Network Analysis of a System Supporting the Formation of Cyber-communities

Kunihiko FUJITA
k.fujita@ntt.com
NTT Communications
Tokyo, Japan 100-8019

Koji KAMEI
kamei@cslab.kecl.ntt.co.jp
NTT Communication Science Labs.
Kyoto, Japan 619-0237

Eva JETTMAR
true@stanford.edu
Department of Communication
Stanford University

Sen YOSHIDA
yoshida@cslab.kecl.ntt.co.jp
NTT Communication Science Labs.
Kyoto, Japan 619-0237

Kazuhiro KUWABARA
kuwabara@cslab.kecl.ntt.co.jp
NTT Communication Science Labs.
Kyoto, Japan 619-0237

Abstract : *Network analysis is a method for discovering the structure of the communications flow between group members. In this paper, we introduce network analysis first. Next, we introduce* ***Community Organizer*** *a system designed to support network communities, and report results of an experiment applying network analysis.*

1. Introduction

With the advances in computer networks, people are getting connected more than ever. New types of *cyber societies* based on the communications of people on networks are beginning to emerge.

Our focus is on *network communities* as a form of communications for people in cyber societies. The emphasis of this focus is the necessity of computer systems to support the formation of network communities. We have already constructed a prototype system named ***Community Organizer***. We have also conducted evaluations of ***Community Organizer*** in our laboratories. In these evaluations, statistical analysis was applied, and it was found that the spatial representation offered by ***Community Organizer*** is effective at encouraging communications between users [1]. However, statistical analysis cannot detect aspects of communications between users. Accordingly, in this paper, we summarize a ***Community Organizer*** experiment, and describe the results of the experiment, in which *network analysis* was applied to detect aspects of communications.

2. *Community Organizer* Experiment

The objective of ***Community Organizer*** is to support the formations of network communities. Unlike traditional communities where geographical and institutional properties define the boundaries of the communities, a network community consists of people who share common interests.

The experiment we conducted was a control experiment. Two versions of the ***Community Organizer*** software were used: the "full" version which provides meaningful spatial representations in terms of proximity based on similarities and differences in the specifications of user interests, and the "listing" version, which was prepared for the experimental control condition (Fig.1). The "full" version was the previously proposed ***Community Organizer***, and the "listing" version was prepared specifically for this study. Both versions had the same look and feel. A total of 42 subjects participated in the experiment. The subjects were randomly assigned to one of two conditions; however, the conditions were balanced for gender, age, and online experience, based on data gathered from a pre-test questionnaire. Each group used one version of the ***Community Organizer*** software for one week at their workplace. In order to provide anonymity, each subject used a "handle" name and icon.

Full version

Listing version

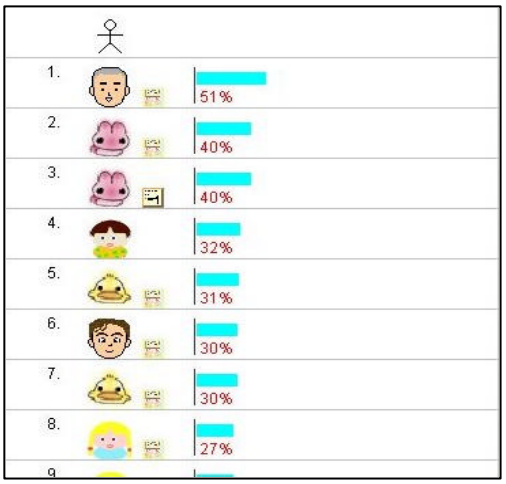

Fig 1: Two Versions of ***Community Organizer***

3. Network Analysis

Network analysis is a method for discovering the structure of the communications flow between group members[2]. The analysis is based on patterns of network data. The network data includes matrix data, which represents who is communicating with who in numerical value.

We used the utterance log of a public message board, i.e., one of the communication tools of ***Community Organizer***, as the target of the network analysis. This is because the subjects used the public message board the most over the duration of the experiment.

The public message board has almost the same function as a bulletin board system. All of users can make a public message board when they want to talk about a certain topic with others. The public message board is appeared as an icon on each viewer. All of the other users can "talk" on that public message board.

3.1 The Measures of Network Analysis

In the network analysis, the relationships between users are represented as a graph. Using graph theory, a number of characteristics of networks can be determined.

(1) *Centrality:* A measure of the extent to which a given entity (user) is directly or indirectly "related" via links to other entities (users). The member who gets highest value of centrality has the possibility to be the most important person in the group. Centrality is calculated by the formula shown below. x_{ij}

$$Centrality_i = \frac{\sum_{j=1}^{N} x_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}}$$

in this formula is the number of links from member *i* to member *j*. *N* is the number of all members.

