This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE CCNC 2010 proceedings

Mobile Social Closeness and Communication Patterns

Santi Phithakkitnukoon and Ram Dantu
Department of Computer Science & Engineering
University of North Texas
Denton, TX 76203, USA
Email: {santi, rdantu}@unt.edu

***Abstract*—As mobile networks expand rapidly to facilitate the rising number of mobile phone population, more mobile social services are being developed and offered. To create an efficient social functionality, characteristics of mobile social network must be studied. Social closeness is one of the basic fundamentals of any kind of social networking. In this paper, closeness in mobile social network is the subject of the study, from which social grouping scheme is proposed and validated against the feedbacks of human subjects. Based on the proposed grouping scheme, a study of the impact of the mobile social closeness to the similarity in calling patterns and reciprocity is presented.**

I. Introduction

With a rapidly growing population of mobile device users, more mobile social services are expected. Thus, research and development in mobile social computing have intensified. We believe that understanding the mobile social network is a first and essential step towards creating an intelligent functionality that enhances quality of life through a system that comprehends behavior and context of its user(s). In this paper, we quantify mobile social closeness and study its correlation to calling patterns and reciprocity. To our knowledge, no scientific research has been reported in mobile social closeness and grouping.

The rest of the paper is organized as follows. Section 2 presents a model for computing mobile social closeness and a grouping scheme with dataset description and validation of the model. Section 3 describes a method for characterizing calling patterns and examines the similarity in calling patterns and reciprocity, with respect to the social tie. Section 4 briefly reveals some related work. The paper is concluded in section 5 with a summary and an outlook on future work.

II. Mobile Social Closeness and Grouping

In social science, the social closeness of people has been discussed and found that it can be based on the amount of time and the intensity (frequency) of communication [1][2]. Granovetter [1] suggests that the time spent in a relationship and the intensity along with the intimacy and reciprocal services form a set of indicators for social tie. Marsden and Cambell [2] evaluate the indicators and predicators of strength (tie) described by Granovetter [1] and conclude that "social closeness" or "intensity" provides the best indicator of strength or tie.

In mobile social network, the amount of time and the intensity of communication can be measured by the call duration (talk time) and the call frequency (number of phone calls).

In our daily life, we communicate with people in the mobile network at different instances. These people constitute our mobile social network. Based on amount of time and intensity of communication with these people, our mobile social network can be divided into three broad groups:

Group 1: Socially Closest Members – These are the people with whom we maintain the highest socially connectivity. Most of the calls we receive, come from individuals within this category. We receive more calls from them and we tend to talk with them for longer periods. Typically, the face-to-face social tie of these people is family member, friend, and colleagues.

Group 2: Socially Near Members – People in this group are not as highly connected as family members and friends, but when we connect to them, we talk to them for considerably longer periods. Mostly, we observe intermittent frequency of calls from these people. These people are typically neighbors and distant relatives.

Group 3: Socially Distant Members – These individuals have less connection with our social life. These people call us with less frequency. We acknowledge them rarely. Among these would be, for example, a newsletter group or a private organization with whom we have previously subscribed. This group also includes individuals who have no previous interaction or communication with us. We have the least tolerance for calls from them *e.g.,* strangers, telemarketers, fund raisers.

We quantitatively define the social closeness between user i and user j from the user i's perception ($S(i,j)$) by Eq. 1.

$$S(i,j) = \sqrt{(1-F(i,j))^2 + (1-T(i,j))^2}, \qquad (1)$$

where $F(i,j)$ is the normalized call frequency (normalized to the maximum call frequency among all users with whom user i communicate) between user i and user j, which is given by Eq. (2), and $T(i,j)$ is the normalized call duration or talk time (normalized to the maximum talk time among all users with whom user i communicate) between user i and user j, which is given by Eq. (3).

$$F(i,j) = \frac{f(i,j)}{\max\limits_{k \in U_i}\{f(i,k)\}}, \qquad (2)$$

978-1-4244-5176-0/10/$26.00 ©2010 IEEE

$$T(i,j) = \frac{t(i,j)}{\max\limits_{k \in U_i}\{t(i,k)\}}, \quad (3)$$

where $f(i,j)$ is the total number of calls or call frequency between user i and user j, $t(i,j)$ is the total call duration or talk time between user i and user j, and $U_i = \{1, 2, ..., N\}$ is the set of all users associated with user i (*i.e.,* all users who have made/received calls to/from user i with total of N users).

