IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON URBAN COMPUTING AND INTELLIGENCE

Received March 20, 2019, accepted April 7, 2019, date of publication April 11, 2019, date of current version April 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2910602

A Fast Algorithm for Community Detection of Network Systems in Smart City

FANGYU LIU[1,2] AND GANG XIE[2,3,4,5]

[1]College of Physics and Optoelectronics, Taiyuan University of Technology, Taiyuan 030024, China
[2]College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China
[3]Shanxi Key Laboratory of Advanced Control and Intelligent Information System, Taiyuan University of Technology, Taiyuan 030024, China
[4]Engineering Research Center for Key Technologies of Flat Panel Display Intelligent Manufacturing Equipment, Taiyuan University of Technology, Taiyuan 030024, China
[5]School of Electronic and Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China

Corresponding author: Gang Xie (xiegang@tyut.edu.cn)

This work was supported in part by the Key Research and Development Plan of Shanxi Province under Grant 201703D111027, in part by the Shanxi International Cooperation Project under Grant 201803D421039, and in part by the National Natural Science Foundation of China-Shanxi Coal-Based Low-Carbon Joint Fund under Grant U1810121.

ABSTRACT In this paper, a novel algorithm is designed to detect the community structure of network systems in the smart city based on the biogeography-based optimization (BBO) algorithm and the Newman, Moore, and Watts (NMW) small-world network. We have incorporated the NMW small-world network to the BBO algorithm to enhance the ability of migration of the habitat by using the connection mechanism of the NMW small-world network. With the help of small-world network information sharing, the convergence speed of the BBO algorithm has significantly improved. The first step of the algorithm design is to generate an NMW small-world network containing nodes equal to the number of habitats with good connectivity, which facilitates better information exchange between the nodes. In the second step, the habitat in the BBO algorithm is dynamically assigned to the small world network, and then, the BBO algorithm migrates and mutates according to the connection relationship of the NMW small-world network. Finally, the new designed NMW-BBO algorithm is evaluated for community detection via four real networks and computer-generated networks, and one of them is exhibited the characteristics of a large network. The numeric simulations are also employed to demonstrate that the new algorithm exhibits better accuracy and robustness.

INDEX TERMS BBO algorithm, NMW small world network, smart city, complex network, community detection.

I. INTRODUCTION

The complex topological properties of network systems in smart city can affect the faults spreading behavior. The identification of critical nodes in the network through community detection algorithm, and protecting these nodes can effectively suppress the spreading of large-scale faults in the network [1], [2], such as cascading failure in a smart grid. More efficient district topological properties indicate stronger fault resistance ability [3]. Therefore, studies on community detection can have significant importance on securing smart cities [4]. Last decade has witnessed a substantial increase in scientific research focusing the vast areas of complex networks. Complex networks have been extensively studied and successfully applied in many sub-disciplines such as electricity grids, neural networks, social networks and internet [6]. And it was found that the distribution of edges is not random, but there are some intrinsic rules for degree distribution in complex networks. The most significant rule exists commonly in networks is "community structure" [7], which means that nodes in the same community are more closely connected, while nodes between the communities are relatively less closely connected.

The purpose of community detection is to find an appropriate method to expose the inner connections that exist in the network. After the detection of network community structure, it becomes easier to precisely elucidate the deeper structure, features and functions of the network. Due to the importance of the network community structure, the extensive and laborious studies have been conducted to

The associate editor coordinating the review of this manuscript and approving it for publication was Rongbo Zhu.

propose a large number of effective algorithms, including the hierarchical clustering algorithms [8], graph partitioning algorithms [11], [13], the heuristic algorithms [14] and the modularity-based algorithms [10].

By considering the effectiveness of heuristic algorithms and the modularity-based algorithms, a novel algorithm is proposed and designed in this study by using a combination of the BBO algorithm and NMW small world network. The traditional heuristic algorithm follows some evolution rules to solve the community detection problems. Although those optimization methods can obtain the optimal solution, it does not maximize the use of the information interaction between different individuals and makes the convergence time of the algorithm longer with poor robustness [18]. In this study, we propose a new community detection algorithm that combines a small world network with a heuristic algorithm (BBO algorithm). The new algorithm can take advantage of the connectivity of small world networks to accelerate information sharing. Compared with general community recognition algorithms, the new algorithm has higher accuracy and faster speed. And because of the nature of the small world itself, the new algorithm is more suitable for large networks.

After precise introduction, this paper is further organized into four sections as follows: In Section 2, some preliminaries are introduced to highlight the theoretical basis of newly designed algorithm. Section 3 presents the insights to the NMW small world network and BBO algorithm, including how to integrate and apply these algorithms for network community detection. Section 4 reveals application of our novel algorithm on several real networks and a computer-generated network. Finally, Section 5 includes the concluding remarks.

II. PRELIMINARIES

The theoretical basis of this work includes graph theory, community structure and modularity. A brief view of these important terms is given below:

$G = (V, E)$ is usually used to represent an undirected graph consisting of n nodes and the edges connecting these nodes. Then, the set of nodes can be represented as $V = \{v_1, v_2, v_3, \ldots, v_n\}$, and the set of edges can be represented as $E \subseteq V \times V$. $e_{ij} = (v_i, v_j)$, that denotes an edge between node v_i and node v_j. The adjacency matrix A is used to represent the connection relationship of nodes in the network, i.e. if there is an edge connection between nodes i and j, and the value of A_{ij} represents the weight of edge e_{ij}. For a network with a weight of 1, its adjacency matrix can be expressed by the following equation (1):

$$A_{ij} = \begin{cases} 1 & e_{ij} \in E \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Additionally, all the elements on the main diagonal of matrix A are zero $A_{ij} = 0$.

