Evaluating community structure in bipartite networks

Xin Liu
Department of Computer Science
Tokyo Institute of Technology
Tokyo, Japan
tsinllew@ai.cs.titech.ac.jp

Tsuyoshi Murata
Department of Computer Science
Tokyo Institute of Technology
Tokyo, Japan
murata@cs.titech.ac.jp

***Abstract*—Communities in unipartite networks are often understood as groups of nodes within which links are dense but between which links are sparse. Such communities are not suited to bipartite networks, as there is only one-to-one correspondence between communities of different types. Recently, B. Long et al. introduced the link-pattern based community, which allows many-to-many correspondence between communities. In this paper, we propose a measure for evaluating the goodness of different partitions of a bipartite network into link-pattern based communities. Such a measure is useful for both comparing various community detection methods and devising new community detection algorithm based on optimization. We demonstrate the effectiveness of the proposed measure using the famous Southern women bipartite network.**

Keywords-community structure; modularity; link mining; bipartite network

I. Introduction

Networks are a natural representation for various kinds of complex systems in biology [1], economy [2], sociology [3], and other fields [4]. A key feature of many networks is community structure [5-7]: the tendency for network nodes to be divided into distinct groups, or *communities*, such that nodes within the same community are, in some sense, similar to each other, whereas nodes from different communities are dissimilar. Community detection, the identification of these communities, has broad applications. It helps uncover unknown functional modules, such as Web pages with related topics on World Wide Web [8], groups of individuals sharing common interests in social networks [9], or sets of genes conducting biological processes together in biological networks [10, 11]. Community structure also has strong consequence for dynamic processes on networks and thus can be exploited to find new efficient context-aware routing algorithms [12, 13].

Most of the research on community detection is in unipartite networks (please refer to [14, 15] for extensive overviews), where communities are often understood as groups of nodes within which links are dense but between which links are sparse.

There are many real world systems that can be naturally modeled as bipartite networks, whose nodes are divided into two disjoint sets X and Y (corresponding to two different types), and only the link between two nodes in different sets is allowed. For example, the DBLP computer science bibliography can be modeled as an author-paper bipartite network, where edges connect authors and the papers they have authored. Other examples include the human sexual bipartite network consisting of men and women; the collaboration bipartite network consisting of actors and events; the online shopping bipartite network consisting of consumers and products. Although community structure in unipartite networks has been well investigated, a thorough analysis of this equivalent problem in bipartite networks has not been made yet.

One important issue is the evaluation of community structure, or a partition of a bipartite network into communities. Before addressing this issue, we have to answer the fundamental but controversial question--what is the definition of community in bipartite networks? If we accept the definition in unipartite networks that communities consist of densely linked nodes, a community in an author-paper bipartite network should contain both authors and papers, since there is no author-author edge or paper-paper edge. Consequently, there is only one-to-one correspondence between author-communities and paper-communities, as shown in Fig. 1(a).

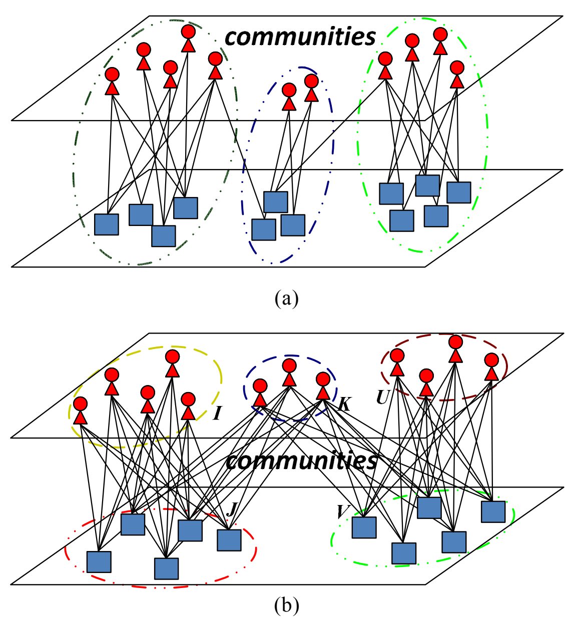

Figure 1. Communities in bipartitie networks. (a) The traditional link-density based communities. (b) The link-pattern based communities.

