Community Detection Using Discrete Bat Algorithm

Anping Song, Mingbo Li, Xuehai Ding, Wei Cao, Ke Pu

***Abstract*—Community detection is very important for understanding the function of complex networks. The traditional community detection algorithms such as spectral clustering and FN (Fast Newman) algorithm tend to fall into local optimum or cannot identify the number of communities. In this paper, based on BA (Bat Algorithm), a new swarm intelligence optimization algorithm, we proposed DBA (Discrete Bat Algorithm) for community detection. DBA can automatically identify the number of communities and easily search the global optimal solution to overcome the shortcomings of traditional algorithms. Compared with traditional algorithms, experimental results show that DBA has higher accuracy and efficiency. We applied DBA to the community detection of bidding networks and achieved good results, which shows that our algorithm has a strong practical value.**

***Index Terms*—complex network, community detection, Bat Algorithm, bidding detection**

I. Introduction

Complex networks originated in our lives. For example, the World Wide Web, social networks, electricity networks, neural networks and protein interaction networks can all be seen as a complex network. Study of complex networks involving physics, mathematics, biology and sociology, has become one of the most important interdisciplinary fields [1], [2]. With the in-depth study of complex networks, scholars have found that many complex networks have a common nature: community structure. Like the small world [3], scale-free [4], the community structure of complex networks is one of the most popular and important properties of the topology. Community detection can be applied to the web mining, social network analysis, semantic-analytics [23] and many fields of biology and has important significance in research.

Girvan and Newman are pioneers in the research of community detection algorithm. They proposed the Girvan-Newman (GN) algorithm [5] in 2001. GN algorithm iteratively computes the edge betweenness and removes the edge of maximum betweenness. A hierarchical clustering tree is built by top-down approach. Because of the high computational demands, GN algorithm is only suitable for small networks. But GN algorithm plays a very important role in the field of community detection. Community structure was considered as a ubiquitous property of complex network for the first time, which inspired in-depth study on this issue and set off a wave of research on community detection [6]. Maximum flow community (MFC) algorithm [7] and GN algorithm are similar in the way of cutting network. Theoretical basis of MFC algorithm is the Max Flow-Min Cut theorem. MFC algorithm continuously removes Min Cut edge, so network is divided into communities iteratively. The principal disadvantage of MFC algorithm is the high computational demands it makes.

Fast Newman (FN) [8] algorithm was proposed by Newman in 2004. It considered community detection as a global optimization problem and the objective function is Q function (modularity function). By local search strategy, FN algorithm repeatedly joins communities together in pairs, choosing the join that result in the maximum ΔQ at each step, which build a hierarchical clustering tree from the bottom up. FN algorithm has a considerable advantage over GN algorithms in time complexity, but the result obtained by FN is a local optimum and lacks diversity. GA (Guimera-Amaral) algorithm [9], based on simulated annealing algorithm, was proposed by Guimera and Amaral in 2005. The objective function of GA algorithm is the same with FN algorithm. However, GA algorithm is slow in convergence and sensitive to parameters.

Spectral clustering [10]-[12] transforms community detection to the approximate optimal solution of constrained quadratic form, which is strict in mathematical theory, but does not automatically recognize the number of communities.

Basing on optimization method, Ronghua Shang *et al* proposed an improved genetic algorithm using modularity [13]. Genetic algorithm is very slow in convergence, because the mutation and crossover operations are random. Many scholars try to apply swarm intelligence algorithm in community detection. H Chang *et al* applied Ant Colony Optimization to community detection [14] with the help of a new kind of heuristic information. Qing Cai *et al* applied particle swarm optimization (PSO) algorithm to the community detection of signed network [15] in 2014, and have achieved good results. Bat Algorithm (BA), proposed by Yang in 2010, is a new swarm intelligence algorithm [16]. The original BA can only be used to solve the continuous optimization problem. In this paper, we discretized the original algorithm and proposed a novel discrete bat

Anping Song is associate professor with the School of Computer Engineering and Science, Shanghai University, Shanghai, 200444 China (e-mail: apsong@shu.edu.cn).

Mingbo Li is with the School of Computer Engineering and Science, Shanghai University, Shanghai, 200444 China (e-mail: lmb@shu.edu.cn).

Xuehai Ding is with the High-Performance Computing Center, Shanghai University, Shanghai, 200444 China (e-mail: dinghai@shu.edu.cn).

