of the users. We do not see a very significant amount of tie growth between the second and third period, in contrast to the growth experienced between the first and second period. This appears to correlate with the amount of phone calls during the period. The increase and decrease of phone calls and network position can be partly explained by the MIT's calendar cycle. Showing that during holiday period, tie acquisition and phone call frequency are relatively low. Extrapolating, more interactions and common tasks between actors facilitate network growth.

Another explanation for the differing rate of growth during the two periods is that an actor that has a high degree centrality has more access to resources, and thus should be able to gain more ties and consequently enhance its position. Within a closed network, such as the one we have here, the actors can gain up to a certain number of ties until the network is saturated, where the rate of new tie development would slow down. Therefore the ties accumulated between period one and two would be larger than the number of ties accumulated between period two and three.

As we are examining the accumulative development of the network overtime, it is important that the network would exhibit a trend of increased ties and density before performing any detailed analysis. The number of ties and density of the network has been computed using UCINET. The network at T=0 consists of 96 ties, at the end of the period; the number of ties has increased to 210. The density of the network has increased from 0.0103 to 0.0226. This shows that the network saturates over time. At a high level, this indicates that actors within the network are acquiring more ties and would consequently increase their accumulated centrality within the network.

The social network structure is computed using centrality measures. The summarized measures are presented below. Over the period, the network has increased in all of its centrality measures.

Table 1: Network centrality at T=0

	Degree	Closeness	Betweenness	Eigenvector	dwReach
Mean	1.031	1.141	0.105	4.457	3.11
Std Dev	1.375	0.072	0.263	13.65	2.76
Sum	100	53.617	10.219	432.333	301.78
Variance	1.89	0.005	0.069	186.32	7.61
Minimum	0	1.042	0	0	1
Maximum	5.208	1.215	1.228	69.346	9.08

Table 2: Network centrality at T

	Degree	Closeness	Betweenness	Eigenvector	dwReach
Mean	2.255	1.562	0.742	6.185	9.85
Std Dev	2.156	0.26	1.552	12.959	6.08
Sum	218.75	118.706	72.018	599.897	955.07
Variance	4.646	0.067	2.407	167.937	37
Minimum	0	1.042	0	0	1
Maximum	10.417	1.792	8.122		21.24

The UCINET output also computes each of these five measures for all participants. The results have failed the normal distribution test, thus we turn to non parametric methods to analyze the dataset

Before performing statistical analysis, the result sets have been graphed using scatter plot such as below showing a relationship between initial degree and the resultant degree. There is a very obvious trend that the resultant network would depend on its initial network. Interpreting, the more degree a person has at inception, the more degree the person is likely to acquire. Which makes sense, as a person has more connection is likely to encounter more people.

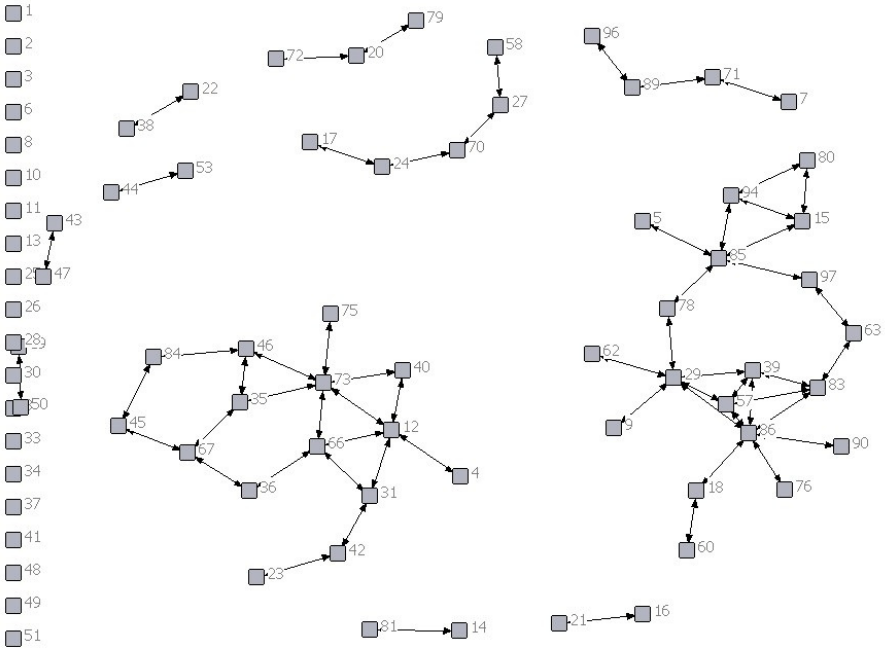
The Network Development Pattern:

Based on the network diagram produced in NetDraw, we are able to make some preliminary observations on the network structure and how it develops over time.

