

Connecting Productivity with Social Capital via Daily Mobile Phone Logs

B. Padmaja¹, V. V. Rama Prasad², K. V. N. Sunitha³

¹Department of CSE, JNTUH, Hyderabad, India

²Department of CSE, Sree Vidyanikethan Engineering College, Tirupati, India

³BVRIT Hyderabad College of Engineering for Women, Hyderabad, India

Email: b.padmaja@gmail.com, vvramaprasad@gmail.com, k.v.n.sunitha@gmail.com

Received 16 March 2016; accepted 18 April 2016; published 21 April 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Human behavior and their social interactions can be quantified and modeled with the use of smart phones and any wearable badges which senses and captures real-life interactions. In traditional social sciences, such information was gathered by conducting surveys. However in digital era, smart phones are regarded as a popular tool which automatically senses much human information to quantify our lives. Reality mining gives a clear picture of a human being and its social relations. Social Network Analysis (SNA) is a powerful research tool which provides a comprehensive analysis on ego-alter communications with their individual productivity within a community. In this paper, various popular measures of social network analysis have used to study a closed community through their mobile call logs for a period of time. We experimented various social network measures both on daily basis and also over a period of time. The pattern shows that the relationships and interaction between ego-alter ties have more productive benefits. Using Pearson-correlation analysis, we observed that significant (positive) correlation exists between various network properties and their productivity. Results showed that degree (size) has the strongest positive correlation with average productivity, followed by effective size, efficiency, constraint, hierarchy, and k-core of an individual. Density and betweenness centrality have a weak, negative correlation with productivity. Hence social capital has a significant influence on human productivity.

Keywords

Reality Mining, Social Network Analysis (SNA), Social Capital

1. Introduction

Humans are social by nature and they start their daily activities by interacting with others. Social interactions

Note for the authors: ¹Research Scholar; ²Professor; ³Principal.

How to cite this paper: Padmaja, B., Prasad, V.V.R. and Sunitha, K.V.N. (2016) Connecting Productivity with Social Capital via Daily Mobile Phone Logs. *Social Networking*, 5, 62-74. <http://dx.doi.org/10.4236/sn.2016.52007>

come in many forms such as having lunch with colleagues, playing sports with friends and travelling with family. Online communications, Bluetooth proximity and citations networks are some of the examples of social networks, which are being analyzed to identify important social connections [1] [2]. But these data do not correspond to daily life interactions. Bluetooth has been used as a platform for social interaction and communication among groups of people [3]. By analyzing the Bluetooth proximity network like GroupUs, one can discover multiple group activities from time-stamped social interaction links [4]. Several works have been conducted on data collected from a real organization over several weeks with Bluetooth (BT), infrared sensors (IR), RFID devices [5] where interaction links were localized by their proximity to fixed stations. Another strategy for behavioral data collection is to resort to image and video processing based on cameras placed in the environment [6]. This approach provides rich datasets, but in turn, is computationally complex.

Eagle and Pentland [7] coined the term Reality Mining and used smart phone Bluetooth transceivers, phone logs, and cell tower ids to find the social network structure, identify social patterns in daily user activity and infer relationships.

Today, there are more than a billion smart phones in use and millions of users are using sensors for capturing their individual activity [8]. In this paper, we use the communication data collected by smart phones to monitor individual's social behavior and how it is related to their productivity. In other words, we focus on finding relationship between daily productivity of individuals with their mobile call log data.

In social network analysis, actors and the events are viewed as independent units and the ties between actors are used as channels for flow of resources. In social network analysis the focus is not on an individual, but mostly on an entity comprising of a collection of individuals and the linkages among them. Many network methods focus on dyads, triads, subgroups or entire networks. The set of actors on which measurements are taken is called an actor set and the set of ties among the actors is called a relation. Usually a social network consists of a set of actors and the relationships among them [9].

Social capital theory states that social networks play an important role in everyone's life. They are almost indispensable. Moreover, the productivity of individuals and groups is affected by social contacts due to the social connections that are developed and the trustworthiness that arises. Overall, the core idea of social capital theory is that social networks have value [10].

Social networks have value and social capital refers to the resources available in the social networks [5]. These resources include job opportunities, trust and emotional support, information finding, any new ideas on chosen subject. These resources reside in networks of relationships [11]. Social capital focuses on the individual's accessibility and use of embedded resources in social networks to gain benefits or profit to the individual. Individuals engage in interactions within a network, which facilitates the flow of information in order to produce benefits [12].

In this paper, individual's productivity is measured from the reality mining dataset. The dataset contains user's daily communication logs from their smart phones, within a residential community, for a period of 3 months. From the daily call logs, an individual's social relations are understood by analyzing their day-to-day interactions or communication with others. Social capital contributes to productivity through interactions between people [13]. According to Robert D. Putnam, there is an association between social contacts and productivity. Social capital shows contribution from people through interaction with others within their community and at their work place.

2. Related Work

Ghoshal *et al.* defined social capital by measuring the sum of potential resources which are embedded, available and derived from the social network relationships by a person or a social network. The structural dimension of social capital is bothered about the patterns that exist among a bunch of individuals [14].

Burt 2001, Social ties provides useful information about opportunities and choices to every individual. Strong and weak ties are two major types of social structures. Some social ties due to their strategic locations and positions make an individual productive and help to carry more valued resources and power [15].

