# IEICE Proceeding Series

Emergence of Hierarchical Small-world Property in SNS for College Students

Atsushi Tanaka, Masaki Tomochi

Vol. 1 pp. 98-101
Publication Date: 2014/03/17
Online ISSN: 2188-5079

2012 International Symposium on Nonlinear Theory and its Applications
NOLTA2012, Palma, Majorca, Spain, October 22-26, 2012

# Emergence of Hierarchical Small-world Property in SNS for College Students

Atsushi Tanaka† and Masaki Tomochi‡

†Graduate School of Science and Engineering, Yamagata University
4-3-16 Jonan, Yonezawa City, Yamagata, 992-8510 Japan
‡College of Economics and Environmental Policy, Okinawa International University
2-6-1 Ginowan, Ginowan City, Okinawa, 901-2701 Japan
Email: tanaka@yamagata-u.ac.jp, mtomochi@okiu.ac.jp

**Abstract**—A number of studies on social networks like SNS have been focused on their statistical properties. In order to analyze dynamics of social networks, we have constructed a special SNS site only for college students in several areas in Japan and investigated the emergence of new relationship between those students. Since our network is based on seminar groups that belong to each college, the emergence of nested small-world property is naturally observed in each hierarchical level. To clarify the community dynamics, we neglect their personal affiliation attributes and analyze the network based on community analysis techniques, especially overlapping clustering. As a result, we observe dynamics of nested hierarchical structures and the fusion and disruption among communities.

## 1. Introduction

Recent development of social media like SNS, Twitter, Facebook, etc. is remarkable. Networks in SNS that are created in virtual world based on the connection in real world are growing, and they have attracted attention because it is suitable for studying the dynamics of community formation and disruption. For example, in the study of Mixi, the largest SNS in Japan, the size of communities is categorized into 3 areas and its distribution has a power law dependence, though, there is a large gap in the intermediate area [1]. On the other hand, in the study of categorizing large amount of SNS of various size with some network characteristics, it is categorized into 5 types [2]. Moreover, if the size of communities is categorized into 4 types according to the features of users' behavior, differences in activities among the users can be observed, and the mechanism is validated by the multiagent model [3]. In order to analyze the network dynamics in social media, we have constructed an original SNS and observed the transformation of the network for two years. In the research of this SNS, some small-world communities connect each other because of the users' social and geometric distances, thus the nested structure of small-world is observed. Hence the structure proposed by fractal $\beta$ model [5] is confirmed. In this paper, through analyzing more objective network data with community analysis, it is confirmed there exists nested structure of small-world.

## 2. Small-world property

We all sometimes find "it's a small world" in our daily lives. Such a small-world property has been the subject in social science for a long time, and above all, the chain-letter experiment, that is called small-world method, by Milgram et al. is famous [6–8]. Their experiment was to send a letter to an acquaintance to reach a targeted person who is an unknown to a starter, and through this experiment, they obtained the famous result that is called "six degree of separations". Kleinfeld raised questions about their result because there seemed to be some problems in their experiment [9, 10]. Those problems are as follows.

1. There is a bias in the selection of examinees, thus results are biased.
2. The number of completed chain latters is too few.

From these considerations, suggesting that there exist effects of spatial and social distance on a friendship network, she proposed an idea, that is, "the ʻlumpy oatmealʼ theory", that we live in a world with many small worlds possibly but not necessarily connected, might be viewed as the "weak" form of the small world phenomenon, for which we do have evidence. Hence Kleinfeld gave a very suggestive idea that the network in human society consists of the combination of small-worlds which need not to be always connected. We empathize her idea very much, and it is advocated in this paper as a result.

Speaking of the verification by models, $\beta$ model by Watts et al. was epoch-making [11]. They started from $k$-connected regular graph, and randomly rewired the links with probability $p$, then succeeded in decreasing the average path length $L$ drastically with small $p(0.01 \leq p \leq 0.1)$ keeping their cluster property. What make $L$ small is the effect of shortcuts emerged by rewiring and they must be needed for small-world property. Though several derivative models have been proposed after that, they all supposed the uniformity of nodes basically, thus the social distance and geometric effects are ignored.

As opposed to it, Tomochi proposed fractal $\beta$ and its derivative model which reflected the influences of geometric and social distances between nodes for edge rewiring probability [5]. As a result, the small-world network which consists of several connected small-worlds is created, and the nested structure of small-worlds is discussed.