978-0-7695-4211-9/10 $26.00 © 2010 IEEE
DOI 10.1109/SocialCom.2010.91

Recently B. Long, et al. introduced the link-pattern based community, a generation of the traditional "link-density based community" [16]. A link-pattern based community is a group of nodes which have the similar link patterns, i.e. the nodes within a community link to other nodes in similar ways (in the words of network science, they are structurally equivalent). Fig. 1(b) shows an illustration of such communities in an author-paper bipartite network. For example, the nodes in the author-community K have the similar link patterns: they all strongly link to nodes in paper-community J, and at the same time strongly link to nodes in paper-community V; The nodes in author-community I also have similar link-patterns, but different from that in author-community K: they only strongly link to nodes in paper-community J. Obviously the many-to-many correspondence between author-communities and paper-communities is allowed. Thus, the link-pattern based community is better suited to bipartite networks.

In this paper, we propose a measure for evaluating the goodness of different partitions of a bipartite network into link-pattern based communities. Experiments on the famous Southern women bipartite network show that our measure successfully provides consistent evaluations for different partitions as the objective evaluations given by social scientists, while previously proposed measures fail to do that. Our measure is useful for both comparing various community detection methods and devising new community detection algorithm based on optimization.

II. RELATED WORKS

The study of community structure is closely related to the graph partitioning problem in graph theory. This problem is to divide a connected graph into a given number of pieces based on some measures, such as the min cut [17, 18], ratio cut [19, 20], normalized cut [21], and min-max cut [22]. In recent years, the study of community structure has attracted a great deal of interest, especially in the realm of statistical physics [14]. One difference from graph partitioning is that the number of communities is not given, but should be automatically searched.

Communities in unipartite networks are often understood as groups of nodes within which links are dense but between which links are sparse. To evaluate such a community structure, Newman and Girvan proposed the *modularity* measure [23, 24]. It is based on the idea that a random network is not expected to have community structure, so the possible existence of community structure of a given network is revealed by comparison between the number of intra-community edges in this given network and that number in a random network. Though modularity suffers from the resolution limit problem [25], it has been widely accepted as a *de facto* standard.

F. Radicchi et al. introduced the definition of strong and week community [26] in unipartite networks, based on comparing the numbers of edges pointing inside and outside the community. Following this idea, Medus and Dorso proposed a measure for evaluating community structure in unipartite networks [27].

Due to the success of modularity in unipartite networks, some researchers extended it to bipartite networks and proposed *bipartite modularity*. According to different understandings of bipartite network community, there are different versions of bipartite modularity, such as the one proposed by Barber [28], the one proposed by Murata and Ikeya [29], the one proposed by Suzuki and Wakita [30], and the one proposed by Guimerà et al. [31]. The first three bipartite modularities seek dense links between specific communities, so the sprit of them is to compare the number of dense intra-community links in the given network and that number in a random bipartite network. The last one focuses only on community structure of nodes of one type. Take the community structure of actors in an actor-event bipartite network as an example. The idea is that the more events that two actors co-participate in, the more probable they are in the same actor-community. So this bipartite modularity actually compares the number of co-participated events for each pair of actors within a community in the given network and that number in a random actor-event bipartite network.

There are also measures for evaluating overlapping community structure (a node can belong to more than one community) in unipartite networks [32-34].

In addition, when the *correct* partition (the *true* community structure) is available, the most widely used measure for evaluating a given partition is *normalized mutual information* [35], which is based on information theory and calculates the agreement between the two partitions.

