A Study of Social Behavior in Collaborative User Generated Services

Pengyan Yao
Key Laboratory of Universal Wireless Communications (Beijing University of Posts and Telecommunications), Ministry of Education
100876,Beijing, China
yaopyfly@gmail.com

Zheng Hu
Key Laboratory of Universal Wireless Communications (Beijing University of Posts and Telecommunications), Ministry of Education
100876,Beijing, China
huzheng.red@gmail.com

Zhenzhen Zhao, Noël Crespi
Institut Télécom, Télécom SudParis
91000, Evry, France
{zhenzhen.zhao, noel.crespi}@it-sudparis.eu

ABSTRACT

User-generated content has become more and more popular. The success of collaborative content creation such as Wikipedia shows the level of user's accomplishments in knowledge sharing and socialization. In this paper we extend this research in the service domain, to explore users' social behavior in Collaborative User-Generated Services (Co-UGS). We create a model which is derived from a real social network with its behavior being similar to that of Co-UGS. The centrality approach of social network analysis is used to analyze Co-UGS simulation on this model. Three Co-UGS network actors are identified to distinguish users according to their reactions to a service, i.e. ignoring users, sharing users and co-creating users. Moreover, six hypotheses are proposed to keep the Co-UGS simulation. The results show that the Co-UGS network constructed by the sharing and co-creating users is a connected group superimposed on the basis of the social network of users. In addition, the feasibility of this simulation method is demonstrated along with the validity of applying social network analysis to the study of users' social behavior in Co-UGS.

Categories and Subject Descriptors

G.3 [**Mathematics of Computing**]: Probability and Statistics – *Random number generation, Statistical computing;* H.2.8 [**Information Systems**] Database Applications – *Data mining*; J.4 [**Computer Applications]** Social and Behavioral Sciences – *Sociology;*

General Terms

Algorithms, Measurement, Performance, Design, Human Factors, Experimentation.

Keyword

User Generated Services; social network analysis; data mining; centrality.

ICUIMC'12, February 20–22, 2012, Kuala Lumpur, Malaysia.

1. INTRODUCTION

Web 2.0 is a set of economic, social, and technological trends characterized by user participation, openness, and their network effects [1]. The great success of the UGC (User Generated Content) phenomenon raises the new concept of enabling end-users to create their own services. In contrast to UGC, which focus on creation for itself, for the fun of it, for socialization and for its knowledge sharing aspects, user generated service (UGS) puts weight on the service personalization and service reuse. UGS is distinguished by the fact that end-users need to handle the processes of creating services that are not supported in existing content-oriented web 2.0 platforms [2].

Collective intelligence, as the key characteristic observed from the content creation process, is inspiring us to enable users to create services together [3]. The emergence of the social web has made the concept of collective intelligence a real possibility. In the social web, "people socialize or interact with each other throughout the World Wide Web; social interactions lead to the creation of explicit and meaningfully rich knowledge representations" [4]. Collaborative user-generated services (Co-UGS) put the User-Generated Services system in the Social Network to allow end-users to discover the required services through searching and retrieving, and to interact with the created service through sharing and co-creation. interact with the created service through sharing and co-creation.

Social Network studies are becoming increasingly popular and have been applied to several fields such as law enforcement, marketing, analyzing the spread of disease, as well as in the improvement of organizational performance However, little research has been done on user behavior in the process of collaborative user-generated services. Significantly, till now, no practical Co-UGS platform has been established. For this paper, a Co-UGS simulation platform was built, based on data obtained from a real social network. This was followed by the designing of a group of service generation rules. A service collaborative generating experiment was executed on the platform according to those rules. The paper ends with an analysis and suggestions for service recommendation.

Some key results are presented as follows:

Firstly, the Co-UGS network constructed by the sharing and co-creating users is a connected group, which possesses the same power-law distribution as a social network of users.

Secondly, the results revealed that the number of Co-UGS network users and the number of connections of an initial node is not linear relationship, and that with the increasing degree indexes of initial nodes the number of Co-UGS network users first goes upwards and then trends downwards.