Wei Cao is with the School of Computer Engineering and Science, Shanghai University, Shanghai, 200444 China (e-mail: caowei@shu.edu.cn)

Ke Pu is with the School of Computer Engineering and Science, Shanghai University, Shanghai, 200444 China (e-mail: frankpu @shu.edu.cn).

This study is financially supported by National Natural Science Foundation of China (Grant No.91330116) and (Grant No.61303094). Their support is greatly appreciated.

algorithm (DBA) to detect the community structure in complex network. Experimental results show that compared with several algorithms described above, DBA has obvious advantages on the accuracy of community partition and efficiency. Then we applied DBA to the bidding networks and achieved good results.

The rest of this paper is organized as follows. Section II focuses on related work and BA algorithm; Section III shows the discretization of BA and the processes of DBA; Section IV is experiment results in artificial network and real-word networks; Section V is the conclusion and prospect.

II. Related work

A. The definition of community

Community structure is an important feature that many complex networks have in common, but there is no strict definition yet. In literature [17], communities are regarded as sub-graphs of network which have dense intra-links and sparse inter-links. In reality, the communities correspond to specific feature or attribute of a network. For example, in a social network, a community may represent a circle of friends; in a literature mutual reference network, a community may represent a relevant scientific field.

Thus, community detection is to divide complex network into several sub-graphs, the mathematically description is as follows:

Given a complex network *G* (*V, E*), where *V* is the vertices set, *E* is the edges set. To find a division $C = C_1C_2......C_k$, where $C_1 \cup C_2 \cup ......C_k = V$ and $C_i \cap C_j = \varnothing \ \forall i, j$. A good division of *C* is to make intra edges of C_i as many as possible and edges between C_i and C_j as few as possible.

B. Modularity function

Modularity function was proposed by Newman in 2004 [18], which is currently the most widely used evaluation function of community detection.

Given a network, assume that the network was divided into *k* communities. Define a k-order symmetric matrix *e*, let e_{ij} be one-half of the fraction of edges in the network that connect vertices in community *i* to those in community *j*. so that the total fraction of such edges is $e_{ij} + e_{ji}$. The trace $Tre = \sum_i e_{ii}$ is the total fraction of edges that fall within communities. A division is good if the *Tre* is large. But the *Tre* cannot be a good measure for community detection. If all vertices are put into a single community, we will get maximal value of 1, which does not reflect any community structure of the network. Thus, let a_i be the fraction of edges that connects to the community *i*, we define the modularity function as follows:

$$Q = \sum_i (e_{ii} - a_i^2) = Tre - \left\| e^2 \right\| \tag{1}$$

Where $\|x\|$ is the sum of all elements in the matrix *x* and a_i^2 is the fraction of edges that fall within communities when all edges are connected together at random. Q is the difference between e_{ii} and a_i^2. If the edges of a network are connected at random, this means no community structure is formed, the value of Q approximately equal to 0. Maximum Q is 1 and the actual situation is in the range of 0.3 and 0.7. $Q > 0.3$ indicates significant community structure in the network.

C. The fraction of vertices identified correctly

If the community structure of a network is known, we use *FVIC* (the fraction of vertices identified correctly) [19] to evaluate the result of community detection. We define *FVIC* as follows:

$$FVIC = \sum_{i=0}^{k} \frac{maxC_i}{N} \tag{2}$$

Given a network, *C'* is the correct community division, including *k* communities. *C* is the community detection result. $\forall C_i \in C'$, find $C_j \in C$ that has the most common vertices with C_i, denoted as $maxC_i$. Where *N* is the number of vertices in the network. The larger *FVIC* is, the closer the result is to the correct division.

D. Bat Algorithm

BA was proposed by famous scholar Yang in 2010. BA, which combines the features of PSO algorithm and simulated annealing algorithm, is strong in convergence and global search capability.

Basing on the echolocation behavior of bats, BA simulates the frequency, emission rates and the loudness when bats forage. Bats change the wavelength λ by adjusting the pulse frequency $f\,(\lambda f = \upsilon)$. When the wavelength coincides with the size of insects, bats are able to locate the target. The position and velocity updating rule of BA is similar with PSO algorithm. BA generates a random solution by a random flight operation to avoid falling into local optimum, which is the defect of PSO. The local search operation of BA is similar to simulated annealing algorithm, which makes the BA algorithm converge more quickly. BA combines the advantages of PSO algorithm and simulated annealing algorithm, which makes BA more outstanding.