One direct application of the above measures is to evaluate the goodness of different partitions provided by various community detection methods and compare their performances. Another application is to devise community detection methods. One chooses a measure as a quality function, and devises heuristics which works by searching over possible partitions for one that optimizes the quality function. For instance, Modularity optimization is the most popular community detection method in unipartite networks. Numbers of modularity optimization algorithms [36-44] were proposed and were successfully employed to analyze the community structure of real-world systems.

B. Long et al. introduced the link-pattern based community, which generalizes the various kinds of communities. They also proposed a model to learn the link-pattern based community structure, with the number of communities given beforehand [16]. C. Y. Lin et al. alternatively used K-means algorithm to learn the link-pattern based community structure [45].

In this paper, we focus on evaluating the goodness of different partitions of a bipartite network into link-pattern based communities. Different from the one proposed by B. Long [16], our measure does not require the user to specify the number of communities beforehand. That is, our measure favors the most natural partition of a bipartite network into link-pattern based communities.

III. THE PROPOSED MEASURE

In this section, we propose a measure for evaluating the goodness of different partitions of a bipartite network into link-pattern based communities. Our idea is to use the link-patterns hidden in the partition to construct an approximate model network, and compare it with the original network: the higher of the similarity between the model network and the original network, the better of the partition.

For simplicity, we term the nodes of two types in a bipartite network as X-nodes and Y-nodes; and the communities of two types, which are composed of X-nodes and Y-nodes respectively, as X-communities and Y-communities. In the following, we use small Latin letters i, k, u and j, l, v, respectively, for indices of X-nodes and Y-nodes, and capital Latin letters I, K, U and J, L, V, respectively, for indices of X-communities and Y-communities.

Now assume a bipartite network $G = (V_X, V_Y, E)$, where V_X and V_Y are X-node set and Y-node set, and $E \subseteq (V_X \times V_Y)$ is the edge set. The numbers of X-nodes, Y-Nodes and edges are denoted as N_X, N_Y and M, i.e. $N_X = |V_X|$, $N_Y = |V_Y|$, and $M = |E|$. The link structure of this bipartite network can be represented as a N_X-by-N_Y adjacency matrix $\mathbf{A}$:

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Suppose a partition divides X-nodes and Y-nodes into C_X and C_Y link-pattern based communities (in this paper we only focus on non-overlapping communities, i.e. each X-node/Y-node belongs to one and only one X-community/Y-community). Without loss of generality, assume that the rows and columns of the adjacency matrix $\mathbf{A}$ are rearranged, such that X-nodes within the same X-communities are put together and Y-nodes likewise. Then $\mathbf{A}$ is blocked:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{(1,1)} & \mathbf{A}_{(1,2)} & \cdots & \mathbf{A}_{(1,C_Y)} \\ \mathbf{A}_{(2,1)} & \mathbf{A}_{(2,2)} & \cdots & \mathbf{A}_{(2,C_Y)} \\ \cdots & \cdots & \ddots & \cdots \\ \mathbf{A}_{(C_X,1)} & \mathbf{A}_{(C_X,2)} & \cdots & \mathbf{A}_{(C_X,C_Y)} \end{pmatrix}, \tag{2}$$

where $\mathbf{A}_{(I,J)}$ represents the link structure between X-community I and Y-community J.

In the model network, the same nodes and the community memberships of the original network are preserved. Suppose the adjacency matrix of the model network is denoted as $\mathbf{A}^*$, which can be blocked in the same way as $\mathbf{A}$:

$$\mathbf{A}^* = \begin{pmatrix} \mathbf{A}^*_{(1,1)} & \mathbf{A}^*_{(1,2)} & \cdots & \mathbf{A}^*_{(1,C_Y)} \\ \mathbf{A}^*_{(2,1)} & \mathbf{A}^*_{(2,2)} & \cdots & \mathbf{A}^*_{(2,C_Y)} \\ \cdots & \cdots & \ddots & \cdots \\ \mathbf{A}^*_{(C_X,1)} & \mathbf{A}^*_{(C_X,2)} & \cdots & \mathbf{A}^*_{(C_X,C_Y)} \end{pmatrix}, \tag{3}$$