Thirdly, the results demonstrate the feasibility of the simulation method and the validity of applying social network analysis in the study of Co-UGSs, which will have a significant impact.

The rest of the paper is organized as follows: in section II we present the related work on Co-UGS, the social network analysis method, and introduce a real online social network. Section III proposes a set of hypotheses and definitions as the principles on which to base the design of a Co-UGS simulating method. Section IV describes the simulation process and analyzes the results. Section V concludes the paper and provided some ideas for future extensions.

2. RELATED WORK

2.1 User Generated Service

In 2009, Z. Zhao and N. Crespi first clarified the concept of user generated service[2], and described the social service co-creation concept [3], emphasizing service creation by and for 'ordinary people' (non-technical users). The basic UGS process is based on the concept of service composition, which allows creating/assembling a new enhanced service from existing services or a set of basic pieces of services. The basic pieces of services are functional service building blocks with intrinsic functions that are reusable at runtime by various services.

We can use an example to explain the service co-creation process. Suppose an end user has created a service using readily available service building blocks. After it has been exposed to the community and discovered by other users, there will generally be three responses from users: ignoring, sharing and co-creating. Ignoring means that even though users may use the created service, they have little interest in sharing or co-creating. Sharing means that users think the created service is interesting and valuable, so they use it and pass it on to friends without making any improvement. Sharing can make more people discover the created services earlier and faster. Co-creating refers to those users who are not only interested in the created services, but ho are also interested in improving it by adding or exchanging some other service enablers, and then share it afterwards. The co-creating operation allows different users to collectively create more personalized, intelligent, and convenient services. It is the key and core process of collective user-generated services, since it embodies the collective wisdom.

2.2 Social Network Analysis

Social network analysis (SNA) is the mapping and measuring of relationships and flows between people, groups, organizations, computers, URLs, and other connected information/knowledge entities. The nodes in the network are the people and groups while the links show relationships or flows between the nodes. SNA provides both a visual and a mathematical analysis of human relationships. The analysis of social networks focuses on how to establish a model to represent the connections and tries to describe the structures of group relations, while simultaneously researching the influence of those structures on group relations or individuals in those groups.

Why use SNA to model the Co-UGS process? On one hand, since service co-creating or sharing is passed on the social network composed of users, the structure of social relations between users will affect the creation of the cooperation relationship between actors in the group. On the other hand, Co-UGS means that people create a service to solve a problem by collective means. Therefore, it is necessary to analyze the structure of users with social network analysis.

There are many metrics (measures) in social network analysis. One of the most important social network analysis mechanisms is centrality, which was introduced by Freeman [5-6] and consists of degree, betweenness and closeness.

The degree aspect indicates the relative importance and the location of a particular node in the network. In a social network, a node that has directly connected with many other nodes actually sees itself and can be seen by others in the network as an indispensable source of information. A general measure of centrality based on degree is:

$$C_D(n_i) = d(n_i) \qquad (1)$$

where $d(n_i)$ is the degree of node i.

Betweenness measures to what extent a node can play the role of intermediary in the interaction between the other nodes. Nodes located on many shortest paths (geodesics) between other nodes have higher betweenness compared to other nodes. For a graph $G = (V, E)$ with $\boldsymbol{n}$ nodes, the betweenness $CB(k)$ for vertex i is:

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk} \qquad (2)$$

where g_{ik} is the number of shortest geodesic paths from j to k, and $g_{ik}(n_i)$ is the number of shortest geodesic paths from j to k that pass through node i.

Another, more sophisticated centrality measure is closeness based on geodesic distance, which is the mean geodesic (i.e., shortest path) distance between a node and all other nodes reachable from it. From retrospective view, closeness can provide information about node independence. The simplest mathematics for closeness centrality is:

$$C_C(n_i) = \frac{1}{\sum_{j=1}^{g} d(n_i, n_j)} \qquad (3)$$

where d is the geodesic distance between respective nodes; for all those nodes that are not connected, the geodesic distance is infinity.