III. Discrete bat algorithm

A. Define of bat location and velocity

The original BA algorithm is used to solve continuous problem, but community detection is a discrete problem. Therefore we propose the DBA for community detection.

First, we define the position of bat. In the original BA, if the solution space is n-dimensional, a position is an n-dimensional vector. For community detection, we use the decimal code. The definition of a position is $X = X_1X_2......X_n$, where n is the number of vertices of the given network, $X_i \in X$ is a decimal integer. For any *i* and *j*, $X_i = X_j$ means vertex *i* and vertex *j* are divided into the same community.

Then we define the bat velocity. The velocity of original algorithm is the difference between the position of the current individual and the current best individual. But for discrete code, we cannot get the difference directly by subtraction. Therefore, we propose a new velocity formula, the operation between the current individual and the current best individual becomes *XOR*. In fact, bat will adjust its velocity by learning from the current best. The learning process is actually a comparison between the positions. So the *XOR* operation actually reflects the difference between two network divisions, as a bat on behalf of a network division.

If $X_i = X_i^*$, where X_i^* is the current best individual, then $X_i \oplus X_i^* = 0$; If $X_i \neq X_i^*$, then $X_i \oplus X_i^* = 1$. It can be seen that *XOR* well reflects the difference between the current individual and the current best individual. The current individual will adjust itself by learning from the current best individual.

B. Discrete formula

The position and velocity update formula of DBA algorithm are defined as follows:

Pulse frequency formula:

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \tag{3}$$

Where β is a random number, $0 < \beta < 1$, $f_{\max}$ is maximum frequency, $f_{\min}$ is minimum frequency. The velocity at step t is given by:

$$V_i^t = V_i^{t-1} + (X_i^t \oplus X_i^*) f_i \tag{4}$$

Where V_i^t is the new velocity, V_i^{t-1} is the velocity of the previous generation, X_i^t is the current position, X_i^* is the current best position, f_i is the current frequency, $\oplus$ is *XOR*.

Velocity discrete formula:

$$Sig(V_i^t) = 1/(1 + exp(-V_i^t)) \tag{5}$$

$$V_{id}^t = \begin{cases} 1 & if\ rand() < Sig(V_i^t) \\ 0 & if\ rand() \geq Sig(V_i^t) \end{cases} \tag{6}$$

Where *Sig* is the sigmoid function. *Sig* mapped the velocity to the range (0, 1). *rand*() generates a random number in the range of (0, 1).

We can get the discrete velocity V_{id}^t by the above formula. Then we will calculate the new position of the bat. In the original algorithm, we can get the new position vector by adding the position of the previous generation to the current velocity vector, but not for discrete issues. In this regard, we proposed a new position update function.

For the current position $X^t = X_1^t X_2^t......X_n^t$:

$$X_i^t = \begin{cases} X_i^t & if V_{id}^t = 0 \\ X_i^t & if\ V_{id}^t = 1, rand() \geq r_i \\ X_{inew}^t & if\ V_{id}^t = 1, rand() < r_i \end{cases} \tag{7}$$

Where r_i is current emission rate, X_{inew}^t is the new position of X_i^t after adjustment, the adjusting method is described as follows:

For each vertex v_i, $0<i<n$, if $V_{id}^t = 1$, then calculate:

$$Connect_j = f(v_i, C_j) \quad C_j \in C, 0 < j < k \tag{8}$$

Where $C = C_1 C_2......C_k$ is a community detection result; k is the number of communities. $f()$ calculates the edges between vertex v_i and $C_j, 0 < j < k$.

If community $C_{\max} (C_{\max} \in C)$ corresponds to the max

TABLE I
THE PSEUDO-CODE OF DBA

The pseudo-code of DBA
Input: adjacency matrix of complex networks, population and other parameters
Output : communities
1) Initialize the population *m*, the max and min pulse frequency $f_{\max}$ and $f_{\min}$, max emission rates r_0, max loudness A_0, frequency increase factor *γ*, sound attenuation coefficient *α*, the max number of iterations *step*
2) Initialize the position and velocity of each bat randomly, calculate the *fitness* of each bat, select the current best X^*
3) while (*t<step*)
4) Update the current velocity, position and fitness of each bat
5) if ($rand() \geq r_i$)
6) Generate a local solution X_{local} for the current bat by local search
7) if ($f(X_{local}) > f(X_i)$)
8) Replace the current solution with X_{local}
9) Generate a new random solution X_{new}
10) if ($rand() < A \,\&\, f(X_{new}) > f(X^*)$)
11) Accept the new solutions X_{new}, update A and r
12) Get the current best by rank all bats
13) end while
14) Decode the optimal solution is and output the communities.

