

A study on reliable social network metrics

Sanghyun Noh, Sangyong Han

Abstract—Social Networks express information flows between individuals or groups using relationships between users. These social networks are becoming gradually huge and many related problems are coming to the fore. In this study, problems that become issues among these problems of social networks are examined and mentioned and the necessity of solving the problems is dealt. In addition, measures to solve the problems and future study directions are presented in order to establish reliable social networks.

Index Terms—Social network, Home network, Access control, Trust model, Sparsity problem

I. INTRODUCTION

SOCIAL networks express the concept of psychological and social relationships between individuals or groups in the form of graphs using actors and relations[1],[2]. Since people use these social networks based on Internet to manage their interests or profiles or share those with people with close relations, Internet based social networks are gradually becoming huge [2],[3]. As social networks become huge, many problems occur and they are pointed out. However, most studies of social networks are focusing on analyzing graph based social networks rather than solving these problems [4].

Recently, as Social Network Service (SNS) has been spotlighted, diverse studies have been conducted in relation to recommendation systems or information retrieval systems using Social Networks [5]. However, these systems have not been able to reliable information as the reliability or accuracy of the recommender or the informant has been low and unwanted information leaks have been naturally occurring through reliability based social networks.

The first causes of the problems occurred is the inaccuracy of the trust scores of relations between actors connected directly or indirectly. On social networks, trust scores are universally used as a yardstick to judge others' reliability for online transactions. Therefore, accurate trust management models are necessary for online transactions to be as reliable as offline transactions.

The second cause includes the sparsity problem and data deluge problem in social networks established in the real world that occur because the relations between actors are sparse or

excessive. If the relations are sparse, channels to obtain information are insufficient and this will threaten the diversity of information while if the relations are excessive on the contrary, information not related to the information wanted will overflow to reduce the accuracy of information.

The final cause is the information access problem that even information not wanted to be shared between individuals or groups is easily leaked through social networks. If a user provides information to a subject trusted by the user, that may not be the end of the information flow but the subject may share the user's personal information with another subject trusted by the first subject based on his criteria.

In this paper, the problems in social networks presented above and incidental problems that may consecutively occur due to the problems will be described and future study directions intended to solve these problems will be mentioned.

This paper is composed as follows. In Section2-A, a problem of reduced reliability in social networks arising from the inaccuracy of trust scores will be mentioned and in Section2-B, problems of acquisitions of wrong information and reduced diversity of information resulting from the Data deluge problem and sparsity problem. In Section2-C, the problem of information leaks in social networks will be covered. Finally, in Section 3, future study directions to solve these problems will be covered.

II. PROBLEMS IN SOCIAL NETWORKS AND THE NECESSITY TO SOLVE THE PROBLEMS

A. The problem of reduced reliability of social networks due to the inaccuracy of Trust Scores

By utilizing Social Networks, a user can obtain necessary information through acquaintance or acquaintance's acquaintance. This is advantageous in that the user can obtain more accurate information than receiving information from a person who is not known to the user at all. As shown in fig.1, although the user can accurately judge the trust score of the acquaintance known to user. However, there are difficulties in defining the trust score of the acquaintance's acquaintance.

Sanghyun Noh is with School of Computer Science and Engineering, Chung-Ang University, 221, Heuk seok-dong, Dongjak-gu, Seoul 156-756, Korea(corresponding author to provide phone: 02-824-1187; fax: 02-824-1187; e-mail: shnoh@ec.cse.cau.ac.kr).

Sangyong Han is with School of Computer Science and Engineering, Chung-Ang University, 221, Heuk seok-dong, Dongjak-gu, Seoul 156-756, Korea, He is Corresponding Author (e-mail: hansy@cau.ac.kr).

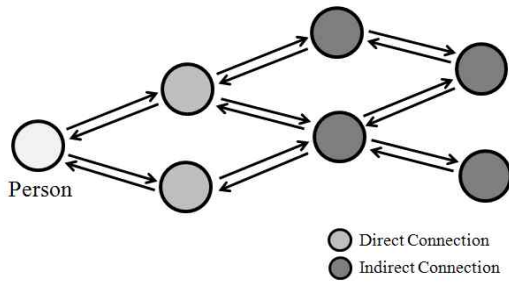


Fig. 1. Schematization of direct connections and indirect connections in social networks

Although the trust scores of directly connected relations are accurate, there exist diverse problems occurring due to the inaccuracy of the trust scores of indirectly connected information sources. Trust Models in social networks normally become yardsticks to judge the other parties of transactions whether they can be trusted or not in recommendation systems or online transactions. Indeed, in the case of trust score based online transactions; the transactions may be accomplished or cancelled depending on trust scores. Therefore, studies of models to calculate trust scores that will guarantee reliability and accuracy are urgently needed. Existing trust score calculation models calculate the trust scores of indirect connections simply using the trust scores of the subject directly connected to the user as a factor and thus errors frequently occur in the process of calculation and therefore those models are not very suitable for reliable online transactions.