In our algorithm, we determined whether the algorithm converged via calculating the tightness of nodes between communities. Therefore, a function that accurately described the tightness between communities was a prerequisite for ensuring the accuracy of the algorithm. Importantly, it is established that the modularity function Q was an excellent evaluation function to measure the strength of community structure, i.e. a larger value of modularity Q implied a stronger community structure [17]. For a network that divided into k communities, its modularity was defined by the following equation (2):

$$Q = \sum_{s=1}^{K} (q_s - a_s^2) \tag{2}$$

where q_s represented the fraction of edges, which connected the nodes in community S, and a_s was fraction of the ends of edges joined to the nodes of community S.

The positive modularity value suggests that there are more edges in the community, and the network community structure is more obvious. Otherwise, the negative value indicates that there are fewer inner edges. The simplified form of equation (2) is shown below:

$$Q = \sum_{i=1}^{K} [\frac{l_i}{M} - (\frac{d_i}{2M})^2] \tag{3}$$

where the l_i represents the sum of weights of the inner edges of the i th community. Usually, the weight of the edge is 1. At this situation l_i is also the total number of the inner edges of the community i. d_i expresses the sum of all weights of membership in the i th community. M and K are respectively represents the edges number and the community number in the network.

In addition to the modularity, which can detect the quality of community division, the normalized mutual information (NMI) [20] and the D-function comparison method [21] can measure the standard of network division. However, these two measurement methods need to know the real partition of the network.

III. NMW-BBO ALGORITHM FOR COMMUNITY DETECTION

This section highlights the composition and integration of NMW (Newman, Moore, & Watts) small world network and BBO algorithm to devise a novel community detection algorithm, by following the principles related to the small world network that have been described in the previous studies [23]. The NMW small world network is obtained by randomly adding the side on the basis of the WS (Watts & Strogatz) [25] small world network. In this method, the number of nodes of the NMW small world network is equal to the number of habitats in the BBO algorithm. Then, the NMW small world network is subjected to modification that dynamically limits the migration and mutation operation of the BBO algorithm. Following are several important aspects to design a new algorithm.

A. BUILDING NMW SMALL COMMUNITY NETWORK

The NMW model was proposed by Newman, Moore, & Watts [23]. They added edges to the WS small world network to create NMW small world network. An NMW small world network $G = C(n, k)$ was obtained by randomly increasing the connected edges at a probability P on a ring with n number of nodes. Here, n is the number of nodes present in the network and k is the number of neighbors to the left and right of the nodes. In this way, the number of average edges of the NMW small world network model is $nk + nkP$. Newman et al. derived the characteristic length of the network as follows [23].

$$l = \frac{n}{k} f(nkP) \tag{4}$$

In equation (4), the function $f(x)$ is defined as follows.

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1} \frac{x}{\sqrt{x^2 + 2x}} \tag{5}$$

In our algorithm, we needed to create a NMW small world network $G = C(n, k)$ with n nodes, and each node exhibited a fixed number of k left and right neighbors. In order to display the NMW small world network, we generated a NMW small world network $G = C(20, 4)$, that contained 20 nodes and the minimum number of edges of each node was kept as 4. At the probability $P = 0.5$, the topology of the NMW network corresponds to the following Fig.1.

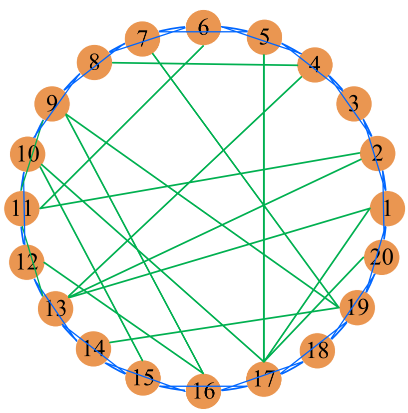


FIGURE 1. Topological diagram of the NMW small world network model.

In Fig. 1, the NMW small world network model was obtained by connecting edges on the basis of a ring network. Consequently, the connectivity of the network was relatively strong. The blue edges in Fig. 1 represents the edges of the initial network and green edges mean the increased edges with probability P. The diameter of the network is related to k and P. Generally, larger k and P values close to 1 are selected to generate a small world network. Therefore, the nodes in the network can be connected directly or through a few intermediate points.

B. PRELIMINARY BBO ALGORITHM

The BBO algorithm is a heuristic optimization algorithm, which is a bionic intelligent optimization algorithm [26] that imitates the migration of species in habitats. It is similar to other heuristic algorithms, such as Genetic Algorithm (GA) [26], the Ant Colony Optimization (ACO) algorithm based on species genetic mechanism [27] algorithm based on ant foraging, and the Particle Swarm Optimization (PSO) algorithm based on bird flight [29]. Compared with GA algorithm, BBO algorithm has more diversity and individual updating mechanism, so it has better performance than GA algorithm.

The biological model of BBO algorithm originates from a group [34] of isolated islands (habitats), in which species can migrate between habitats in search of better living environment [31]. A set of Suitability Index Variables (SIVs) is used to characterize each island, including habitat characteristics SIVs, such as rainfall, geographical conditions, climate, and sunlight. Specifically, a matrix $X\left(X_{ij} \in X | 1 \leq i \leq H,\ 1 \leq j \leq V\right)$ is used to represent a set of habitats. H and V are respectively stand for the number of habitants and the SIV number in habitants. The Habitat Suitability Index (HSI), an index directly associated with SIVs, was introduced to describe the habitat suitable for species to survive. The number of species inhabited in a habitat was closely related to HSI value. Generally, the islands with more species tend to have larger HSI values, while the low HSI values corresponded to fewer island species. However, the island species with larger HSI values were under greater pressure. So, they were more inclined to migrate to islands with the lower HSI values. As a result, the diversity of habitat species and the HSI value increased. However, if the migration of species did not lead to increase the HSI, the migrated species tend to extinct or migrate to other habitats.

C. BBO ALGORITHM COMBINED WITH DYNAMIC NMW MODEL

The convergence of the BBO algorithm could be improved by using the connectivity of NMW small world network to accelerate the information transmission of BBO algorithm. This study seeks a dynamic combination method for the NMW model and the BBO algorithm, which is named as the NMW-BBO algorithm. The process of dynamic integration of the NMW model into the BBO algorithm includes various steps, such as initialization, migration, and mutation of the algorithm, which are discussed in the following section.