Connect, adjust v_i to $C_{\max}$ with a certain probability.

C. Local search

The local search strategy of DBA is similar to the simulated annealing algorithm. Advantage of local search is that it can easily search for the optimal solution and accelerate the convergence. The method is outlined below.

Given a solution $C = C_1C_2......C_k$, $\forall C_i \in C$, calculate the *Connect* of vertex v and community C_i , see equation (8), where $v \in C_i$. For the vertex $v_{\min}$ corresponding to the minimum *Connect*, calculate its *Connect* with all other communities C_j , where $C_j \in C$ *and* $i \neq j$. Finally, we adjust vertex $v_{\min}$ to $C_{j\max}$, where $v_{\min}$ and $C_{j\max}$ have the maximum *Connect*.

D. Work flow of DBA

We use a decimal code in this paper. If two vertices have the same code, they are divided into the same community. In the initialization, we generate each code randomly. However, this will divide some unconnected vertices into the same community, which is obviously of no sense. Therefore, we test each code in the initialization, so as to avoid nonsense initialization.

The objective function of DBA is the Q function. We use two termination conditions. First, manually set the number of iterations; second, if the Q function value of the optimal solution remains unchanged for 10 iterations, the algorithm stops.

Table I shows the pseudo-code of DBA.

IV. Experiments and Results

In this paper, experiments were performed in desktop computer with the operating system Windows7, CPU i3-3240, clocked 3.4GHz, memory 4G. Our algorithm was coded in Java, JDK1.7. Compiling environment is MyEclipse.

We applied DBA to artificial network, real-word networks and bidding networks. The difference between artificial network and real-word network is that the community structure of the former is known in advance.

A. Artificial network

First, we used a symmetrical network with known community structure proposed by Girvan and Newman in the paper [5] to validate the performances of DBA. The network generated by computer, constructed with 128 vertices, is divided into four communities in average. The edges are randomly generated by a fixed probability p_{in} and p_{out} . Where p_{in} is the possibility that both vertices of an edge belong to the same community, and p_{out} for different communities, $p_{in} > p_{out}$. The probabilities were chosen so as to keep the average degree of a vertex equal to 16.

In the experiments, we changed the dense between communities by adjusting p_{in} and p_{out} . When p_{out} increases, p_{in} decreases, the intra-links decrease and inter-links increase. This makes community division more difficult.

The proposed algorithm was used in the community detection of the above network, and several other classical algorithms for comparison. Initial p_{out} is 0, which means no connection between the communities. The experimental results of FN algorithm, spectral clustering, DPSO algorithm and DBA are shown in Fig 1.

In Fig 1, the horizontal axis represents p_{out} , the vertical axis represents *FVIC*. Original DPSO algorithm is used for community detection of signed network, so we changed it and applied it to undirected network. We ran each algorithm 10 times for each p_{out} , taking the best result.

We can see from Fig 1 that *FVIC* obtained by DBA was 1 with $p_{out} = 6$. This means that if $p_{out} \leq 6$, community division of DBA is completely correct. p_{out} of FN algorithm and DPSO algorithm in this case are 5 and 5.5 respectively, which is significantly lower than DBA. Only for $p_{out} \geq 7$ does *FVIC* of DBA start to fall off substantially. In other words, the algorithm performs very well almost to the point at which each community has as many intra-links as inter-links. The curve of DBA is above DPSO and FN all the time, which means accuracy of DBA is better than the other two algorithms.

