

# Modularity Density for Evaluating Community Structure in Bipartite Networks

Yongcheng Xu<sup>1,\*</sup> and Ling Chen<sup>1,2</sup>

<sup>1</sup>Department of Computer Science, Yangzhou University, Yangzhou, 225009, China

<sup>2</sup>State Key Lab of Novel Software Tech, Nanjing University, Nanjing, 210093, China

**Abstract:** Bipartite networks are an important category of complex networks in human social activities. Newman and Girvan proposed a measurement called modularity to evaluate community structure in unipartite networks called modularity. Due to the success of modularity in unipartite networks, bipartite modularity is developed according to different understandings of community in bipartite networks which all contains an intrinsic scale that depends on the total size of links and ignores the number of nodes in the bipartite network. In addition, the size heterogeneity of communities and degree of nodes often affects the measure of community. In this work, we propose a quantitative measure for evaluating the partition of bipartite networks into one-to-one correspondence between different type communities based on the concept of average bipartite modularity degree. Unlike the bipartite modularity measures previously proposed, the new measure can overcome the resolution limits. Experiments on the artificial and real-world bipartite networks validate the accuracy and reliability of our bipartite modularity density.

**Keywords:** Bipartite networks, bipartite modularity, community structure, resolution limits.

## 1. INTRODUCTION

Various kinds of real-world complex systems could be modeled as complex networks [1], where nodes (or vertices) represent the objects and edges represent the interactions among these objects. Bipartite network is an important category of complex networks in naturally real world systems, where the nodes (vertices) are divided into two types, and no two nodes of the same type are connected. Many real-world networks are naturally bipartite networks, such as scientists-papers cooperation network [2, 3], the actor-films network [4], disease-gene network [5], club members-activities network [6], audience-songs network [7], computer terminals-data networks in P2P system [8], and investors-company network [9, 10] and so on.

An important feature of the real world networks is typically the community structure [11-13]. Networks can often be divided into communities, and nodes within the same community are densely connected, but connections between nodes of different communities are much sparser [14]. Since many networks demonstrate such a community structure, the description and detection of such a community structure have great practical importance. Because communities are relatively independent of one another structurally, it is believed that each of them may correspond to some fundamental functional unit. Identifying and analyzing such communities in a large network provided an approach for functional

dissection of the network and sheds light on its organizational nature.

In order to evaluate the qualities of detected communities from unipartite networks, Newman and Girvan [15] introduce the measurement called modularity which reflects the extent, relative to a null model network, to which edges are formed within modules instead of the between modules. Fortunato and Barthélemy [21] recently claimed that modularity contains an intrinsic scale that depends on the total size of links in the network. The modules which are smaller than this scale may not be resolved even in the extreme case that they are complete graphs connected by single edges. To solve this problem, Zhenping Li *et al.* [16] presented a modularity measurement called modularity density, and claimed that their modularity density is superior to the widely used modularity and also prove its equivalence with the objective function of the kernel  $k$  means. Though modularity suffers from the resolution limit problem, it has been widely accepted as a de facto standard measurement.

In recent years, the problem of evaluating community structure in bipartite network has drawn much attention of the researcher. On account of the success of modularity in unipartite networks, some researchers extend it to bipartite networks and propose bipartite modularity. According to different understandings of bipartite network community, there are different versions of bipartite modularity measures, such as the ones proposed by Guimera *et al.* [17], the one proposed by Baber [18], the one proposed by Murata [19] and so on etc. These bipartite modularity measures are based on different requirement for the communities of bipartite

networks. Therefore, their quantification is still a subject of debate. Two aspects greatly complicate this problem. In general, the size heterogeneity of communities often greatly affects the measure of community [20]. Another aspect is that, even in a specific network, the generation mechanism or link degree may vary greatly.

Although the resolution limit problem raised by Fortuna to and Barthélemy aims at the unipartite networks originally, but the problem also exists in the existing modularity measures for bipartite networks. Using those modularity measures, the modules which are smaller than the scale of bipartite networks may not be resolved even in the extreme case that they are complete graphs connected by single edges is also exist in bipartite networks.

In this work, we propose a quantitative measure for evaluating the partition of bipartite networks into one-to-one correspondence between different type communities based on the concept of average bipartite modularity degree. We call this quantitative measure the bipartite modularity density. In addition to its simple form, we show that the proposed criterion also improves the resolution limit in bipartite community structure detection. We show the effectiveness of the bipartite modularity density by theoretical analysis and numerical tests of on both artificial and real-world bipartite networks.