Putnam 1995, established that social capital has an inclination towards mutually beneficial collective action and it is derived from the degree of associations among people within a particular group or community. It improves an individual's social credentials because of individual's accessibility to resources through social networks [16].

Lin 2001, experimented that the notion behind social capital is very simple and straight forward: investment in social relations gives expected returns in the marketplace. In this way, social capital works as a basis of productivity in any business organization also [15].

Pennings *et al.* 1998, done various studies to measure human capital at various levels and these studies demonstrate the impact on performance of the human resource behavior [17].

Coleman 1988, Social capital may build human capital by increasing interaction between individuals, so as they learn from each other. Social capital is a property of social relationships and a resource actor's possess and share. Therefore every individual has different levels of social capital as well as different ability to take advantage of it [18].

Krackhardt 1992, there are two important aspects to productivity which contributes to social capital. First one is that the social relations are used to mobilize people to give their contribution to a project. The second one is using team members' social capital to supplement and balance the knowledge of the team [19].

Burt 1992a, the value of social capital does not depend on the number of contacts an individual has, but on the structure of relationship within the network and the type of contacts they maintain with the network. Social capital is associated with productivity because getting a job is one of the best-known uses of networks. Many people find jobs through personal contacts than by any other means. People with rich social capital are paid better, promoted faster at young ages when they create value. This advantage of social capital arises when a network contains many "structural holes" or gaps [20].

Granovetter 1983, a closely-unified network of friends who all communicate with each other may not be very productive because of redundancy. Therefore, both indirect ties as well as direct ties are considered when we evaluate social capital. Few people may have direct ties, but they may be connected to several others, and may contribute more than a large group of people who are all tied to each other. Thus, the productive value of social capital depends not only on direct contacts, but also on indirect contacts and this is clearly reflected in the structure of the network [21].

Mark Granovetter 1973, introduced the most influential concept of tie strength. It refers to a sense of closeness with another person [12]. A tie can be either strong, weak or absent. Loose acquaintances are known as weak ties (Burt 2004), and trusted friends or family are called strong ties (Scharfer *et al.* 1981). A model of tie strength allows us to understand important questions in social networks like: 1) Who are you close to? 2) Who are your acquaintances? 3) How many times have you talked to them in the last month?

Kalimeri *et al.* discovered various social activities from human-human interaction logs captured by mobile sensing networks and then grouped them by the set of involved people, time and location of the occurring event [22].

The stronger is the tie strength, the more is the communication in a relationship and vice versa [12] [23]. Strong ties hold communities together, where as weak ties provide the bridges that make communities viable over the long term. Frequency and duration of communication has been cited as a major aspect of tie strength.

Numerous studies have examined the maximum number of close ties a person may have. A survey of 3000 Americans showed that people have an average of four strong ties, with most having between two and six. Some studies have also used proximity as a substitute for quantity of social interaction between pairs [24], signifying that communication frequency falls exponentially as members of a pair move farther apart [25]. Granovetter identified four dimensions of tie strength: duration, intimacy, intensity, and reciprocity [2].

Only a handful of studies relate social capital to individual level productivity. In this paper, we have used the mobile call logs details of 131 individuals staying in a community. We observed the communication patterns of mobile phone users and found a coupling between their interaction strength and individual productivity. We have defined strong ties as contacts with whom the participant has frequent communication and it is used as an effective proxy for the strength of a relationship.

We base our research on smart phone based call logs as they allow the users to virtually maintain a personal network diary on their phones for significantly long duration with very little or almost no effort. This passive collection of data—with one time permission—allows us to capture rich detailed information about user networks with no cognitive load or interruption to user's daily activity. It also allows the data to be objective rather than suffering from recall/cognitive/social biases in eliciting network information from comparable methods.

Table 1 shows the uniqueness of this research lies in the measure used to reach the conclusion. It also explains the related work done by researchers in terms of both social and temporal granularity. Earlier most of the work was done by collecting data from mostly surveys and very few using Bluetooth and Socio metric badges.

Table 1. Summary of related work.

Work	Temporal Granularity	Social Granularity	Measure Productivity	Measure Social Capital	Type of Network Data
Putnam	Yearly	Communities (Group/ Individual)	Survey	Surveys, city census	None
Lin	Yearly	Individual, Organization and Community	Large Scale National Surveys in Organizations and Communities and Observations	Surveys	Community
Kalimeri	Over a period of time	Communities	Surveys	Surveys	Sociometric Badges Bluetooth
Burt	Over a period of time (Yearly)	Communities (Both on Individual and Groups)	Surveys	Surveys	Interview
Granovetter	Both on daily and over a period of time	Both Individuals and Social Groups	Surveys	Surveys	Interview
This work	Daily	Individual	Surveys	Social Network Analysis	Call Logs

But in this paper, we have used phone call logs of smart phones of individuals to study human productivity both on daily basis and over a period of time. Such method of finding productivity has never been explored before.

3. Measuring Social Capital Using Call Log Data

3.1. Major aspects of Social Capital

There are two major aspects of social capital.

1) Long term component: It is the social capital which is acquired over a period of time and gives productive benefits. Here it uses the social structure as capital in its own right.