Period 1:

The first network captured from 20th September 2004 to 20th October 2004 shows that the initial network consist of many sub groups and loners at the start of the period. Not a significant amount of actors are connected initially Analyzing at the node level, we see 5 separate groups of different size have been formed depending on their position within the school (We consider only groups with 3 or more actors). Many dyads are also observed. This appears to agree with the notion that actors form networks based on their proximity.

Figure 2: Network diagram 1 as at the 20th October, 2004

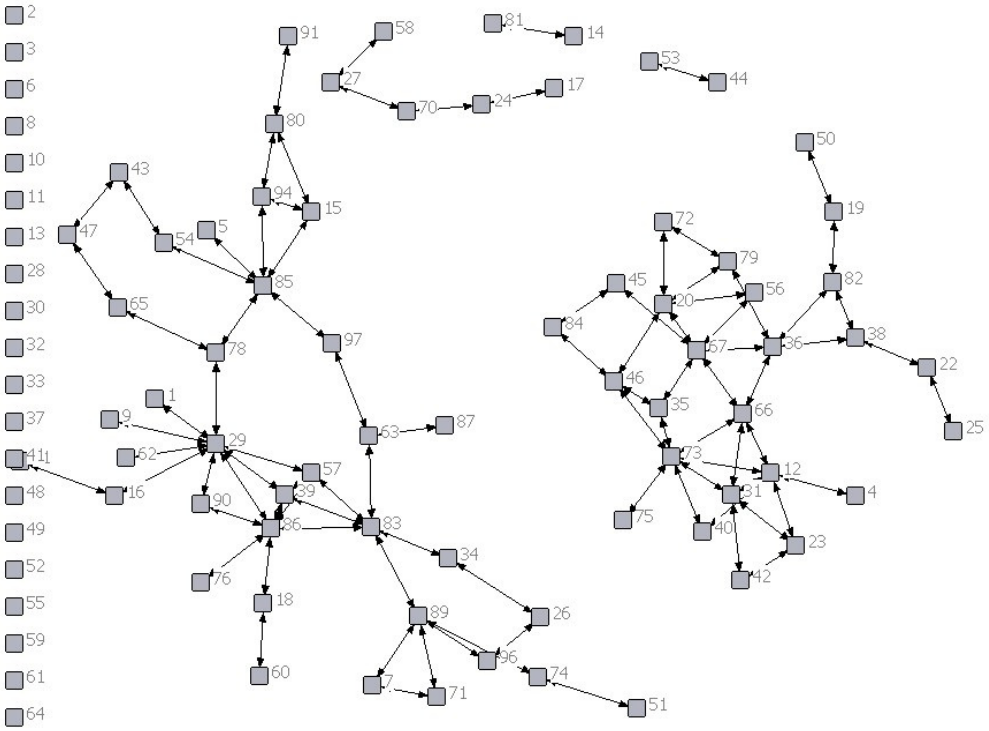


Period 2:

During the second time segment, we notice more connectivity between the actors within the network. The network topology consists of both star network and line network. The disconnected sub groups have reduced to three, showing that the actors are now better connected. Most of the connections have been established during this

period. We can see that in both clusters – the Sloan school and the Media Lab, there is a significant amount of accumulated tie growth.

Figure 3: Network Diagram 2 as at 20th January 2004



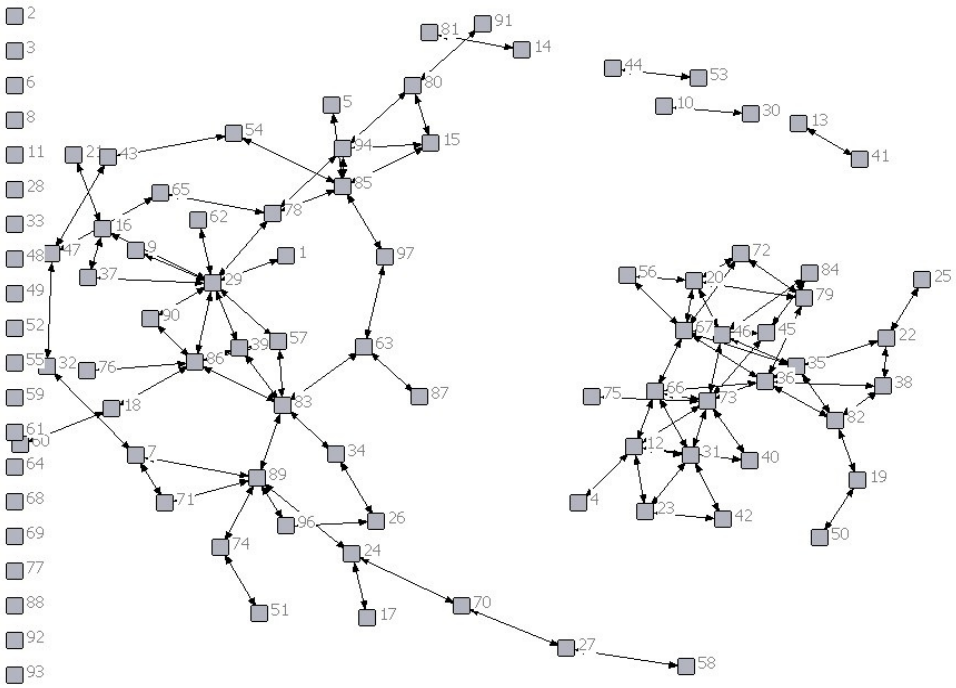
Period 3:

The resultant network is captured at the end of the period. As we assume that the actors that never made any phone calls during the period are due to their non involvement in the study, or not using the phone as their primary contact phone, we will ignore the loners for now, and concentrate on the groups. The network diagram 3 now further shows that the typical connections formed over time depend on actors' physical proximity. Two distinct groups are observed: The Sloan school group, and the Media lab group. Intra group activities is significant and inter group activity is none.

Comparing the actors' position in the three periods, we can see that the dominant actors generally remained in the same position. Actors that have a star topology, such as actor 29, 73, 83, have generally remained in star network over the three periods, they have the highest centrality measures. At the same time, they also have the highest call activities throughout the study period. Actor 40 experiences a larger centrality change as it has evolved form not being connected, being connected, to being connected to a central actor.

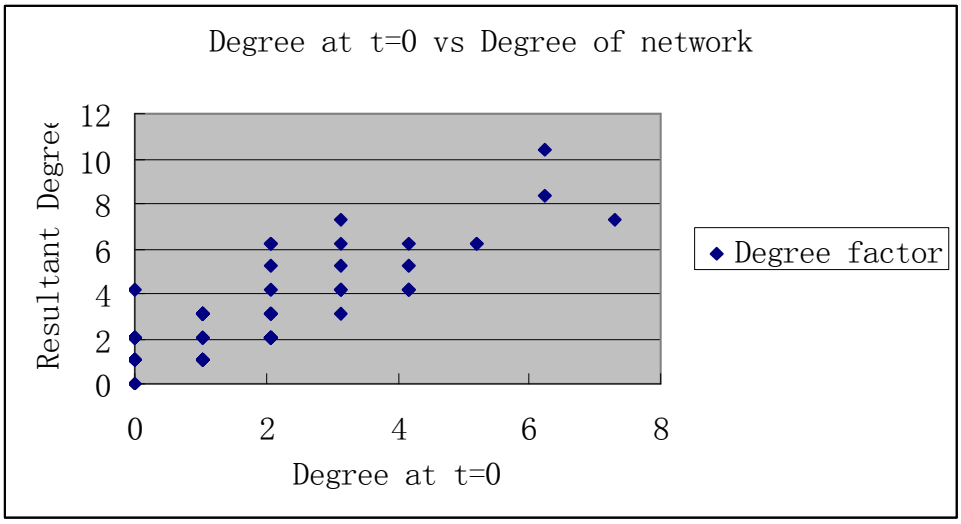
Interpreting the diagram further, and in coincidence with Eagle's [17] study, the Sloan school network appears to be denser than the Media lab. This is due to three reasons: one as suggested by Eagle, this is because that business students are more active; secondly, the business school consist of students that have a common attribute in comparison to the Media lab students. I.e. they are all new graduates coming into the school in contrast to the Media lab students who has a sample across different years as well as school staff. Lastly because that the business school students have a smaller population thus the connections required between actors for a network to be saturated is smaller. Further, as we know the population mix for between the two groups are different, i.e.: new students for the Sloan school, and mixed population for the Media Lab, we would like to know if such aggregation pattern of the network over time would also be influenced by the group's established age.

Figure 4: Network Diagram 3 as at end of the period



One of the most obvious observations from network diagram is the increase of degree or connections. To extrapolate the relationship between the initial degree centrality and the resultant degree centrality, a scatter plot is constructed to confirm the initial finding. Strong correlation has been observed between an actor's initial degree and resultant degree.