Then we place edges between each X-community I and Y-community J at random. Specifically, the number of edges between $i \in I$ and $j \in J$ is[a]:

$$A^*_{ij} = \frac{\sum_{v\in J} A_{iv} \cdot \sum_{u\in I} A_{uj}}{\sum_{u\in I} \sum_{v\in J} A_{uv}} \quad \forall i \in I, j \in J\,, \tag{4}$$

where $\sum_{v\in J} A_{iv}$, $\sum_{u\in I} A_{uj}$ and $\sum_{u\in I} \sum_{v\in J} A_{uv}$ are, respectively, the number of edges linking i and J, the number of edges linking j and I, and the number of edges linking I and J in the original network. (A direct implication is that each node has the same degree as in the original network [40].)

According to the definition of link-pattern based community, community I and J are linked following some link-pattern in the original network. Thus, one can expect that the randomly placed edges between I and J in the model network would resemble the corresponding edges in the original network. In other words,

$$\| \mathbf{A}_{(I,J)} - \mathbf{A}^*_{(I,J)} \|^2 = \sum_{i\in I} \sum_{j\in J} (A_{ij} - A^*_{ij})^2 \tag{5}$$

should be small. Summing over all I and J, we should have a small value for

$$\| \mathbf{A} - \mathbf{A}^* \|^2 = \sum_I \sum_J \| \mathbf{A}_{(I,J)} - \mathbf{A}^*_{(I,J)} \|^2 \,. \tag{6}$$

But the above measure favors partitions into small communities and it takes its minimum value of 0 when each X-node or Y-node is divided into its own single community. Thus a denominator is added to penalize too small communites:

$$\sum_I \sum_J \frac{\| \mathbf{A}_{(I,J)} - \mathbf{A}^*_{(I,J)} \|^2}{|I| \cdot |J|}\,, \tag{7}$$

where $|I|$ and $|J|$ are the number of nodes in I and J respectively. We use δ_i and θ_j to indicate the X-community and Y-community of i and j. Finally, our measure becomes:

$$\sum_i \sum_j \frac{\left(A_{ij} - \sum_{v\in\theta_j} A_{iv} \cdot \sum_{u\in\delta_i} A_{uj} / \sum_{u\in\delta_i} \sum_{v\in\theta_j} A_{uv}\right)^2}{|\delta_i| \cdot |\theta_j|}\,. \tag{8}$$

Based on the above discussion, a smaller value of the measure indicates a better partition of a bipartite network into link-pattern based communities. Though the measure takes its minimum value of 0 for the trivial partition when each X-node is the sole member of one of Nx X-communities, or when each Y-node is the sole member of one of N_Y Y-communities[b], it can be used to evaluate other non-trivial partitions.

IV. EXPERIMENTAL RESULTS

In this section we present experiments on comparing our proposed measure with some other measures for evaluating community structure in bipartite networks. They are, in order, the bipartite modularity proposed by Barber (Barber-measure) [28], the bipartite modularity proposed by Murata and Ikeya (Murata-measure) [29], the bipartite modularity proposed by Suzuki and Wakita (Suzuki-measure) [30], and the bipartite modularity proposed by Guimerà et al. (Guimerà-measure) [31].

The data set used should be a bipartite network with a priori known community structure, as the experiments are to

[a] Suppose this number to be 0, if the denominator is 0.

[b] We can replace the denominator of Eq. (11) with |δ-1|·|θ-1| to prevent the trivial partition getting the minimum value of 0.

compare different measures. Here we use the famous Southern women [46] data set. This data set was collected by Davis et al. in the town of Natchez, Mississippi during the 1930s as part of an extensive study of class and race in black and white society in the Deep South. It describes the participation of 18 women in 14 social events. A bipartite network whose nodes are women and social events, and whose edges are the participation of the women in the events can then be derived.