2) Short term component: It is the social capital which is acquired over a short period of time (like daily basis) and gives productive benefits. Here it describes a network as your access to people with specific resources, which creates a correlation between theirs and yours. Here the embedded resources are accessed and/or mobilized in purposive actions.

When an individual comes in contact with his community members, there will be an accumulation of social capital, which may immediately satisfy his social needs and it may be sufficient for the substantial improvement of life in the whole community (Hanifan, 1920).

Social capital is not just a single entity, but a collection of different entities having two common characteristics: they all consist of some aspect of social structures, and they facilitate certain actions of individuals who are within that structure. Social capital is productive, making possible the achievement of certain things that would not be attainable in its absence. Social capital inherits the structure of relations between persons and among persons [3].

3.2. Measuring Social Capital via Social Network Analysis

1) Size/Degree: Size/Degree is nothing but the number of alters that an ego is directly connected to possibly weighted by tie strength. From the mobile call log data it can be easily computed as the number of social contacts each ego (individual) has.

2) Effective Size: It is defined as the number of people an individual or ego is connected to, minus the redundancy in the network, that is, it reduces to the unique connections of the network.

Effective size = Size – Redundancy

It gives the count of number of non redundant ties (unique contacts) that the ego has.

3) Heterogeneity: It is the variety of actors they are connected to alter with respect to relevant dimensions. Social capital is more from stronger homogeneous groups because of bonding social capital, where as building connections between heterogeneous groups are likely to be more fragile. We have not considered this feature as our mobile call log data is from a community which is a homogeneous group. We have not considered this feature as our mobile call log data contains bonding social capital rather than bridging social capital.

4) Betweenness Centrality: The number of times that ego occupies a between position on the geodesics connecting many pairs of other actors in the network. Normally in the shortest path connecting two other nodes, a between actor might control the flow of information or the exchange of resources.

5) Eigenvector centrality: This approach is an effort to find the most central actors in terms of the global or overall structure of the network, and pay not as much of attention to patterns that are more local. Here factor analysis is used to identify dimensions of the distances among actors, and the location of each actor with respect to each dimension is stored in an eigenvector. We have not considered this feature as eigenvector centrality is a simple reputation based centrality metric which is unsuitable with respect to our mobile call log data.

6) Density: It is defined as the sum of the values of all ties divided by the number of possible ties. That is, with valued data, density is typically defined as the average strength of ties across all possible (not all actual) ties. We have used this popular feature in our analysis because dense networks are supposed to have more social capital and it is highly suitable for our dataset.

7) Closeness Centrality: This approach emphasizes the distance of an actor to all others in the network by focusing on the distance from each actor to all others. It can be measured using path distances, reach or eigenvector centrality. We have not considered this feature as we observed betweenness centrality is more appropriate with respect to our dataset.

8) Closeness: It is the total graph theoretic distance from ego to all others in network. We have not considered this feature because the more the distance to other nodes, the less the chance of receiving information in a timely way.

9) Reciprocity: It is a measure to specifically characterize directed networks. Link reciprocity measures the tendency of vertex pairs to form mutual connections between each other. We have not considered this feature because in mobile call networks a person may not always have mutual links, so reciprocity studies may introduce errors.

10) Homophily: It is the tendency of individuals to associate and bond with similar others. In homophilic relations, individuals share common characteristics which make communication and relationship formation easier [26].

11) Tie Strength: The willingness of alters to provide resources to ego is usually equated to the intensity of social relationships or tie strength. It is used to indicate how accessible social resources are through a certain relationship.

12) Weak Ties: Mark Granovetter uncovered the strength of weak ties as more novel information flows to individuals through weak rather than strong ties. We have used this feature because of its popularity and importance with regard to our dataset.

13) Structural holes: Structural hole is defined as lack of connection between two nodes that is bridged by a broker (Burt, 1992). Structural holes measures can be computed on both valued and binary data. We have used binary data *i.e.* a relation is present or not.

a) Effective size of the network (EffSize): It is the number of alters that ego has, minus the average number of ties that each alter has to other alters.

b) Efficiency: It normalizes the effective size of ego's network by its actual size.

c) Constraint: It is a summary measure that taps the extent to which ego's connections are to others who are connected to one another.

d) Hierarchy: It describes the nature of the constraint on ego. If the total constraint on ego is concentrated on a single actor, then the hierarchy measure will have a higher value.

14) K-Core: In social network, it is defined as the (unique) maximal induced sub-graph with minimum degree at least K. Alternatively, the K-Core is the result of iteratively deleting nodes that have degree less than K, in any order.

3.3. Features Identified and Explanation

In this paper, **Table 2** shows the social network measures taken based on its popularity and suitability with respect to our dataset. Most of the features are computed using UCINET 6.0 software.

Table 3 shows two types of users. All the numbers which start with SP or FA are community users whose call logs taken and also participated in the survey and the other numbers are external or infrequent users. We have concentrated only on community users and preprocessed data of non-community users.

Social capital is defined as resources embedded in individual's social networks, resources that can be accessed or mobilized through ties in the networks (Lin 2001a: Chapter 2). Social capital should be assessed in

Table 2. List of social network features.