Figure 5: Degree Scatter plot



5. Overview of Preliminary Findings

Drawing from the preliminary analysis, we observe that

1. The social network to change overtime
2. A certain degree of network change would depend on an actor's initial position within the network
3. Groups tend to form due to actor commonality
4. The network would become more dense over time
5. Sloan business school users are more integrated than the Media lab students. This could be a consequence of the age of the established group or the attributes of the group. I.e. business school students vs. science students.

Hypothesis one: Initial network structure

To test this hypothesis, the network data at T=0 and T is extracted. Spearman correlation is performed to explore the correlation strength between the two sets of network position data on centrality measures. The correlation test has been performed at the 0.05 significance level; the results that are significant at this level have been highlighted in bold. We are not concerned with the correlation between the initial network versus the initial network, or the end network versus the end network. The findings are tabulated below.

Table 3: Initial Network Spearman Correlation Matrix

Variables	Degree at T=0	Closeness at T=0	Betweenness at T=0	Eigenvector at T=0	Reach at T0	Degree at T	Closeness at T	Betweenness at T	Eigenvector at T	Reach at T
Degree at T=0	**1**	**0.602**	**0.851**	**0.450**	**0.889**	**0.784**	**0.371**	**0.665**	**0.309**	**0.690**
Closeness at T=0	**0.602**	**1**	**0.592**	**0.834**	**0.665**	**0.347**	**0.691**	**0.385**	-0.185	**0.691**
Betweenness at T=0	**0.851**	**0.592**	**1**	**0.385**	**0.784**	**0.723**	**0.391**	**0.702**	**0.245**	**0.638**
Eigenvector at T=0	**0.450**	**0.834**	**0.385**	**1**	**0.453**	**0.263**	**0.685**	**0.322**	**-0.369**	**0.583**
Reach at T=0	**0.889**	**0.665**	**0.784**	**0.453**	**1**	**0.681**	**0.348**	**0.604**	**0.270**	**0.599**

The correlation matrix provides correlation values that are very significant at the 0.05 level. To interpret the effects the initial network has on the resultant network, we also use non parametric regression to ensure that the relationship is valid. The results show that there are significant correlations between the initial network structure and the end network structure. Such that even though a network changes over time, they change with respect to their initial network position. This result conforms to the notion that a network structure is developed based on its initial network structure.

Hypothesis Two: Position or Group Age

This hypothesis compares social network measures against the group existence age. I.e. newly established group versus an existing group. The duration of how long the group is existent depends on the position information provided by the participants.

We expect the groups that have been established longer have a stronger social network position as they have had more time for network development, but experiences a slower rate of change as a result of their stabilized network.

To test the hypothesis, the actors' position indicated in survey is mapped to the group establishment age, and this is then compared to 1) the resultant network, 2) the amount of change between T and T=0.

Spearman correlations have been computed to explore the correlation relation. The correlation test have been performed at the 0.05 significance level, the results that are significant at this level have been highlighted in bold. The findings are tabulated below.

Table 4: Outcome 1) the age of the group versus the resultant network

Correlation matrix (Spearman):

Variables	Position	Degree	Closeness	Betweenness	Eigenvector	dwReach
Position	**1**	-0.214	-0.067	-0.207	-0.131	-0.152

Table 5: Outcome 2) the age of the group versus the change experienced in the resultant network

Correlation matrix (Spearman):

Variables	position	Degree	Closeness	Betweenness	Eigenvector	Reach
Position	**1**	0.034	-0.134	0.018	0.244	-0.077

Hypothesis Three: Socialness

This hypothesis tests the relation between an actor's social activeness and the development of the network position. This hypothesis analyzes the data in two dimensions: the Socialness of the actor within the network versus the resultant network, and the Socialness of the actor versus the amount of position change experienced in the network.

We are expecting the more socially active actors to form more strategic relationship overtime, such that they will end up with more central positions within the network. They would also experience faster changes in their centrality as they actively seeks and participates in the network

To test the hypothesis, the actor's Socialness is constructed using the surveyed results of how close they are to the community and how much they'd prefer to meet new people. This is then compared to 1) the resultant network, 2) the amount of change between T and T=0.

Spearman correlations have been computed to determine the correlation strength. The correlation test have been performed at the 0.05 significance level, the results that are significant at this level have been highlighted in bold. The findings are tabulated below.

Table 6: The actors Socialness versus the resultant network

Correlation matrix (Spearman):

Variables	Socialness	Degree	Closeness	Betweenness	Eigenvector	dwReach
Socialness	**1**	**0.445**	**0.454**	**0.371**	-0.096	**0.505**

Table 7: The actors Socialness versus the change in the actor's network position between T to T=0

Correlation matrix (Spearman):

Variables	Socialness	Degree change	Closeness change	Betweenness change	Reach change	Eigenvector change
Socialness	**1**	**0.267**	0.160	**0.350**	**0.490**	-0.247

The results indicate that both dimensions of the network development correlate strongly with the actors' Socialness.