1) NMW-BBO INITIALIZATION

When the NMW-BBO algorithm is initialized, the number of habitats should be equal to the number of nodes in the NMW small world network. Thus, each habitat in the NMW-BBO algorithm can be projected onto a node in the NMW network one by one. In this way, the initialized habitats are represented by matrix X, in which the elements are represented by $X_{ij}\,(1 \leq i \leq H,\ 1 \leq j \leq V)$. Here, i and j respectively represent the i th habitat and the j th dimensional SIV value. In addition, H represents the number of habitats, and V means the number of SIVs that a habitat contains. According to the previous description, the number of nodes contained in the

NMW small world network model is H. Then, the NMW small world network is shown as $G = C\,(H, k)$.

2) DYNAMIC MIGRATION STRATEGY OF NMW-BBO ALGORITHM

After the NMW-BBO algorithm is initialized, a node position in the NMW small world network dynamically allocated for each habitat. Moreover, the allocation rule is to find a node in the NMW small world network model for each habitat according to the fitness function of the habitat prior to the dynamic migration. Afterwards, it is necessary to determine whether the migration operation is performed among the habitats according to the connection relationship of the nodes in the NMW small world network. Since habitats are mapped to NMW small world network, the migration strategy needs to be based on the constraints of this NMW network model. At the same time, the immigration rate λ and number of species S have the opposite trends. While, there is a positive correlation between emigration μ and species number S. The relationship between them satisfies the following two equations (6) and (7)

$$\lambda = I_{max}(1 - \frac{S}{S\max}) \tag{6}$$

$$\mu = \frac{S}{S\max} E\max \tag{7}$$

where I_{max} and $E\max$ represent the maximum immigration rate and the maximum emigration rate respectively; and $S\max$ is the maximum number of species that a habitat can survive. Note that the equations (6) and (7) are just one of the ways to calculate λ and μ. More similar methods can be found in the previous study [30].

The migration condition of NMW-BBO algorithm is not only related to the migration rate, but also to the topology of NMW small world network. The migration of two habitats should not only meet the restriction of immigration and emigration, but also be connected by two nodes in the NMW small world network. This constraining relationship indicates that data exchange between habitats is restricted, but in fact, it uses the connectivity of small-world network to terminate unnecessary migration between habitats without affecting the final result. The direct result of this is to accelerate the convergence of the algorithm, which has advantages in large-scale network community detection. In addition, it is not easy to get into the local optimal solution, because the unnecessary intermediate process is removed.

To understand the dynamic migration process of NMW-BBO, we selected an example to show the process. After habitat initialization, it was sorted out correspond to a node in the NMW small world network. Therefore, the correspondence relationship before dynamic migration was shown in the following Fig. 2, when $G = C\,(20, 4)$ and $P = 0.5$. Here, the number of nodes in the small world network is 20, which is equal to the number of habitats. The number of nodes in a small-world network may not be 20, as long as it is consistent with the number of habitats in the algorithm. Probability p represents the probability of random increase of edge in small-world network. If p = 1, there are edge connections between all nodes and migration operations can be carried out between any two corresponding habitats. In this case, small-world network is ineffective. When p = 0, it means that the network node of the small world only has edges with the neighbor node, and the network connectivity is the worst. Generally, when p = 0.5, network connectivity is very good. Therefore, p = 0.5 is selected in this paper.

In Fig. 2, the corresponding habitat is $X = (X_1, X_2, X_3, \ldots, X_{20})^T$, and each habitat has already correspondence to a node in the network. Before assigning those nodes to habitats, the habitats are sorted out according to the fitness function value. And then map the sorted habitats to the corresponding NMW small world topological nodes. For example, the nodes 2, 6, 9, 10, 11, 12, 13 are related to habitats X_2, X_6, X_9, X_{10}, X_{11}, X_{12}, X_{13} after sorting based on fitness function. At the same time, the connection relationship of nodes affects the relationship between habitats. For example, node 11 is connected to nodes 2, 6, 9, 10, 12 and 13, which is represented by thick pink edges in Fig. 2. As a result, the habitat X_{11} can only migrate with habitats X_2, X_6, X_9, X_{10}, X_{12} and X_{13}.

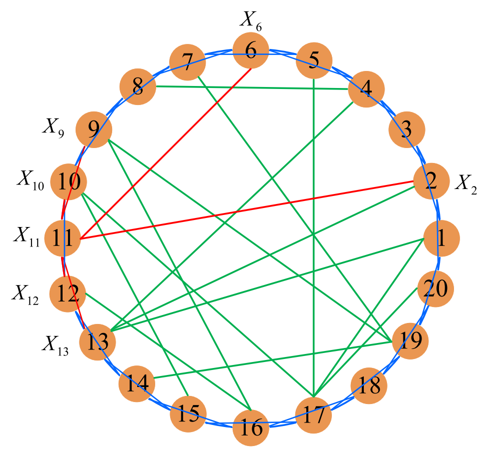


FIGURE 2. The congruent relationship between habitat and the NMW small world network.

3) NMW-BBO MUTATION

Mutation is a dramatic change in the habitat, which changes environmental conditions of the habitat. In natural environment, the habitat mutation occurs as a result of drastic changes in rainfall, temperature, humidity, and geochemical reactions. The mutation rate is used to describe the speed dramatic change in habitat. Moreover, the mutation rate is inversely proportional to the species existence probability. If the habitat species has a higher survival rate Ps, the mutation probability will be lower and vice versa. In general, the habitats with higher HSI values have lower mutation rate and vice versa.

The relationship between the mutation rate M and the species existence probability Ps is shown in the follow equation (8):

$$M(X_i) = M_{max}(\frac{1 - P_s}{P_{max}}) \tag{8}$$

where $M(X_i)$ is the mutation rate of the i habitats; M_{max} is the maximum mutation rate determined by the actual optimization problem; P_{max} is the maximum survival rate.