## 2. RELATED WORKS

Currently, the method of detecting community structure in complex networks can be divided into two main classes. One class is the *global methods*, of which the most notable example is the modularity introduced by Newman and Girvan [15]. Global methods regard community detection as a global optimization problem, where the objective functions are different various methods. Due to the complexity of such optimization problems, the global methods are typically stochastic in nature. The other class is the *local methods*, among which the best known example is the  $k$ -clique method proposed by Palla *et al.* [22, 23] which utilizes local structural information to reveal the community structure of a network. The local methods are usually deterministic.

In unipartite networks, communities are often modeled as sets of nodes closely connected. Nodes in different communities are sparsely connected. To evaluate the quality of a particular community partitioning of a network, Newman and Girvan introduces a qualitative measure called modularity [15]. A widely used and quite successful method for the identification of communities in unipartite networks is maximizing the modularity function. Using modularity as the objective function, the community identification problem becomes a combinatorial optimization which is analogous to the identification of the ground state of a disordered magnetic system.

For bipartite networks, Guimera *et al.* [17] proposed a bipartite modularity which focuses on the connectivity of only one node type, and employ simulated annealing for modularity optimization. Consequently, their algorithm aims at partitioning communities for nodes of one type at one step. As we know, in some cases two types of node could not be treated symmetrically.

Barber [18] extended the definition of Newman's modularity in unipartite network to be appropriate for bipartite networks and presents a bipartite modularity based on the assumption that there is a one-to-one correspondence between communities of different node type. He also proposed an algorithm called *Adaptive Brim* for detecting community structure by maximizing this bipartite modularity. However, the main weakness of Barber's bipartite modularity is that the number of communities has to be determined in advance. It should be noticed that in most of the existing community detection algorithms, the number of communities must be specified in advance, and cannot be found by the method itself. This is not practical in many real applications.

Murata proposed another bipartite modularity [19], which gives consistent result as the modularity of Newman and Girvan when apply to unipartite networks. Unlike previous proposals, his proposal treats two types of nodes in a uniform framework. Murata's bipartite modularity favors the case when edges coming from or going out of a cluster connect with the same cluster on the opposite side. Murata claimed that his bipartite modularity overcomes the weakness of the constraint of one-to-one correspondence between communities of different types. That is to say, a community of one type nodes may correspond to many communities of the other type of nodes, and vice versa. Although Murata's bipartite modularity allows one-to-many correspondence between communities of different node types, experiments on different datasets show that using Murata's measurement as the objective function can obtain one result: there is one-to-one correspondence between communities of two node types. Unlike previous bipartite modularity, Murata's bipartite modularity can measure the qualities of one-to-many correspondence between communities of different vertex types.

## 3. BIPARTITE MODULARITY DENSITY

As we mentioned in the last section, several bipartite modularity measures for identifying communities in bipartite networks are proposed. Guimera's bipartite modularity focuses on the connectivity of only one type of nodes [17], and Barber's bipartite modularity is based on the assumption that there is one-to-one correspondence between communities of different vertex types [18]. Murata's bipartite modularity, which gives consistent result as Newman's modularity applied to unipartite networks, is one-to-many correspondence between communities of different vertex types [19].

Bipartite network can be represented by a bipartite graph  $G=(V,E)$ , where  $V=V^X \cup V^Y$ . Here,  $V^X$  and  $V^Y$  are respectively the two types of vertices in  $G$ .  $E$  is the set of its edges. There is no edge between the vertexes in the same set of  $V^X$  ( $V^Y$ ), namely, every edge  $(v_i^X, v_j^Y) \in E$  satisfies  $V_i^X \in V^X$ , and  $V_j^Y \in V^Y$ . Assuming that  $V^X$  has  $n$  vertices, and  $V^Y$  has  $m$  vertices, the adjacency matrix of  $G$  has a block off-diagonal form as:

$$A = \begin{bmatrix} 0_{n \times n} & A_{n \times m} \\ A_{m \times n}^T & 0_{m \times m} \end{bmatrix} \quad (1)$$

where  $0_{n \times n}$  and  $0_{m \times m}$  are all-zero  $n \times n$  and  $m \times m$  matrixes respectively,  $A_{n \times m}$  and  $A^T_{m \times n}$  are  $n \times m$  and  $m \times n$  non-zero matrixes. Since the adjacency matrix is symmetric, we can just use the matrix  $A_{n \times m}$  to represent the bipartite graph  $G$ , where each row represents a vertex in the set of  $V^X$ , each column represents a vertex in the set of  $V^Y$ .