Hypothesis 4: Amount of Calls

There has been no prior research conducted on how phone calls would affect the change with respect to the actor's position within the social network. One would assume that as the number of phone calls increase, either one of the two would happen: more contacts are reached, or the strength of the ties grows. A preliminary observation made on the phone calls frequency, by ranking the actors in order of how many phone calls they make during the two periods indicates that the actors' ranks stayed the same. Therefore the actors are consistent in their phone usage.

The total number of calls of the participants have been tabulated into excel, Spearman correlation test is applied at the 0.05 significance level. The following result has been found. The numbers in bold indicates the variables are related to each other at the significance level.

Table 8: Phone call vs. Social Network

Correlation matrix (Spearman):

Variables	Totalw	Degree	Closeness	Betweenness	Eigenvector
Totalw	**1**	**0.778**	**0.257**	**0.493**	**0.296**

The result indicates that the total number of calls an actor makes significantly correlates with the network centrality position of the actor. The results of Hypothesis tests are summarized in the table below:

Table 9: Summary of Hypothesis Testing Outcome

	Degree	Closeness	Betweenness	Eigenvector	Reach
H1: Initial Network	Yes	Yes	Yes	Yes	Yes
H2: Group Age	No	No	No	No	No
H3: Socialness	Yes	Yes	Yes	No	Yes
H4: Call Amount	Yes	Yes	Yes	Yes	N/A

N.B: 'Yes' Indicates that the hypothesis is true, 'No' indicates the hypothesis is rejected. N/A has not been measured.

Discussion

This paper makes a contribution towards understanding the factors that cause the social network to develop in a certain way. This helps the actors to understand and predict their future centrality position in the network, forecasting who the powerful actors will be in forthcoming times. By looking at the impact on the network position triggered by the differences in underlying factors, it allows people to manipulate their behaviour so that they can be in a better central position over time.

The existing studies indicate that a social network would generally change over time due to actor composition and underlying dynamism. Majority of the studies points in the direction that the initial social network structure would dictate the future structural position of the actors within the network. Some of those studies recognise that the network structure at inception time is the only explanatory variable for future network development. There are however, a few studies that suggest that other factors could potentially impact how a network would be structured.

As social networks would also need to be studied within specific context, we recognise that there has been no prior study done in the Reality Mining Mobile communication domain, thus our study is carried out to explore the network formation determinants in this temporal campus communication environment. The study is conducted by performing social network analysis on the data set provided by MIT. A lot of attention has been paid to the quality of the data set when performing analysis. Few short comings of the data which have not been recognised in previous studies on the same dataset, have now been identified and resolved using correct SQL statements, data errors have been minimised. This allows future work on the same data set to be carried out with more accuracy. Social network analysis methodologies are applied to extract each node's centrality position information within their social network. These centrality measures are then compared to the suggested measurable factors. Spearman correlation and Regression have been performed to test the validity of our hypothesis and the actual dataset.

Our results show that in the Reality Mining mobile communication context, social network development process does not only depend on its initial network position, but also other factors such as communication behaviours, and individual social characteristics. These factors are guidance to how and what people may be able to

manipulate to construct a required social network structure, or to grow into a stronger social network position.

Our result set tests both views. It assesses the impact of the initial network on the resultant network, and goes further to test how some suggested attributes affect the rate of change on centrality measures. On a general level, our findings indicate that although at this stage, the initial network structure has the most explanatory power over the other variables proposed, it is evidential that more than just the initial network structure details are required to explain the development of the social network.

The first hypothesis tests the impact the initial structural positions of the participants in the network have on the resultant network. We found that the resultant degree, Betweenness, closeness, eigenvector, reach is impacted by the very same measures at the inception of the network. This relationship has been proposed in previous studies in face to face interaction domains as most real life social networks would have a degree of durability in its structure. However this is also true in our Mobile communication context. The repeated recognition on the explanatory power of the initial network to the end network, leads us to believe that this is perhaps the most important determinant in forecasting actors' structural position in forthcoming time. This notion is confirmed by the results found through the high r-square values through regression analysis. The results allow us to predict that, under normal circumstance, an actor that is in a stronger network position in the mobile communication context is likely to maintain its position rank within the network. The development traits of the actors within the network are one that follows an incremental pattern. It is unlikely that an actor that is in a low centrality position, for example a singleton person to expand its network rapidly. Consequently if we have measured an actors' centrality at the initial time, we can make some predictions about how central the actor will be as time progresses. This allows us to find the powerful actors within the network in future times by using existing network data.