(2) *Density:* A measure of the number of other entities (users) to which an entity (user) is directly related. The group that gets higher value of density can be said the relationships between group members would be a dense and the activity of each group members would be standardized. On the other hand, the group that gets lower value of density can be said the relationships between group members would be thin and the activity of each group members would be various. The value of z_{ij}

$$Density = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} z_{ij}}{N(N-1)}$$

in this formula is 1 if there is at least one relationship between the member *i* to member *j*.

(3) *Cohesion:* A measure of the degree of "interconnectedness" of entities (users). This measure is similar to the density, but the numerator is incremented when there is

$$Cohesion = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} \left(z_{ij} + z_{ji}\right)}{N(N-1)/2}$$

interconnected relationship between two entities (users). The "+'' in this formula would

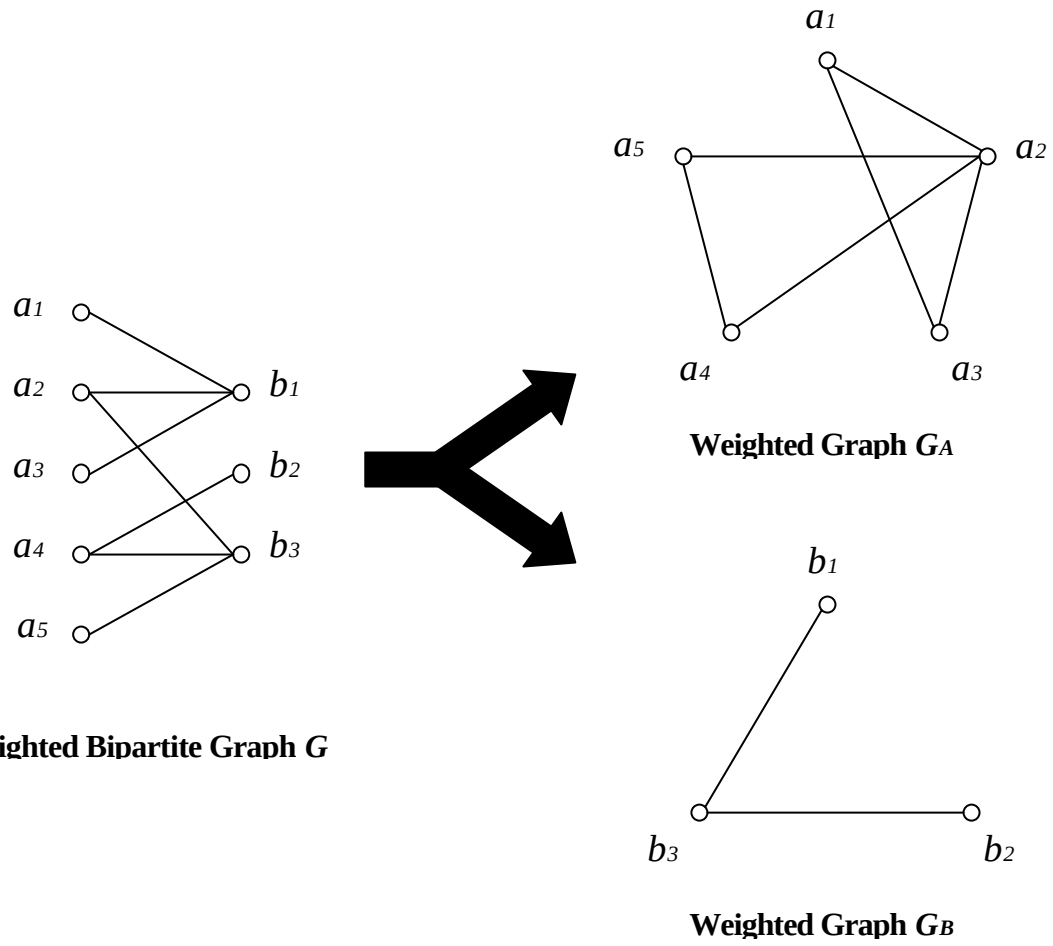

Fig 2: Division Bipartite Graph into Two Graphs

be "1'' if both values of z_{ij} and z_{ji} are "1''.

3.2 Network Analysis for Heterogeneous Environment

So far, we explained about ordinary network analysis applying to the groups that include only a kind of entities in a group (*homogeneous environment*). In this section, we describe about network analysis applying to the groups that include plural kinds of entities in a group (*heterogeneous environment*). We introduce an example of a sales record of e-commerce site to explain.