Southern women data set and its bipartite network have been much studied by social scientists. Davis used ethnographic knowledge and divided the 18 women into two groups: women 1 through 9 in the first group and women 9 through 18 in the second (woman 9 is a secondary member of both groups) [46]. A similar grouping that put women 1 through 9 in the first group and women 10 through 18 in the second has been identified by Freeman as the consensus from 21 different studies [47]. Taking these two statements together, we believe that the partition that divides women into two communities--{1-9} and {10-18}, is the *correct* partition.

Fig. 2 (a)-(d) describe 4 different partitions of the Southern women bipartite network (referred as partition (a), (b), (c), and (d) in the following), which are obtained by various different methods. Each partition gives the partitions of both women and events. If we only look at women, for whom we know the correct partition, we can make an objective evaluation for these 4 partitions using the *normalized mutual information* (NMI) [35]. NMI, which is regularly used in papers about community detection in networks, is an information-theoretic measure that calculates the amount of common information between a given women partition and the correct partition. When the two partitions match completely, we have a maximum value of 1, and when the given partition is totally independent of the correct partition, we have a minimum value of 0. A larger value of NMI indicates a better partition.

Next, we use different measures to evaluate partitions (a)-(d), and compare their evaluation results with the objective evaluation discussed above. All of the evaluation results are summarized in Table I (Since Guimerà-measure is only for partition of nodes of one type, it is calculated based on the women partitions. For our measure, the smaller of the value, the better of the partition; while for the other measures, the larger of the value, the better of the partition).

It should be noted that comparing the absolute values between measures is meaningless because different measures concern different features. To make it clear, in Table II we rank these partitions, based on the values in Table I (each row lists the rankings of different partitions, from 1 to 4, as evaluated by the corresponding measure).

We can find that our measure is the most reliable, as the corresponding rankings coincide with that of the objective evaluation. Murata-measure can identify partitions (a) and (b) as the best and the second best partitions. It reverses the rankings of partitions (c) and (d). Compared with Murata-measure, our measure favors partitions with many-to-many correspondence between communities of the two types, while Murata-measure prefers community pairs with dense links. As claimed by the authors of [30], partition (d), to some extent, reflect the many-to-many correspondence between women-communities and event-communities, with secondary women members being successfully separated [30]. So our measure rates it higher than partition (c).

Other measures are more or less biased towards specific kinds of partitions. For example, Barber-measure can only handle partitions where the numbers of women-communities and event-communities are the same. For this reason, there is no record for partition (d) as evaluated by Barber-measure.

To sum up, this example shows the effectiveness of our measure for evaluating community structure in a real-world bipartite network.

