



Analyzing Human Communities by Interaction and Role Indices

M. S. Siddiqui¹, Z. A. Shaikh², A. R. Memon³ and M. Shah⁴

¹GSESIT, FEST, Hamdard University, Karachi, Pakistan,

²Center of Research for Ubiquitous Computing, National University of Computer and Emerging Sciences, Karachi, Pakistan

³FEST, Hamdard University, Karachi, Pakistan,

⁴Lancashire Business School, University of Central Lancashire, United Kingdom,

Emails: ²zubair.shaikh@nu.edu.pk, ³armemon2001@yahoo.com, ⁴mhshah@uclan.ac.uk

*Corresponding author: M. S. Siddiqui, e-mail: shahab.siddiqui@ieee.org

Received 12 th May 2011 Revised 13th June 2011

Abstract: A Human community (network) is formed when human eProfiles (nodes) are linked (edges) and interact with each other. We propose two indices namely, Interaction Index (II) and Role Index (RI), which will help in analyzing human community. Interaction Index will estimate the intensity of communication of one human on other. Role Index will determine the participation status of each individual in a human community. This analysis will be used in our ongoing research and development of Community Algorithm. It will help in classifying individuals depending on their interaction in any human community based on human eProfiles.

Keywords: Interaction Index, Role Index, Human Community Ontology, Community Ontology, Community Algorithm

INTRODUCTION

The field of Computational Intelligence (Reynolds, 2009) has grown from infancy to adulthood. Start being artificial intelligence powered by genetic algorithm, neural networks, fuzzy logic, and machine intelligence and emerged as swarm intelligence, having algorithms like Ant Colony Optimization, Particle Swarm Optimization and Cultural Algorithm. Procedural programming in Pascal, C/C++, and LISP has been transformed to genetic programming. Community Algorithm (Siddiqui *et al.*, 2009a, 2009b) is a variant of Genetic Algorithm (GA) (Wang, 2007) developed by John Holland and works on same footings of Cultural Algorithm developed by Reynolds *et al.*, (1997). This research will help in formalizing the concept of human community (Siddiqui *et al.*, 2009a) in Community Algorithm (Siddiqui *et al.*, 2009a, 2009b).

Some exciting principles were found while studying social networks. First one was clique, which is a subset of individuals in which every person is connected to every other person, sharing common interests, views, purposes, patterns of behavior, or

ethnicity. Similarly, peer group can be seen as a social (informal primary) group consisting of humans, share a similar or equal status and having same age and interact within the social aggregate. Another concept of small world phenomenon (Auber *et al.*, 2003, Inaltekin *et al.*, 2009 and Watts 1999) was studied, which resulted into six degree of separation principle. Duncan Watts used this concept by in his books (Watts 1999, 2003) on Social Networks. Clustering is also been used in various studies of social networks.

The rest of the paper is organized as follows. In section 2, the parameters of human eProfiles were discussed, which are used in formation of human communities (Siddiqui *et al.*, 2009a). Also, interaction and role indices were defined. Results based on analysis of human communities by statistical operations were discussed in section 3, Followed by conclusion.

MATERIAL AND METHOD

EXPERIMENT

Formation of Human Communities

In this section, we firstly describe the characteristics of human eProfiles, which can be seen in Table 1. Secondly, the human communities will statistically analyze. Lastly, we apply two indices – Inter-

action and Role on human communities, which was created on the basis of number of parameters of human eProfiles. We can define a community as a graph $G = \langle V, E \rangle$, where, $V = \{v_i\}$ is the set of human eProfiles and $E = (e_{ij})$ is the communication link between elements of V . In (Table 1), there exists a pool of human eProfile parameters from different domains, which can extract human communities. Common

features among human eProfiles will link one human with other to form a human community (Siddiqui *et al.*, 2009a). The resultant (sample) human community of 10 nodes can be seen in (Fig. 1). In Computer Science, only web communities (Flake *et al.*, 2000) were discussed, which are also known as FLG Community (Flake *et al.*, 2000).

