

# An Integrated Method for Social Network Extraction

Masahiro Hamasaki  
National Institute of Advanced  
Industrial Science and  
Technology (AIST)  
2-41-6 Aomi, Tokyo, Japan  
hamasaki@ni.aist.go.jp

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

Keisuke Ishida  
National Institute of Advanced  
Industrial Science and  
Technology (AIST)  
2-41-6 Aomi, Tokyo, Japan  
ksk-ishida@aist.go.jp

Tom Hope  
National Institute of Advanced  
Industrial Science and  
Technology (AIST)  
2-41-6 Aomi, Tokyo, Japan  
tom-hope@aist.go.jp

Takuichi Nishimura  
National Institute of Advanced  
Industrial Science and  
Technology (AIST)  
2-41-6 Aomi, Tokyo, Japan  
taku@ni.aist.go.jp

Hideaki Takeda  
National Institute of  
Informatics (NII)  
2-1-2 Hitotsubashi,  
Chiyoda-ku, Tokyo, Japan  
takeda@nii.ac.jp

## ABSTRACT

A social network can become bases for information infrastructure in the future. It is important to extract social networks that are not biased. Providing a simple means for users to register their social relation is also important. We propose a method that combines various approaches to extract social networks. Especially, three kinds of networks are extracted; user-registered *Know link* network, Web-mined *Web link* network, and face-to-face *Touch link* network. In this paper, the combination of social network extraction for communities is described, and the analysis on the extracted social networks is shown.

## Categories and Subject Descriptors

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

## General Terms

Algorithms, Experimentation

## Keywords

social network, Web mining, user interaction

## 1. INTRODUCTION

This paper presents an integrated method for social network extraction. Social networks play important roles in our daily lives. Our lives are enormously influenced by social networks without our knowledge of the implications, and there are many applications using social networks [6]. In the context of the Semantic Web, social networks are crucial to realize a web of trust, which enables the estimation of information credibility and trustworthiness [2]. Ontology construction is also related to a social network [4]. Also in

studies on computer-mediated communication, social networks have been examined with keen interest.

Social networking services (SNSs) have become popular. Friendster and Orkut are among the earliest and most successful SNSs. An interesting point of SNSs is that a user can see not only others' personal attributes (e.g., name and affiliation, hobby) but also social networks (acquaintance lists). Acquaintance lists reveal information about users' personalities. On the other hand, acquaintance lists are proof that acquaintances can track activity within the SNS. We can expect that the purview of acquaintances restrict deviant behavior within a community (e.g., assume a false name, abusive language). An SNS that manages and stores social networks can become a base of information infrastructure in the future. Social networks should reflect actual social networks in that community and should be unbiased if we use SNS as an information infrastructure of one community.

There are several ways to obtain social networks: FOAF (Friend-of-a-Friend) is a vocabulary to describe information on a person and the relation to others. Both SNS data and FOAF data is created by users themselves. On the other hand, automatic detection of relation is also possible from various sources of online information such as e-mail archives, schedule data, and Web citation information [1]. There is another stream of work to obtain social networks; observing persons' behaviors in ubiquitous and wearable devices [5].

Whichever method we will take for obtaining a social network, it suffers from some degree of flaws: For example, SNSs data and FOAF data, which is based on self-report survey, suffers from data bias and sparsity. Automatically obtained network, e.g., Web-mined social network, brings us a good view of prominent persons, but does not work well for novices, students, and other "normal" people. Social networks observed using wearable devices is constrained by device-specific characteristics; it may have detection errors, limitation of detection scopes, and the bias of usage by users.

This paper describes our attempt to combine three ways to extract social networks and analyzes obtained social networks.

## 2. INTEGRATION OF SOCIAL NETWORK EXTRACTION

Figure 1 shows our model to extract and combine social networks. We take three approaches: (1) The first is based on web mining techniques. It can create initial networks automatically from available web information. (2) The second is based on real-world user interaction (e.g. face-to-face communication) in communities. It extracts social network with capturing user interactions. (3) The last is based on user interaction on the Web system, similarly as in SNSs. Users can describe own social network by themselves.