Social Network Features	Reason
Size/Degree	In mobile call log data, if an individual (ego) is having more social contacts then there is high chance that the individual (ego) will get the desired resources. Hence there is likely chance that an ego with higher degree may be productive.
Effective Size	We have used this feature as it is highly suitable with respect to our dataset, because a person with more effective size is considered to be having more unique social contacts and hence may be productive.
Betweenness Centrality	This is a popular feature in social network analysis and we have computed this feature from our mobile call log data. An actor with high betweenness centrality plays a powerful role in the network.
Density	We have used this popular feature in our analysis because dense networks are supposed to have more social capital and it is highly suitable for our dataset.
Tie Strength	We have used this feature because of its popularity and suitability with respect to our dataset where contact frequency and contact duration are indicators of tie strength.
Weak Ties	We have used this feature because of its popularity and importance with regard to our dataset.
Structural holes	We have used this feature because of its importance in ego networks. Ronald Burt has developed a number of measures to examine the position of each actor in their neighborhood.
K-Core	We have used this feature because it shows active participation of an individual in a social network if his/her at least K friends are also actively participating. This feature is intuitively appealing in our application.

Table 3. Sample dataset.

SOURCE	DESTINATION	DATE AND TIME OF CALL	TYPE OF CALL
FA10-01-12	FA10-01-11	11/11/2010 0:19	INCOMING
FA10-01-11	FA10-01-12	11/11/2010 0:19	OUTGOING
FA10-01-11	FA10-01-12	11/11/2010 0:20	OUTGOING
FA10-01-12	FA10-01-11	11/11/2010 0:20	INCOMING
FA10-01-11	SP10-01-21	11/11/2010 0:20	OUTGOING
SP10-01-21	FA10-01-11	11/11/2010 0:20	INCOMING
FA10-01-07	2013346763	11/11/2010 0:47	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:48	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:48	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:49	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:50	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:51	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:52	OUTGOING
FA10-01-07	2013346763	11/11/2010 0:53	OUTGOING
SP10-01-25	6173195257	11/11/2010 2:50	OUTGOING
SP10-01-52	8883002305	11/11/2010 4:42	OUTGOING
SP10-01-52	8883002305	11/11/2010 4:43	OUTGOING
SP10-01-52	8883002305	11/11/2010 4:43	OUTGOING
SP10-01-52	8883002305	11/11/2010 4:44	OUTGOING
SP10-01-52	8883002305	11/11/2010 4:45	OUTGOING
SP10-01-53	2549996859	11/11/2010 5:45	INCOMING
SP10-01-33	8886613502	11/11/2010 7:10	OUTGOING
SP10-01-33	7236427232	11/11/2010 7:12	INCOMING
FA10-01-58	8056377243	11/11/2010 7:31	OUTGOING
SP10-01-11	FA10-01-18	11/11/2010 7:40	INCOMING

terms of two parameters; one is embeddedness of resources and its accessibility (mobilization). The more the degree of access to embedded resources, the better the return to the actor. Mobilized social capital helps an actor in their career development and it imitates the usage of a particular tie (strong/weak tie) and its resources in production. It reveals the actual process of linkage between capital and its attainment [23].

To measure the effectiveness of mobilization, we have explored an individual's contact with others and their frequency of communication. In this paper, the mobilized social capital has exposed consistent results—the frequency of communication affects the individual's productivity. Lai 1998 gave the joint effects of accessible and mobilized social capital. According to Lie, neither mobilization nor accessible capital theories can operate totally independently of one another.

To elucidate the difference between embeddedness and accessibility of resources, we have divided our experimentation work into two categories. We have conducted our experimentation both on daily basis and also over a period of time. Some of the social network features like in-degree, out-degree, tie strength, strong and weak tie percentage etc. we have computed daily basis and correlated with their average productivity scores of an individual. And some other social network features like betweenness centrality, k-core, density, structural hole etc. we have computed for the overall period and correlated with their average productivity.

4. Hypothesis Generation

We have analyzed our dataset both on daily basis and also over a period of time and we have formulated the following hypothesis. **H1** represents newly formulated hypothesis and **H0** represents null hypothesis.

H1: A person having more contacts in a social network will be more productive.

H0: Persons social contacts are not related to productivity.

People during interaction learn from each other. Social capital contributes to productivity through interactions between people [13]. The association between people's social contacts and their productivity is mentioned in Robert D. Putnam's national bestseller "**Bowling Alone: The collapse & Revival of American Community**". Hence, we hypothesize that people with more social contacts have higher productivity.

To test the hypothesis, we analyzed the correlation between overall degree (overall basis) of nodes/actors with their average productivity scores. We calculated $r\text{-value} = 0.2541148$ and p-value from Pearson (r) score which yielded a $p\text{-value} = 0.006136$ [confidence level = **0.95**, significance level $\alpha = 0.05$] and we found that $p\text{-value} < \alpha$, it indicates that the observed data is inconsistent with the assumption that the null hypothesis is true; hence the null hypothesis must be rejected.

H2: People with more betweenness centrality will have higher productivity than people with less betweenness centrality.

H0: Person's betweenness centrality score doesn't improve their productivity.

Betweenness centrality is an indicator of a node's centrality in a network. A node with high betweenness centrality has a large influence on the transfer of resources through the network, under the assumption that resource transfer follows the shortest path. According to Linton Freeman, a person with higher betweenness centrality score is more powerful in a communication network.