D. NMW-BBO ALGORITHM FOR COMMUNITY DETECTION

The key of NMW-BBO algorithm for network community detection is to integrate network community partition information into the algorithm during initialization. Then, modularity function Q is used as the fitness function of NMW-BBO algorithm. In this study, we fused the community index of the nodes in the network into the individual element of the NMW-BBO. For example, the i th habitat $X_i = [x_{i1}, x_{i2}, x_{i3}, x_{i4}]$ represents the community detection result of 4 nodes complex network. When $x_{i1} = x_{i2}$, the nodes 1 and 2 reside in the same community. For instance, the initialized habitats X can be expressed by the following matrix.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1V} \\ x_{21} & x_{22} & \cdots & x_{2V} \\ \cdots & \cdots & \cdots\cdot & \cdots \\ x_{H1} & x_{H2} & \cdots & x_{HV} \end{bmatrix} \tag{9}$$

where H represents the number of habitats and V is the SIV number in a habitat. The element x_{ij} in the initialized matrix X represents the community index of the j th node in the i th community detection result. In particular, this initialization method is consistent with the initialization method previously used by Pizzuti *et al.* [36]. In the network $G = (V, E)$, the number of nodes in the network is $n = |V|$, and the number of edges is $m = |E|$. After initialization, the number of nodes in the network is equal to the matrix dimension $n = V$.

The methods to combine the NMW-BBO migration and mutation strategy have been described earlier. The NMW-BBO algorithm flow was obtained through initializing the relevant parameters, as shown in the following Table 1.

TABLE 1. The NMW-BBO algorithm flow.

Algorithm 1 NMW-BBO algorithm
Input： habitats $X = X_1, X_2, X_3, \ldots, X_H{}^T$, each habitat $X_i = [X_{i1}, X_{i2}, \ldots, X_{iV}]$
Output： The best optimization result
1：Habitat initialization, calculate the fitness function.
% Migration operation
2： For habitat $X_{i,}\ i \in 1, H$
3： For SIV value $X_{ij,}\ j \in 1, V$
4： If (Select X_i with probability λ_i)
5： {Select X_k with probability μ_k, replace the corresponding variable $X_{ij} \leftarrow X_{kj}$}
6： End if
7： End for
8： End for
% Mutation operation
9： For habitat $X_{i,}\ i \in 1, H$
10： For SIV value $X_{ij,}\ j \in 1, V$
11： If (meet mutation requirements)
12： {replace the SIV values of the original habitat with randomly generate new SIV}
13： End if
14： End for
15： End for
16： Determine whether the stop condition is satisfied, otherwise continue to step 2.
17：End the optimization and output the result

IV. EVALUATING THE PERFORMANCE OF THE NMW-BBO ALGORITHM

In this study, four real networks (Zachary's karate club, Dolphin sociality, American college football and a large network The Protein-protein Interaction Network in Budding Yeast (PINBY)) and a computer-generated network were employed to evaluate the community detection performance of the newly designed algorithm. Moreover, the new algorithm was compared with traditional community detection algorithms, such as Fast Newman [10], Label Propagation Algorithm (LPA) [33] and BBO algorith6. Here, we initialized some important parameters before proceeding with the test, such as the maximum mutation rate $M_{max} = 0.05$, the maximum immigration rate $I_{max} = 1$, emigration rate $E_{max} = 1$ and the number of habitat $H = 50$. The corresponding NMW Small World Network was $G = C\,(50, 2)$ at the selected probability $P = 0.5$. Modularity function Q was chosen as the fitness function of the NMW-BBO algorithm. In order to further improve the accuracy of our test, the normalized mutual information (NMI) [35] was employed to measure, if the proportion of the nodes were correctly detected. The range of NMI was between 0 and 1. The value of NMI = 1 expressed that the result of the division was exactly the same as the real result, otherwise the value of NMI < 1.

A. COMPUTER-GENERATED BENCHMARK NETWORKS

The computer-generated networks are artificial virtual networks, often used to test the community detection algorithms. Girvan and Newman's [10] planted l-partition model was employed in this study to further evaluate the performance of the NMW-BBO algorithm. In the computer-generated network, there are l groups and each group has g nodes. The connection probability of the same group nodes is p_{in}, whereas the connection probability of nodes between groups is p_{out}. The larger p_{in} than p_{out} revealed that the network has more obvious community characteristics. Then, we selected most widely used models in the computer-generated networks, i.e. $l = 4$ and $g = 32$, where each node was consisted of 16 edges. The number of connected edges between nodes in the same group was Z_{in}, and the number of connected edges between nodes and nodes outside the group was Z_{out}, satisfying the relationship $Z_{in} + Z_{out} = 16$.

The computer-generated network was identified by NMW-BBO, LPA and Fast Newman algorithms. Afterwards, one hundred experiments were performed on a computer and the mean NMI value was calculated. When $Z_{out} \in (1, 10)$, the results for various algorithms are plotted in Fig. 3.

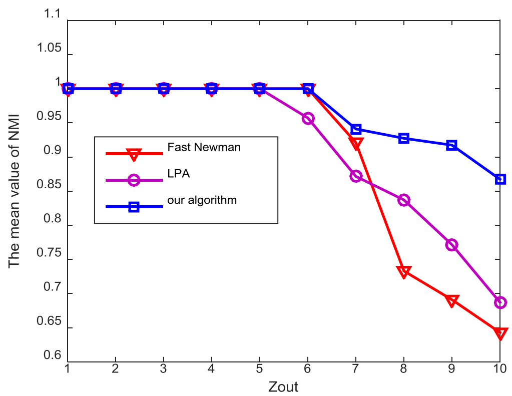


FIGURE 3. The mean value of NMI in computer-generated network when applying Fast Newman, LPA and NMW-BBO algorithms.

In Fig. 3, Fast Newman algorithm and our algorithm (NMW-BBO) have achieved very high precision when $Z_{out} \leq 6$, and the NMI value is equal to 1. In addition, the NMI value corresponding to NMW-BBO algorithm is higher than that of LPA and Fast Newman algorithm. We can see from the curve in Fig. 3 that our algorithm has higher stability. Finally, we draw the network partition topology diagram of NMW-BBO that is shown in the following Fig. 4.