Our third hypothesis found that the more active actors are more likely to develop into stronger positions within the social network. This part of the finding agrees with the existing work in the area. Taking a step further, we found these actors are also more likely to develop more centrality in a shorter period, at a faster rate. This finding allows us to derive different information. When we are given a social network structure, but no information regarding the actor's personality trait, we can use the

network structure information to find out who is most socially active within the network. Conversely, with the socialness of the actors known, we are able to predict how central these same actors will be in future time and at what rate would they grow their position in comparison to actors that are not as active. This finding begins to confirm our belief that network development is not only formed based on initial social network, but also on other factors.

Hypothesis 4 found that context specific behaviour pattern is also an influencing factor to network development. A person that has a high communication frequency, are more likely to grow into a more central position. Or if we consider it conversely, given a network structure, we will be able to find the actors that communicate more frequently. This finding is reasonable. If an actor has more connections, in order to attempt to maintain all relations in comparison to an actor that has half of the connections, effectively this actor has to make more communications to pass information through in the mobile domain.

One of the dictating factors for the development of network is the density of the existing network. A closed network that is more saturated is not as likely to grow as fast as a network that has new actors joining constantly. Given that the other variables are constant, an actor that enters a network early is more likely to have obtained a more central position than those who enters later. Thus the change experienced in the actor's network position development depends on his age within the network. While testing hypothesis two, we observed a tendency that as duration of the person within the network decreases, the growth of the person's centrality is faster. Unfortunately although we observed this trend, statistically this is only valid at a p value of 0.2. This may be due to a few reasons. One of the main reasons is data limitation. The group age in the sample data is not equally distributed. More than half of the subjects in the studied spreadsheet file are new students to the faculty, while there are only a few participants that are from the senior years, for instance 7^{th} year and 6^{th} year student. This severely impacts the outcome of our study, as we are assuming that these few students in the senior years represent the whole population of their group. If one of those students is an outlier, or differs from the typical behaviour of the student within their represented group, this would have adverse effect on the outcome of the study. This type of error is data selection problem where the sampled data deviates from real data. Another thing is, we have already seen in the results from other hypothesis indicate that social network position depends also on initial social network structure,

socialness and communication frequency. As we are using different sets of data to compute all the factors against the resultant social network, we cannot draw inference in our study whether there are relationships between the other factors and the group age factor. One improved way to analyse this is to divide actors into their respective groups according to socialness rating group, or social network group, this reduces the impact imposed by other factors. However, as we have data limitation due to the distribution of the actor's positions, it would be better to perform this analysis on a larger dataset with more even distribution across the different actor groups.

References

1. Adamic, L., Adar, E., 2005: *How to Search a Social Network*, Science Direct, HP Labs, USA
2. Arnold D., 2008: *The Stages of a Group: Forming, Norming, Storming, Performing and Adjourning*, Associated Content, <http://www.associatedcontent.com/article/885095/the_stages_of_a_group_forming_norming.html?cat=4> Retrieved 01/10/08
3. Backstrom.L, Huttenlocher.D. Lan.X.,kleinberg.J., 2006: *Group formation in large social networks: Memership, growth and evolution*, KDD 2006: pp44-54
4. Berdayes, V., 2008: *Temporal Communication, Course notes*, Saint Mary's college, Indiana, <http://www.saintmarys.edu/~berdayes/vincehome/courses/comm200/notes/chronemicsf02.html> Retrieved 01/09/08
5. Bidart, C. Degene, A., 2005: *Introduction: The dynamics of personal networks.* Social Networks 27 pp. 283-287
6. Burgelman, J.-C. 2001: *Traveling with communication Technologies in Space, Time and Everyday Life: An Exploration of Their Impact.* <http://www.firstmonday.dk/issues/issue5_3/burgelman/> Retrieved 02/09/08
7. Buskens V., 2008: *Concent or conflict: Coevolution of Coordination and Networks,* Journal of Peace Research, Vol. 45, No. 2, 205-222 (2008)
8. Callon, M., 1987: *Society in the Making: The Study of Technology as a Tool for Socio-logical Analysis.* In Bijker, W. E., Hughes, T. P. and Pinch, T. J. (eds.) The Social Construction of Technological Systems. Cambridge/MA.
9. Carley, M., Borgatti, S., Krackhardt, D., 2005: *On the robustness of centrality*