The Bavelas article, "**A Mathematical Model for Group Structures**" (1948) remains one of the high water marks in the social network literature. He showed also that the structural location of an actor in a network makes a difference, as the actors view themselves, and the group task, differently depending on whether they are at the center or the periphery of the small group communication structure.

To test the hypothesis, we analyzed the correlation between Freeman's Betweenness centrality scores of nodes/actors with their average productivity scores. We calculated $r\text{-value} = 0.07517343$ and p-value from Pearson (r) score which yielded a $p\text{-value} = 0.4246$ [confidence level = **0.95**, significance level $\alpha = 0.05$] and we found that $p\text{-value} > \alpha$, it indicates that the observed data is inconsistent with the assumed hypothesis; hence the null hypothesis is true and must be accepted.

H3: Individuals in cohesive subgroups engage in reciprocity transactions, hence it improves the productive benefits of individuals within the group.

H0: Cohesive subgroups contains more redundant information, hence it produces less productive individuals.

From the social-psychological perspective, individuals are most strongly influenced by members of their primary groups *i.e.* people with whom they engage in frequent interactions (Cooley 1909, Epstein 1961, Festinger,

Schachter, and Back 1950, Kadushin 1966) and anthropologists have argued that primary groups are integral to understanding people within the context of their communities (Barnes 1972, Bott 1971). Actors are more likely to enforce each others' trust and engage in reciprocity transactions, if they are members of the same subgroup. In terms of social structure, subgroup boundary affects the types of actions in which actors engage.

To test the hypothesis, we analyzed the correlation between K-Core of nodes/actors with their average productivity scores. We calculated *r-value* = **0.237107** and p-value from Pearson (r) score which yielded a *p-value* = **0.01073** [confidence level = **0.95**, significance level α = **0.05**] and we found that p-value < α , it indicates that the observed data is inconsistent with the null hypothesis; hence the assumed hypothesis is true and must be accepted.

H4: Denser networks are more likely to promote the sharing of resources hence produces more productive individuals.

H0: Dense networks have redundant information hence produces less productive individuals.

Network closure or density is a distinctive advantage of social capital because it maintains and enhances trust, norms, authority and sanctions. These may ensure that individuals can mobilize network resources. In dense networks, resources can be preserved and reproduced (Bourdieu 1986), but on the other hand, if a person is looking for a job or better job, then bridges in the network will be more useful (Lin 1999, Marsden and Hurlbert 1988, De Graaf and Flap 1988, Burt 1992).

To test the hypothesis, we analyzed the correlation between densities of nodes/actors with their average productivity scores. We calculated *r-value* = **-0.02364713** and p-value from Pearson (r) score which yielded a *p-value* = **0.834** [confidence level = **0.95**, significance level α = **0.05**] and we found that p-value > α , it indicates that the observed data is inconsistent with the assumed hypothesis; hence the null hypothesis is true and must be accepted.

H5: Effective size of a person shows its strength of connection with others, hence improves productivity.

H0: Effective size of a person has no relation with their productivity.

Conceptually the effective size is the number of people ego is connected to, minus the redundancy in the network, that is, it reduces to the non-redundant elements of the network. In other words, it gives the count of number of non redundant ties (unique contacts) that each ego has. It is an important measure of ego-centric network and a part of structural hole measurement proposed by Burt (1992). The effective size (non-redundant ties) of the network of a subject i is calculated as

$$NR_i = \sum_j (1 - \sum_q p_{iq} m_{jq}) \text{ with } q \neq i, j.$$

To test the hypothesis, we analyzed the correlation between effective sizes of each node with their average productivity scores. We calculated *r-value* = **0.2572613** and p-value from Pearson (r) score which yielded a *p-value* = **0.005511** [confidence level = **0.95**, significance level α = **0.05**] and we found that p-value < α , it indicates that the observed data is inconsistent with the null hypothesis; hence the assumed hypothesis is true and must be accepted.

H6: The greater the efficiency of an ego network, the greater the role of broker an individual plays, hence individual productivity improves.

H0: Efficiency of an ego network plays no role in improving individual productivity.

The measure of the efficiency of an ego network is very closely linked to its effective size. It can be defined as

$$ER_i = NR_i / N_i$$

where ER_i is the measure of efficiency for a subject i , NR_i is the effective size of i 's network, N_i is the no of ties. This measure is positive indicator of the presence of structural holes within social networks.

To test the hypothesis, we analyzed the correlation between efficiency of each node with their average productivity scores. We calculated *r-value* = **-0.1919905** and p-value from Pearson (r) score which yielded a *p-value* = **0.04072** [confidence level = **0.95**, significance level α = **0.05**] and we found that p-value < α , it indicates that the observed data is inconsistent with the null hypothesis; hence the assumed hypothesis is true and must be accepted.

H7: The constraint measure the level of closure (influence) of an individual within its support network, hence exerts more productivity.

H0: Constraint expresses less degree of involvement of an individual with others, hence reduces individual productivity.