Table 1. Human eProfile parameters in different domains

eProfile Parameter	Social Network		System S/W		Email	Genealogy	
	Orkut	FaceBook	NT /XP	Linux/ Unix	Hotmail/ GMail/ Yahoo	Geno Pro	Family Tree
Username	X	X	√	X	√	X	X
Firstname	X	X	√	√	√	√	√
Middlename	X	X	X	X	X	√	√
Lastname	X	X	√	√	√	√	√
Suffix	X	X	X	X	X	√	√
Nick Name	X	X	X	√	X	√	√
Gender	X	√	X	X	X	√	√
Birthday	√	√	X	X	X	√	√
Place of Birth	X	X	X	X	X	√	√
Address	√	X	√	X	X	√	√
City	√	X	√	X	X	√	√
State	√	X	√	X	X	√	√
Country	√	X	√	√	√	X	X
ZIP Code	√	X	√	X	√	√	√
Email Address	√	√	X	X	√	√	√
Home Phone	√	X	√	X	X	√	√
Mobile Phone	√	X	√	X	X	√	√
Religion	√	X	X	X	X	X	X
Mate	X	X	X	X	X	√	√
Father	X	X	X	X	X	√	√
Mother	X	X	X	X	X	√	√
Sibling	X	X	X	X	X	√	√
Childless	√	X	X	X	X	√	√
Relation Status	√	X	X	X	X	X	X
Discipline	X	X	X	X	X	√	√
Institution Name	X	X	X	X	X	√	√
Study Type	X	X	X	X	X	√	√
Edu. Summary	X	X	X	X	X	√	√
Company Name	X	X	√	X	X	√	√
Designation	X	X	√	√	X	√	√
Industry Name	X	X	√	X	X	√	√
Photograph Name	X	√	X	X	X	√	√
# Emails Send	X	X	X	X	√	X	X
# Emails Received	X	X	X	X	√	X	X

During the research, a project **GAHC** (Asrar *et al.*, 2010) was developed, which generates community graphs based on selected parameters (First Name, Gender, Religion, City, Country and Designation) from the human eProfiles given in (Table 1 and Fig. 1) shows human community graph produced by GAHC tool (Asrar *et al.*, 2010) on the basis of religion

having 10 nodes.

Defining Indices

In this section, we will like to define two indices, namely, interaction index and role index, which will classify the role and level of interaction of each user in different human communities.

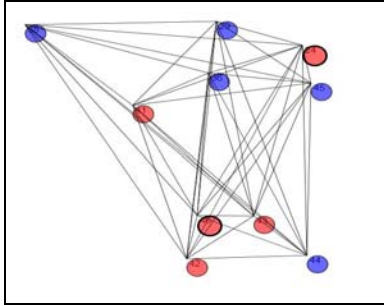


Fig. 1. A sample 10 nodes human community (Dark outlined nodes are the start and end nodes)

Interaction Index

We define an index for several levels of interaction in different communication mediums. Different levels of Interaction Index (II) are defined in (Table 2).

Table 2. Defining the Interaction Index

Interaction Index	Matrix Values
H = High	Max (Matrix Operation)
M = Medium	> Min (Matrix Operation) and < Max (Matrix Operation)
L = Low	Min (Matrix Operation)
No Interaction	0.00

S =

0.00	0.02	0.03	0.03	0.05	0.14	0.02	0.11	0.05	0.55
0.17	0.00	0.17	0.00	0.02	0.01	0.05	0.14	0.44	0.00
0.00	0.06	0.00	0.05	0.06	0.07	0.13	0.33	0.04	0.25
0.01	0.04	0.09	0.00	0.07	0.69	0.09	0.00	0.01	0.01
0.01	0.16	0.00	0.00	0.00	0.06	0.17	0.53	0.00	0.06
0.05	0.30	0.02	0.05	0.04	0.00	0.06	0.07	0.11	0.30
0.15	0.05	0.01	0.07	0.27	0.00	0.00	0.05	0.27	0.14
0.17	0.10	0.02	0.03	0.05	0.01	0.05	0.00	0.51	0.07
0.00	0.06	0.06	0.06	0.61	0.00	0.19	0.02	0.00	0.00
0.02	0.53	0.17	0.14	0.07	0.02	0.00	0.03	0.03	0.00

E =

0.00	0.15	0.02	0.11	0.06	0.58	0.07	0.00	0.01	0.00
0.05	0.00	0.05	0.14	0.44	0.00	0.05	0.06	0.07	0.13
0.06	0.02	0.00	0.14	0.02	0.10	0.05	0.54	0.07	0.00
0.31	0.18	0.04	0.00	0.02	0.01	0.04	0.10	0.31	0.00
0.00	0.12	0.13	0.12	0.00	0.05	0.06	0.07	0.12	0.32
0.01	0.32	0.10	0.08	0.04	0.00	0.01	0.06	0.03	0.34
0.11	0.28	0.00	0.07	0.00	0.23	0.00	0.28	0.00	0.04
0.08	0.00	0.05	0.26	0.02	0.45	0.07	0.00	0.01	0.05
0.14	0.02	0.07	0.14	0.01	0.00	0.08	0.38	0.00	0.15
0.15	0.00	0.03	0.58	0.02	0.02	0.06	0.02	0.11	0.00

0.00	0.12	0.06	0.63	0.08	0.00	0.01	0.01	0.03	0.07
------	------	------	-------------	------	------	------	------	------	------

Role Index

Roles of human in different communities can be represented using Role Index. (Table 3) defines the Role Index (RI) that will highlight the distinction of leaders and followers in human communities.