For spectral clustering, when $p_{out} \geq 1$, *FVIC* is less than 1. The process of spectral clustering is the process to relax the

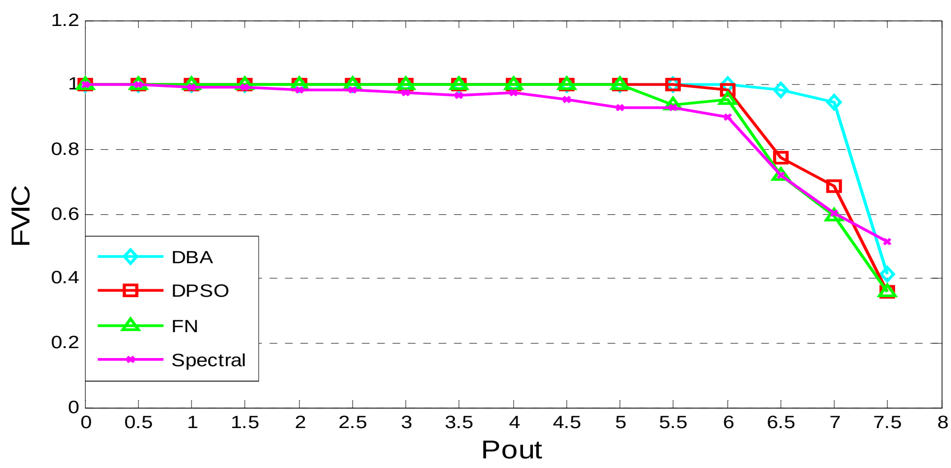

Fig. 1. Experiment results of artificial network

accuracy of community division, so the results will always have a little error. But when $p_{out} = 7.5$, *FVIC* obtained by spectral clustering is the highest. The reason is that the number of communities is given in advance for spectral clustering, so the communities are divided more evenly. Therefore, a higher *FVIC* does not mean that spectral clustering has advantages compared with other algorithms.

TABLE II
SPECIFICATIONS OF REAL-WORD NETWORKS

Network	Football	Karate Club	Political Books
Vertices	115	34	105
Edges	615	78	441

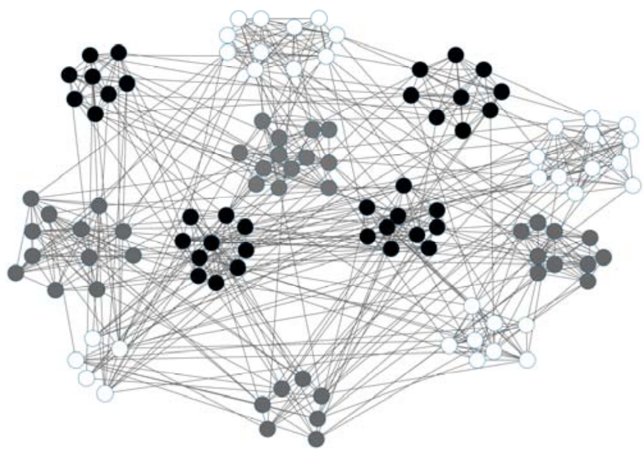

Fig. 2. Community detection result of Football

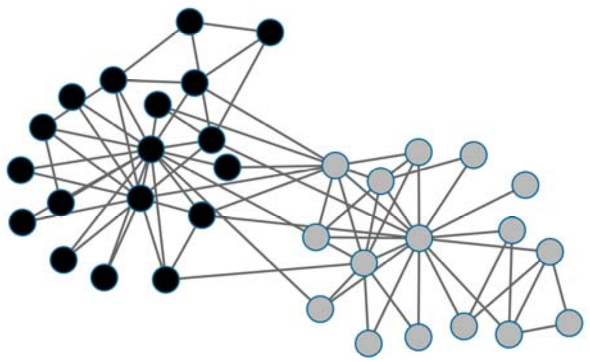

Fig. 3. Community detection result of Karate Club

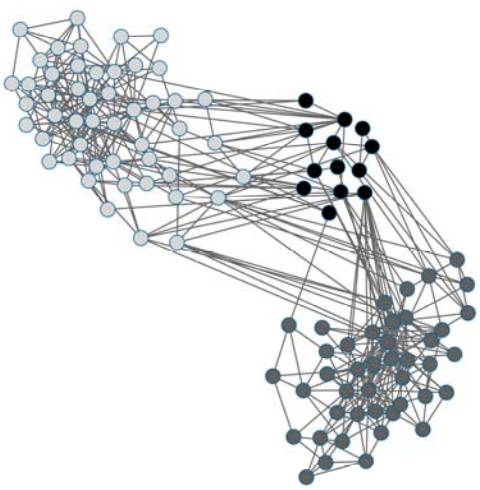

Fig. 4. Community detection result of Political Books

In summary, for the artificial network community detection, DBA algorithm performed better than the other three algorithms.

