1H3-3

A Network Generation Model by Layer Aggregation based on the Multilayer Network

Chao Fan[*1] Fujio Toriumi[*1]

[*1] Graduate School of Engineering, The University of Tokyo

Multilayer network has attracted considerable attention recently and its research has now become one of the most important directions in network science, but network generation model based on multilayer network does not catch many researchers' eyes in the past few years. In this paper, we propose a new network generation model by layer aggregation in the background of the multilayer network. The idea presented here is quite simple. Each community is regarded as a single layer of multilayer network and one node can exist in many layers, because each person belongs to several communities, such as family, school, interest group, business organization, etc. The network appearing on social network service (SNS) is a network generated from combination of all communities. That is to say, social network is generated from a multilayer network. In this paper, we proposed a new network generation model to represent such features. We applied the proposed model to a number of experiments to imitate a real-world network. Finally, experimental result confirms the effectiveness of our model.

1. Introduction

Social networking sites are gaining increasing popularity in people's daily life over the last few years. Many online social networks such as Facebook, Twitter, and LinkedIn have emerged as a new form of media (aka. social media) that facilitate interaction among people and dissemination of public information. Consequently, an increasing number of researchers are focusing on the structure and generation model of such social network [Newman 04, Newman 06, Pasta 13, Shinoda 08].

However, classical analysis and generation model reckon social network as a simple graph with nodes and links, which may escape many important features and much information behind a real-world social network. Because a person often interacts with others with different relationship and each relationship could form a layer of connectivity. In a word, a good network model should include "multi-layer" structure to represent a real-world social network.

Multilayer network's research is enjoying a boom and now has become one of the most important directions in network science [Kivelä 14]. Nevertheless, a network generation model utilizing multilayer network is rare in recent research. In this paper, we propose a new network generation model by aggregating different layers based on multilayer network. Four parameters of model are adjusted to display different performance. Furthermore, we acquire a network existing most similar features of the real-world network with our model.

2. Network Model

Social Network is regarded as a multilayer network. Each community that one person belongs to can be assumed as a single layer of multilayer network. Each node can exist in many layers, for a person often belongs to different communities, such as family, school, interest group, business organization, etc. Thus, the network obtained as social network is a superimposed network of each community.

To represent such features of network, we propose a multilayer-network model to obtain a network, which is shown in Figure 1. The upper ones are different communities of social network, each of which can represent a layer of multilayer network. One person often belongs to several communities, so we can use two nodes to represent the same entity in different layers as Figure 1. By aggregating all layers into one network, we are able to obtain an integrated network (the inferior one in Figure 1), which is often a general form of real-world data in online social network service, such as Facebook, Twitter, LinkedIn and so forth.

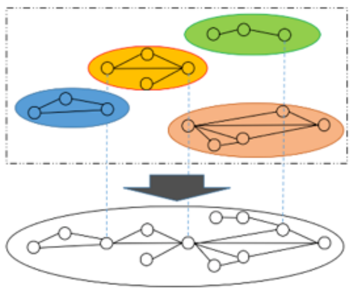

Figure 1: The generating process of model

3. Model Parameters

There are four parameters in our model: layer size distribution, belonging layer distribution, inner layer network model, and inter layer degree correlation.

3.1 Layer size distribution

Layer size distribution is a functional relationship between two quantities: layer's size (number of node in a layer) and the frequency of same size. We hypothesize that a layer is equivalent to a community, so layer size distribution will be the same as community size distribution.

Contact: Chao Fan, School of Engineering, The University of Tokyo, 517, Eng. Bldg.8, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan, E-mail: fanchao333@gmail.com

In our model, the distribution of Layer's size is selected from following two distributions:

- Power-law distribution
- Normal distribution

3.2 Belonging layer distribution

Belonging layer distribution represents a functional relationship between two quantities: the layer number that a person belongs to and its frequency.

Just like layer size distribution, the distribution of the number of belonging layers are also selected from followings:

power-law distribution and normal distribution.