Fig. 4 reveals that the nodes in the computer-generated network are completely and correctly divided into 4 communities, indicating that the algorithm test is successful.

B. THE ZACHARY'S KARATE CLUB NETWORK

The Zachary's karate club network is a real network that commonly used to test the recognition ability of community detection algorithms. It was designed by an American Sociology Professor Zachary who observed the social relations among 34 members of a karate club in an American University [37]. Due to the divergence of interests of the club, the supervisor (node 34) and the coach (node 1) became central figures in both factions. In the network, the faction where node 1 located represents a group, while the node 34 locates

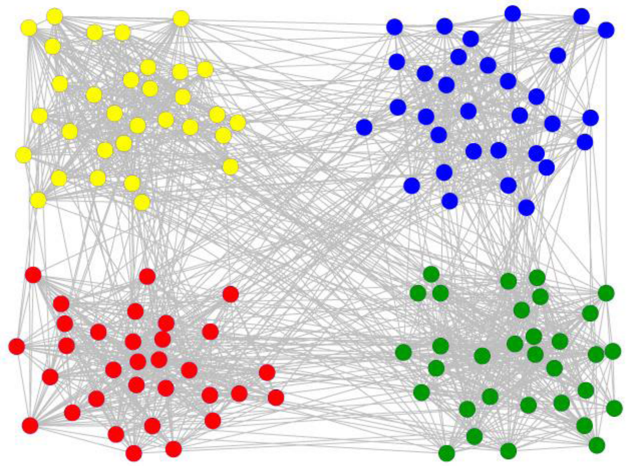


FIGURE 4. The community structure of the computer-generated networks by NMW-BBO algorithm, when $Z_{out} = 6$.

in another group. Fig. 6 shows topology of Zachary's karate club network.

Because the Zachary's karate club network has 34 nodes, the generated habitat matrix is $X\left(X_{ij} \in X | 1 \leq i \leq 50, 1 \leq j \leq 34\right)$. Therefore, we used newly designed NMW-BBO algorithm to identify the communities in the network, and drew the convergence curve of the fitness function in the algorithm optimization process as showed in Fig. 5.

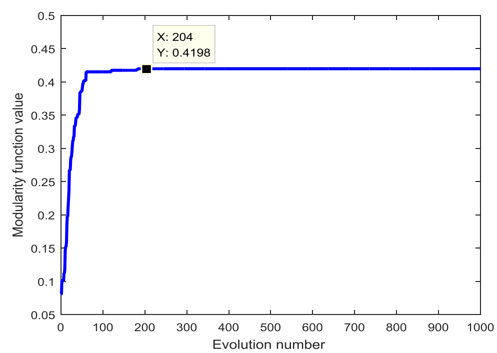


FIGURE 5. The modularity function convergence curve of the Zachary's karate club network with NMW-BBO algorithm.

Fig. 5 clearly shows that the NMW-BBO algorithm can effectively identify the Zachary's karate club network. Here, after 204 evolutions, the identified network has the highest modularity value of 0.4198, which is the same with the highest modularity function value of 0.4198 reported in the previous study [38]. Further, the NMW-BBO algorithm converges very quickly, and the actual simulation observation also shows that the algorithm converges faster. In order to study the topology of the network, the simulation results of Fig. 5 are selected, and the network topology is drawn in Fig. 6.

Fig. 6 showed that the network was divided into four communities when the modularity reached to its maximum

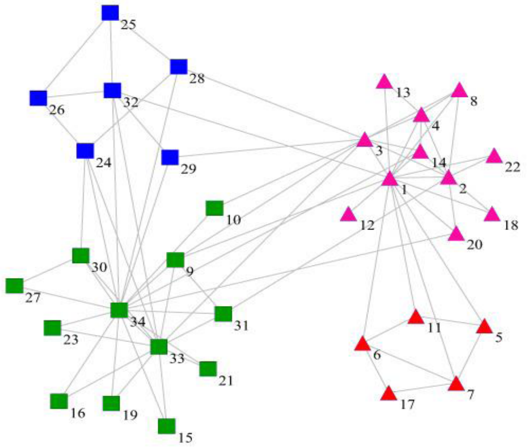


FIGURE 6. The community detection result of the Zachary's karate club network by the NMW-BBO algorithm.

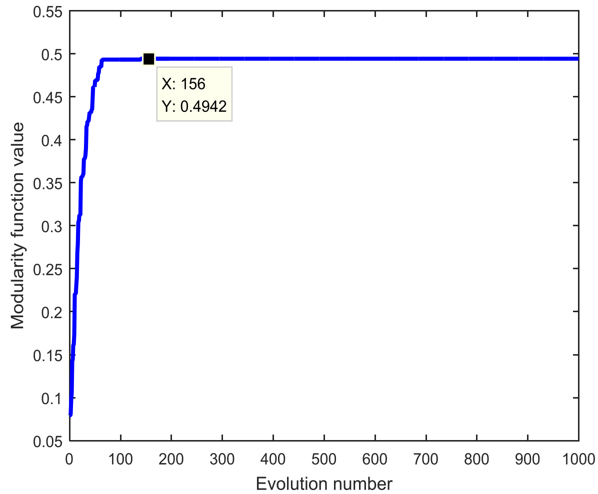


FIGURE 7. The modularity function convergence curve of the dolphin social network with NMW-BBO algorithm.

value. The network was divided into 4 groups represented by 4 medium colors and 2 large groups denoted by squares and triangles. The two large communities were consistent with the general knowledge, indicating that our algorithm exhibited the high accuracy.