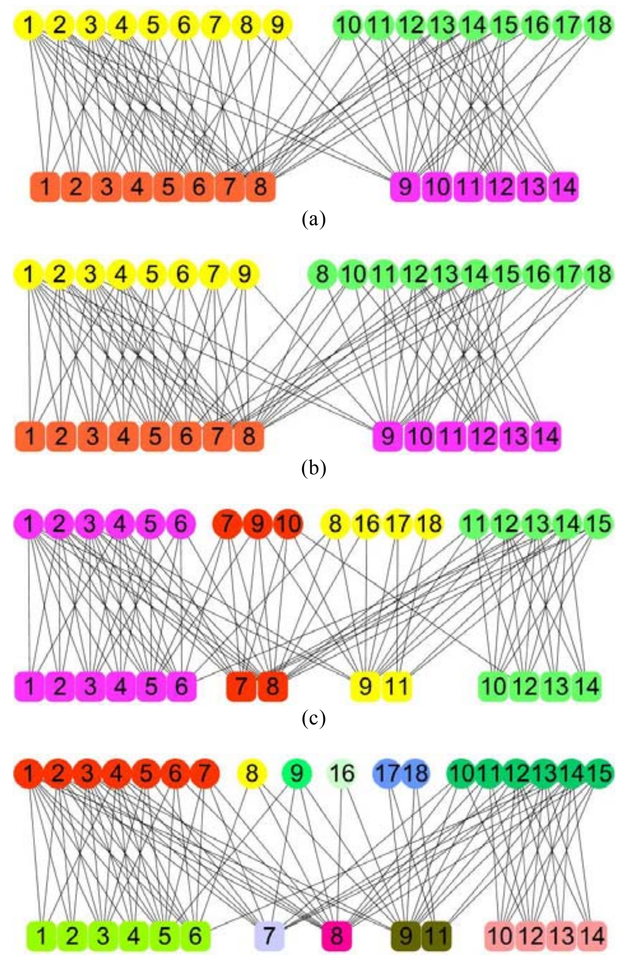

Figure 2. Partitions of the Southern women bipartite network. Women nodes are indicated as circle symbols located at the top side, while event nodes are indicated as square symbols located at the bottom side. Nodes of the same community are painted the same color. (a) The partition obtained by optimizing Murata-measure [29] using a simple agglomerative (bottom-up) algorithm. (b) The partition obtained by optimizing Guimerà-measure using simulated annealing [31]. (c) The partition obtained by BRIM, an algorithm for optimizing Barber-measure [28]. (d) The partition obtained by optimizing Suzuki-measure using a simple agglomerative (bottom-up) algorithm [30].

TABLE I. EVALUATIONS OF THE PARTITIONS (a), (b), (c) AND (d) IN FIG. 2 BY DIFFERENT MEASURES.

Measure	(a)	(b)	(c)	(d)
NMI	1.0000	0.7428	0.4513	0.6440
This paper	0.3672	0.3737	0.8042	0.4187
Barber-measure	0.3184	0.3159	0.3455	--
Murata-measure	0.5749	0.5680	0.4789	0.4770
Suzuki-measure	0.2111	0.2068	0.2186	0.2367
Guimerà-measure	0.2154	0.2169	0.1354	0.1537

TABLE II. RANKINGS OF THE PARTITIONS (a), (b), (c) AND (d) IN FIG. 2, AS EVALUATED BY THE DIFFERENT MEASURES.

Measure	(a)	(b)	(c)	(d)
NMI	1	2	4	3
This paper	1	2	4	3
Barber-measure	2	3	1	--
Murata-measure	1	2	3	4
Suzuki-measure	3	4	2	1
Guimerà-measure	2	1	4	3

V. CONCLUSION AND FUTURE WORK

A new measure for evaluating community structure in bipartite networks is proposed in this paper. Different from the previous measures, our measure is based on link-patterns, which allows many-to-many correspondence between communities and is thus better suited for bipartite networks. Experiments on the famous Southern women bipartite network validate the effectiveness of our measure over previous ones.

Although we do not introduce an algorithm for community detection in bipartite networks, one can use the proposed measure as a quality function, and develop a proper algorithm that searches over possible partitions for one that optimizes the quality function. The online shopping bipartite networks such as eBay and Amazon are fairly large, on the order of several million edges at the very least. Developing scalable algorithm to analyze their structure is left for our future work.

There are tripartite networks in the Web. For example, social tagging systems can be represented as tripartite networks composed of three types of nodes (users, URLs and tags). There are attempts for analyzing tripartite networks using projection method [48]. However, projection generally loses some information of the original network. It is possible to extend our measure to tripartite networks and develop algorithms to identify communities directly. Evaluating and identifying community structure of heterogeneous networks is one of the important and challenging topics of Web mining.

REFERENCES

[1] A. -L. Barabási and Z. N. Oltvai, "Network biology: understanding the cell's functional organization," Nature Reviews Genetics, vol. 5, pp. 101-113, 2004.