Table 3. Defining the Role Index

Role Index	Probability Values
Most Active	Most High Interaction Index
Active	Most Medium Interaction Index
Passive	Most Low Interaction Index
No Role	0.00

RESULTS AND DISCUSSION

After generating human communities, analysis has been made on the interaction frequency based on number of emails and SMS and chat session hours among different humans living within a community. Consider a human community from a University environment with N=10 (nodes) and providing data for each users for different medium of communication i.e. email (E), chat (C) and SMS (S). Data for these communications will be modeled through matrices.

$$C = \begin{bmatrix} 0.12 & 0.00 & 0.15 & \mathbf{0.45} & 0.06 & 0.02 & 0.03 & 0.04 & 0.12 & 0.05 \\ 0.06 & 0.15 & 0.00 & 0.14 & 0.04 & \mathbf{0.42} & 0.05 & 0.02 & 0.04 & 0.11 \\ \mathbf{0.63} & \mathbf{0.45} & 0.14 & 0.00 & 0.01 & 0.04 & 0.07 & 0.20 & 0.21 & 0.01 \\ 0.08 & 0.06 & 0.04 & 0.01 & 0.00 & 0.05 & 0.09 & \mathbf{0.23} & 0.09 & 0.18 \\ 0.00 & 0.02 & \mathbf{0.42} & 0.04 & 0.05 & 0.00 & 0.06 & 0.02 & 0.08 & 0.03 \\ 0.01 & 0.03 & 0.05 & 0.07 & 0.09 & 0.06 & 0.00 & 0.08 & \mathbf{0.22} & 0.10 \\ 0.01 & 0.04 & 0.02 & 0.20 & \mathbf{0.23} & 0.02 & 0.08 & 0.00 & 0.10 & 0.05 \\ 0.03 & 0.12 & 0.04 & 0.21 & 0.09 & 0.08 & \mathbf{0.22} & 0.10 & 0.00 & 0.17 \\ \mathbf{0.07} & 0.05 & 0.11 & 0.01 & \mathbf{0.18} & 0.03 & 0.10 & 0.05 & 0.17 & \mathbf{0.00} \end{bmatrix}$$

We analyzed and verified the results on the basis of statistical operations on these three matrices. In each of the matrices, the values marked in bold are at the maximum level of interaction in different communication mediums for (i, j) users, where i is the sender and j is the receiver user. In other words, it is the level of communication between user_i and user_j.

When weighted average is applied on the matrices to each of the medium of communication i.e. email (E), chat (C) and SMS (S) for analyzing human interaction in a University community, we get 6 combination matrices in result. The average of weighted average (AWA) can be seen in the following equation:

$$AWA_{ij} = \frac{(C.2E.3S)_{ij} + (S.2E.3C)_{ij} + (E.2S.3C)_{ij} + (S.2C.3E)_{ij} + (C.2S.3E)_{ij} + (E.2C.3S)_{ij}}{6} \quad (1)$$

Following are the different equations from (2) to (6) for statistical operations of Average (AV), Biased Weighted Average for SMS (BWAS), Email (BWAE), and Chat (BWAC) and Average of all Biased Weighted Average (ABWA) on the same matrices for each of the medium of communication:

$$AV_{ij} = \frac{E_{ij} + S_{ij} + C_{ij}}{3} \quad (2)$$

$$BWAE_{ij} = \frac{3 * E_{ij} + S_{ij} + C_{ij}}{5} \quad (3)$$

$$BWAS_{ij} = \frac{E_{ij} + 3 * S_{ij} + C_{ij}}{5} \quad (4)$$

$$BWAC_{ij} = \frac{E_{ij} + S_{ij} + 3 * C_{ij}}{5} \quad (5)$$

$$ABWA_{ij} = \frac{WAVE_{ij} + WAVS_{ij} + WAVC_{ij}}{3} \quad (6)$$

The summary of all of the operations defined in equations from (1) to (6) applied on the three matrices can be seen in (Table 4). Values in bold are the maximum ones and values in italic and underline are the minimum ones. It can be concluded statistically in table 4, that eProfile 4 has the maximum level of communication in almost every medium of interaction and eProfile 7 has the minimum level of communication in every medium of interaction.