C. THE DOLPHIN SOCIALITY NETWORK

The dolphin social network is another benchmark network for testing the community detection algorithms, which was constructed from the social relationship of 62 bottlenose dolphins in New Zealand by Lusseau *et al.* [41]. Due to the dolphins foraging, the Dolphin network was divided into two large groups, one containing 42 bottlenose dolphins and the other containing 20 bottlenose dolphins. To detect the community structure of dolphin social network by NMW-BBO algorithm, we generated a habitat matrix $X\left(X_{ij} \in X | 1 \leq i \leq 50,\ 1 \leq j \leq 62\right)$. Then, the modularity function convergence curve of the network was obtained.

When the algorithm converges, the topology structure of the divided network is shown in Fig. 8.

It can be seen from Fig. 8 that the dolphin social network is divided into 3 communities, which are denoted with 3 different colors. The green and red groups represent a large community, which is consistent with the actual network division. The sub community contains the nodes 4, 5, 9, 12, 16, 19, 24, 25, 30, 36, 40, 46, 52, 56, and 60, which are represented by green color in the community (Fig. 8).

D. COLLEGE FOOTBALL NETWORK

We further evaluated of our NMW-BBO algorithm by using college football network, which was based on the American football games during year 2000. This network was designed by Girvan and Newman [36]. The network contained 115 nodes and 613 edges that implied 115 teams played 613 games. The network was divided into 12 conferences, because teams were divided into 12 groups. Games

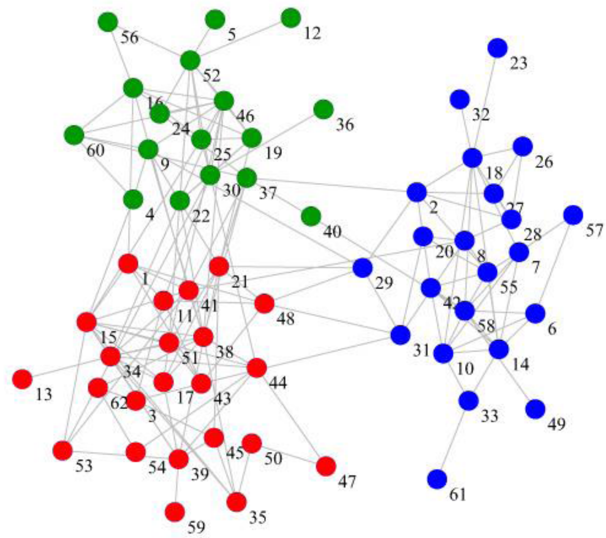


FIGURE 8. Topological structure of dolphin social network with NMW-BBO algorithm.

were played more frequently within the same group than between the members of the different groups.

Analyzing the college football network by using the NMWBBO algorithm, we firstly initialized the habitats matrix $X\left(X_{ij} \in X | 1 \leq i \leq 50,\ 1 \leq j \leq 115\right)$ of the algorithm, and then calculate the modularity function value of the network in the evolution process. The convergence curve of modularity value is shown in Fig. 9.

According to Fig. 9, we selected the network partition result when the modularity function converges and at this point the topology figure of the partition network was obtained, as shown in Fig. 10.

Our algorithm accurately identified 11 communities, which were all divided into correct communities except some nodes. Since, the number of in-group and component games

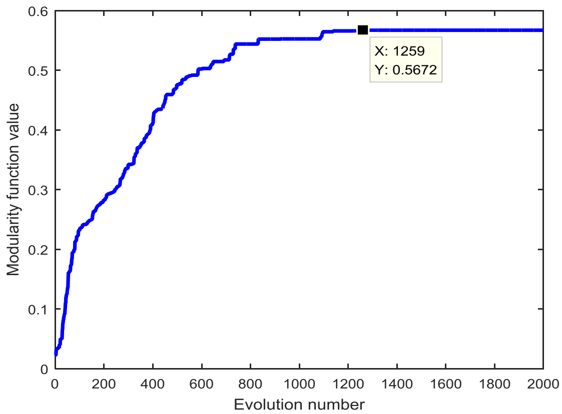


FIGURE 9. The modularity function convergence curve of the dolphin social network with NMWBBO algorithm.

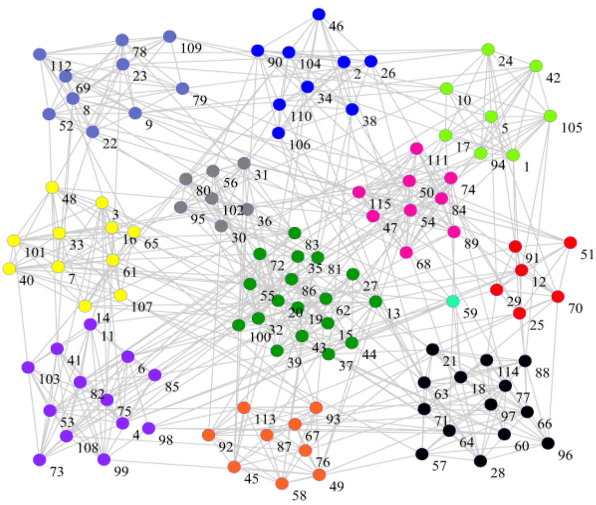


FIGURE 10. Topological structure of college football network extracted by the NMW-BBO algorithm.

are closes, these points are not clearly divided into any one community. For these points, other algorithms were also difficult to partition accurately, as reported previously [40].

E. THE PINBY NETWORK

Finally, we used a large real-world network PINBY to test the performance of the NMW-BBO algorithm. The PINBY network is a relatively complex network of protein interactions, which contains 2361 nodes and 7182 edges [41]. The PINBY is a single-cell eukaryotic protein that is widely used in the field of biological research. And it is mainly used to study the structure and function of the DNA. Many of the protein-protein interactions are hidden in the network of protein-protein interactions and one-third of the protein functions are not shown. However, PINBY's topological structure can be used to predict its functions.

By applying NMW-BBO algorithm to PINBY network, we chose the initialized habitats matrix X $\left(X_{ij} \in X | 1 \leq i \leq 50, 1 \leq j \leq 2361\right)$. The number of algorithm evolutions has been increased from 1000 times to 10000 times. The convergence curve of network modularity value is shown in Fig. 11 and the final network topology obtained by our algorithm is shown in Fig. 12.