## 3. Research of SNS and community analysis

The research of SNS, one of the social media, deals with the growing networks extending from real across virtual world, and it is suitable for studying the network dynamics. The networks in SNS strongly depend on human relationship, thus the community analysis is indispensable. The community analysis itself has been studied for a long time especially in social science, and several kinds of methods have been proposed so far. As typical examples, there are the method using betweenness centrality by Girvan et al. [12], the one using *Q* value by Newman [13], and its developed version for larger-scale networks with more speed by Clauset et al. [14]. However, they are all for partitioning the large community into some distinct smaller communities. If we consider human relationship in a real world, there often happens that some people gets deeply involved in several communities at the same time. Therefore, if we consider the network in social media, it is rather natural to think of overlapping communities, and hence we use overlapping community analysis. Among several kinds of these techniques that have already proposed, the Clique Percolation Method (CPM) is the most popular and its implemented software has been available. Hence we analyze the network mainly with this method.

## 4. Results and discussions

In order to investigate the network dynamics in SNS, we have constructed SNS site a.k.a. “tomocom.jp”. This SNS is only for college students and completely invitation-based where seminar students are invited by college teachers in several areas in Japan. It possesses higher reliability thanks to the guardian system. This site was opened in 2009, and captured more than 400 users by the end of 2010. Though interaction within their own seminars is basic, the connections between different seminars have been created by writing and browsing blogs. Since some revitalizations of the site were held, the connections became denser. The network structure of this SNS at the end of 2010 is shown in Fig.1.

As we mentioned, since this SNS network is based on seminars, there exist premised communities. In each seminar or community, students who belong to different years and sexes are mixed. As matter of course, the connection between same years tend to be stronger than between different years and the same goes for the same and different sexes, and thus the hierarchical network structure in each level can be expected. In fact, it is found that these connections are small-world like and it is also advocated by the simplified model of friendship creation [5].

However, one might consider this result as a natural consequence because of geometric and social distances between the students in the SNS. Hence, if the community analysis of a simple network data set without their attributes is conducted, then we can study the dynamics between communities. At first, the result of analysis using CNM method [14] which is a representative community analysis method is shown in Fig.2. The solid line shows the change of *Q* values in the community partitioning, which gives us the estimation of partitioning. The dotted line is the change of the rates of concordance between extracted and real communities, which give us effective estimation. In this method, those nodes that belong to two or more communities are not permitted. However, in the real world, there are some nodes which have active connections to several communities at the same time, and it is not clear which community they should belong to by this simple data analysis. Such nodes could belong to other community using other partitioning method. This is the limitation of this method.

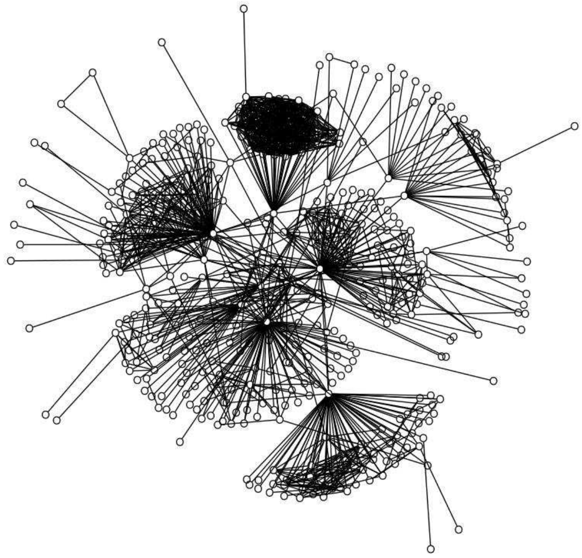

Figure 1: The network of tomocom.jp. We can find each seminar creates its own community.

Hence it is better to use community partitioning methods which allow nodes to belong to several communities, and we use CPM method, one of such a representative methods [15]. This method allows us not only to have overlap nodes but also to obtain the number of communities through data analysis. Since it is based on finding and expanding cliques, even though the whole number of communities decreases in accordance with increasing *k*, it is still possible to find denser communities(Fig.3).

If we fix *k* and investigate the number of communities over time, it is not constant and we can see communities which fuse and disrupt repeatedly(Fig.4). Now we define a connection between two communities iff there is one or more edges between these two communities. It allows us create new network in community level, and some clusters between communities can be observed, that indicates the nested structure of small-world property(Fig.6). In fact, *L* and *C* are almost constant over time, that means the small-world property in community level is confirmed(Fig.7,8).

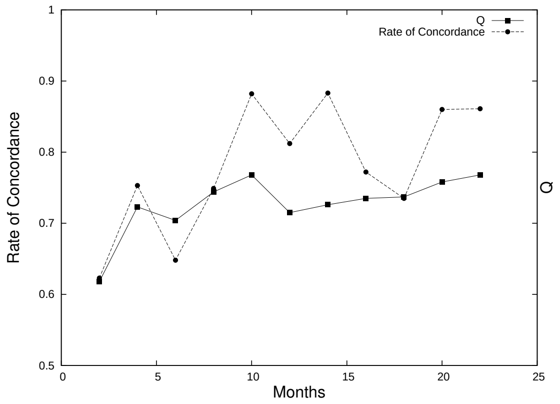


Figure 2: Result of community analysis using CNM method.The solid and dotted lines show the changes of $Q$ value and the rate of concordance between extracted and real communities, respectively.