Community detection in a bipartite graph  $G = (V^X \cup V^Y, E)$  is to partition into  $C$  sub-graphs  $G_i = (V_i^X, V_i^Y, E_i)$ , where  $C$  is the number of the partitioning communities,  $V_i^X \subset V^X$ ,  $V_i^Y \subset V^Y$ ,  $\bigcup_{i=1}^C V_i^X = V^X$  and  $\bigcup_{i=1}^C V_i^Y = V^Y$ . Let  $|V_i^X|$  and  $|V_i^Y|$  respectively be the number of nodes in the set of  $V_i^X$  and  $V_i^Y$ . We further fine  $D(V_i^X, V_i^Y) = \sum_{j \in V_i^X} \sum_{k \in V_i^Y} A(j, k)$ ,  $D(V_i^X, \bar{V}_i^Y) = \sum_{j \in V_i^X} \sum_{k \in \bar{V}_i^Y} A(j, k)$  and  $D(\bar{V}_i^X, V_i^Y) = \sum_{j \in \bar{V}_i^X} \sum_{k \in V_i^Y} A(j, k)$  where  $V_i^X \cup \bar{V}_i^X = V^X$  and  $V_i^X \cap \bar{V}_i^X = \emptyset$ . The well-known Murata's bipartite modularity  $Q_B$  is defined as follows:

$$Q_B = \sum Q_{B_i} \sum \frac{D(V_i, V_m)}{D(V, V)} - \frac{D(V_i, V)}{D(V, V)} \times \frac{D(V_m, V)}{D(V, V)} \quad (2)$$

$$m = \underset{k}{\operatorname{argmax}}(e_{ik}) \quad (2)$$

Modularity optimization for  $Q_B$  seems to be an effective method to detect communities in bipartite networks. However, Fortunato and Barthélemy [21] recently pointed out the serious resolution limit of Newman's modularity applied to unipartite networks, and claimed that the size of a detected module depends on the size of the whole network. In section V, we will empirically show that this resolution limit also exists in Murata's bipartite modularity. This is mainly because Murata's bipartite modularity measure does not contain information on the number of nodes in a community and the choice of partition is highly sensitive to the total number of links in the network [24].

To overcome such limit, we introduce a new quantitative measure, which is related to the density of sub-graphs. We first define the degree of a bipartite network equal to the number of edges of bipartite network:  $D(G) = E(G) = M$ . Let  $G_i = (V_i^X, V_i^Y, E_i)$  be a community of  $G$ , then we define the average bipartite modularity degree of bipartite sub-graph  $G_i = (V_i^X, V_i^Y, E_i)$  as follows:

$$Q(G_i) = D_{in}(G_i) - D_{out}(G_i) \quad (3)$$

where  $D_{in}(G_i)$  is the average inner degree of the sub-graph  $G_i$ , which is equal to the number of edges in bipartite sub-graph  $G_i$  divided by the total number of nodes in sets  $V_i^X$  and  $V_i^Y$ .  $D_{out}(G_i)$  is the average outer degree of sub-graph  $G_i$ , which is equal to the number of edges in bipartite sub-graph  $G_i$  divided by the total number of nodes in sets  $V_i^X$  and  $V_i^Y$ . Then the average degree of bipartite sub-graph  $G_i = (V_i^X, V_i^Y, E_i)$  can be easily formulated as follows:

$$Q(G_i) = \frac{D(V_i^X, V_i^Y) - D(V_i^X, \bar{V}_i^Y) - D(\bar{V}_i^X, V_i^Y)}{|V_i^X| + |V_i^Y|} \quad (4)$$

Here, the intuitive idea is that  $Q(G_i)$  should be as large as possible for a valid community in bipartite networks. Then we define the bipartite modularity density of a partition as

the summation of the average bipartite modularity degrees of all the communities  $G_i$  for  $i=1, \dots, c$ . Suppose a bipartite network  $G$  is partitioned into communities  $G_1, \dots, G_c$ , the bipartite modularity density of such a partition can be calculated as follows:

$$Q = \sum_{i=1}^c Q(G_i) = \sum_{i=1}^c \frac{D(V_i^X, V_i^Y) - D(V_i^X, \bar{V}_i^Y) - D(\bar{V}_i^X, V_i^Y)}{|V_i^X| + |V_i^Y|} \quad (5)$$