Therefore, $S(i,j)$ has values in the range $[0, \sqrt{2}]$, which indicates the mobile social closeness between user i and user j from user i's perspective where 0 implies the closest and $\sqrt{2}$ implies the farthest relation. Based on this quantity, we can categorize all users associated with user i into three social groups using a simple grouping algorithm as follows.

Let R denote the Euclidean distance from coordinate (μ_F, μ_T) to (1,1) where μ_F and μ_T are the means of $F(i,j)$ and $T(i,j)$, respectively and $j \in U_i$. If $S(i,j) \leq R/2$, then user j belongs to Group 1, if $R \geq S(i,j) > R/2$, then user j belongs to Group 2, and if $S(i,j) > R$, then user j belongs to Group 3. Graphical representation is shown in Fig. 1.

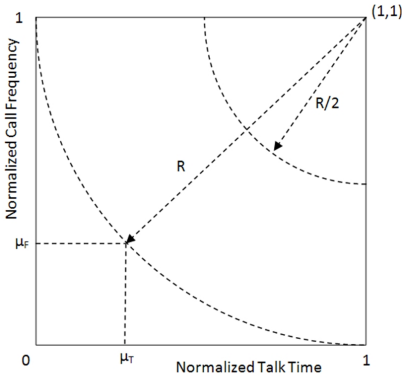

Fig. 1. Graphical representation for identifying boundaries of mobile social groups.

As social closeness and social group are defined according to the perception of user i, therefore using analogy of the circle, user i can be referred to as a *Center User*, where the distance from the center of the circle (Center User) represents the closeness of social relationship to other *Associated Users*.

Property 1. *Social closeness and social group are not symmetric, i.e., $S(i,j)$ may or may not equal to $S(j,i)$ and thus $G(i,j)$ may or may not equal to $G(j,i)$.*

Social group is based on social closeness which is measured by amount of time and intensity of communication between the Center User and Associated User. Social closeness is computed according to the Center User's perception of each Associated User compared to all other Associated Users. Since different Center Users may have different Associated Users with different amount of time and intensity of communication, thus social closeness is not symmetric. For example, user j is perceived by user i as a member of group 1, however user i is perceived as a member of group 2 of user j since user j has other Associated Users to whom user j communicate more than user i.

From Eq. (1), Eq. (2), and Eq. (3), $S(i,j)$ can be defined as

$$S(i,j) = \sqrt{(1 - \frac{f(i,j)}{\max\limits_{k \in U_i}\{f(i,k)\}})^2 + (1 - \frac{t(i,j)}{\max\limits_{k \in U_i}\{t(i,k)\}})^2}, \quad (4)$$

and $S(j,i)$ can be defined as

$$S(j,i) = \sqrt{(1 - \frac{f(j,i)}{\max\limits_{m \in U_j}\{f(j,m)\}})^2 + (1 - \frac{t(j,i)}{\max\limits_{m \in U_j}\{t(j,m)\}})^2}. \quad (5)$$

Since $f(i,j) = f(j,i)$ and $t(i,j) = t(j,i)$, Eq. (5) can be rewritten as

$$S(j,i) = \sqrt{(1 - \frac{f(i,j)}{\max\limits_{m \in U_j}\{f(j,m)\}})^2 + (1 - \frac{t(i,j)}{\max\limits_{m \in U_j}\{t(j,m)\}})^2}. \quad (6)$$