2.3 Data Source and Network Analysis

To make the simulation results come out closer to the actual situation and thereby make the platform more credible, real network data is obtained to build as the context of the Co-UGS platform environment. At the same time, we use the central degree analysis method to analyze user's network attributes, which is a reference to assess user's personal behavior characteristics as well as the user group's behavior characteristics.

In this paper, social relationship network data was downloaded from the Stanford large network dataset collection [7]. The whole network contains 7115 nodes and 103689 edges. A node represents a user, and each user is assigned a natural number from 1 to 7115as its user ID.

Firstly, the connections of the nodes in the network are investigated. It was found that the node connections obey a linear trend in a log-log plot, which demonstrates that nodes are not uniformly connected and that some key roles exist in the network. Therefore, the social network being studied is revealed as a scale-free network.

Based on the social network analysis method, centrality indexes of the network were studied by analyzing the network data with UCINET [8-9]. The outcomes are displayed in Table 1 and Table 2. According to the centrality results, the 2573rd node was deemed to be the most significance in degree, betweenness and closeness, while the 1780th node shows the least significance in degree and betweenness.

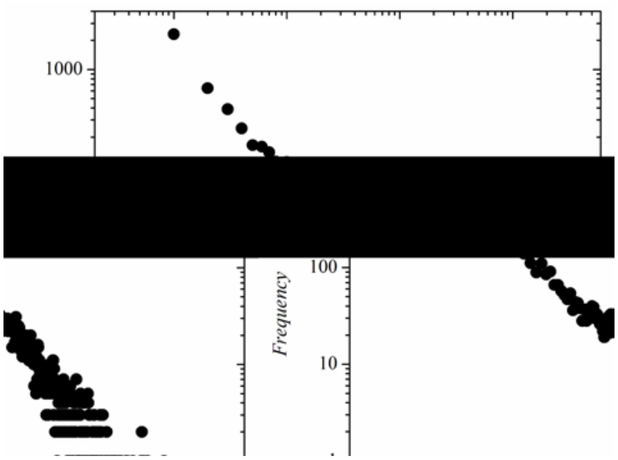

Figure 1 Node connection distribution of the real online social network

TABLE I. CENTRALITY DEGREE OF EACH NODE IN THE NETWORK

User ID	degree	betweenness degree	closeness degree
1	54	0.015	1.923
2	29	0.002	1.915
3	23	0.001	1.916
4	322	0.732	1.941
5	24	0.006	1.918
6	226	0.306	1.934
7	81	0.03	1.922
……	……	……	……

TABLE II CENTRALITY ANALYSIS OF NODES IN THE NETWORK

	Degree	Betweenness	closeness
Mean	28.324	7884.652	717580.375
Std Dev	57.574	37227.445	4154963
Sum	201524	56099300	5105584640
Variance	3314.793	1385882624	17263716466688
Min	1	0	363030
label (min)	1780	1780	2573
Max	1065	1549872.875	50613748
label (max)	2573	2573	2116

3. HYPOTHESIS AND DEFINITION

3.1 Hypotheses

The following hypotheses are formulated in the system of collective generated service:

Supposing there is a smart platform of collaborative user service creation on the above-mentioned real online social network, which makes up a collaborative service creation environment. The original social network's 7115 members are the potential end users of the platform. End users and the collaborative service creation environment together form a system of collaborative generated service.

Hypothesis 1: In the visual service creation environment, every user (node) can freely create/assemble a new enhanced service, either from existing services or from a set of basic pieces of services, and can only share services to direct contacts.

Hypothesis 2: In the network, only one node will be the original node for one originally created service at the same time. The node then shares the originally created service to its directly connected nodes.

Hypothesis 3: In the network, each node only have three possible reactions: sharing, co-creating and ignoring, when it discover one service (originally created service or co-created service) from its directly connected nodes.

