Social Network Extraction of Conference Participants

Hironori Tomobe
The University of Tokyo
Hongo 7-3-1, Tokyo 113-8656,
Japan
tomobe@miv.t.u-
tokyo.ac.jp

Yutaka Matsuo
National Institute of
Advanced Industrial Science
and Technology
Aomi 2-41-6, Tokyo 135-0064,
Japan
y.matsuo@carc.aist.go.jp

Kôiti Hasida
National Institute of
Advanced Industrial Science
and Technology
Aomi 2-41-6, Tokyo 135-0064,
Japan
hasida.k@carc.aist.go.jp

ABSTRACT

In a ubiquitous computing environment, it is desirable to provide a user with information depending on a user's situation, such as time, location, user behavior, and social context. At conventions, such as academic conferences and exhibitions, where participants must register in advance, the social context of participants can be extracted from the Web using their names and affiliations without asking them many questions. In our study, we attempt to extract the social network of participants from the Web, where a node represents a participant and an edge represents the relationship of two participants. The network shows the position of each participant and some participant clusters. Furthermore, this network can be used in many services, such as finding an appropriate person to introduce or negotiate with someone, and who one should talk to in order to expand his/her network efficiently.

Categories and Subject Descriptors

H.4.m [**Information Systems**]: Miscellaneous; D.2 [**Software**]: Software Engineering

Keywords

social network, search engine, personalized service

1. INTRODUCTION

In a ubiquitous computing environment [6], much information about users' behavior can be obtained by a sensor network. We seek to provide users with personalized information depending on the situation: time, location, and user behavior. Especially at conventions such as academic conferences, the social context of each user is very important because they gather to experience new encounters and exchange knowledge face-to-face.

Assume a participant at a conference wants to make friends with researchers with similar interests near his current location. A future ubiquitous environment might detect the user's location and recommend that the user talk to a certain person. However, without background knowledge about the social network, the system may recommend the user's colleague or supervisor because they share the same interests.

Utilizing background knowledge about participants' social network, many potential applications can be considered. Assume a user wants to talk to a certain person and wants someone to introduce her. With the help of social network knowledge, the system

can determine who is appropriate to introduce her. Conversely, one can find the path from herself to anyone with whom she might be talking. Another example might be efficient networking. A weak tie, which in social network theory is a connection between groups that don't ordinarily interact, play an important role in getting valuable information [2]. The system can suggest who may be a candidate for this weak tie; that is, one who shares the similar interests, but is in a different social group. Also, if one wishes, one could find who is appropriate to make a tie with in order to move you more centered in the network [1].

At academic conferences such as WWW2003, a participant must register a profile (at least name and affiliation) prior to the conference. In such cases, it is reasonable to assume that we have a list of participants and time to gather information about those participants from the Web. Referral Web [3] is a project to discover a social chain from an individual to the target person from the Web; however, in our case, fortunately we have a list of the names in advance, and try to discover the whole network structure among participants from the Web. Digital services for social events are not rare [5]. But our approach is different in that the system is conscious of the social network generated from information on the WWW.

Our system will serve at the 16th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI 2003), installed as a location-based information support system [4]. In this poster, we briefly overview the social network detection algorithm.

2. SOCIAL NETWORK EXTRACTION

We assume that names and affiliations of participants are given beforehand. (Although some participants appear without prior registration, we currently ignore these cases.) Therefore, nodes of the social network are invented first.

Next, edges between nodes are added utilizing Web information. The most simple approach is to measure relevance of two nodes based on the number of retrieved results by a search engine. For example, assume we are to measure the relevance of two names 'Yutaka Matsuo' (denoted X) and "Hironori Tomobe" (denoted Y). We first put a query "X and Y" to a search engine and get a documents including those words in the text. (We only need the number of matched documents, not the whole contents of the matched documents.) Also, we put a query "X or Y", and get b matched documents. The relevance of "Yutaka Matsuo" and "Hironori Tomobe" is approximated by the Jaccard coefficient $\#(X \cap Y)/\#(X \cup Y)$, say a divided by b.