[2] N. Economides, "The economics of networks," International Journal of Industrial Organization, vol. 16, no. 4, pp. 673-699, October 1996.

[3] L. C. Freeman, The Development of Social Network Analysis: A Study in the Sociology of Science. Vancouver, BC: Empirical Press, 2004.

[4] A. -L. Barabási, Linked: The New Science of Networks. Cambride, MA: Perseus Publishing, 2002.

[5] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," Proc. Natl. Acad. Sci. USA, vol. 99, pp. 7821–7826, 2002.

[6] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," Proc. Natl. Acad. Sci. USA, vol. 101, no. 9, pp. 2658-2663, March 2004.

[7] M. A. Porter, J. P. Onnela, P. J. Mucha, "Communities in networks," Notices. Amer. Math. Soc. vol. 56, pp. 1082-1097, 2009.

[8] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the Web," Computer Networks, vol. 33, pp. 309-320, 2000.

[9] B. Wharf and M. Clague, Community Organizing: Canadian Experiences. Toronto: Oxford University Press, 1997.

[10] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," Nature, vol. 435, no. 7043, pp. 814-818, 2005.

[11] N. Gulbahce, S. Lehmann, "The art of community detection," BioEssays, vol. 30, no. 10, pp. 934-938, 2008.

[12] A. Arenas, A. D.- Guilera, and C. J. P.- Vicente, "Synchronization reveals topological scales in complex networks," Phys. Rev. Lett. vol. 96, 114102, 2006.

[13] J. G. Restrepo, E. Ott, and B. R. Hunt, "Characterizing the dynamical importance of network nodes and links," Phys. Rev. Lett. vol. 97, 094102, 2006.

[14] S. Fortunato, "Community detection in graphs," Physics Reports, vol. 486, pp. 75-174, 2010.

[15] L. Danon, J. Duch, A. D.- Guilera, and A. Arenas, "Comparing community structure identification," J. Stat. Mech. P09008, September 2005.

[16] B. Long, X. Wu, Z. Zhang, and P. S. Yu, "Community learning by graph approximation," in Proceedings of the 7th IEEE International Conference on Data Mining, October 2007, pp. 232-241.

[17] B. Kernighan and S. Lin, "An efficient heuristic procedure to partition graphs," Bell Syst. Tech. J., vol. 49, pp. 291-307, February 1970.

[18] C. M. Fiduccia and R. M. Mattheyses, "A linear-time heuristic for improving network partitions," in Proceedings of the 19th conference on Design automation, January 1982, pp.175-181.

[19] C.- K. Cheng and Y.-C. A. Wei, "An improved two-way partitioning algorithm with stable performance," IEEE Trans. Comput.-Aided Des. vol. 10, no. 12, pp. 1502-1511, December 1991.

[20] L. W. Hagen and A. B. Kahng, "New spectral methods for ratio cut partitioning and clustering," IEEE Trans. Comput.-Aided Des. vol. 11, no. 9, pp. 1074-1085, September 1992.

[21] J. Shi and J. Malik, "Normalized cuts and image segmentation," IEEE Trans. Pattern Anal. Mach. Intell. vol. 22, no. 8, pp. 888-905, August 2000.

[22] C. H. Q. Ding, X. F. He, H. Y. Zha, M. Gu, and H. D. Simon, "A min-max cut algorithm for graph partitioning and data clustering," in Proceedings of the 1st IEEE International Conference on Data Mining, November 2001, pp. 107-114.

[23] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," Phys. Rev. E, vol. 69, 026113, 2004.

[24] M. E. J. Newman, "Modularity and community structure in networks," Proc. Natl. Acad. Sci. USA, vol. 103, no. 23, pp. 8577-8582, June 2006.

[25] S. Fortunato, M. Barthélemy, “Resolution limit in community detection,” Proc. Natl. Acad. Sci. USA, vol. 104, no. 1, pp. 36-41, January 2007.