Hypothesis 4: In the network, the service sharing and co-creating will follow the directed-acyclic-graph dissemination pattern, i.e. each nodes will only experience (create, share, co-create or ignore) just one time for one same service.

Hypothesis 5: If a user creates, shares or co-creates a service, all of his/her directly connected users, except for the users who have used or co-created the service already, will discover it eventually.

Hypothesis 6: A user can discover the created service only by means of the sharing with his/her direct contacts.

Users and service attributes and what they represent

To better simulate the collaborative service process, personal attributes, group properties, and service attributes are described as follows:

We have already mentioned that an end-user's participation in the Co-UGS process mainly depends on two aspects: the user's individual attributes and that user's social network's attributes. A user's response to a service is mainly determined by the former. User's individual attribute contains many factors, for example, a user's gender, age, education, profession, hobby, location. Different factors have varying degrees of influence on a user's reaction to a service. It is impossible to calculate the degree of effect generated by each factor. Meanwhile, a factor imposes different impacts on different services. However, since our purpose is to study the whole co-creating process in the network and the transmission features of a Co-UGS, we do not have to take into account each sub-factor's degree of effect. We only need to know the comprehensive interest of each user in certain types of services, which ultimately determines the user's service reaction. In addition, there are various users in a large social network who have different comprehensive interests in the same service. Therefore, if we can quantify a user's comprehensive interest, we can use the random number that best represents it.

Interest value (V): quantifies the integrated interest of a user to a service. Expressed by the letter V, the range is (0=<V<=1.0), 0 is completely not interested, and 1.0 is an infinite strong interest.

Sharing threshold (S): the lowest interest value that a user must attain when he/she decides to share a service (0<S<0.8).

Collective creating threshold (C): the lowest interest value that a user must attain when he/she decides to co-create a service (S<C<1).

Service created level (L): describes the co-creating extent of a service, and $L_i(j)$ means how many times the service initially generated by user *i* has been co-created when user *j* uses it.

We set each user's service created level at 0 as the initial value, and the service-created level of initial created service user *i* is 1. For example, when a service which has been co-created two times by two end-users is discovered by user *j*, if user *j* continues to improve it then the service-created level will add one and the service level of user *j* can be written as $L_i(j)=4$; if user *j* just shares it, then the service-created level is unchanged and the service level of user *j* can be written as $L_i(j)=3$; and we if user *j* ignores it, then the service-created level is unchanged and the service level of user *j* can be written as $L_i(j)=0$ (L>-1).

If user *j* does not know the starting point of the service, the service-created level of user j is written as L(*j*). According to the different *j* values, the largest $L_i(j)$ is written as L_i. L_i represents the co-creating extent of a service that was initially created by user *i*.

Number of sharing users (SN): the number of users who share that service in the Co-UGS network**.**

Number of creating users (CN): the number of users who share a service in the Co-UGS network**.**

Number of Co-UGS network users (N): the total number of users constituting the Co-UGS network**. The relationship is: N=SN+CN**

SN_i, CN_i, and N_i represent SN，CN and N of the Co-UGS network that are constructed when user *i* is the initial creator of a generated service. (1=<*i*<=7115).

3.2 User reaction judging rule

Currently, data is only used to define the context environment; therefore, to further construct a simulation platform with Co-UGS character, rules should be given to realistically constrain this platform.

The user reaction judging rules are defined as follows:

User *i* is directly connected to user *n* who has just created, shared or co-created a service. If user *i* discovers the service, then for user *i*:

(1) If L(*i*) = 0 exists, then

a If S(*i*)>V(*i*), user *i* chooses to ignore the service. Set L(*i*)=0 and keep SN and CN unchanged ;

b If C(*i*)>V(*i*)>=S(*i*), user m shares the discovered service. Set L(*i*) = L(*i*) and add one to SN ;

c If V(*i*)>C(*i*), user *i* chooses to co-create the discovered service. Both CN and the service created level will add 1.

(2) If L (*i*) = 0 does not exist, it means user *i* has judged before.