Constraint expresses each individual's degree of involvement with other directly connected members in the network. It offers the level of closure of ego within its support network. The constraint C_i that influences a subject i is

$$C_i = \sum_j c_{ij}$$

where c_{ij} expresses the amount of constraint exerted on i by j and is:

$$c_{ij} = \left(p_{ij} + \sum_q p_{iq} p_{qj} \right)^2 \text{ with } q \neq i, j.$$

To test the hypothesis, we analyzed the correlation between constraints of each node with their average productivity scores. We calculated $r\text{-value} = -0.2812441$ and p-value from Pearson (r) score which yielded a $p\text{-value} = 0.002436$ [confidence level = **0.95**, significance level $\alpha = 0.05$] and we found that $p\text{-value} < \alpha$, it indicates that the observed data is inconsistent with the null hypothesis; hence the assumed hypothesis is true and must be accepted.

H8: An individual's hierarchical position enables him/her to obtain more opportunities and helps to make best choices, hence it improves individual productivity.

H0: An individual's hierarchical position in the network has nothing to do with productivity.

Lin's theory of social capital says that the more you do investment in social relations, the more chances of gaining benefits through the shared resources. The embedded resources in social networks help to improve the productivity through flow of information within the community. And the strategic and hierarchical position of an individual obtains more useful information about the opportunities and able to take effective decisions.

To test the hypothesis, we analyzed the correlation between hierarchies of each node with their average productivity scores. We calculated $r\text{-value} = -0.3120463$ and p-value from Pearson (r) score which yielded a $p\text{-value} = 0.0007253$ [confidence level = **0.95**, significance level $\alpha = 0.05$] and we found that $p\text{-value} < \alpha$, it indicates that the observed data is inconsistent with the null hypothesis; hence the assumed hypothesis is true and must be accepted.

H9: Structural hole plays a major role in individual's career development and advancement, hence makes an individual productive.

H0: Structural hole plays least role in making an individual productive.

An actor with a network rich in structural holes can add value to a community through new ideas and opportunities. This in turn, helps an individual's career development and advancement. Structural hole is used for the separation between non-redundant contacts. Non-redundant contacts are connected by a structural hole. A structural hole is a relationship of non-redundancy between two contacts. Actors who develop ties with disconnected groups gain access to a broader array of ideas and opportunities than those who are restricted to a single one (Granovetter 1973). Teams which draw members from diverse demographic categories benefit because such teams generate links between people with different skills, information, and experience. Such ties within the team bridge structural holes in the larger organization, and thereby enhance its capacity for creative action.

To test the hypothesis, we analyzed the correlation between average productivity scores and structural hole [Effective Size, Efficiency, Constraint, Hierarchy]. We calculated $r\text{-value} = 0.2572613$ (*Effective Size*), $r\text{-value} = -0.1919905$ (*Efficiency*), $r\text{-value} = -0.2812441$ (*Constraint*) and $r\text{-value} = -0.3120463$ (*Hierarchy*) and p-value from Pearson (r) score which yielded $p\text{-value} = 0.005511$ (*Effective Size*), $p\text{-value} = 0.04072$ (*Efficiency*), $p\text{-value} = 0.002436$ (*Constraint*) and $p\text{-value} = 0.0007253$ (*Hierarchy*) [confidence level = **0.95**, significance level $\alpha = 0.05$]. We found that all the p-value is lower than the significance level α , it indicates that the observed data is inconsistent with the null hypothesis; hence the assumed hypothesis is true and must be accepted.

5. Testing with Data

Our study is based on mobile phone call logs dataset. The data consists of 111,444 voice calls between 131 adult members of a young family living community for a period of 3 months *i.e.* between 11 November, 2010 to 21 March, 2011. The dataset contains continuous collection of their call logs including the date, time and duration of call. Here the call type is incoming, outgoing or missed between two individuals. This data helped us to build

a phone communication network of the community where each node is an individual person (actor) and links are the type of calls made by them.

In addition to this, we have used the survey data which contains daily polling of happiness, stress, productivity, healthy eating, sleep and hangout related information which is manually entered data by the participants.

In this paper, we have performed Pearson correlation analysis to analyze the correlation between average productivity score of an individual (collected from survey data) with their long term (overall period) and short term (daily basis) social network features. From the results, we have formulated a set of hypothesis which is given in the **Table 4**. **Figure 1** shows correlation result between indegree, outdegree with their average productivity.

5.1. Combining Features into a Predictive Model

In this paper, multiple associations were found to be true, so we used Attribute-Selected Random Forest classifier with 66 - 34 split in Weka. This yielded 65.79% accuracy or a 0.735 ROC area. Given that the baseline

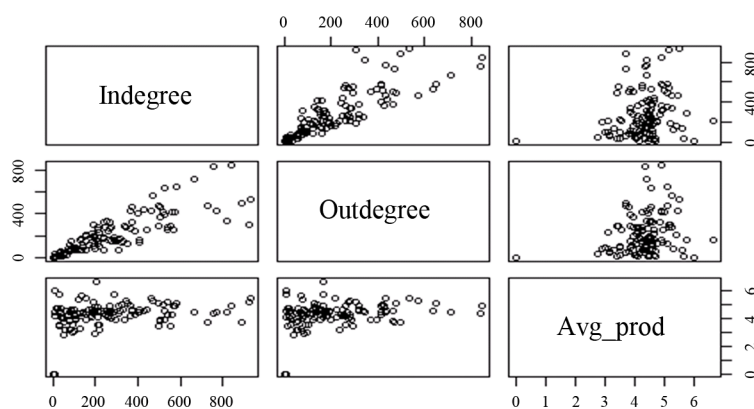


Figure 1. Correlation result between indegree, outdegree vs average productivity.