B. Real-world networks

In this section we applied our algorithm to three real-word networks. They are the American football network [5], Zachary karate club network [20] and the American political books networks [21].

American football network is a representation of the game schedule of American college football in 2000 season. These teams are from 12 leagues. Vertices represent teams and edges represent regular-season games between the two teams they connect. Games are more frequent between members of the same league than between members of different league. Zachary karate club network represents the relationship of the members. Because of dispute within the organization, the network is divided into two distinct communities. Vertices of American political books network represent political books sold on Amazon. If two books were bought by the same buyer, there is an edge between the corresponding vertices. Specifications are shown in Table II.

We ran DBA, DPSO algorithm and spectral clustering 10 times on each network, recorded the number of communities obtained by the best results, the corresponding $Q_{\max}$, and the average value Q_{avg} . Community detection results are shown in Table III.

We use a graphical tool Cytoscape to show the best results got by DBA. The community detection results of Football, Karate Club and Political Books are shown in Fig 2, Fig 3 and Fig 4 respectively.

DBA and spectral clustering algorithms detected all 12 communities in the football network; DPSO and FN were 10 and 9 respectively. The number of communities is set in advance for spectral clustering, but DBA algorithm detected it automatically. As can be seen from Table III, FN algorithm obtained the largest Q_{avg} in the football network. But all results obtained by FN algorithm are exactly the same, so $Q_{avg} = Q_{\max}$. $Q_{\max}$ of DBA algorithm is not only far greater than FN algorithm, but also higher than DPSO algorithms and spectral clustering.

Results of the Karate Club network and the Political Books network are similar to the results of the football network. FN algorithm can get the max Q_{avg} . Because of the oneness of result, $Q_{\max}$ of FN algorithm is much smaller than that of DBA and DPSO. What's more, DPSO algorithm is tends to fall into local optimum, $Q_{\max}$ and Q_{avg} are smaller than those of DBA.

Thus, we can see the advantages of DBA. It can not only automatically detect the number of communities, but also obtain more accurate results than the traditional algorithms.

TABLE III
EXPERIMENT RESULTS OF REAL-WORD NETWORKS

Network	Algorithm	Number of communities	Q_{avg}	$Q_{\max}$
Football	DBA	12	0.531	0.596
	DPSO	10	0.532	0.563
	FN	9	0.543	0.543
	Spectral	12	0.447	0.531
Karate Club	DBA	2	0.362	0.392
	DPSO	2	0.345	0.369
	FN	2	0.376	0.376
	Spectral	2	0.332	0.351
Political Books	DBA	3	0.463	0.479
	DPSO	2	0.445	0.470
	FN	4	0.467	0.467
	Spectral	3	0.394	0.439