- Power-law distribution
- Normal distribution

3.3 Inner layer network model

There are many ways to connected nodes within a layer. One simple idea is to connect all nodes in one layer with each other by links as a complete network. But in social interaction of human society, it is well-known that not all members in a group are friends. Therefore, we could construct inner layer network structure by using the idea of network growth model.

In our model, we use following network models to connect nodes in a layer:

- Complete Network
- BA model [Barabási 99]
- CNN model [Vázquez 03]

3.4 Inter layer degree correlation

In multi-layer network, one node can belong to many layers. In different layer, the node has different number of links, or rather different degree. When carrying on layer aggregation, we consider whether each node has a degree correlation in each layer. Because different inter layer degree correlation may lead to different network structure.

We implement two correlation methods.

- Correlated: Each node has degree correlation on each layer. For example, the node with higher degree in one layer is also a high-degree node in the other.
- Random: A degree of each node is selected randomly on each layer.

4. Measurement

Modularity is proposed by Newman [Newman 04, Newman 06] and measures the strength of division of a network into communities. A high modularity implies dense intra-community links and less inter-community links. Let $G = (V, E)$ be an undirected graph, $|V| = n$, and $|E| = m$. A_{xy} is an element of the adjacent matrix of G.

$$A_{xy} = \begin{cases} 1 & (x, y) \in E \\ 0 & otherwise \end{cases}$$

The degree of a node x can be defined as $k_x = \sum_y A_{xy}$. Thus, the modularity Q is defined as follows.

$$Q = \frac{1}{2m}\sum_{xy}\left[A_{xy} - \frac{k_x k_y}{2m}\right]\phi(C_x, C_y) = \frac{1}{2m}\sum_{xy}\left[A_{xy} - P_{xy}\right]\phi(C_x, C_y)$$

Where C_x represents the vertex x belongs to community C_x. The $\phi(C_x, C_y)$ is 1 if x and y are in the same community and 0 otherwise. P_{xy} is the probable vertex degrees between vertices x and y at random.

As Modularity only measures the quality of a division from a global perspective, in this paper we propose a new measurement we call "ego-network's modularity" to measure the feature of generated network in a microscopic way. Accordingly, we can calculate modularity for each node's ego-network by using this new measurement and acquire the distribution.

5. Simulation

Combining previous four parameters will result in different pattern of generation model. We have generated many networks by simulation. Some distribution of ego-network's modularity is shown in Figure 2, 3, 4 & 5.

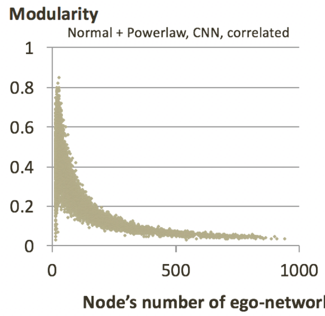

Figure 2: Modularity distribution of ego-network (Parameter: Normal, Powerlaw, CNN, correlated)

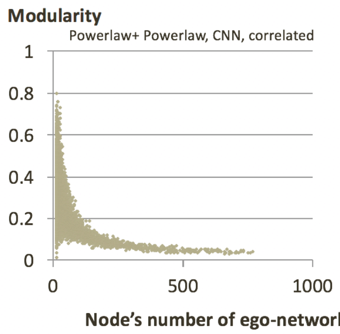

Figure 3: Modularity distribution of ego-network (Parameter: Powerlaw, Powerlaw, CNN, correlated)

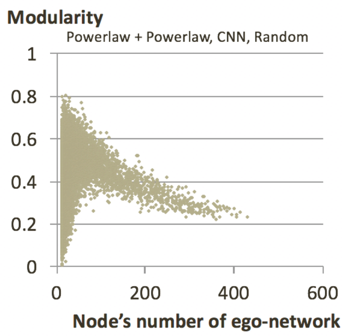

Figure 4: Modularity distribution of ego-network (Parameter: Powerlaw, Powerlaw, CNN, random)

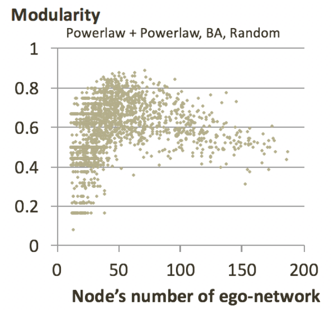

Figure 5: Modularity distribution of ego-network (Parameter: Powerlaw, Powerlaw, BA, random)

To validate the proposed model, we try to imitate weibo network (see weibo.com), which is a twitter like social media, whose features of network are shown in Table 1. Modularity of ego-network is shown in Figure 6.