The modularity density of a community partition is the summation of the modularity degrees of all communities  $G_i$  of in the given partition. Notice that this measure provides a way to determine if a certain microscopic description of the graph is accurate in terms of communities. The larger the value of  $Q$ , the higher quality a partition is. Therefore the community detection problem in bipartite networks can be treated as a problem of finding a partition of a network such that its bipartite modularity density  $Q$  is maximized. Since our approach is to maximize the modularity density  $Q$ , every term  $Q(G_i)$  must be non-negative. Therefore, the communities of bipartite networks by optimizing  $Q$  are consistent with the strong and weak community definition in unipartite networks suggested by F. Radicchi *et al.* [25] which is also suited to bipartite networks.

#### 4. SUPERIORITY - BIPARTITE MODULARITY DENSITY

To overcome the resolution limit in the existing modularity measurements, we introduce a new quantitative measure called modularity density which is related to the density of sub-graphs. In this section, we use example to illustrate the superiority and reliability of the bipartite modularity density.

##### A Complete Bipartite Graph

We now define a  $K_{n,m}$  clique as a complete bipartite sub-graph with  $n$  nodes of the  $V^X$  type and  $m$  nodes of the  $V^Y$  type. Given such a complete bipartite graph  $K_{n,m}$  with  $n$  nodes of the  $V^X$  type and  $m$  nodes of the  $V^Y$  type, community detecting methods by maximizing Murata's bipartite modularity divides  $K_{n,m}$  into two or more bipartite sub-graphs as communities. But it is obvious that the best solution is treating  $K_{n,m}$  itself as one community since it is a completed sub-graph. Therefore, Murata's bipartite modularity optimization obviously is not reasonable in this situation. But our bipartite modularity density can obtain the optimal solution and avoids partitioning it into two or more bipartite sub-graph.

Now we prove this superiority of our bipartite modularity density. Suppose a partition scheme  $P$  divides the complete bipartite graph  $K_{n,m}$  into two communities  $G_1$  and  $G_2$ , and the numbers of nodes in  $G_1$  and  $G_2$  are  $n_1$  of the  $V^X$  type,  $m_1$  of the  $V^Y$  type,  $n_2$  of the  $V^X$  type,  $m_2$  of the  $V^Y$  type, respectively. Another partition scheme  $G$  treats the complete bipartite graph  $K_{n,m}$  as one community. Obviously, scheme  $G$  is the optimum and better than scheme  $P$ . Let  $Q_G$  and  $Q_P$  be the

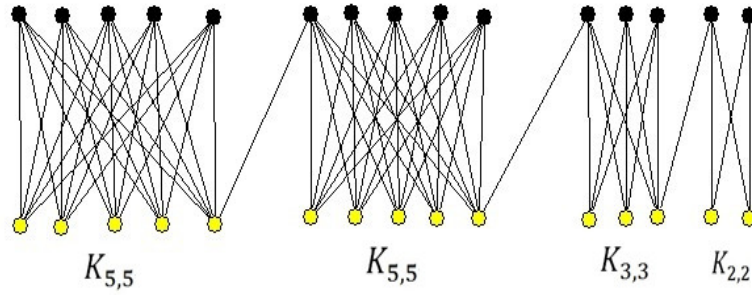


Fig. (1). A bipartite network with small complete bipartite sub-graph.

bipartite modularity densities of  $G$  and  $P$  respectively. By (2) and (5), we can get the value of  $Q_G$  and  $Q_P$  as follows:

$$Q_G = \frac{n \times m}{n + m} \Rightarrow 1$$

$$Q_P = \frac{n_1 m_1 - n_1 m_2 - m_1 n_2}{n_1 + m_1} + \frac{n_2 m_2 - n_2 m_1 - m_2 n_1}{n_2 + m_2}$$

It is easy to verify that when  $2 \leq n, m$  and  $1 \leq n_1, m_1 \leq n, m$ ,

$$Q_P < 0$$

Because  $Q_G > Q_P$ , optimizing bipartite modularity density can reach the optimal solution  $G$ , and does not divide the complete bipartite graph into two or more parts. If we optimize Murata's bipartite modularity as objective function, we cannot obtain such optimal solution  $G$  since Murata's bipartite modularity for  $G$  is zero. Therefore our bipartite modularity density is superior to Murata's bipartite modularity.