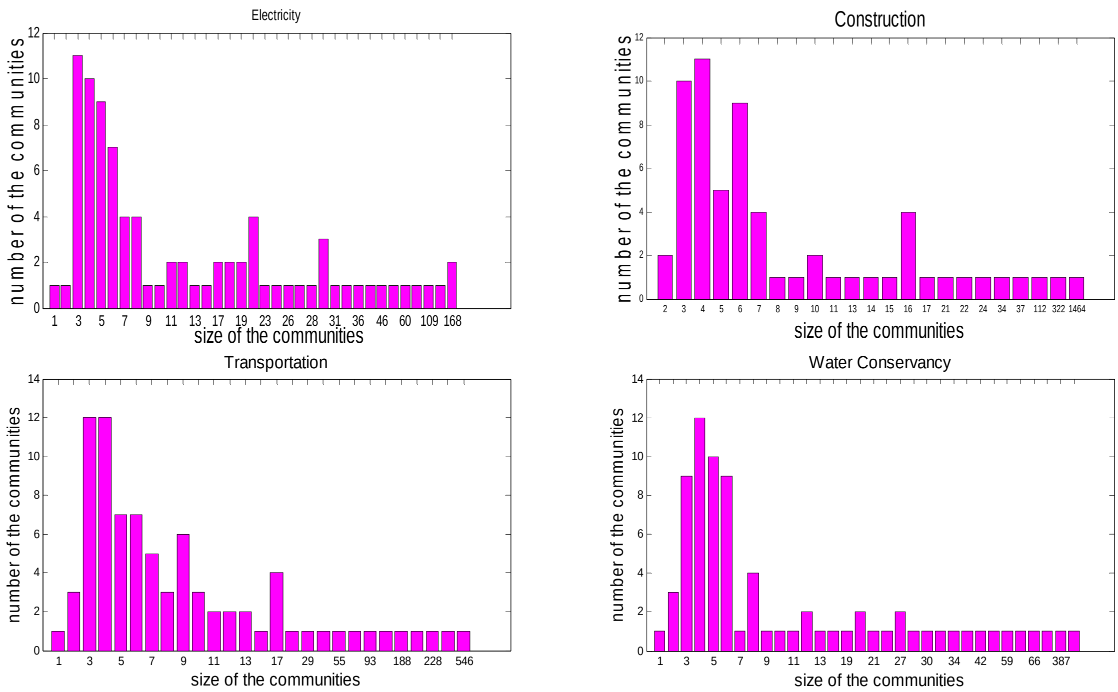

Fig. 5. Community detection results of bidding networks

C. Bidding networks

Experiment results of the above two sections show that DBA algorithm has advantages in detecting the small scale networks compared with the traditional algorithms. In this section we applied DBA to detect communities in a large complex networks. Experimental data is the 2010-2014 bidding data of a Chinese province, provided by the public resources trading center of the province.

There will be “group” phenomenon when companies bid, which means some companies always appear in the same project. These companies have suspicion of surround-bidding according to some experts. We can find out these companies by community detection.

First we classified the data according to type of project. Then for each class, we constructed networks of the companies participated in the bidding. Vertices represent companies. If two companies bid in the same project, there is an edge between them. If they bid *n* times together, weight of the between edge is *n*.

We chose bidding networks of four categories for community detection, electricity (1686 vertices), construction (2392 vertices), transportation (2384 vertices) and water conservancy (2039 vertices). The results were shown in Fig 5.

The horizontal axis represents the size of the communities and the vertical axis represents the number of the communities. Bidding experts pointed out that because of some reasons, the scale of a “group” is generally between 3 and 15. As can be seen from Fig 5, the size of communities concentrated in the range of 3 and 15. The percentage of communities with size between 3 and 15 is above 70%. That is, 70% of the communities detected by DBA are in line with the actual situation, which is of great practical value. We can regard it as a judgment of surround-bidding.

We summarize the above four categories (electricity, construction, transportation and water conservancy). These communities have a broad distribution of sizes from 1 to nearly 1500. The distribution is shown in cumulative form in Fig 6. We observe that the distribution of the community size is approximately power law in form.

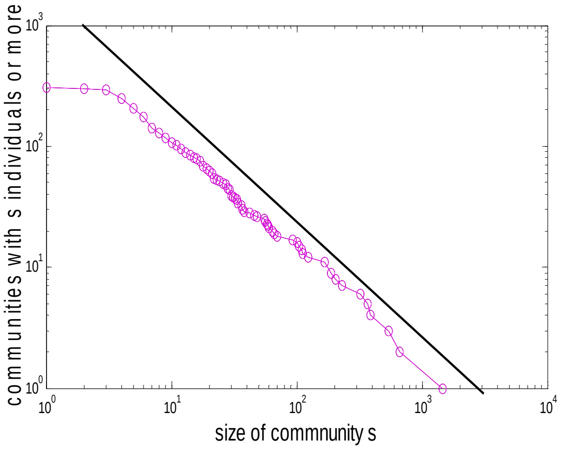

Fig. 6. Cumulative distribution of the summary result of four categories: electricity (85 communities), construction (62 communities), transportation (82 communities) and water conservancy (76 communities). A total of 305 communities were found.

In Fig 6, the black line represents the slope the plot would have if the distribution followed a power law with exponent roughly equaling to -1. The same kind of experiment was done by Newman in paper [8], using a network of collaborations between physicists. Newman got the result that the distribution is approximately power law with exponent −1.6. Compared with Newman's experiment, our experiment result is more close to a perfect power law [22], which means our result is even better.

V. Conclusions

In this paper we proposed a new community detection algorithm DBA based on swarm intelligence. DBA automatically detects the number of communities and is able to jump out of local optimum, which overcomes the shortcomings of traditional algorithms. Artificial network and real-word networks experiments have proved DBA is superior to the traditional algorithms such as FN algorithm, DPSO algorithms and spectral clustering in all respects. Furthermore, we applied DBA to the community detection of bidding networks. The results are consistent with the prediction of some experts, which also validates the practicality of DBA. We use only one objective function, Q function, in this paper and multi-objective should be considered in further studies. In addition, we will focus on the parallelization of DBA to meet the challenges of larger networks.