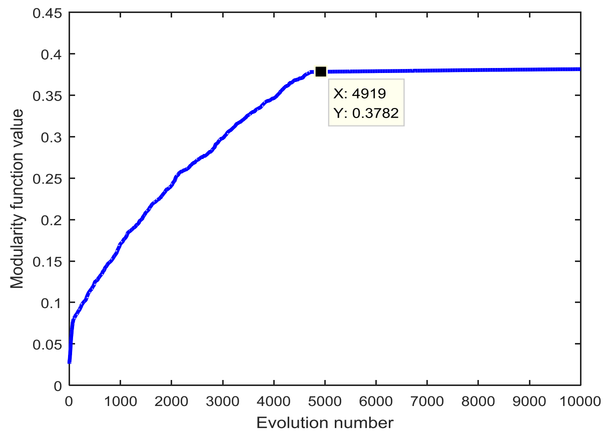


FIGURE 11. The modularity function convergence curve of the PINBY network with NMWBBO algorithm.

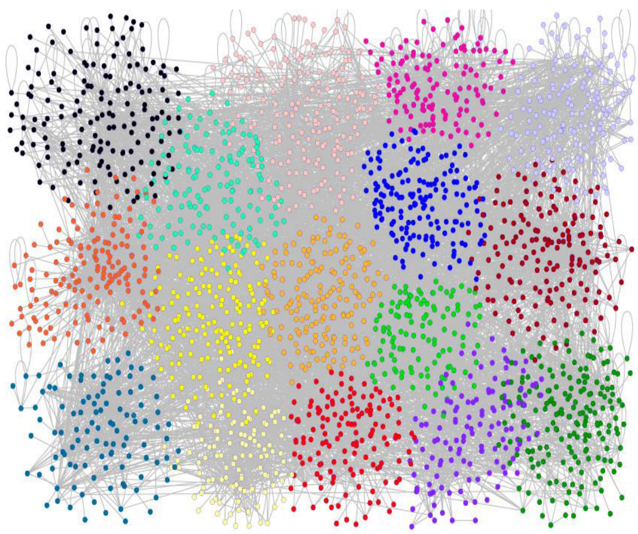


FIGURE 12. Topological structure of the PINBY network with NMWBBO algorithm.

By using our algorithm, we detected 16 communities that represented by 16 colors in Fig. 12, and each community contained more than 100 nodes. Since protein is a macromolecular network, its sub modules contain a large number of nodes, making the community relatively large.

At the same time, our algorithm found that the largest community contained 516 nodes, which was very close to the results of protein research in the previous study [44]. In addition, the network modularity value in Fig. 11 reaches 0.3782, indicating the network has strong community structure.

V. CONCLUSIONS

This study presents a novel community detection algorithm based on the BBO algorithm and the NMW small word network. We associated the NMW-BBO algorithm with the community structure. Moreover, the modularity function is selected as the standard to measure the detection of community. Thus, an individual in the NMW-BBO algorithm can represent a community detection result. Several real networks were used to test the performance of the algorithm and the results showed our algorithm has high effectiveness and precision for community detection is worthy to mention here that our algorithm has the ability of network stratification because it selects modularity function as its fitness function. This could be attributed to the maximum value of the network modularity, which may didn't necessarily express the result of the real partition, but the subdivision on the real network. Therefore, our NMW-BBO algorithm has potentials to detect hierarchanaical networks and can employ for multiple applications in urban computing, such as urban traffic network and power network [45], [46]. But NMW-BBO algorithm is a kind of heuristic algorithm, and its fitness function selects the modularity function. Because the modularity function cannot measure the community characteristics of all networks, that is, the high modularity function of the network does not necessarily mean that the community characteristics of the network are the best. Therefore, the new algorithm also has some limitations, and the later research should find a more suitable community characteristics measurement function as the algorithm's fitness function.

REFERENCES

[1] Y. Yang, G. Xie, and J. Xie, "Mining important nodes in directed weighted complex networks," *Discrete Dyn. Nature Soc.*, vol. 2017, Mar. 2017, Art. no. 9741824.

[2] Y. Yang and G. Xie, "Efficient identification of node importance in social networks," *Inf. Process. Manage.*, vol. 52, no. 5, pp. 911–922, 2016.

[3] S. Hong, B. Wang, X. Ma, J. Wang, and T. Zhao, "Failure cascade in interdependent network with traffic loads," *J. Phys. A, Math. Theor.*, vol. 48, no. 48, 2015, Art. no. 485101.

[4] H. Samani and R. Zhu, "Robotic automated external defibrillator ambulance for emergency medical service in smart cities," *IEEE Access*, vol. 4, pp. 268–283, 2016.

[5] L. Cui, G. Xie, Y. Qu, L. Gao, and Y. Yang, "Security and privacy in smart cities: Challenges and opportunities," *IEEE Access*, vol. 6, pp. 46134–46145, 2018.

[6] M. Boguñá, *The Structure and Dynamics of Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2011.

[7] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[8] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, p. 47, 2001.

[9] Y. Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[10] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 2, Art. no. 026113, 2004.

[11] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 70, Dec. 2004, Art. no. 066111.

[12] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, Jun. 2003, Art. no. 066133.

[13] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: W. H. Freeman and Company, 1979.

[14] B. W. Kernighan and S. Lin, "An efficient heuristic procedure for partitioning graphs," *Bell Syst. Tech. J.*, vol. 49, no. 2, pp. 291–307, Feb. 1970.

[15] E. R. Barnes, "An algorithm for partitioning the nodes of a graph," in *Proc. IEEE Conf. Decis. Control Including Symp. Adapt. Processes*, Dec. 1981, pp. 303–304.

[16] D.-Q. Zhou, X. Wang, S.-Y. Cheng, and Y. Chen, "Community detection algorithm via discrete PSO," *Syst. Eng. Electron.*, vol. 38, no. 2, pp. 428–433, 2016.