measures under conditions of imperfect data. Social networks, Elsevier B.V

10. Chung, K.S.K., Hossain, L., Murshed., S.T.H, 2007: *Exploring temporal communication through social networks,* PhD thesis, School of Information Technologies, University of Sydney, Australia
11. Clauset, A, Eagle., N. 2007: *Persistence and periodicity in a dynamic proximity network*, DIMACS Workshop on Computational Methods for Dynamic Interaction Networks
12. Congleton, B., Nainwal, S. 2007: *Mining the mine: Exploratory social network analysis of the reality mining dataset*, University of Michigan
13. Cushman, W., Rosenberg, J., 1991: *Human Factors in Product Design.* Elsevier, Amsterdan.
14. Demange, G., Wooders, M. (eds), 2004: *Group Formation in Economics: Networks, Clubs and Coalitions*. Cambridge University Press, Cambridge, UK, pp. 11-57.
15. Doreian, P., Stokman F., N., 1997: *The dynamics and evolutions of social networks,* Gordon and Breach Publishers, Amsterdam
16. Doring, N. 2002: *Have you finished work yet?- Communicative functions of text messages (SMS).* Receiver#6. The mobile Self: <http://www.receiver.vodafone.com/06/articles/index05.html> Retrieved 02/09/08
17. Eagle, N., N., 2005: *Machine Perception and Learning of Complex Social Systems*, PhD Thesis, Massachusetts Institute of Technology, United States.
18. Eclipse, 2008: Eclipse IDE for Java Developers. Downloaded from <http://www.eclipse.org/downloads/>
19. Feld, S.L., 1981: *The focused organization of social ties.* American Journal of Sociology, 86: 1015-1035.
20. Flap, H., 1999: *Creation and returns of social capital. A new research program*. La Revue Tocqueville 20, 1-22.
21. Freeman, L.C, Everett, M.G, Borgatti, S.P., 2002: *UCINET 6 for Windows: Software for Social Network Analysis,* Analytic Technologies, USA
22. Freeman., l., C., 2004: *The Development of Social Network Analysis: A Study in the Sociology of Science*. Booksurge Publishing, Vancouver, Canada
23. Gant, D. & Kiesler, S. (forthcoming). *Blurring the boundaries: Cell phones, mobility, and the line between work and personal life.* In B. Brown, N. Green,

& R. Harper (Eds.). Wireless world: Social and interactional aspects of the mobile age. NY: Springer Verlag.

24. Gesellschafts, V., F., et.al, 2003: *Understanding the mobile telephony usage patterns, The rise of the mobile communication 'dispositif'*, Germany
25. Gloor, P.A., 2005: *Capturing team dynamics through temporal social surfaces*. Proceedings of the ninth International conference on information Visualisation
26. Gloor, Peter A. Zhao, Y., 2008: *Visualising time in social networks with TeCFlow*, <http://www.ickn.org/JoSS_subm/TeCFlow4JoSS.htm> retrieved 01/10/08
27. Goyal, S., Redondo, F., V., 2002: *Network formation and social coordination*, *Games and Economic Behavior,* Vol. 50, No. 2. (February 2005), pp. 178-207
28. Granovetter, M., 2005: *The impact of social structure on economic outcomes*. Journal of Economic Perspectives 19, pp33-50.
29. Gruenfeld, D.H., Mannix, E.A., Williams, K.Y., and Neale, M.A. 1996: *Group composition and decision making: how member familiarity and information distribution affect process and performance.* Organizational Behavior and Human Decision Process, 67, pp1-15
30. Haddon, L. 2000: *The Social Consequences of Mobile Telephony*. In Ling, R. and Trane, K. (eds.). The Proceedings from a Seminar About Society, Mobile Telephony and Children (pp. 2-6). Telenor R&D N38/2000
31. Hanneman, R.A., 2004: *Introduction to social network methods*, University of California
32. Hawe. P and Ghali, L., 2007: *Use of social network analysis to map the social relationship of staff and teachers at school*, Oxford Journals, Volume 23, pp 62-69
33. Haythornhwaite, C., 1998: *A social network study of the growth of community among distance learners,* Informaiton Research, Vol.4 No.1, July 1998
34. Hossain, L. Chung, K. and Tan, A. 2006: *Exploring temporal communication through social networks*, PhD thesis, Univesrity of Sydney, Australia
35. Jackson, M., 2005: *A survey of models of network formation. Stability and efficiency.*
36. Jin, E., M., Girvan, M., Newman, M., E., J., 2001: *The structure of growing social networks*, Physical Review Letters E 64, 2001
37. Katz, J. E. (ed.) 2003: *Machines that become us: The social context of*