### Small Complete Bipartite Sub-graph

Given a bipartite network consisting of four cliques, two of which are  $K_{n,m}$  and the others are  $K_{p,q}$ , where  $2 \leq p \leq n$  and  $2 \leq q \leq m$ . Since the resolution limits of Murata's bipartite modularity, it has a tendency to merge small complete bipartite sub-graph and small bipartite modules. In the following we prove the bipartite modularity density based optimization does not have such a problem. Let  $P_1$  be the partitioning scheme which divides the network into four communities, each clique is a community. Let  $P_2$  be the partitioning scheme which divides the network into three communities: each of the large clique  $K_{n,m}$  is a community, and the two small bipartite cliques  $K_{p,q}$  are merged into one community.

Obviously, scheme  $P_1$  is the optimum and better than scheme  $P_2$ . Let  $Q_1$  and  $Q_2$  be the bipartite modularity densities of  $P_1$  and  $P_2$  respectively. By (2) and (5), we can get the value of  $Q_1$  and  $Q_2$  as follows: are as follows

$$Q_4 = \frac{nm - 1}{n + m} + \frac{nm - 3}{n + m} + \frac{pq - 2}{p + q} + \frac{pq - 2}{p + q}$$

$$Q_3 = \frac{nm - 1}{n + m} + \frac{nm - 3}{n + m} + \frac{pq + pq + 1 - 2}{(p + p)(q + q)}$$

It is easy to verify that when  $2 \leq p \leq n$  and  $2 \leq q \leq m$ ,

$$Q_4 - Q_3 = \frac{pq - 3.5}{p + q} > 0$$

Because  $Q_1 > Q_2$ , optimizing bipartite modularity density can reach the optimal solution  $P_1$ . If we optimize Murata's bipartite modularity as objective function, we cannot obtain such optimal solution  $P_1$ , and it will merge the two small bipartite cliques into one community. We will empirically illustrate this fact by examples in Section V. The experimental results show that our bipartite modularity density is superior to Murata's bipartite modularity.

From those examples, we can see that correct community partitioning results in bipartite network can be achieved by maximizing the bipartite modularity density. Such a fact demonstrates the effectiveness of our bipartite modularity density acting as a quantitative function for community structure.

## 5. EXPERIMENT RESULTS

We conduct experiments on both artificial network and well-studied real networks to compare our bipartite modularity density with other measurements for evaluating community structure in bipartite networks, such as the modularities proposed by Guimera [17], Baber [18], and Murata [19], we adopt these measures to evaluate the partitioning community structure. We will empirically illustrate the effectiveness of our bipartite modularity density and demonstrate that our bipartite modularity density is more effective and reasonable than other ones in detecting communities in bipartite networks.

### Artificial Bipartite Network

To show the superiority of our bipartite modularity density, we use an artificial bipartite network as shown in Fig. (1). This bipartite network has 30 nodes where each type has 15 nodes. In the network, there are four bipartite cliques with different sizes:  $K_{5,5}$ ,  $K_{5,5}$ ,  $K_{3,3}$  and  $K_{2,2}$ .

Let  $P_0$  denote the partition which has 4 communities with different size:  $K_{5,5}$ ,  $K_{5,5}$ ,  $K_{3,3}$  and  $K_{2,2}$ , and let  $P_1$  denote the partition which divides the network into 3 communities where  $K_{3,3}$  and  $K_{2,2}$  are merged into one community. Then we use our modularity and Murata's modularity to evaluate the

**Table 1. Evaluations of the partitions  $P_0$  and  $P_1$ .**

Partitions Measure	$P_0$	$P_1$
Our-Modularity	6.167	5.999
Murata-Modularity	0.789	0.793

two different partitions  $P_0$  and  $P_1$ , and the evaluation results are summarized in Table 1.

From the Fig. (1), we can directly know the communities structure of the artificial bipartite network. Since two adjacent bipartite cliques are connected by only one edge, obviously  $P_0$  is the best solution since each community it detected is a clique. From Table 1, we can see that our modularity get higher value for the partition of  $P_0$ , while Murata's modularity get higher value for the partition of  $P_1$ . This result proves that Murata's modularity has a tendency to merge small complete bipartite sub-graph and small bipartite modules, and his modularity is obviously unreasonable in this example of community detecting. Through the comparison and analysis, we can know that our modularity is more reasonable and believable than Murata's.

### Southern Women Network

To verify the accuracy and reliability of our bipartite modularity density, we test on the southern women network [26]. Southern women network is a well-known benchmark of network analysis, and its community structure is already known. This dataset was collected by Davis *et al.* in the town of Natchez, around Mississippi during the 1930s as part of an extensive study of class and race in black and white society in the Deep South. The network describes the participation of 18 women in 14 social events. If a woman attended an event, there will be an edge linking their nodes.