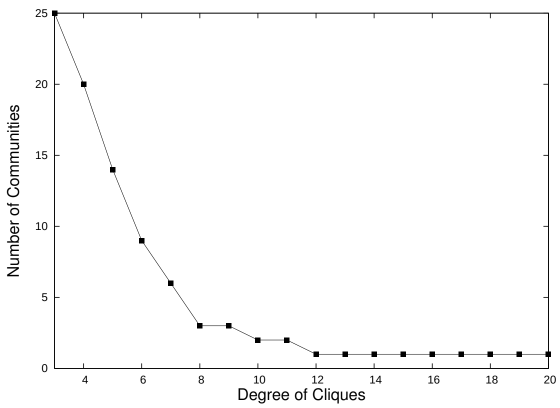


Figure 3: Change of the number of communities on the clique degree $k$. The larger $k$ becomes, the smaller the number of communities becomes. It agrees with the real number at $k = 6$.

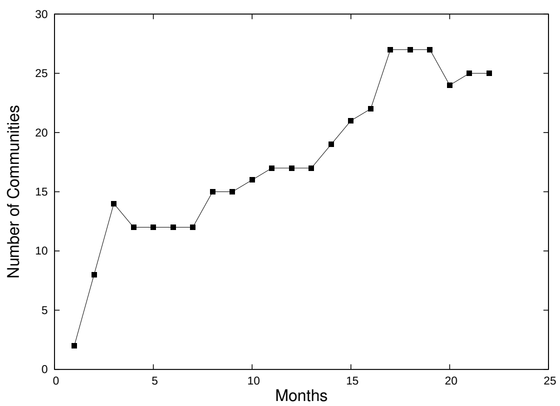


Figure 4: The number of communities over time. This result is obtained using CPM method with $k = 3$.

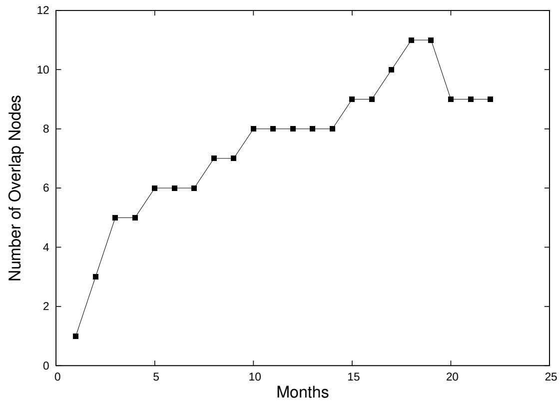


Figure 5: The number of overlap nodes over time. It increases with time, but decreases around 20 months.

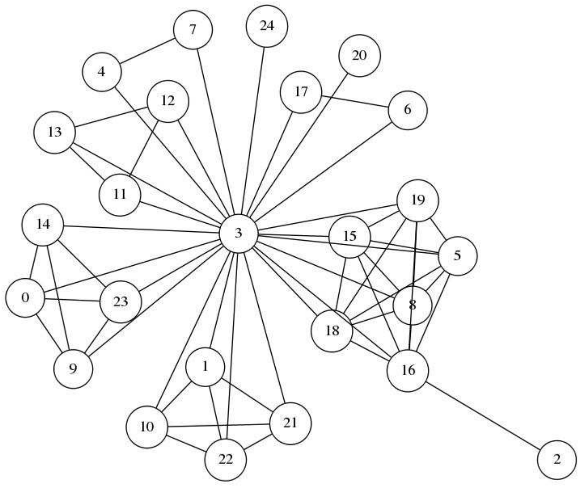


Figure 6: The network structure in community level. If we create a network regarding one community as one node, a similar network can be observed.

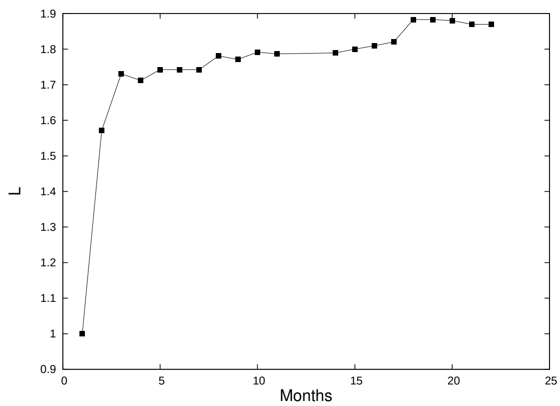


Figure 7: The values of $L$ in community level over time. They are almost constant regardless of the growth of the network.