If social closeness is symmetric, *i.e.,* $S(i,j) = S(j,i)$, then

$$\sqrt{(1 - \frac{f(i,j)}{\max\limits_{k \in U_i}\{f(i,k)\}})^2 + (1 - \frac{t(i,j)}{\max\limits_{k \in U_i}\{t(i,k)\}})^2} = \sqrt{(1 - \frac{f(i,j)}{\max\limits_{m \in U_j}\{f(j,m)\}})^2 + (1 - \frac{t(i,j)}{\max\limits_{m \in U_j}\{t(j,m)\}})^2}, \quad (7)$$

where equality holds if and only if $\max\limits_{k \in U_i}\{f(i,k)\} = \max\limits_{m \in U_j}\{f(j,m)\}$ and $\max\limits_{k \in U_i}\{t(i,k)\} = \max\limits_{m \in U_j}\{t(j,m)\}$.

Property 2. *Social closeness and social group change over time.*

In our daily life, relationships inevitably change over time. Meeting new people with whom the closer relationships established and not keeping in touch with whom the relationships become further are part of our social life. It is inherently true in mobile social network that social closeness changes over time. Situations bring people together and take them apart. These situations can be work, school, hobby, or any event in life. As soon as the phone numbers have been exchanged or given, a new social member may arise and possibly gain closer relationship as time progresses. Thus social closeness and social group change over time.

A. Real-life Datasets

In this study, we use two sets of real-life call logs of 30 combined users with nearly 3,00 associated callers/calees and over 46,000 call activities. Our first dataset consists of three-month call logs of 20 individual mobile phone users, which were collected at University of North Texas (UNT) during summer of 2006. These 20 individuals were faculty, staff, and students. These call logs were collected as part of the Nuisance Project [3], where Kolan *et al.* [3] studied the nuisance level associated with each phone call. Our second

dataset consists of three-month call logs of ten mobile phone users, which were collected during summer of 2008 at UNT. These ten subjects were also faculty, staff, and student. In addition, during our second dataset collecting process, we interviewed the subjects about the social closeness for all of his/her Associated Users by having the subjects identifying the perceived social group for each Associated User. As the result, our second dataset includes an additional information on social group corresponding to each Associated User. The details of the data collecting process are described in [4].

As part of the data collecting process for both datasets, each user downloaded three months of detail telephone call records from his/her online accounts on the mobile phone service provider's website. Each call record in the dataset had 5-tuple information; Date, Start time, Type (Incoming or Outgoing), User ID, and Talk time.

B. Validation of Social Grouping

To validate the accuracy of our social closeness/group computation, we use the second set of our data, which contains social group information. *We are able identify social groups correctly with the overall accuracy rate of 93.8%*. The detailed result is shown in Table I, which presents the number of correct classification (Hit), the number of incorrect classification ($Miss$), and the accuracy rate ($\frac{Hit}{Hit+Miss}$) for each Center User.

Based on the follow-up interviews with these ten subjects, most of "$Miss$" are caused by confusion between the face-to-face social closeness and mobile social closeness. For example, one of the subjects indentifies his roommate as a group 1 member but since the subject sees his roommate quite often thus the subject does not make/receive many phone calls to/from him. As the result, his roommate is classified to group 2 based on our calculation (Eq. (1)) but identified as group 1 member by the subject. To avoid the biased feedbacks from the subjects, we do not provide any information about our social closeness computation or much more details about the three social groups than the description provided earlier in this section. Nevertheless, we believe that we have a good result in accuracy rate and, in addition, we do not have a single incorrect classification that misses more than one level of social group.

Furthermore, as stated by Property 2 that social relationships change over time. With our real-life datasets, we thus further experimentally validate Property 2 by showing an example of an actual social-group plot of a randomly selected Center User from our datasets in Fig. 2, from which we can see that the Associated User 8 used to be a member in group 1 (Fig. 2(a)) but as time progresses, he/she has changed calling behavior towards the Center User (or the Center User changes his/her calling behavior towards the Associated User 8) by which furthers relationship apart and leads the user 8 to become a member of group 2 at 30 days later (Fig. 2(b)).