4. SIMULATION AND RESULT ANALYSIS

A simulation was carried out according to the information presented in parts II and III. The results are analyzed here to explore the relationship between the different centrality indexes of initial nodes and the number of users participating in the Co-UGC process and the features of the participating group.

We used an online social network containing 7115 users, assigning each user a natural number as user ID from 1 to 7115. Next, we randomly generated an interest value between 0 and 1, a sharing threshold (0, 0.8) and a co-creating threshold in (S,1) for each user, and set the service created level of each user as zero. For user *i*, these are written as V(*i*),S(*i*),C(*i*), and L(*i*). The original number of Co-UGS nodes (N) was set as zero. According to increasing user ID numbers, different users in turn were selected as the Co-UGS transfer process simulation starting node from which the Co-UGS network was built outward. The process is detailed below:

First, user *i* creates a service $L_i(i)=1, SN_i=0$ $CN_i=0$.

Second, according to hypothesis 4, each user j directly connected to that initial user will discover the created service. The reaction of user *j* is determined as SN_i, CN_i and $L_i(j)$ according the judging rule.

Third, the second procedure is repeated until both SN_i and CN_i stop increasing, which means that a collective co-creation process is finished.

Fourth, increasing *i* from 1 to 7115 in turn, the three steps above are repeated.

The experimental data are presented in Table 3:

TABLE III THE SIMULATION OF Co-UDGS FROM EACH USER

User ID	SN_i	CN_i	N_i	L_i
1	3547	282	3829	4
2	3534	282	3816	4
3	3562	287	3849	4
4	3429	291	3720	4
5	13	1	14	2
6	3383	262	3645	3
7	3490	278	3768	5
……	……	……	……	……

Generally, the more users that participate in the collective service creation, the better the Co-UGS. The Co-UGS network scale is not only related to the service, but also has an important relationship with the properties of the social network, and especially to the initial member's level of importance. Figures 3 to 5 demonstrate how N varies by the initial member's degree, by the betweenness degree and the closeness degree. It appears that N is not linear with a member's importance. In Figure 3, when a member's degree keeps growing, N first increases rapidly from zero to a high level, and then begins to decrease slowly. The turning point is around 200. That case is similar to the one in Figure 4. When a service starts to transfer from a lower betweenness degree member, the N stays smaller, but when the betweenness degree increases to a certain level the N will maintain larger. The results in Figures 3 and 4 indicate that an initial member's connections and betweenness centrality have a significant influence on the spreading of a Co-UGS . As the member's connections represent the people's relationship with others, so the more connections a member has the better a Co-UGS can be made. Betweenness centrality, however, reflects how important a node is in influencing members' communications. So the higher the betweenness centrality, the more a member plays a more important role in affecting information transferring in a network, and a better Co-UGS could be accomplished at a higher betweenness level.

For the closeness degree, even though its variation with SN is not very similar with the degree and betweenness degree, the basic pattern is still identical. As the graph in Figure 5 shows, when a member's normalized closeness is larger, the distance of this member to all the other members is shorter. If a service is created and begins to be transferred from a high closeness centrality member, the service could be discovered by his/her short-distance neighbors and then spread to the whole network within a limited number of steps.