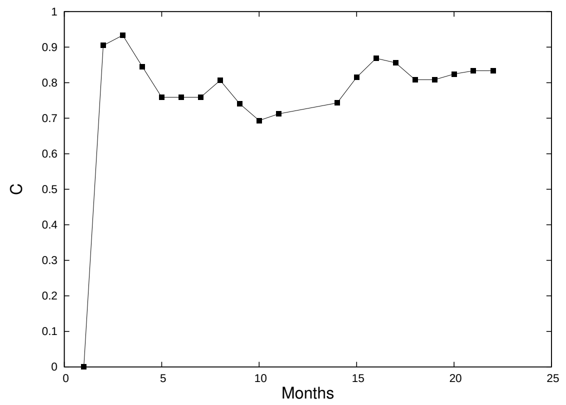


Figure 8: The values of $C$ in community level over time. They are almost constant just like $L$.

## 5. Conclusion

We have constructed SNS site for college students in several areas in Japan and investigated the network dynamics of human relationship. In that network, the emergence of small-world property was observed in each level between different sexes, years and seminars. Since this result is considered to be a natural consequence of geometric and social distances determined by their attributes, we ignored these attributes and conducted analysis on simple network data. It gives us an important suggestion for analyzing the dynamics of general SNS networks.

As a result, with the ordinal community analysis, we observed some nodes which belong to different communities in different time and so communities which change over time. We also carried out the community analysis that includes overlap nodes to clarify the dynamics, and it revealed the changes of the numbers of overlap nodes and communities in detail. Moreover, the small-world property in community level was confirmed, thus the nested structure of small-world property was also confirmed in data level.

For the future works, firstly, we will observe dynamics of the network over longer periods of time, secondly, explain the cause of diffusion and disruption of communities clearly by eliminating super nodes, i.e. teacher nodes, and thirdly, analyze them from the viewpoint of the information propagation. This research is supported by Grant-in-Aid for Scientific Research on Innovative Areas(20200042).

## References

[1] K. Yuta, N. Ono and Y. Fujiwara, "Structural Analysis of Human Network in Social Networking Services."*IPSJ Journal*, Vol.47, No.3, pp.865-874, 2006.

[2] F. Toriumi, H. Yamamoto, H. Suwa, I. Okada, K. Izumi and Y. Hashimoto, "Comparison Analysis among Large Amount of SNS Sites."*Transactions of the Japanese Society for Artificial Intelligence*, Vol.25, No.1, pp.78-89, 2010.

[3] F. Toriumi and K. Ishii, "Simulation of Encouragement Methods for SNS based on User Behavior Model."*Proc. of 3rd World Congress on Social Simulation*, 2010.

[4] M. Tomochi, A. Tanaka and T. Shichijo, "Stratification and Nested Structure of Small World in a Friendship Network. -Data Analysis, Modeling, and Simulation on a Social Networking Service a.k.a."Tomocom"."*Sociological Theory and Methods*, Vol.26, No.1, pp.83–97, 2011.

[5] M. Tomochi, "A model of a nested small-world network."*Sociological Theory and Methods*, Vol.25, No.1, pp.19-29, 2010.

[6] S. Milgram, "The Small World Problem."*Psychology Today*, Vol.1, pp.61–67, 1967.

[7] J. Travers and S. Milgram, "An experimental study of the small world problem."*Sociometry*, Vol.32, pp.425–443, 1969.

[8] C. Korte and S. Milgram, "Acquaintance links between White and Negro populations: Application of the small world method."*J. Personality and Social Psychology*, Vol.15, No.2 pp.101-108, 1970.

[9] J. S. Kleinfeld,"Six Degrees of Separation: An Urban Myth?" *Psychology Today* Vol.35, No.2, pp.74, 2002.

[10] J. S. Kleinfeld,"The small world problem." *Society*, Vol.39, No.2, pp.61–66, 2002.

[11] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks." *Nature*, Vol.393, pp.440–442, 1998.

[12] M. Girvan, M. and M. E. J. Newman, "Community structure in social and biological networks." *Proc. Natl. Acad. Sci. U.S.A.*, Vol.99, No.12, pp.7821–7826, 2002.

[13] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, Vol.69, 066133, 2004.

[14] A. Clauset, M. E. J. Newman and C. Moore, "Finding community structure in very large networks," *Physical Review E*, Vol.70, 066111, 2004.

[15] G. Palla, I. Derényi, I. Farkas and T. Vicsek, "Uncovering the overlapping modular structure of protein interaction networks." *FEBS JOURNAL* 272: pp.434–434 Suppl. 1, 2005.