[26] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, “Defining and identifying communities in networks,” Proc. Natl. Acad. Sci. USA, vol. 101, no. 9, pp. 2658-2663, March 2006.

[27] A. D. Medus and C. O. Dorso, “Alternative approach to community detection in networks,” Phys. Rev. E, vol. 79, 066111, 2009.

[28] M. J. Barber, “Modularity and community detection in bipartite network,” Phys. Rev. E, vol. 76, 066102, 2007.

[29] T. Murata and Tomoyuki Ikeya, “A new modularity for detecting one-to-many correspondence of communities in bipartite networks,” Advances in Complex Systems, vol. 13, no. 1, pp. 19-31, 2010.

[30] K. Suzuki, K. Wakita, “Extracting multi-facet community structure from bipartite networks,” in Proceedings of 2009 International Conference on Computational Science and Engineering, vol. 4, pp. 312-319, August 2009.

[31] R. Guimerà, M. S.- Pardo, and L. A. N. Amaral, “Module identification in bipartite and directed networks,” Phys. Rev. E, vol. 76, 036102, 2007.

[32] V. Nicosia, G. Mangioni, V. Carchiolo, and M. Malgeri, “Extending the definition of modularity to directed graphs with overlapping communities,” J. Stat. Mech. P03024, March 2009.

[33] H. W. Shen, X. Q. Cheng, and J. F. Guo, “Quantifying and identifying the overlapping community structure in networks,” J. Stat. Mech. P07042, July 2009.

[34] A. Lázár, D. Ábel, and T. Vicsek, "Modularity measure of networks with overlapping communities," arXiv:0910.5072, October 2009.

[35] A. Fred and A. Jain, “Robust data clustering,” in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 2003.

[36] M. E. J. Newman, “Fast algorithm for detecting community structure in networks,” Phys. Rev. E, vol. 69, 066133, 2004.

[37] A. Clauset, M.E.J. Newman, and C. Moore, “Finding community structure in very large networks,” Phys. Rev. E, vol. 70, 066111, 2004.

[38] J. Duch, A. Arenas, “Community detection in complex networks using extremal optimization,” Phys. Rev. E, vol. 72, 027104, 2005.

[39] A. Medus, G. Acuna, C.O. Dorso, “Detection of community structures in networks via global optimization,” Physica A, vol. 358, pp. 593-604, 2005.

[40] M. E. J. Newman, “Finding community structure in networks using the eigenvectors of matrices,” Phys. Rev. E, vol. 74, 036104, 2006.

[41] P. Schuetz and A. Caflisch, “Efficient modularity optimization by multistep greedy algorithm and vertex refinement,” Phys. Rev. E. vol. 77, 046112, 2008.

[42] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” J. Stat. Mech. P10008, 2008.

[43] X. S. Zhang, R. S. Wang, Y. Wang, J. Wang, Y. Qiu, L. Wang, and L. Chen, “Modularity optimization in community detection of complex networks,” Europhys. Lett. vol. 87, 38002, 2009.

[44] Xin Liu and Tsuyoshi Murata, “Advanced modularity-specialized label propagation algorithm for detecting communities in networks,” Physica A, vol. 389, no. 7, pp. 1493-1500, 2010.

[45] C. Y. Lin, J. L. Koh, and A. L. P. Chen. “A Better Strategy of Discovering Link-Pattern based Communities by Classical Clustering Methods,” in Proceedings of the 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining, June 2010.

[46] A. Davis, B. B. Gardner, and M. R. Gardner, Deep South. Chicago: University of Chicago Press, 1941.

[47] R. Breiger, K. Carley, and P. Pattison, Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. Washington, DC: The National Academics Press, 2003.

[48] N. Neubauer and K. Obermayer, “Towards community detection in k-partite k-uniform hypergraphs,” Workshop on Analyzing Networks and Learning with Graphs, December 2009.