Therefore, studies to develop trust management models intended to enhance the accuracy of the trust scores of indirect connections should be conducted for reliable online transactions.

B. Data deluge & Sparsity problems in social networks

Ideal social networks include ample relations between actors and diverse kinds of information. However, social networks established in reality have a problem that users become to refer to information even when trust scores between actors are very low to the extent to be negligible for the reason that relations exist. This problem becomes the cause of the data deluge problem of being provided with even data with little association with the data being searched for. In addition, those social networks also have a problem that the number of information sources and the amount of information obtained are decreased because relations between actors are sparse on the contrary as shown in Fig.2. These problems are called sparsity problems.

Data deluges in social networks cause problems of being impossible to find actually necessary data or taking much time to find necessary data because all kinds of information including even those not much related with users are provided. A measure to solve this problem is filtering relations between actors.

The sparsity problem in social networks is a concept opposite to data deluges and this is a factor that hinders the diversity of information because sources of information are limited due to the sparsity of relations. For users to obtain diverse and accurate information on the Web, information sources should be diverse

while the creditability and accuracy of information should be guaranteed. To this end, diverse relations on social networks are necessary. Methods to enrich relations on establish social networks are necessary to expand relations. Existing studies to solve sparsity problems create relations simply by using the values of similarity between actors [6][7]. A limitation of these studies is that they don't have any clear definition and range of the relations created as such.

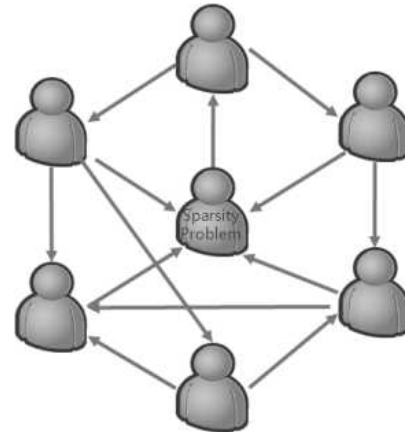


Fig.2. Sparsity problem on social networks

Therefore, unlike existing studies, we should study not only similarity values between users but also measures to reduce or expand relations including factors that show the characters of users such as users' activities on the Web.

C. Problems of information leaks through social networks

Sharing information on social networks involves a risk that the information may be implicitly shared by even people not wanted by the user. This occurs because the information is delivered through the subject trusted by the user and given the information to other subjects trusted by the first subject.

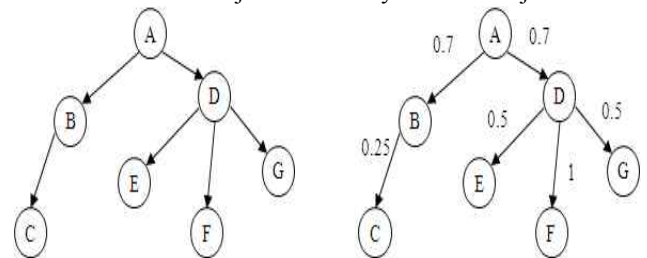


Fig. 3. Degrees of the risk of information leaks in relation to reliability on social networks

Fig.3. shows trust scores on a social network. A trusts B and D to the same degree of 0.7. B has relations with only user C and thus even if A's information is leaked, it will be leaked to only C. However, D has relations with three users, E, F and G and thus, in the worst case, information may leak toward unwanted three indirectly connected users and therefore, it can be said that D has a higher risk of information leaks than B.

Therefore, the transfer and sharing of information through social networks should be judged based on a combination of the risk of information leaks and trust scores rather than being judged by simply using trust scores.

Later, it will become crucial to use the calculation models to calculate the risk of information leaks used in access control as well as trust models and combine the results in order to derive finally reliable values.

III. STUDY DIRECTIONS TO SOLVE THE PROBLEMS

Researches should be conducted in order to enable establishing initial social networks by establishing relations and relation trust scores based on the similarity of the attributes of actors. After that, methods to newly compose relations through filtering for solving data deluge problems and relation enrichment for solving sparsity problems should be studied. Once these problems have been solved, social networks will be able to guarantee enhanced reliability based on more accurate trust scores.