Differences between e-commerce site shopping and traditional shopping is as follows; 1) Sellers of e-commerce site can identify customers 2) Sellers of e-commerce site have sales records 3) Customers of e-commerce site can evaluate goods. This can represent as weighted bipartite graph shown as Fig.2.

In ordinary network analysis, we can't treat two types of entities (sales goods and customers) simultaneously. Accordingly, we divide a bipartite graph into two graphs. Each graph includes only one type of entities. After dividing, we can use ordinary network analysis for them.

4. Discussion

4.1 Analyzing Relationships between Users

The standard deviations of centrality based on the in-degree suggest that an uneven distribution of the "full" version (0.034) is lower than that of the "listing" version (0.051). The standard deviations of centrality based on the out-degree show the opposite result ("full": 0.067, "listing": 0.060), however, the difference in values is smaller than that of the in-degree (Table 1).

Both cohesion and density of the "full" version is higher than that of the "listing" version. This means the users of the "full" version communicated more closely (Table 2).

From the measures shown above, we can conclude that the "full" version of ***Community Organizer*** can support group communications more than the "listing" version.

4.2 Analyzing Relationships between Public Message Boards

Table 1: Centrality (relationships between users)

Full's subjects	In	Out	Listing's subjects	In	Out
F01	0.110	0.024	L01	0.105	0.211
F02	0.085	0.207	L02	0.092	0.158
F03	0.024	0.024	L03	0.026	0.053
F04	0.098	0.146	L04	0.079	0.026
F05	0.110	0.073	L05	0.105	0.013
F06	0.061	0.195	L06	0.197	0.079
F07	0.098	0.049	L07	0.079	0.053
F08	0.037	0.000	L08	0.013	0.013
F09	0.037	0.000	L09	0.092	0.013
F10	0.073	0.049	L10	0.079	0.026
F11	0.085	0.098	L11	0.000	0.092
F12	0.049	0.000	L12	0.013	0.026
F13	0.000	0.000	L13	0.053	0.158
F14	0.049	0.098	L14	0.013	0.039
F15	0.024	0.037	L15	0.013	0.000
F16	0.049	0.000	L16	0.013	0.039
F17	0.012	0.000	L17	0.026	0.000
			L18	0.000	0.000

Table 2: Density and Cohesion (relationships between users)

	Full	Listing
Density	0.238	0.195
Cohesion	0.092	0.052

There are no direct relationships between public message boards. However, we can reveal the relationships between boards using method as we mentioned at section 3.2.

There are 26 boards made by Full version users, and 22 boards made by Listing version users. We calculate centrality and density (Table 3, 4).

5. Conclusion

In this paper, we introduced ***Community Organizer***, and then examined its effectiveness in supporting communities by performing a field experiment and network analysis.

Acknowledgements

We thank our colleagues at NTT Communication Science Laboratories for participating in the experiment reported in this paper. We also thank Advanced Business Works at NTT Communications for the software implementation.

Table 3:Centrality (relationships between public message boards)

Full		Listing	
Tf01	0.033	Tl01	0.041
Tf02	0.024	Tl02	0.051
Tf03	0.013	Tl03	0.017
Tf04	0.017	Tl04	0.034
Tf05	0.068	Tl05	0.092
Tf06	0.017	Tl06	0.048
Tf07	0.050	Tl07	0.022
Tf08	0.042	Tl08	0.026
Tf09	0.024	Tl09	0.015
Tf10	0.040	Tl10	0.028
Tf11	0.027	Tl11	0.065
Tf12	0.012	Tl12	0.034
Tf13	0.050	Tl13	0.124
Tf14	0.004	Tl14	0.068
Tf15	0.004	Tl15	0.021
Tf16	0.054	Tl16	0.029
Tf17	0.043	Tl17	0.016
Tf18	0.044	Tl18	0.015
Tf19	0.030	Tl19	0.041
Tf20	0.065	Tl20	0.034
Tf21	0.077	Tl21	0.055
Tf22	0.064	Tl22	0.125
Tf23	0.076		
Tf24	0.063		
Tf25	0.031		
Tf26	0.026		

Table 4: Density and Cohesion (relationships between public message boards)

	Full	Listing
Density	0.600	0.710

References

[1] Koji Kamei, Eva Jettmar, Kunihiko Fujita, Sen Yoshida and Kazuhiro Kuwabara : Community Organizer: Supporting the Formation of Network Communities through Spatial Representation, In *the 2001 Symposium on Applications and the Internet (SAINT 2001)*, pages 207-214, 2001.

[2] Yuki Yasuda : Network Analysis, Shinyo-sha, 1997.