Table 4. Summary of hypothesis results.

Hypothesis No	Hypothesis: Granularity	Hypothesis: Description	Result: Significant	Direction
H1	Overall Period	A person having more contacts in a social network will be more productive.	Yes	Positive
H2	Overall Period	People with more betweenness centrality will have higher productivity than people with less betweenness centrality.	No	Positive
H3	Overall Period	Individuals in cohesive subgroups engage in reciprocity transactions, hence it improves the productive benefits of individuals within the group.	Yes	Positive
H4	Overall Period	Denser networks are more likely to promote the sharing of resources hence produces more productive individuals.	No	Negative
H5	Overall Period	Effective size of a person shows its strength of connection with others, hence improves productivity.	Yes	Positive
H6	Overall Period	The greater the efficiency of an ego network, the greater the role of broker an individual plays, hence individual productivity improves.	Yes	Negative
H7	Overall Period	The constraint measure the level of closure (influence) of an individual within its support network, hence exerts more productivity.	Yes	Positive
H8	Overall Period	An individual's hierarchical position enables him/her to obtain more opportunities and helps to make best choices, hence it improves individual productivity.	Yes	Negative
H9	Overall Period	Structural hole plays a major role in individual's career development and advancement, hence makes an individual productive.	Yes	Positive
H1	Daily Basis	The number of strong ties is positively related to productivity.	Yes	Negative
H2	Daily Basis	An individual with high indegree and outdegree is considered to be influential, hence productive.	Yes	Positive

model for classification (Zero-R) was able to obtain 50.45% accuracy and 0.50 ROC area, this marks a relative improvement of 30.38% in terms of accuracy and 47% in terms of ROC area.

==== Evaluation on test split ====

==== Summary ====

Correctly Classified Instances	25	65.7895%
Incorrectly Classified Instances	13	34.2105%
Kappa statistic	0.2882	
Mean absolute error	0.3554	
Root mean squared error	0.4635	
Relative absolute error		70.7897%
Root relative squared error		92.2395%
Total Number of Instances	38	

==== Detailed Accuracy By Class ====

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.81		0.529	0.654	0.81	0.723	0.735 0
0.471		0.19	0.667	0.471	0.552	0.735 1
Weighted Avg.		0.658	0.378	0.66	0.658	0.647 0.735

We ran a similar Linear Regression Model for the numeric values. It shows that the holistic model is significantly predictive and it can predict 26.5% of the variance in the productivity level of the individuals.

5.2. Correlation Results

```
>cor.test(d$Avgprod,d$Indegree, method = "pearson")
Pearson's product-moment correlation
data: d$Avgprod and d$Indegree
t = 3.5259, df = 123, p-value = 0.0005937
alternative hypothesis: true correlation is not equal to 0
95 percent con_dence interval:
0.1345290 0.4544093
sample estimates:
cor
0.3029788

>cor.test(d$Avgprod,d$Outdegree, method = "pearson")
Pearson's product-moment correlation
data: d$Avgprod and d$Outdegree
t = 3.5144, df = 123, p-value = 0.0006177
alternative hypothesis: true correlation is not equal to 0
95 percent con_dence interval:
0.1335543 0.4536214
sample estimates:
cor
0.3020771

>cor.test(d$Avgprod,d$strongtiescallper, method = "pearson")
Pearson's product-moment correlation
data: d$Avgprod and d$callper
t = -S4.3112, df = 4764, p-value = 1.656e-05 (0.00001656)
alternative hypothesis: true correlation is not equal to 0
95 percent con_dence interval:
-0.09057168 -0.03400882
sample estimates:
```

cor
-0.06234031

6. Conclusion

Every social network has value and individual's social contacts influence the productivity of individuals and groups [10]. The uniqueness/newness of this paper lies in the selection of reality mining dataset (*i.e.* call log dataset), which is collected from the Smartphone's of a group of community users, because of the popularity of Smartphone's as a SNA tool to maintain social contacts. We have identified and tested a number of important features of social network measures, which have impact on productivity. The hypothesis testing results show that social network measures like degree, structural hole (effective size, efficiency, constraint and hierarchy) and K-core have influence on individual's productivity over a period of time, whereas features like betweenness centrality and density have negative influence on their productivity. In the similar manner, indegree, outdegree and strong ties have weak significance (or negative correlation) on a daily basis, but surprisingly, it shows positive significance with their productivity over a period of time. This work can be extended to correlate social network measures with other human behaviors.

7. Inferences

- 1) The long term associations/contacts with friends/colleagues help an individual to improve their productivity.
- 2) Every individual's strength of relationship with others helps them to improve their productivity.
- 3) An individual who acts as a broker across structural holes adds more value than others, hence more productive than others.