In addition, models to calculate trust scores with indirectly connected actors should also be studied. It is considered that these models will be able to derive accurate results only by using both “the degree of trusting” and “the degree of not trusting” and giving weighted values on the trust scores of directly connected relations.

Reliable access control methods that will blend the trust scores obtained as such with the risk of information leaks should be also studied.

IV. CONCLUSION

In this paper, problems arising in social networks that are gradually becoming huge along with the development of Internet are described and future directions of studies intended to solve the problems are presented. Although these problems have already been pointed out in many papers, studies to solve these problems are still insufficient. Specifically, in social networks exist the problem of ambiguity of indirect connection trust score calculation models, the problem of inaccuracy arising from data deluge problems or sparsity problems and the problem of individuals' or groups' information leaks.

If the problem of ambiguity of trust scores, data deluge problems and sparsity problems can be controlled in order to solve occurring problems, it will become possible to establish more reliable social networks. In addition, it is expected that if the combination of the trust scores established as such and the risk of information leaks is utilized in access control, information leak problems could be improved.

In future, studies in the direction of dealing and solving these problems should be conducted and it is considered that the relevant studies can bring about the development of related areas through solving representative problems in social networks.

ACKNOWLEDGMENT

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the HNRC (Home Network Research Center)—ITRC (Information Technology

Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2009-C1090-0902-0035) and also partially supported by National Research Foundation of Korea Grant funded by the Korean Government (2010-0015799).

REFERENCES

- [1] Adar, E.; Re, C. "Managing Uncertainty in Social Networks" *Data Engineering Bulletin* 2007, 30, 23-31.
- [2] Borgatti, S.P.; Mehra, A.; Brass, D.J.; Labianca, G. "Network Analysis in the Social Sciences". *Science* 2009, 323, 892-895.
- [3] Stefan Weiss, "Online Social Network and the Need for New Privacy Research in Information and Communication Technology", Third International Summer School organized by IFIP WG 9.2, 9.6/117, 11.6, 6th-10th August, 2007, Sweden
- [4] Golbeck, J. "Trust and Nuanced Profile Similarity in Online Social Networks". *Acm T Web* 2009, 3, -.
- [5] Golbeck, J. "FilmTrust: Movie Recommendations using Trust in Web-based Social Networks", CCNC'2006, IEEE 2006 3rd publication, 2006
- [6] Huang, Z.; Zeng, D.; Chen, H., A Link analysis approach to recommendation under sparse data. In *the Tenth Americas Conference on Information Systems*, New York, 2004; pp 1-9.
- [7] Papagelis, M.; Plexousakis, D.; Kutsuras, T. Alleviating the sparsity problem of collaborative filtering using trust inferences. *Trust Management, Proceedings* 2005, 3477, 224-239.

Sanghyun Noh received B.S degrees in computer science and engineering from ChungAng University, South Korea, in 2009. After get a B.S degree, he is currently taking a graduate course at ChungAng University in South Korea.

He takes part in home network research under the HNRC (Home Network Research Center)—ITRC (Information Technology Research Center) supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2009-C1090-0902-0035) and also join in social network research supervised by National Research Foundation of Korea Grant funded by the Korean Government (2009-0076290).

Dr. Sangyong Han is a professor of the School of Computer Science and Engineering of the Chung-Ang University in Seoul, Korea. He received a Bachelor of Science degree in Engineering from the Seoul National University and a Doctorate in Computer Engineering from the University of Minnesota. In 1984, he joined IBM Laboratory in Kingston, NY and later became a member of IBM Watson Research Center when he left a company in 1995. During his tenure in IBM, he served as an adjunctive professor of RPI in Troy, NY for three years and was a visiting scholar of UC Berkeley for one year.

He joined the Chung-Ang University in 1995, where he is currently a professor of the School of Computer Science and Engineering. He served as a chair person of the Society of Korea Software Appraisal and Evaluation for two years and a director of the Information and Telecommunication Research Institute of the Chung-Ang University. He has published more than 100 papers in scientific journal and conferences including more than 30 journal papers indexed by SCI (or SCIE). He also registered 10 domestic and international patents. His research interests are in the area of next generation web technologies including web services and semantic web, information retrieval and optimization algorithms.

Dr. Han is vigorously involved in a series of international conferences. He served as General Co-Chair of NWeSP'05, NWeSP'06, NWeSP'07, NWeSP'08, NWeSP'09, NWeSP'10, CiCLing2004 and was a General Chair of the first Conference of KSEA. He also served as an Editor-in-Chief of International Journal of Web Services and Practices and was a guest editor of Journal of Digital Information Management. Dr. Han is also included in Marquis "Who's Who 2009 edition".