Southern women dataset and its bipartite community structure have been much studied by social scientists. In the original investigation, Davis *et al.* [26] used general ethnographic knowledge of the community to assign the women to two communities. The communities consisted of women numbered 1-9 and of women numbered 9-18, and woman 9 is a secondary member of both communities. A similar grouping that put women 1-9 in the first community and women 10-18 in the second one has been identified by Freeman [27] as the consensus from 21 different studies. Taking these two statements into account, we believe that the partition that divides women into two communities: {1-9} and {10-18}, is the more reasonable partition.

As we know, when the correct community structure is available, the most widely used measure for evaluating a given partition is normalized mutual information (NMI) [28], which is based on information theory and calculates the agreement between the two partitions. If we only look at the part of women, for whom we know the correct partition, we can make an objective evaluation for different partition using NMI in contrast to the correct partition. For example, when the two partitions match completely, we have a maximum

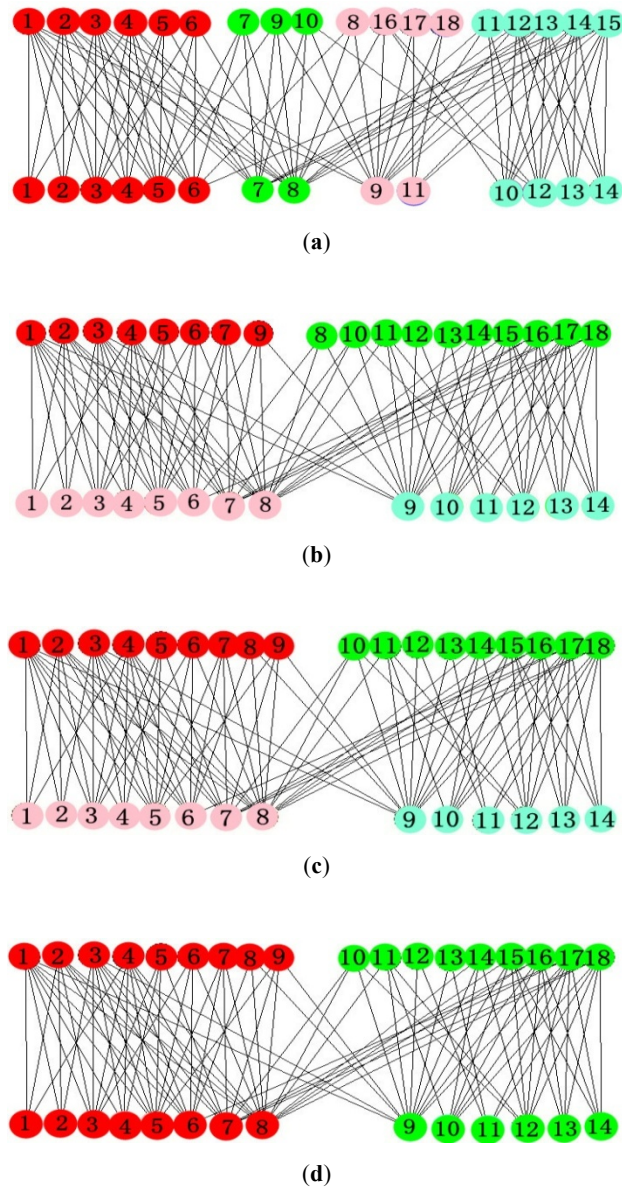
value of 1, and when the given partition is totally independent of the correct partition, we have a minimum value of 0. A larger value of NMI indicates a better partition.

Firstly, we use our bipartite modularity density as an objective function to detect the community structure on the Southern women bipartite network, and the optimization method is based on ant colony optimization for detecting community structures from bipartite networks proposed by Yongcheng Xu *et al.* [29]. And the best partition is as shown in Fig. (2d), and here the communities of different node type are one-to-one correspondence.

Then we using previously mentioned bipartite modularities as an objective functions to optimization can obtain different partitions of the Southern Women bipartite network, as shown in Fig. (2a-c). Firstly we calculate 4 different partitions of the Southern Women bipartite network with the measure of NMI. Then we use different bipartite modularity to evaluate partition (a)-(d), and compare their evaluation results with the objective evaluation discussed above. All of the evaluation results are summarized in Table 2.