TABLE I
VALIDATION OF SOCIAL GROUPING

User	Hit	$Miss$	Accuracy Rate (%)
1	60	5	92.31
2	57	6	90.48
3	48	5	90.57
4	141	13	91.56
5	127	8	94.07
6	188	11	94.47
7	88	3	96.70
8	80	6	93.02
9	62	1	98.41
10	87	4	95.60
Overall	938	62	93.80
Mean	93.80	6.20	93.72
Std. Dev.	44.82	3.61	2.64

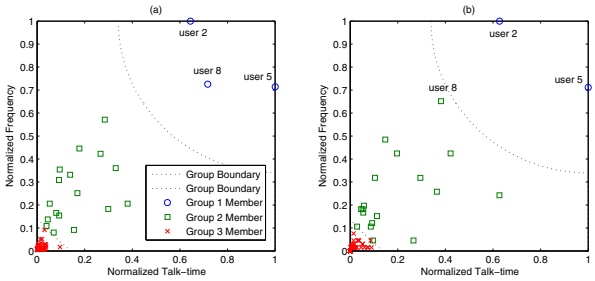

Fig. 2. (a) Social relationship at time T and (b) social relationship at 30 days later ($T + 30$)

III. SIMILARITY IN CALLING PATTERNS

Daily, we receive calls from our social group members. Every person exhibits a unique calling pattern. We have shown in our previous work [5] that the calling patterns (of Associated Users) can be characterized by arrival time of the calls using kernel density estimator. We believe that not only the calling pattern of each Associated User is unique but the calling pattern from Center User to each Associated User is also unique. We also believe that there exists some similarity in calling patterns between an Associated User and the Center User, which can lead to drawing a correlation between similarity in calling patterns and social closeness.

Calling pattern from user *i* to user *j* can be represented by Gaussian kernel estimation as

$$C_{i,j}(t) = f(h_{i,j}[n]), \tag{8}$$

where $f(\cdot)$ is the Gaussian kernel estimator, and $h_{i,j}[n]$ is a histogram function of arrival time of calls from user i to user j where $n = \{1, 2, 3, \ldots, 24\}$ is the hour slot.

For a given Center User i, Fig. 3 and Fig. 4 show three different calling pattern pairs; Fig. 4(a) shows calling pattern from Center User i to Associated User a (outgoing calls to a), $C_{i,a}(t)$ and calling pattern from Associated User a to Center User i (incoming calls from a), $C_{a,i}(t)$ where user a is member of social group 1, Fig. 4(b) shows $C_{i,b}(t)$ and $C_{b,i}(t)$ where user b is a member of group 2, and Fig. 4(c) shows $C_{i,c}(t)$ and $C_{c,i}(t)$ where user c belongs to group 3. By visual inspection, one can observe that there is more similarity

in calling patterns between Center User i and member of group 1 than Center User i and member of group 2, and even more as to compare to similarity between user i and member of group 3.

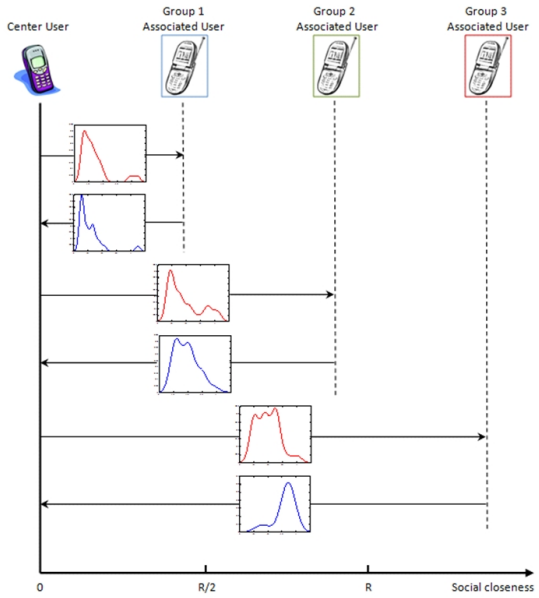

Fig. 3. Calling patterns (outgoing patterns are in red and incoming patterns are in blue, the direction of the calling pattern can also be determined by the arrow) between a Center User i and three different Associated Users who are members of social group 1, 2, and 3.