If the Jaccard coefficient of a node pair is larger than the given threshold, an edge is added with its weight equal to the Jaccard coefficient. Some problems are:

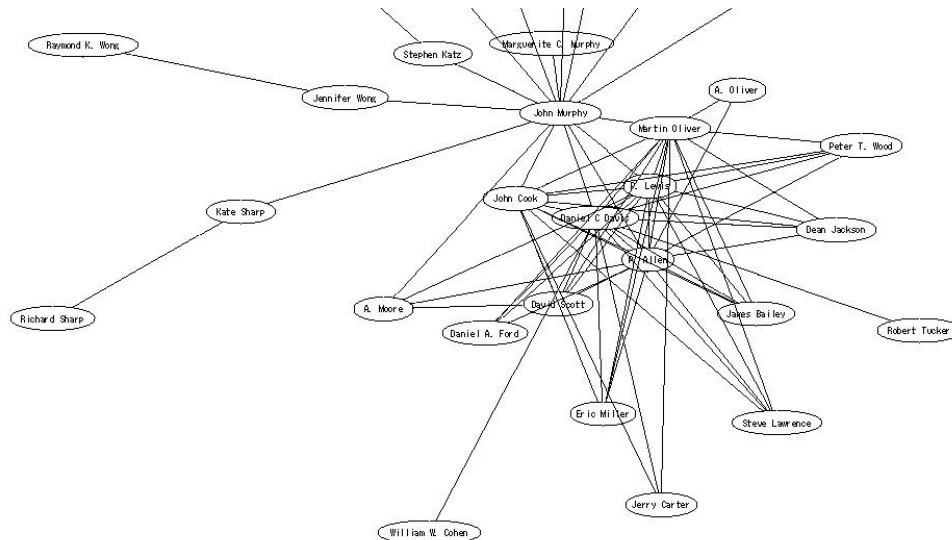- There can be more than one person with the same family and

Figure 1: Social Network of Contributors in WWW2002.

given name. Adding affiliation to the query will alleviate this problem, but degrade the coverage.

- The Jaccard coefficient generally gives a famous person few number of edges because denominator b is very large compared to numerator a. We can modify denominator b to $\min(\#X, \#Y)$. In some cases, it is more appropriate to employ a directed network representation instead of an undirected network, assigning conditional probability $\#(X \cap Y)/\#X$ to each directed edge.

It is more useful if information services can "label" each edge. For example, two nodes have the relation of "colleageus of the same company," "former supervisor – student," "members of the same board," and so on. To enable this, we use a machine learning approach to find the relationship of two nodes. We collect Web pages which show a certain relationship between two people and learn characteristic features of these pages.

3. EXAMPLE AND APPLICATION

Aside from providing a network structure for information service software, the current system outputs the network in SVG format. We can view the network using a SVG viewer.

Figure 1 is a part of the social network among contributors to WWW2002. Around 650 people contributed to the conference, therefore we have 650 nodes in the network.

We can see each participant's location in the network: clusters of participants and connecters of clusters. In the future, we may be able to show the distribution of research topics superimposed on the social network, how one can efficiently expand one's network, and the difference and characterization of conferences from the viewpoint of a social networks.

A various information services aware of the social network are possible, such as recommendation of a person who is distant from you in the network but has similar interests, and searching for an appropriate introducer or negotiator. Furthermore, location-based and text-based application will be enhanced with the help of social network knowledge, e.g., by refraining from suggesting a social network "neighbor" or associate because they already know each other well.

4. CONCLUSION

Japanese people are relatively sensitive to hierarchical social relations with others. Therefore, background knowledge of the social network is essential for providing more personalized information. However, asking questions about the relationship with others is very intrusive. Our approach seems promising in that the social network can be obtained from the Web, i.e, an open source of knowledge. This poster is a preliminary report; however, we will attempt to further develop an information system with consciousness of an individual's social context.

APPENDIX

A. ADDITIONAL AUTHORS

Additional authors: Mitsuru Ishizuka (The University of Tokyo, email: ishizuka@miv.t.u-tokyo.ac.jp).

B. REFERENCES

[1] L. C. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1:215–239, 1979.

[2] M. Granovetter. Strength of weak ties. *American Journal of Sociology*, 78:1360–1380, 1973.

[3] H. Kautz, B. Selman, and M. Shah. Referral web: Combining social networks and collaborative filtering. *Communications of the ACM*, 30, 1997.

[4] Takuichi Nishimura, Hideo Itoh, Yoshinobu Yamamoto, and Hideyuki Nakashima. A compact battery-less information terminal (cobit) for location-based support systems. In *Proceedings SPIE*, number 4863B-12, 2002.

[5] Yasuyuki Sumi and Kenji Mase. Digital assistant for supporting conference participants: An attempt to combine mobile, ubiquitous and web computing. In *Proceedings UBICOMP 2001 (LNCS 2201)*, pages 156–175, 2001.

[6] M. Weiser. The computer for the twenty-first century. *Scientific American*, 268(3):94–104, 1991.