personal communication technology. New Brunswick, NJ: Transaction Publishers

38. Kimberley, R., I., 2005: *A Model of Operational Interorganizational Network Formation*, Ph.D thesis, Columbia Univesrity, New York.
39. Klein, K. J., Lim, B. C., Saltz, J. L., Mayer, D. M, 2004: *How do they get there? An examination of the antecedents of network centrality in team networks*. Academy of Management Journal, 47, 925-963
40. Kleinbaum, A., M., Stuart, T., E., Tushman, M., L., 2008: *Communication (and Coordination?) in a Modern, Complex Organization*, working papers, Harvard Business school.
41. Kumar, R., Novak, J., Tomkins, A., 2007: *Structure and Evolution of online social networks*. Research Track, pp611-617, USA
42. Lai, G and Wong, O., 2002, *The Tie effect on information dissemination: The spread of commercial Rumor in Hong Kong*, Social Networks 24, 2002, pp.49-51
43. Leenders, R. T. A. J. 1995: *Structure and Influence: Statistical Models for the Dynamics of Actor Attributes, Network Structure and Their Interdependence.* Amsterdam: Thesis Publishers.
44. Ling, R., Yttri, 2001: *The Mobile Connection: The Cell Phone's Impact on Society*, Elsevier
45. Mandelshtam, O., 2007: *Growth Models of Social Networks,* SURF 2007 Final report
46. Massachusetts Institute of Technology, 2005: *Reality mining dataset, Reality Mining project.* MIT, <http://reality.media.mit.edu/download.php> Retrieved 15/08/08
47. Matthew, G., Marie, D., 2008: *The effect of network structure on dynamic team formation in multi agent systems*, Computational Intelligence, Volume 24, Number 2, May 2008 , pp. 122-157(36)
48. Mayer, A., Puller, S., L., 2008: *The Old Boy (and Girl) Network: Social Network Formation on University Campuses,* Journal of Public Economics, Vol. 92, pp. 329-347,2008
49. Microsoft 2007: *Xlstat, statistical analysis tool, Addinsoft. Free trial copy for win2007,* Downloaded from <http://www.xlstat.com/en>
50. Mitrea, O. 2003: *Understanding Mobile Communication. Theoretical*

Considerations and Empirical Findings. In Buletin tiin ific, seria A, vol. XVI, Baia Mare: Universitatea de Nord, pp. 129-140.

51. Motulsky, H., 1995, Chapter 37: *Choosing a Test, intuitive Biostatistics,* Oxford University Press, New York. ISBN: 0-19-50-8607-4
52. Motulsky, H., 1999: *Analyzing data with Graphpad Prism, Graphad software* inc, San Diego, CA
53. Newcomb, T., M., 1961*: The acquaintance process*. New York: Holt, Reinhart & Winston
54. Nowell, D., L., Kleinberg, J., 2003: *The Link Prediction Problem for Social Networks. Proc*. 12th International Conference on Information and Knowledge Management (CIKM), 2003
55. Nyíri, K., 2006: *Time and Communication*, in: F. Stadler and M. Stöltzner, 2006:., Time and History / Zeit und Geschichte, Frankfurt/M.: ontos verlag, pp. 301-316.
56. Park, I., Ochs, J. 2004: *Overcoming the coordination problem: dynamic formation of networks*. USA.
57. Pearl, H. I., Trickett. E.J. 1981: *Social network formation of college freshmen: personal and environmental determinants*. American Journal of Community Psychology.1988 Vol. 16, No 2, pp. 207 – 250
58. Rummel, R., J., 1967: *Understanding Factor Analysis*, Journal of Conflict Resolution, Dec 1967, pp.444-480.
59. Shi, X., Adamic, L., Strauss, M., 2007: *Networks of strong ties.* Physica A 378 (2007) pp.33-47
60. Skyrms, B., Pemantle, R., 2000: *A dynamic model of social network formation,* University of California, USA
61. Sun Microsystems, *MySQL*, <http://www.mysql.com/> Retrieved 05/09/08
62. Taskar, B., Wong, M., Abbeel, P., Koller, D., 2002: *Discriminative probabilistic models for relational data*. In Proc. UAI
63. Taskar, B., Wong, M., Abbeel, P., Koller, D., 2002: *Link Prediction in Relational Data,* Stanford University
64. Thomas, F., Verlag, V., 2008: *Why Context Matters: Applications of Social Network Analysis*, VS Research
65. Valle, D., Hyman, J.M., Heathcote, H.W, Eubank, S.G., 2007: *Mixing patterns between age groups in social networks*. Social networks 29(2007), pp. 539-

554

66. Wassenman, S., Scott, J., Carrington, P., J., 2005: *Models and Methods in Social Network Analysis*, Cambridge Univesrity Press
67. Weimann, G., 1982, *The Strength of weak conversational ties in the flow of information and influence,* Social networks 5 (1983), pp.245-267