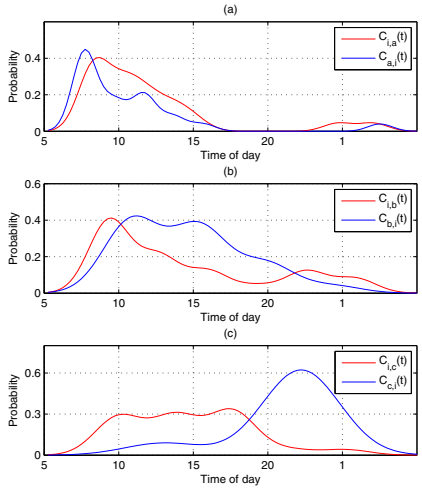

Fig. 4. Calling pattern comparisons between Center User i to (a) a member of social group 1, (b) a member of social group 2, and (c) a member of social group 3; where $C_{i,j}(t)$ is the calling pattern from user i to user j.

We compute the similarity in calling patterns between user i and user j ($Sim(i,j)$) based on Hellinger distance as follows.

$$Sim(i,j) = 1 - d_H^2(C_{i,j}(t), C_{j,i}(t)). \tag{9}$$

Based on Eq. (9), we find that $Sim(i,a)$ = 0.766, $Sim(i,b)$ = 0.452, and $Sim(i,c)$ = 0.125, which confirm our observation.

For each Center User in our datasets, we compute similarity in calling patterns ($Sim(i,j)$) and social closeness ($S(i,j)$) for all Associated Users and then find their averages for each social group. The result is shown in Fig. 5(a) where we can observe that as social closeness becomes more distant, the similarity level in calling patterns decreases. This relationship can also be estimated by a fitting curve of 5th degree polynomial. In addition, Fig. 5(b) shows the average similarity level in calling patterns for each social group. This result is consistent with result shown in Fig. 5(a) where group 1 is clustered around similarity level of 0.8, group 2 is clustered around 0.4, and group 3 clustered around 0.1.

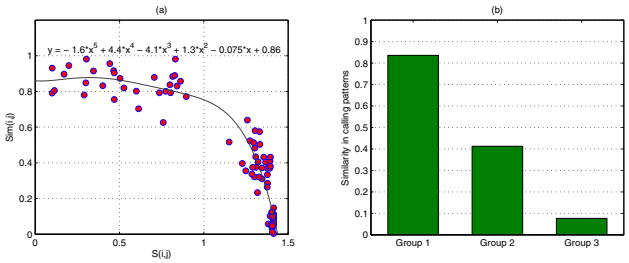

Fig. 5. (a) Similarity level in calling patterns and the corresponding social closeness, (b) Similarity level in calling patterns and the corresponding social groups.

In addition to our analysis of correlation between similarity in calling patterns and social closeness/groups, we quantify *Reciprocity* ($R(i,j)$) as a level of interaction between the Center User i and an Associated User j as follows.

$$R(i,j) = F(i,j) \cdot I(i,j), \tag{10}$$

where

$$I(i,j) = -\frac{f_{in}(i,j)}{f(i,j)} \log_2 \left(\frac{f_{in}(i,j)}{f(i,j)} \right) - \frac{f_{out}(i,j)}{f(i,j)} \log_2 \left(\frac{f_{out}(i,j)}{f(i,j)} \right), \tag{11}$$

$F(i,j)$ is the normalized call frequency defined earlier using Eq. (2), $f_{in}(i,j)$ is the total number of incoming calls from Associated User j to Center User i, $f_{out}(i,j)$ is the total number of outgoing calls from Center User i to Associated User j, and $f(i,j) = f_{in}(i,j) + f_{out}(i,j)$. $R(i,j)$ has value in the range of zero to one, where $R(i,j)$ = 1 implies the highest reciprocity (level of interaction) between user i and user j, and $R(i,j)$ = 0 implies no reciprocity.