We should pay attention that comparing the absolute values between modularity is not enough to show the rationality of the measurements, since different bipartite modularity concern different features. To make it clear, we rank these partitions according to the absolute values as shown in Table 3 where each row lists the ranking of different partitions, from 1 to 4, as evaluated by the corresponding bipartite modularity. As shown in Fig. (2), we can see that (c) which is obtained by optimizing Murata's modularity using a simple agglomerative algorithm and (d) which is obtained by optimizing our modularity using an ant colony optimization algorithm are all the best solution according to at the rankings of NMI.

From Tables 2 and 3, we can see that our modularity and Murata's modularity are the most reliable, because the corresponding rankings coincide with that of the objective evaluation rankings according to NMI. Our modularity can identify partitions (d) and (c) as the best solution, (b) and (a) as the second best and the third best partitions. Compared with Barber's modularity, our modularity coincides with the corresponding rankings according to NMI. Moreover, the main weakness of Barber's modularity is that the number of communities has to be determined in advance. It should be emphasized that in most of the community detection algorithms, the number of communities must be specified in advance not be found by the method itself. This is not practical in many real applications. In contrast to Guimera's modularity which only focuses on the connectivity of only one node type and aims at partitioning communities for nodes of one type at a



**Fig. (2).** Communities of the Southern women bipartite network. Women nodes are indicated as circle symbols located at the top side, while event nodes are indicated as circle symbols located at the bottom side. (a) Final division obtained by BRIM, an algorithm for optimizing Baber-Modularity. (b) Final division obtained by optimizing Guimera-Modularity using simulated annealing. (c) Final division obtained by optimizing Murata-Modularity using a simple agglomerative algorithm. (d) Final division obtained by optimizing Our-Modularity using ACO for community detection on bipartite networks.

**Table 2.** Evaluations of the partitions (a), (b), (c) and (d) in Fig. (2) by different measures.

Partitions Measures	(a)	(b)	(c)	(d)
NMI	0.451	0.743	1.000	1.000
Our-Modularity	-4.07	2.813	2.965	2.965
Guimera-Modularity	0.135	0.217	0.215	0.215
Baber-Modularity	0.345	0.35	-	0.318
Murata-Modularity	0.465	0.579	0.586	0.586

**Table 3. Rankings of the partitions (a), (b), (c) and (d) in Fig. (2) by different measures.**

Partitions Measures	(a)	(b)	(c)	(d)
NMI	3	2	1	1
Our-Modularity	3	2	1	1
Guimera-Modularity	3	1	2	2
Baber-Modularity	1	3	2	2
Murata-Modularity	3	2	1	1

time, our modularity can partition two types of nodes into communities symmetrically.

From this example, we can believe that our modularity is more accurate and reliable than other measurements for evaluating community structure in a real-world bipartite network. Our modularity can be used as an objective function for optimization in detecting communities in a real-world bipartite network.

## CONCLUSION

In this work, we propose a new measurement for evaluating community structure in bipartite networks. As traditional bipartite modularities contain an intrinsic scale that depends on the total size of edges in the bipartite networks and ignores the total size of nodes in the bipartite networks. In addition, when optimizing these traditional bipartite modularities, they tend to merge tiny bipartite communities which are smaller than a scale. Those small communities may not be resolved even in the extreme case that they are complete graphs connected by single edges. Fortunately, our bipartite modularity density can resolve these resolution limits. Our experimental results on the artificial and real-world bipartite networks, it is show that our bipartite modularity density is quiet accurate and reliable.

Moreover, we optimize our bipartite modularity density with ant colony optimization algorithm for detecting communities on the famous Southern Women bipartite networks, the result show us that our bipartite modularity density can be used as the objective function in optimization for community detection in bipartite networks, and can obtain more accurate solution than using other modularity measurements.

## CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

## ACKNOWLEDGEMENTS

This research was supported in part by the Chinese National Natural Science Foundation under grant Nos. 61379066, 61070047, 61379064, Natural Science Foundation of Jiangsu Province under contracts BK20130452,

BK2012672, BK2012128, and Natural Science Foundation of Education Department of Jiangsu Province under contract 12KJB520019, 13KJB520026, 09KJB20013.

