A pattern-match based method for extracting modules in directed networks

Taro Takaguchi*[1,2,†] and Yuichi Yoshida[1,3]

[1] National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, 101-8430 Tokyo, Japan
[2] JST, ERATO, Kawarabayashi Large Graph Project, 2-1-2 Hitotsubashi, Chiyoda-ku, 101-8430 Tokyo, Japan
[3] Preferred Infrastructure, Inc., 2-40-1 Hongo, Bunkyo-ku, 113-0033 Tokyo, Japan
[†] Present Address: National Institute of Information and Communications Technology, 4-2-1 Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan

E-mail: *taro.takaguchi@nict.go.jp

Keyword: Complex networks; Module structure; Directed networks

In many applications from social to economic phenomena, their network representations often appear to contain the directions of links. The direction of a link may represent the direction of messaging activity (i.e., from senders to receivers) or the direction of money transaction (i.e., from lenders to borrowers). The link direction encodes how entities flow on the network. Therefore, one should take into account the link direction when investigating the modules (also called communities) in directed networks [1]. There are multiple methods of module detection specialized for directected networks. However, it is still an open question which methods to be used when we are given a directed network. We point out two roots of this problem. First, a majority of previous methods is based on a single objective function to find modules, while there might be different types of module in a single network. Second, these methods often divide a whole network into modules in an equal sense, while some parts of the network may be not a module in fact.

To move a step forward from this situation, we will propose a method to extract modules in directed networks in this presentation [2]. Our method extracts two types of modules from a directed network on the basis of a unified framework. The core idea of our method is a pattarn matching of directed triangles (i.e., connected subgraph composed of three nodes), similar to the motif analysis [3]. In addition, our method can be regarded as a generalization of truss structure which was originally defined for undirected networks [4]. We developed an efficient algorithm for the method that can process networks with millions of nodes in a reasonable time on laptop computers. Results of the application of the proposed method to real-world networks confirm that most of real networks actually contain the two types of modules we proposed. We will demonstrate that the extracted modules capture meaningfull components in some example networks with node labels and argue that the abundance of two types of modules may be used for categorizing the networks obtained from different research fields.

References

[1] F.D. Malliaros and M. Vazirgiannis. Clustering and community detection in directed networks: A survey. *Phys. Rep.* 533(4):95142 (2013).

[2] T. Takaguchi and Y. Yoshida. Cycle and flow trusses in directed networks. Preprint arXiv:1603.03519

[3] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science* 298(5594):824827 (2002).

[4] J. Cohen. Trusses : Cohesive Subgraphs for Social Network Analysis. Technical report (National Security Agency, Fort Meade, MD, 2008).