[17] Z. Feng, D. Chen, D.-S. Huang, R. Lu, and X. Wang, "Inverse modelling-based multi-objective evolutionary algorithm with decomposition for community detection in complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 513, pp. 662–674, Jan. 2019.

[18] J. Xiao, Y.-J. Zhang, and X.-K. Xu, "Convergence improvement of differential evolution for community detection in complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 503, pp. 762–779, Aug. 2018.

[19] H. DuMarcus W. Feldman, S. Li, and X. Jin, "An algorithm for detecting community structure of social networks based on prior knowledge and modularity," *Complexity*, vol. 12, no. 3, pp. 53–60, 2010.

[20] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 12, pp. 7821–7826, Apr. 2002.

[21] P. Zhang, M. Li, J. Wu, Z. Di, and Y. Fan, "The analysis and dissimilarity comparison of community structure," *Phys. A, Stat. Mech. Appl.*, vol. 367, pp. 577–585, Jul. 2006.

[22] B. Wu, X. Yan, Y. Wang, and C. G. Soares, "An evidential reasoning-based CREAM to human reliability analysis in maritime accident process," *Risk Anal.*, vol. 37, no. 10, pp. 1936–1957, 2017.

[23] D. H. Wolper and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.

[24] M. E. J. Newman and G. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 2, 2004, Art. no. 026113.

[25] M. E. J. Newman, C. Moore, and D. J. Watts, "Mean-field solution of the small-world network model," *Phys. Rev. Lett.*, vol. 84, no. 14, pp. 3201–3204, 2000.

[26] R. Zhu, X. Zhang, X. Liu, W. Shu, T. Mao, and B. Jalaian, "ERDT: Energy-efficient reliable decision transmission for cooperative spectrum sensing in industrial IoT," *IEEE Access*, vol. 3, pp. 2366–2378, 2015.

[27] R. Olfati-Saber, "Ultrafast consensus in small-world networks," in *Proc. Amer. Control Conf.*, 2005, pp. 2371–378.

[28] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *Nature*, vol. 393, pp. 440–442, Jun. 1998.

[29] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA, USA: Addison-Wesley, 1989.

[30] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Comput. Intell. Mag.*, vol. 1, no. 4, pp. 28–39, Nov. 2006.

[31] M. Dorigo and C. Blum, "Ant colony optimization theory: A survey," *Theor. Comput. Sci.*, vol. 344, nos. 2–3, pp. 243–278, 2005.

[32] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," in *Proc. Int. Conf. Neural Netw. (ICNN)*, Nov./Dec. 1995, pp. 1942–1948.

[33] H. Marouani and Y. Fouad, "Particle swarm optimization performance for fitting of Lévy noise data," *Phys. A, Stat. Mech. Appl.*, vol. 514, pp. 708–714, Jan. 2019.

[34] B. Wu, L. Zong, X. Yan, and C. G. Soares, "Incorporating evidential reasoning and TOPSIS into group decision-making under uncertainty for handling ship without command," *Ocean Eng.*, vol. 164, pp. 590–603, Sep. 2018.

[35] D. Simon, "Biogeography-based optimization," *IEEE Trans. Evol. Comput.*, vol. 12, no. 6, pp. 702–713, Dec. 2008.

[36] C. Pizzuti, "GA-Net: A genetic algorithm for community detection in social networks," in *Proc. Int. Conf. Parallel Problem Solving Nature (PPSN)*, Dortmund, Germany, Sep. 2008, pp. 1081–1090.

[37] J. Xie, B. K. Szymanski. (2011). "Community detection using a neighborhood strength driven Label propagation algorithm." [Online]. Available: https://arxiv.org/abs/1105.3264

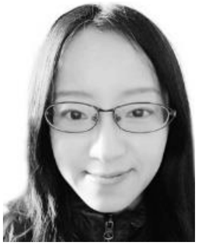
[38] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *J. Mach. Learn. Res.*, vol. 11, pp. 2837–2854, Jan. 2010.

[39] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropol. Res.*, vol. 33, no. 4, pp. 452–473, 1977.

[40] L. I. Yun, G. Liu, and S.-Y. Lao, "A genetic algorithm for community detection in complex networks," *J. Central South Univ.*, vol. 20, no. 5, pp. 1269–1276, 2013.

[41] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations," *Behav. Ecol. Sociobiol.*, vol. 54, no. 4, pp. 396–405, 2003.

[42] B. Yang, H. He, and X. Hu, "Detecting community structure in networks via consensus dynamics and spatial transformation," *Phys. A, Stat. Mech. Appl.*, vol. 483, pp. 156–170, Oct. 2017.
[43] D. Bu *et al.*, "Topological structure analysis of the protein-protein interaction network in budding yeast," *Nucleic Acids Res.*, vol. 31, no. 9, pp. 2443–2450, May 2003.
[44] Z. Huang, G. Shan, J. Cheng, and J. Sun, "TRec: An efficient recommendation system for hunting passengers with deep neural networks," *Neural Comput. Appl.*, vol. 31, pp. 209–222, Jan. 2019.
[45] K. Zhu, R. Zhu, H. Nii, H. Samani, and B. Jalaeian, "PaperIO: A 3D interface towards the Internet of embedded paper-craft," *IEICE Trans. Inf. Syst.*, vol. E97.D, no. 10, pp. 2597–2605, 2014.



FANGYU LIU received the B.S. and M.S. degrees in circuits and systems and control theory from the Taiyuan University of Technology, China, in 2012 and 2003, respectively, where she is currently pursuing the Ph.D. degree in electronic science and technology. Her research interests include machine learning, complex networks, and their applications.

GANG XIE received the B.S. degree in control theory and the Ph.D. degree in circuits and systems from the Taiyuan University of Technology, China, in 1994 and 2006, respectively, where he is currently the Vice President and has been a Professor, since 2008. And he is the Director of the Shanxi Key Laboratory of Advanced Control and Intelligent Information System, Taiyuan, China. He has authored more than 100 papers and holds five invention patents. His main research interests include intelligent information processing, complex networks, and big data. He has received six provincial science and technology awards.