## REFERENCES

- [1] D. Watts and S. Strogatz, "Collective dynamics of small-world networks", *Nature*, vol. 393, no. 6684, pp. 440-442, 1998.
- [2] M. E. J. Newman, "Scientific collaboration networks. I. network construction and fundamental results", *J. Phys. Rev. E.*, vol. 64, pp. 016131, 2001.
- [3] M. E. J. Newman, "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality", *J. Phys. Rev. E.*, vol. 64, pp. 016132, 2001.
- [4] L. Ai-fen, F. Chun-hua, Z. Zeng-ping, C. Hui, H. Da-ren, "An empirical statistical investigation on chinese mainland movie network", *Comp. Sys. Complex. Sci.*, vol. 4, no. 3, pp. 10-16, 2007.
- [5] C. Wen-qin, L. Jun-an and L. Jia, "Research in disease-gene network based on bipartite", *Netw. Project.*, vol. 6, no. 1, pp. 13-19, 2009.
- [6] G. Ergun, "Human sexual contact network as a bipartite graph", *J. Phys. A.*, vol. 308, pp. 483-488, 2002.
- [7] R. Lambiotte, M. Ausloos, "Uncovering collective listening habits and music genres in bipartite networks", *J. Phys. Rev. E.*, vol. 72, p. 066107, 2005.
- [8] S. Le Blond, J. L. Guillaume, M. Latapy, "C lustering in P2P exchanges and consequences on performances", *C CastroM, Renesse R. Peer- to-Peer Systems IV*. Berlin: Heidelberg, pp. 193-204, 2005.
- [9] G. Robins and M. Alexander, "Small worlds among interlocking directors: network structure and distance in bipartite graphs", *J. Comput. Math. Organ. Theory*, vol. 10, pp. 69-94, 2004.
- [10] S. Battiston and M. Catanzaro, "Statistical properties of corporate board and director networks", *Eur. Phys. J. B.*, vol. 38, pp. 345-352, 2004.
- [11] 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.
- [12] 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.
- [13] M. A. Porter, J. P. Onnela, P. J. Mucha, "Communities in networks", *Notices Am. Math. Soc.* vol. 56, pp. 1082-1097, 2009.
- [14] 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.
- [15] M.E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks", *Phys. Rev. E*, vol. 69, no. 026113, pp. 1-15, 2004.
- [16] Z. Li, S. Zhang, R.S. Wang, X.S. Zhang and L. Chen, "Quantitative function for community detection", *Phys. Rev. E*, vol. 77, no. 1, p. 036109, 2008.

- [17] R. Guimera, M Sales-Pardo and L. A Amaral, "Module identification in bipartite and directed network", *J. Phys. Rev. E*, vol. 76, p. 036102, 2007.
- [18] M. J. Barber, "Modularity and community detection in bipartite networks", *Phys. Rev. E*, vol. 76, no. 066102, pp. 1-9, 2007.
- [19] T. Murata, "Community division of heterogeneous networks", *Proceedings of 1st International Conference Complex Sciences: Theory and Applications*, Feb. 2009, pp. 12-23.
- [20] L. Danon, A. Díaz-Guilera, and A. Arenas, "Comparing community structure identification", *J. Stat. Mech.-Theory Exp.* vol. 09, p. 09008, 2005.
- [21] S. Fortunato and M. Barthélemy, "Resolution limit in community detection", *Proc. Natl. Acad. Sci., USA*, vol. 104, pp. 36-41, 2007.
- [22] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering overlapping community structure of complex networks in nature and society", *Nature*, vol. 435, p. 814, 2005.
- [23] G. Palla, A.-L. Barabási, and T. Vicsek, "Quantifying social group evolution", *Nature*, vol. 446, p. 664, 2007.
- [24] M. Rosvall and C. T. Bergstrom, "An information-theoretic framework for resolving community structure in complex networks", *Proc. Natl. Acad. Sci. USA*, vol. 104, p. 7327, 2007.
- [25] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks", *Proc Natl Acad. Sci., USA*, vol. 101, p. 2658, 2004.
- [26] A. Davis, B.B. Gardner, and M. R. Gardner, "Deep South", University of Chicago Press, 1941.
- [27] R. Breiger, K. Carley, and P. Pattison, "Dynamic Social Network-Modeling and Analysis", Workshop Summary and Papers. Washington, DC: The National Academics Press, 2003.
- [28] A. Fred and A. Jain, "Robust data clustering," In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2003.
- [29] Y. Xu, L. Chen, and R. Zou, "Ant colony optimization for detecting communities from Bipartite network", *Journal of Software*, vol. 8, no. 11, pp. 2930-293, November 2013.

---

Received: November 12, 2014

Revised: January 08, 2015

Accepted: January 21, 2015

© Xu and Chen; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.