$R(i,j)$ is a product of a normalized call frequency ($F(i,j)$) and an *Interaction Ratio* ($I(i,j)$). The $F(i,j)$ indicates a level of interaction based on total number of calls between Center User i and Associated User j with respect to all other Associated Users, $I(i,j)$ quantifies the interaction level based on the number of exchanged calls between the two users. This value lies between zero and one ([0, 1]). Fig. 6(a) depicts the graph of the function $I(i,j)$.

After computing $R(i,j)$ for all Center Users in our datasets, we find that the closer the relationship, the higher the reciprocity between the Center User and the Associated User,

Fig. 6(b) indicates that reciprocity increases as social closeness becomes stronger. The average reciprocity is 0.8442, 0.1008, and 0.0035 for groups 1, 2, and 3, respectively. These results also imply that similarity in calling patterns increases with reciprocity.

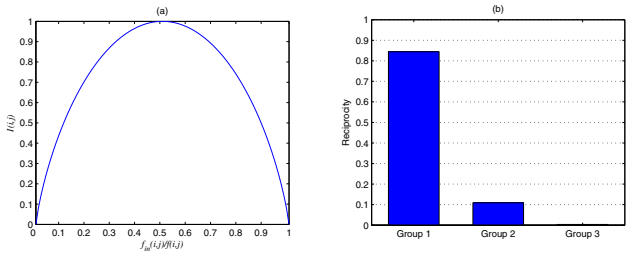

Fig. 6. (a) Graph of function $I(i,j)$ versus $\frac{f_{in}(i,j)}{f(i,j)}$, (b) Integration ratio and the social groups.

IV. Related Work

Social closeness in face-to-face social networks has been studied in psychology from which various definitions [6] [7] [8], components [9] [10], classifications of closeness [11] [12], and social support [13] have been defined.

As online social networking has rapidly become popular, online social characteristics have been extensively studied and the findings have been reported in several literatures among which discussed about social closeness in online communities [14] [15] [16]. To our knowledge, no scientific research has been reported in quantifying closeness in mobile social networks.

Mobile social closeness has been mentioned as an important component of interaction syntax for mobile social software in [17] but never formally defined. A literature that has come close to defining mobile social closeness is [18] in which the authors measured the closeness centrality for mobile phone users based on the definition proposed by Freeman in [19].

V. Conclusion and Future Work

With the ultimate goal of building intelligence into a mobile phone to enhance quality of life, the first and essential step is to understand characteristics of the mobile phone users. As the population of mobile phone users rapidly increases, the analysis of mobile phone users' characteristics must not include only the individual level but also social altitude. In this paper, we develop a mobile social grouping scheme, which allows us to further investigate a relationship between the social tie and the similarity in calling patterns. Our finding shows that the closer the social tie, the higher the similarity. In addition, we find that a closer tie also implies higher reciprocity.

As our future direction, we will continue to investigate the mobile social networks in different aspects such as mobile social group sizes and their successive ratio. There have been research studies in social group sizes and scaling ratio in sociology [20], social anthropology [21] [22], and psychology [23] [24] of face-to-face social networks but not mobile social networks. Thereby it is very motivating and important to investigate it for a better understanding of the mobile social networks and a valuable comparison to the face-to-face as well as online social networks.

Acknowledgment. This work is supported by the National Science Foundation under grants CNS-0627754, CNS-0619871 and CNS-0551694.