Table 4. Summary of all results of matrix operations

User	AWA	AV	BWAE	BWAS	BWAC	ABWA
1	1.00	1.00	1.00	1.00	1.00	1.00
2	1.01	1.01	1.01	1.01	1.02	1.01
3	1.01	1.01	1.00	1.01	1.02	1.01
4	1.25	1.25	1.15	1.15	1.45	1.25
5	0.94	0.94	0.97	0.97	0.90	0.94
6	0.91	0.91	0.95	0.95	0.84	0.91
7	<i>0.90</i>	<i>0.90</i>	<i>0.94</i>	<i>0.94</i>	<i>0.82</i>	<i>0.90</i>
8	0.92	0.92	0.95	0.95	0.86	0.92
9	1.02	1.02	1.01	1.01	1.04	1.02
10	0.92	0.92	0.95	0.95	0.85	0.92

Table 5. Applying Interaction Index and Role Index

User	AWA	AV	BWAE	BWAS	BWAC	ABWA	RI
1	M	M	M	M	M	M	A
2	M	M	M	M	M	M	A
3	M	M	M	M	M	M	A
4	H	H	H	H	H	H	MA
5	M	M	M	M	M	M	A
6	M	M	M	M	M	M	A
7	L	L	L	L	L	L	P
8	M	M	M	M	M	M	A
9	M	M	M	M	M	M	A
10	M	M	M	M	M	M	A

After applying Interaction Index and Role Index on the summary results of the matrix operations in (Table 5), it can be concluded that user 4 has most active role in the human community, while user 7 has

passive role and all other users have active role in the human community. Therefore, the two Indices for classifying interaction and role are highlighting the statistical results achieved in (Table 4).

CONCLUSION

In this paper, human communities were statistically analyzed by using interaction and role indices on the basis of human eProfile parameters. Both the indices classify or grade users, based on their interaction (in terms of Email, Chat and SMS) with other members of the human community. All statistical operations supplement the two indices and helped us in analyzing in the desired direction. This analysis will help in the analyzing human communities and strengthen the ongoing research and development for Community Algorithm.

ACKNOWLEDGEMENT

We are thankful to Higher Education Commission of Pakistan for providing the PhD Scholarship Grant for the research in Computer Science. We are also thankful for utmost help and support of Center of Research in Ubiquitous Computing (CRUC), National University of Computer and Emerging Sciences, Karachi, Pakistan.

REFERENCES

Asrar, M., M. Saleem, F. Ahmed, M.S. Siddiqui, and Z.A. Shaikh, (2010) "GAHC-Graph-based analysis of human communities", *Final Year BS (CS) Project Report (KSPZS05-09F)*, National University of Computer and Emerging Sciences, Karachi, Pakistan.

Auber, D., Y. Chiricota, F. Jourdan, and G. Melançon, (2003) "Multiscale visualization of small world networks", *The Proceedings of the IEEE Symposium on Information Visualization (INFOVIS)*, 75-81.

Flake, G.W., S. Lawrence, and C.L. Giles, (2000) "Efficient identification of web communities", *The Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 150-160.

Inaltekin, H., M. Chiang and H.V. Poor, (2009) "Delay of social search on small-world random geometric graphs", *Social Networks*, August 2009.

Reynolds, R.G., "A genealogy of computational intelligence books", *IEEE Computational Intelligence Magazine*, USA, 56-61.

Reynolds, R.G. and M. Sternberg, (1997) "Using cultural algorithms to support reengineering of rule-based expert systems in dynamic performance environments: A case study in fraud detection", *IEEE Transaction on Evolutionary Computation*, 1 (4): USA, 225-243.

Siddiqui, M.S., Z.A. Shaikh, and A.R. Memon, (2009) "Towards the development of human community ontology", *The Proceedings of the 2009 World Congress on Software Engineering (WCSE 2009)*, Xiamen, China, Vol. (3): May 19-21, 2009a, 8-12.

Siddiqui, M.S., Z.A. Shaikh, and A.R. Memon, (2009) "Towards the development of community algorithm", *The Proceedings of the 2009 International Conference on Information Management and Engineering (ICIME 2009)*, Kuala Lumpur, Malaysia, April 03-05, 2009b, 612-616.

Wang, L.-H., (2007) "A comparison of three fitness prediction strategies for interactive genetic algorithms", *Journal of Information Science and Engineering*, 605-616.

Watts, D. J., (2003) "Six degrees: The science of a connected age", *Norton & Company*, New York, USA.

Watts, D.J., (1999) "Small worlds", *Princeton University Press*.