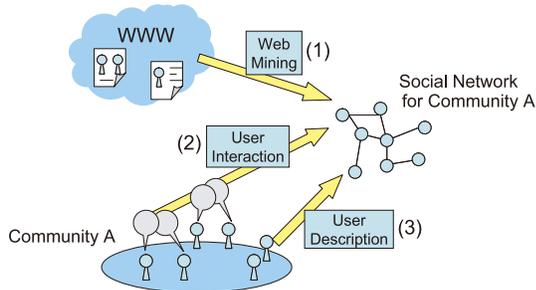


Figure 1: Illustration of the integrated method for social network extraction.

We implement this model for community support system which targets an academic conference. We apply Web mining techniques using a search engine [3] to extract *Web links* (1). We settle information kiosks at the conference site, and deliver name cards (IC cards) for participants. When users put their IC cards at an information kiosk, we regard them as connected by *Touch links* (2). Users can build own acquaintance lists. We regard the registered acquaintances as connected by *Know links* (3).

Figure 2 is a screenshot of our system. When two users put name cards on an information kiosk, they can see a social network including the two.

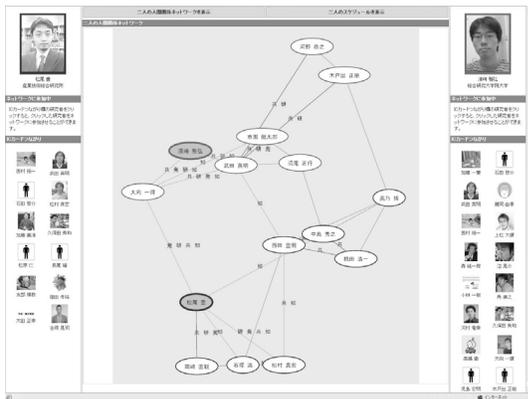


Figure 2: Social network among two persons.

## 3. FIELD TEST

We operated our system at a three-day Japanese domestic conference. That conference had 297 presentations and

579 authors (including co-authors). About 500 participants joined that conference.

One of the interesting findings through our analysis is: Figure 3 shows the number of Web hit (by putting a person's name to a search engine) versus the number of three kinds of links. The more authoritative people (with lots of hit count) tend to have more number of *Web links*. While the most authoritative people do not use *Knows links* the most; active middle-authoritative users use the most. They may know well about the community, and feel interesting. *Touch links* are used by the less authoritative users more; especially, the persons with the same level of authoritativeness are likely to have meets link. It is natural because persons who have fewer acquaintances want more acquaintances, and people are likely to meets people with the same social level.

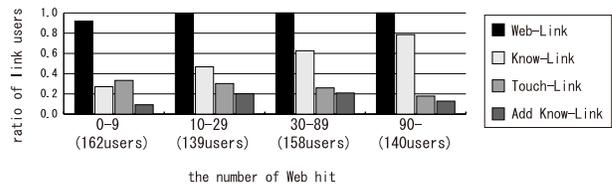


Figure 3: Ratio of Link Users

'Know-Link' means users who have undirected *Know link* and 'Add Know-Link' means users who added *Know link*.

## 4. CONCLUSIONS

In this paper, we proposed an integrated method to extract social networks and implemented it to the conference support system. It can realize extraction of social networks that are not biased and provide a simple means to register social networks for users. It is important to promote communication and discussion among community members. Future works include improvement of network integration and advanced applications.

## 5. ACKNOWLEDGMENTS

This research has been supported by NEDO (New Energy and Industrial Technology Development Organization) as the project ID of 04A11502a.

## 6. REFERENCES

- [1] L. A. Adamic and E. Adar. Friends and neighbors on the web. *Social Networks*, 25(3):211–230, 2003.
- [2] J. Golbeck and J. Hendler. Inferring trust relationships in web-based social networks. *ACM Transactions on Internet Technology*, 2005.
- [3] Y. Matsuo, H. Tomobe, K. Hasida, and M. Ishizuka. Finding social network for trust calculation. In *Proceedings of 16th European Conference on Artificial Intelligence (ECAI2004)*, pages 510–514, 2004.
- [4] P. Mika. Ontologies are us: A unified model of social networks and semantics. In *Proceedings of ISWC2005*, 2005.
- [5] A. Pentland. Socially aware computation and communication. *IEEE Computer*, 2005.
- [6] S. Staab, P. Dmingos, T. Finin, P. Mika, A. Joshi, J. Golbeck, A. Nowak, L. Ding, and R. R. Vallicher. Social network applied. *IEEE Intelligent systems*, pages 80–93, 2005.