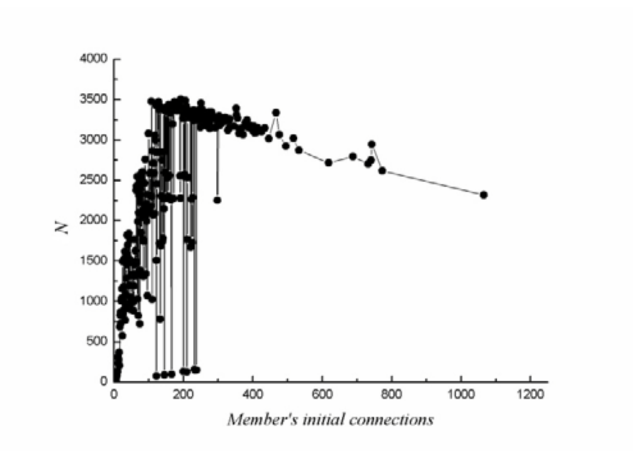

Figure 3. Variation of N by initial member's connections

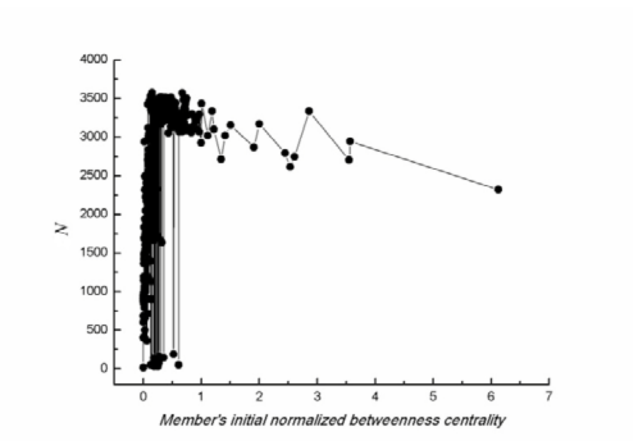

Figure 4. Variation of N by initial member's betweenness centrality. (The unit of X coordination is the normalized betweenness. Normalized flow betweenness centrality of a vertex i is the flow betweenness of i divided by the total flow through all pairs of points where i is not a source or a sink.

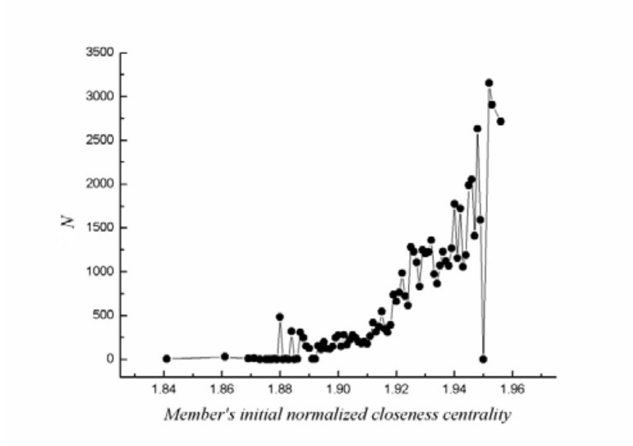

Figure 5. Variation of N by initial member's closeness centrality. (The unitof X coordination is normalized closeness. The normalized closeness centrality of a vertex is the reciprocal of farness divided by the minimum possible farness, expressed as a percentage.)

In addition to the variation of N that shows different patterns as the initial member's centrality changes, it also appears that sharing users and co-creating users are connected reciprocally by co-creating a service, which we called a Co-UGS network. Co-UGS network is a network superimposed on the basis of the social network of end users, constructed by users of the original network connected by service co-creating or sharing. In some sense, the process of collective service creation is actually a process of forming a connected group of a complicated network of users, which means that a number of interested users of a service are gradually connected together around the initial created service as a central node. A connected group is referred to as a sub-graph of a complex network, and the path exists between any two nodes in the sub-graph. There may be several independent groups in a network. In the percolation model, when the system is in a critical state, the size of a connected group exhibits a power-law distribution. Empirical studies have shown that for a large number of scale-free networks, the scale of connected groups also shows a power-law distribution.

To verify that the all of a Co-UGS network is a connected group, the connections of the nodes in the Co-UGS network are investigated. In the Data Source and network analysis section, the social network was proven to be consistent with the power-law distribution. It was found that the node connections of the Co-UGS network obey a linear trend in a log-log plot, demonstrating a power-law distribution.