Table 1: Features of real-world network

	Node	Link	Modularity
Weibo	900	15684	0.1353

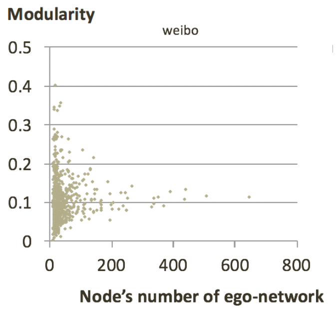

Figure 6: Modularity distribution of ego-network in weibo dataset

A generated network, which shows a most similar features of twitter-like network "weibo", is followings.

- Layer size distribution: *Power-law*
- Belonging layer distribution: *Power-law*
- Inner layer network model: *BA model*
- Inter layer degree correlation: *Correlated*

The distribution of modularities of each ego-network is shown in Figure 7.

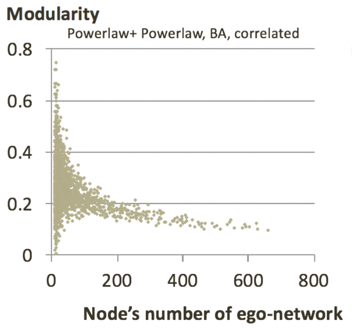

Figure 7: Modularity distribution of ego-network (Parameter: Powerlaw, Powerlaw, BA, correlated)

Two networks shown in Figure 6 and 7 share common features. Higher degree node's ego-networks own a relative low modularity, while lower ones reveal a wide range of this value. Furthermore, too high or low modularity value is less frequent, compared with the middle one.

6. Conclusion

6.1 Summary

In this paper, we have proposed an effective network generation model by layer aggregation based on multilayer network. There are four parameters in our model: layer size distribution, belonging layer distribution, inner layer network model, and inter layer degree correlation.

Rather than utilizing Newman's modularity, we brought forth "ego-network's modularity" to measure the community structure of generated network, which is a neither too macroscopic nor too microscopic evaluation measurement.

Finally, by combining 4 different parameters and doing simulation, we succeed in representing a twitter-like network "weibo" by our method.

6.2 Future work

In future work, we will first investigate the parameter and imitate other social network, like facebook. Second, we want to take other measurement into consideration, e.g. clustering coefficient, average shortest-path length, assortativity, and so forth. Finally, after constructing a network, simulation experiment such as information diffusion will be carried out.

References

[Newman 04] M. E. J. Newman and M. Girvan: Finding and Evaluating Community Structure in Networks, Physical Review E, American Physical Society, 2004.

[Newman 06] M. E. J. Newman: Modularity and community structure in networks, Proceedings of the National Academy of Sciences USA, National Academy of Sciences, 2004.

[Pasta 13] M. Pasta, Z. Jan, A. Sallaberry and F. Zaidi: Tunable and Growing Network Generation Model with Community Structures, 3rd Conference on Social Computing and Applications, SCA 2013, 2013.

[Shinoda 08] K. Shinoda, Y. Matsuo and H. Nakashima: Network Generation Model Based on Multiple Centralities, Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, Japan Society for Fuzzy Theory and Intelligent Informatics, 2008.

[Kivelä 14] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno and M. A. Porter: Multilayer Networks, Journal of Complex Networks, Oxford Journals, 2014.

[Barabási 99] A. L. Barabási and R. Albert: Emergence of Scaling in Random Networks, Science, American Association for the Advancement of Science, 1999.

[Vázquez 03] A. Vázquez: Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, Physical Review E, American Physical Society, 2003.