References

[1] M. Granovetter, "The strength of weak ties," *The American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.
[2] P. V. Marsden and K. E. Campbell, "Measuring tie strength," *Social Forces*, vol. 63, no. 2, pp. 482–501, December 1984.
[3] P. Kolan, R. Dantu, and J. W. Cangussu, "Nuisance level of a voice call," *ACM TOMCCAP*, vol. 5, no. 1, 2008.
[4] S. Phithakkitnukoon and R. Dantu, "UNT mobile phone communication dataset," *http://nsl.unt.edu/∼santi/data_desc.pdf*, 2008.
[5] ——, "Predicting calls - new service for an intelligent phone," in *Proceedings of IFIP/IEEE MMNS '07*. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 26–37.
[6] J. Birtchnell, "Personality set within an octagonal model of relating," *American Psychological Association*, vol. Circumplex models of personality and emotions, pp. 155–182, 1997.
[7] L. Jamieson, *Intimacy, personal relationship in modern societies*. Oxford: Polity Press & Blackwell Publishers Ltd., 1998.
[8] M. Popovic, D. Milne, and P. Barrett, "The scale of perceived interpersonal closeness (PICS)," *Clinical Psychology and Psychotherapy*, vol. 10, pp. 286–301, 2003.
[9] K. Kayser and D. P. Himle, "Dysfunctional beliefs about intimacy," *Journal of Cognitive Psychotherapy*, no. 8, pp. 127–139, 1994.
[10] M. D. Sherman and M. H. Thelen, "Fear of Intimacy Scale: validation and extension with adolescents," *Journal of Marital and Family Therapy*, vol. 13, pp. 507–521, 1996.
[11] J. Orlofsky, J. Marcia, and I. Lesser, "Ego identity status and the intimacy versus isolation crisis of young adulthood," *Journal of Personality and Social Psychology*, vol. 27, pp. 211–219, 1973.
[12] M. T. Schaefer and D. H. Olson, "Assessing intimacy: the PAIR inventory," *Journal of Marital and Family Therapy*, vol. 7, pp. 47–60, 1981.
[13] L. Milne, *Social therapy. A guide to social support interventions for mental health practitioners*. Chichester: John Wiley & Sons, 1999.
[14] R. D. Nolker and L. Zhou, "Social computing and weighting to identify member roles in online communities," *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 87–93, 2005.
[15] G. S. Mesch and I. Talmud, "Online friendship formation, communication channels, and social closeness," *International Journal of Internet Science*, vol. 1, no. 1, pp. 26–44, 2006.
[16] A. V. Zhdanova, L. Predoiu, T. Pellegrini, and D. Fensel, "D.: A Social Networking Model of a Web Community," *Proceedings of the 10th International Symposium on Social Communication*, 2007.
[17] J. Bleecker, "What's your social doing in my mobile? design patterns for mobile social software," *Proceeding of WWW2006 Workshop MobEA IV Empowering the Mobile Web*, pp. 1–6, May 2006.
[18] L. Hossain, K. Chung, and S. Murshed, "Exploring temporal communication through social networks," 2007, pp. 19–30.
[19] L. C. Freeman, "Centrality in social networks: Conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978.
[20] J. S. Coleman, *An Introduction To Mathematical Sociology*. Collier-Macmillan, London, 1964.
[21] C. P. Kottak, *Cultural Anthropology*. Mcgraw Hill, New York, 1991.
[22] R. Scupin, *Cultural Anthropology - A Global Perspective*. Prentice Hall, Englewood Cliffs, 1992.
[23] R. I. M. Dunbar, "Coevolution of neocortical size, group size and language in humans," *Behavioral and Brain Sciences*, vol. 16, no. 4, pp. 681–735, 1993.
[24] W. X. Zhou, D. Sornette, R. A. Hill, and R. I. M. Dunbar, "Discrete hierarchical organization of social group sizes," *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1561, pp. 439–444, 2005.