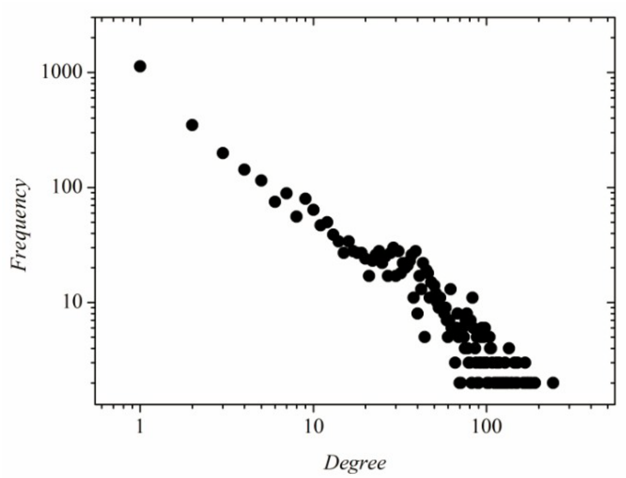

Figure 6. Node connection distribution of the Co-UGS network starting from the 15th node

5. CONCLUSIONS

This paper simulated the process of multi-user collectively-generated services based on real social network data, and has adopted the social network analysis method to study the relationship of the degree of the initial creating service user and the generated Co-UGS network. It has proved that the process of collectively generating services by multiply users is actually a process that, with an initial creating service user as the center, gradually forms a connected group on the social network by connecting some of the users interested in that service. The Co-UGS network possesses the same power-law distribution as the original users' social network.

The number of users in this connected group is the total number of sharing users and co-creators, but the number of co-creators is far less than that of the sharing users. In addition, we found that the total number of Co-UGS network users has a close relation with the initial user's centrality. This relationship can guide us to select the initial service creator to generate the broadest Co-UGS network. The study results demonstrate the feasibility of the simulation method and the validity of applying social network analysis to the study of Co-UGS.

In fact, the simulation combined with the social network analysis method in this paper can analyze the features of Co-UGS and the characteristics of Co-UGS group users from both microcosmic and macrocosmic aspects; producing a valid prediction based on individual behavior and group activities. For quick service delivery and spread, recommend certain key user groups can be recommended to provide the best reference base.

In future work, we will investigate the Co-UGS in terms of two-way communication, considering its heavy load on a network path. In addition, the influence factors on the level of a created service will also be explored. In parallel with that exploration, the prototype platform of Co-UGS will be developed to allow for increased personalization and usability features.

6. ACKNOWLEDGMENTS

This work is funded by the National 3rd Key Program project "Research about Architecture of Mobile Internet" (No. 2011ZX03002-001-01) , and partially supported by National 3rd Key Program project No.2009ZX03002-003-04 and International Sci. & Tec. Program of China No.2010DFA12780.

7. REFERENCES

[1] J.Musser and Tim O'Reilly. 2006. Web 2.0:Principles and Best Practices. O'Reilly Radar.

[2] Z. Zhao, N. Laga, N. Crespi, "A Survey of User Generated Services. IEEE International Conference on Network Infrastructure and Digital Content" (IC-NIDC 2009), Beijing, China, Nov 6-8, 2009. DOI: 10.1109/ICNIDC.2009.5360953

[3] Z. Zhao, N. Laga and N. Crespi, "The Incoming Trends of End-user Driven Service Creation". Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (LNICST), Volume 21, pp. 98-108, 2009, Springer-Verlag. DOI: 10.1007/978-3-642-11532-5_12

[4] B. Gholami, and R. Safavi, "Harnessing Collective Intelligence: Wiki and Social Network from End-user Perspective", International Conference on e-Education, e-Business, e-Management and e-Learning, Sanya, China, January 22-January 24, 2009,pp. 242-246.

[5] Linton C. Freeman. "Centrality in social networks: conceptual clarification" [J]. Social Networks, 1979, 1: 215-239.

[6] Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

[7] Stanford Large Network Dataset Collection, http://snap.stanford.edu/data/wiki-Vote.html

[8] Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

[9] Borgatti, S.P., M.G. Everett, and L.C. Freeman. 1999. UCINET 6.0 Version 1.00. Natick: Analytic Technologie.