References

[1] Satu Elisa Schaeffer, "Graph clustering," Computer Science Review, vol. 1, no. 1, pp. 27-64, Aug 2010.

[2] FD Malliaros, M Vazirgiannis, "Clustering and Community Detection in Directed Networks: A Survey," Physics Reports, vol. 533, no. 4, pp. 95–142 Dec. 2013.

[3] Watts DJ, Strogatz SH, "Collective dynamics of Small-World networks," Nature, vol. 393, no. 6638, pp. 440-442, Jun. 1998.

[4] Barabási AL, Albert R, "Emergence of scaling in random networks," Science, vol. 286, no. 5439, pp. 509-512, Oct. 1999.

[5] Girvan M, Newman MEJ, "Community structure in social and biological networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 99, no. 12, pp. 7821-7826, Dec. 2002.

[6] Yang Bo, Liu Dayou, Liu Jiming, Jin Di, "Complex Network Clustering Algorithms," Journal of Software, vol. 20, no. 1, pp. 54-66, Jan, 2009.

[7] Flake GW, Lawrence S, Giles CL, Coetzee FM, "Self-Organization and identification of Web communities," IEEE Computer, vol. 35, no. 3, pp. 66-71, Mar. 2002.

[8] Newman MEJ, "Fast algorithm for detecting community structure in networks," Physical Review E, vol. 69, no. 6, pp. 279-307, Sep. 2004.

[9] Guimera R, Amaral LAN, "Functional cartography of complex metabolic networks," Nature, vol. 433, no. 7028, pp. 895-900, Feb. 2005.

[10] Shiga M, Takigawa I, Mamitsuka H, "A spectral clustering approach to optimally combining numerical vectors with a modular network," Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, New York, 2007, pp. 647-656.

[11] Ulrike von Luxburg, "A tutorial on spectral clustering," Statistics and Computing, vol. 17, no. 4, pp. 395-416, Dec. 2007.

[12] M. E. J. Newman, "Community detection and graph partitioning," Epl, vol. 103, no. 2, pp. 330-337, May. 2013.

[13] Ronghua Shang, Jing Bai, Licheng Jiao, Chao Jin, "Community detection based on modularity and an improved genetic algorithm," Physica A Statistical Mechanics & Its Applications. Vol.392, no. 5, pp. 1215–1231, Nov. 2013.

[14] H Chang, Z Feng, Z Ren, "Community detection using Ant Colony Optimization," IEEE Congress on Evolutionary Computation, Cancun, Jun. 2013, pp. 3072-3078.

[15] Q Cai, M Gong, B Shen, L Ma, L Jiao, "Discrete particle swarm optimization for identifying community structures in signed social networks," Neural Networks the Official Journal of the International Neural Network Society, vol. 58, no. 10, pp. 4-13, May. 2014.

[16] Xin-She Yang, "A New Metaheuristic Bat-Inspired Algorithm," Nature Inspired Cooperative Strategies for Optimization, pp. 65-74, Apr. 2010.

[17] Radicchi F, Castellano C, Cecconi F, Loreto V, Parisi D, "Defining and identifying communities in networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 101, no. 9, pp. 2658-2663, Mar. 2004.

[18] Aaron Clauset, M.E. J. Newman, Cristopher Moor, "Finding community structure in very large networks," Physical Review, vol. 70, no. 6, pp. 264-277, Dec. 2004.

[19] Newman MEJ, Girvan M, "Finding and evaluating community structure in networks," Physical Review, vol. 69, no. 2, Aug. 2004.

[20] WW Zachary, "An information flow model for conflict and fission in small groups," Journal of Anthropological Research, vol. 33, no. 4, pp. 452-473, 1977.

[21] V.Krebs, unpublished, Available: http://orgnet.com.

[22] Guimerà R, Danon L, Díaz-Guilera A, Giralt F and Arenas A, "Self-similar community structure in a network of human interactions," Phys Rev E Stat Nonlin Soft Matter Phys, vol. 68, no. 6, pp. 93-108, Dec. 2003.

[23] Jose Devezas, Alvaro Figueira, "Finding Language-Independent Contextual Supernodes on Coreference Networks," IAENG International Journal of Computer Science, vol. 39, no. 2, pp. 200-207, 2012.