References

- [1] Backstrom, L., Sun, E. and Marlow, C. (2010) Find Me If You Can: Improving Geographical Prediction with Social and Spatial Proximity. *Proceedings of the 19th International Conference on WWW*, Raleigh, 26-30 April 2010, 61-70. <http://dx.doi.org/10.1145/1772690.1772698>
- [2] Hidalgo, C.A. and Rodriguez-Sickert, C. (2008) The Dynamics of a Mobile Phone Network. *Physica A: Statistical Mechanics and Its Applications*, **387**, 3017-3024.
- [3] Lavelle, B., Byrne, D., Jones, G.J. and Smeaton, A.F. (2007) Bluetooth Friendly Names: Bringing Classic HCI Questions into the Mobile Space. *Proceedings of the 21st British HCI Group Annual Conference on People and Computers*, **2**, 63-66.
- [4] Do, T.M.T. and Gatica-Perez, D. (2013) Human Interaction Discovery in Smartphone Proximity Networks. *Personal and Ubiquitous Computing*, **17**, 413-431. <http://dx.doi.org/10.1007/s00779-011-0489-7>
- [5] Cattuto, C., Broeck, W.V.D., Barrat, A., Colizza, V., Pinton, J.F. and Vespignani, A. (2010) Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLoS One*, **5**, e11596. <http://dx.doi.org/10.1371/journal.pone.0011596>
- [6] Cristani, M., Raghavendra, R., Bue, A.D. and Murino, V. (2013) Human Behavior Analysis in Video Surveillance: A Social Signal Processing Perspective. *Neurocomputing*, **100**, 86-97. <http://dx.doi.org/10.1016/j.neucom.2011.12.038>
- [7] Eagle, N. and Pentland, A. (2006) Reality Mining: Sensing Complex Social Systems. *Personal and Ubiquitous Computing*, **10**, 255-268. <http://dx.doi.org/10.1007/s00779-005-0046-3>
- [8] Song, Y., Morency, L.P. and Davis, R. (2012) Multimodal Human Behavior Analysis: Learning Correlation and Interaction across Modalities. *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, 27-30. <http://dx.doi.org/10.1145/2388676.2388684>
- [9] Wasserman, S. and Faust, K. (1994) *Social Network Analysis: Methods and Applications*. Vol. 506, Cambridge University Press, Cambridge. <http://dx.doi.org/10.1017/CBO9780511815478>
- [10] Fu, Q. (2004) Trust, Social Capital, and Organizational Effectiveness. Major Paper submitted to the Faculty of the Virginia Polytechnic Institute and State University.
- [11] Roberts, S.G.B. and Dunbar, R.I.M. (2011) Communication in Social Networks: Effects of Kinship, Network Size, and Emotional Closeness. *Personal Relationships*, **18**, 439-452. <http://dx.doi.org/10.1111/j.1475-6811.2010.01310.x>
- [12] Granovetter, M.S. (1973) The Strength of Weak Ties. *American Journal of Sociology*, **78**, 1360-1380. <http://dx.doi.org/10.1086/225469>

- [13] Greve, A., Benassi, M. and Sti, A.D. (2010) Exploring the Contributions of Human and Social Capital to Productivity. *International Review of Sociology: Revue Internationale de Sociologie*, **20**, Issue 1.
- [14] Adler, P.S. and Kwon, S.W. (2002) Social Capital: Prospects for a New Concept. *The Academy of Management Review*, **27**, 17-40.
- [15] Lin, N. (1999) Building a Network Theory of Social Capital. *Connections*, **22**, 28-51.
- [16] Putnum, R.D. (2001) Social Capital: Measurement and Consequences. *Canadian Journal of Policy Research*, **2**, 41-51.
- [17] Pennings, J.M., Lee, K. and Van Witteloostuijn, A. (1998) Human Capital, Social Capital, Firm Dissolution. *Academy of Management Journal*, **41**, 425-440. <http://dx.doi.org/10.2307/257082>
- [18] Coleman, J.S. (1988) Social Capital in the Creation of Human Capital. *American Journal of Sociology*, **94**, S95-S120. <http://dx.doi.org/10.1086/228943>
- [19] Krackhardt, D. (1992) The Strength of Strong Ties: The Importance of Philos in Organizations. *Journal of Networks and Organizations: Structure, Form, and Action*, **216**, 239.
- [20] Burt, R.S. (2000) The Network Structure of Social Capital. *Research in Organizational Behavior*, **22**, 345-423. [http://dx.doi.org/10.1016/S0191-3085\(00\)22009-1](http://dx.doi.org/10.1016/S0191-3085(00)22009-1)
- [21] Granovetter, M. (1983) The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, **1**, 201-233. <http://dx.doi.org/10.2307/202051>
- [22] Trinh Minh, T.D., Kalimeri, K., Lepri, B., Pianesi, F. and Gatica-Perez, D. (2013) Inferring Social Activities with Mobile Sensor Networks. *ACM International Conference on Multimodal Interaction (ICMI)*, Sydney.
- [23] Marsden, P.V. and Campbell, K.E. (1983) Measuring Tie Strength. *Social Forces*, **63**, 482-501. <http://dx.doi.org/10.1093/sf/63.2.482>
- [24] Festinger, L. (1950) Informal Social Communication. *Psychological Review*, **57**, 271-282. <http://dx.doi.org/10.1037/h0056932>
- [25] Zipf, K.G. (1949) Human Behavior and the Principle of Least Effort. Addison-Wesley Press, Cambridge.
- [26] Christakis, N.A. and Fowler, J.H. (2009) Connected: The Surprising Power of Our Social Networks and How They Shape our Lives. Little, Brown and Company.